Families of Tuned Scale-Space Kernels *

L.M.J. Florack[1], B.M. ter Haar Romeny[1], J.J. Koenderink[2], M.A. Viergever[1]

[1] 3D Computer Vision Research Group, University Hospital, Room E.02.222,
Heidelberglaan 100, 3584 CX Utrecht, The Netherlands
[2] Dept. of Medical and Physiological Physics, University of Utrecht,
Princetonplein 5, 3584 CC Utrecht, The Netherlands

Abstract.
We propose a formalism for deriving parametrised ensembles of local neighbourhood operators on the basis of a complete family of scale-space kernels, which are apt for the measurement of a specific physical observable. The parameters are introduced in order to associate a continuum of a priori equivalent kernels with each scale-space kernel, each of which is tuned to a particular parameter value.

Ensemble averages, or other functional operations in parameter space, may provide robust information about the physical observable of interest. The approach gives a possible handle on incorporating multi-valuedness (transparancy) and visual coherence into a single model.

We consider the case of velocity tuning to illustrate the method. The emphasis, however, is on the formalism, which is more generally applicable.

1 Introduction

The problem of finding a robust operational scheme for determining an image's differential structure is intimately related to the concept of *resolution* or *scale*. The concept of resolution has been given a well-defined meaning by the introduction of a *scale-space*. This is a 1-parameter family of images, derived from a given image by convolution with a gaussian kernel, which defines a spatial aperture for measurements carried out on the image and thus sets the "inner scale" (i.e. inverse resolution).

The gaussian emerges as the unique smooth solution from the requirement of *absence of spurious detail*, as well as some additional constraints [1, 2, 3]. Alternatively, it is uniquely fixed by the requirement of *linearity* and a set of basic symmetry assumptions, i.c. *translation*, *rotation* and *scale invariance* [4, 5]. These symmetries express the absence of a priori knowledge concerning the spatial location, orientation and scale of image features that might be of interest.

Although several fundamental problems are yet to be solved, the crucial role of resolution in any front end vision system cannot be ignored. Indeed, scale-space theory is gaining more and more appreciation in computer vision and image analysis. Neurophysiological evidence obtained from mammalian striate cortex also bears witness of its vital importance [6]. There is also psychophysical support for the gaussian model [7].

Once the role of scale in a physical observable has been appreciated and a smooth scale-space kernel has been established, the problem of finding derivatives that depend

continuously on the image (i.e. are well-posed in the sense of Hadamard), has a trivial solution [4, 5, 8, 9]: just note that if $\mathcal{D}$ is any linear differential operator, f is a given image and g_σ is the scale-space kernel on a scale σ (fairly within the available scale-range), then the convolution $f * \mathcal{D}g_\sigma$ precisely yields the derivative of f on scale σ, i.e. $\mathcal{D}(f * g_\sigma)$. The 1-parameter family containing the scaled gaussian and its linear derivatives constitutes a complete family of *scaled differential operators* or *local neighbourhood operators* [10].

Despite its completeness, however, the gaussian family is not always the most convenient one. For example, local optic flow in a time varying image can be obtained directly from the output of a gaussian family of space-time filters, at least in principle [11], but it may be more convenient to first *tune* these filters to the physical parameter of interest, i.c. a velocity vector field. This way, the filters have a more direct relation to the quantity one wishes to extract.

To illustrate the formalism we will present an example of filter tuning, i.c. *velocity tuning* [12, 13, 14, 15]. The emphasis, however, is on the formalism based on (Lie-group) *symmetries*, expressing the *a priori equivalence of parameter values*, i.c. velocities. The formalism is readily applicable to a more general class of physical tuning parameters, e.g. frequency, stereo disparity, etc.

2 Filter Tuning

Fourier's powerful method of *dimensional analysis* has received a solid mathematical formulation in the so-called *Pi theorem*. This theorem, together with the introduction of the physical tuning parameter of interest and the symmetry assumptions mentioned in the introduction, provides the main ingredients for the derivation of a family of local neighbourhood operators tuned to that observable. Basically, this theorem states that, given a physical relation $f(x_1, \ldots, x_D) = 0$, there exists an equivalent relation $\tilde{f}(\pi_1, \ldots, \pi_{D-R}) = 0$ in terms of dimensionless variables, which can be found by solving a linear equation. For more details, the reader is referred to [16].

In order to illustrate the tuning procedure, we will now turn to the case of velocity tuning. Our starting point will be a complete family of scaled spacetime differential operators $\Gamma_{\mu_1 \ldots \mu_n}(\underline{X})$:

Definition 1 (The Gaussian Spacetime Family). The gaussian spacetime family is defined as:
$$\left\{ \Gamma_{\mu_1 \ldots \mu_n}(\underline{X}) \stackrel{\text{def}}{=} \sqrt{\pi}^{-(D+1)} \partial_{\mu_1 \ldots \mu_n} \exp\left\{ -\underline{X} \cdot \underline{X} \right\} \right\}_{n=0}^{\infty} \tag{1}$$

in which $\underline{X}$ is a $(D+1)$-vector in spacetime, whose components are given by $X_0 = t/\sqrt{2\tau^2}$ and $X_i = x_i/\sqrt{2\sigma^2}$ for $i = 1 \ldots D$ and in which $\partial_{\mu_1 \ldots \mu_n}$ denotes the n-th order differential operator $\partial^n / \partial X_{\mu_1} \ldots \partial X_{\mu_n}$.

We have the following relationship between the kernels $\Gamma_{\mu_1 \ldots \mu_n}(\underline{X})$, given in dimensionless, scale-invariant coordinates, and the scale-parametrised kernels $G_{\mu_1 \ldots \mu_n}(\mathbf{x}, t; \sigma, \tau)$:

$$\sqrt{2\tau^2}^m \sqrt{2\sigma^2}^{n-m} G_{\mu_1 \ldots \mu_n}(\mathbf{x}, t; \sigma, \tau) \, d\mathbf{x} dt \stackrel{\text{def}}{=} \Gamma_{\mu_1 \ldots \mu_n}(\underline{X}) \, d\underline{X} \tag{2}$$

in which m is the number of zero-valued indices among $\mu_1 \ldots \mu_n$. Although the temporal part of (1) obviously violates temporal causality, it can be considered as a limiting case of a causal family in some precise sense [17]. Sofar for the basic, non-tuned gaussian spacetime family.

The tuning parameter of interest will be a spacetime vector $\underline{\xi} = (\xi_0; \boldsymbol{\xi})$. Apart from this, the following variables are relevant: the scale parameters σ and τ, the frequency variables $\boldsymbol{\omega}$ and ω_0 for addressing the fourier domain, and the variables $\mathcal{L}$ and $\mathcal{L}_0$, representing the scaled and original input image values in fourier space. According to the Pi theorem, we may replace these by: $\Lambda \overset{\text{def}}{=} \mathcal{L}/\mathcal{L}_0$, $\underline{\Omega} = (\Omega_0; \boldsymbol{\Omega}) \overset{\text{def}}{=} (\omega_0\sqrt{2\tau^2}; \boldsymbol{\omega}\sqrt{2\sigma^2})$, $\underline{\Xi} = (\Xi_0; \boldsymbol{\Xi}) \overset{\text{def}}{=} (\xi_0/\sqrt{2\tau^2}; \boldsymbol{\xi}/\sqrt{2\sigma^2})$. Moreover we will use the conjugate, dimensionless variables $\underline{X} = (X_0; \mathbf{X}) \overset{\text{def}}{=} (t/\sqrt{2\tau^2}; \mathbf{x}/\sqrt{2\sigma^2})$. Their dependency is expressed by $\Lambda = g(\underline{\Omega}, \underline{\Xi})$, in which g is some unknown, *scalar* function.

In the $\underline{\Xi} \to \underline{0}$ limit, the $\underline{\Xi}$-tuned kernels should converge to (1). Reversely, by applying a spacetime tuning operation to this underlying family, we may obtain a complete family of spacetime-tuned kernels and, more specifically, of velocity-tuned kernels:

Definition 2 (Spacetime Tuning). A spacetime quantity $Q(\underline{X}, \underline{\Xi})$ is called spacetime-tuned with respect to a given point $\underline{\Xi}$ if $\forall \underline{\Xi}'$ the following symmetry holds: $Q(\underline{X}, \underline{\Xi}) = Q(\underline{X} - \underline{\Xi}', \underline{\Xi} - \underline{\Xi}') \overset{\text{def}}{=} T_{\underline{\Xi}'} Q(\underline{X}, \underline{\Xi})$.

Applying the operator $T_{\underline{\Xi}}$ on a given operand, we obtain a spacetime-tuned version of it: $T_{\underline{\Xi}} Q(\underline{X}) = Q(\underline{X} - \underline{\Xi})$. In this way the gaussian family can be tuned so as to yield:

Result 3 (Family of Spacetime Tuned Kernels). *A complete family of spacetime-tuned kernels is given by:*

$$\left\{ \Gamma_{\mu_1 \ldots \mu_n}(\underline{X}, \underline{\Xi}) \overset{\text{def}}{=} \Gamma_{\mu_1 \ldots \mu_n}(\underline{X} - \underline{\Xi}) \right\}_{n=0}^{\infty} \tag{3}$$

The construction of velocity-tuned kernels from the basic gaussian family is a special case of spacetime tuning, viz. one in which the tuning point is the result of a *galilean boost* applied to the origin of a cartesian coordinate frame. This is a *transvection* in spacetime, i.e. a linear, hypervolume preserving, unit eigenvalue transformation of the following type:

Definition 4 (Galilean Boost). A galilean boost T_γ is a spacetime transvection of the type T_γ: $(\hat{e}_0; \hat{e}_i) \mapsto (\hat{e}_0 + \gamma_k \hat{e}_k; \hat{e}_i)$, in which the $\hat{e}_\mu$ constitute an orthonormal basis in spacetime and γ is an arbitrary D-vector.

T_γ transforms static stimuli (straight curves of "events" parallel to the T-axis) into dynamic stimuli, moving with a constant velocity γ. Note that T_0 is the identity and $T_\gamma^{-1} = T_{-\gamma}$ is the inverse boost. Since a galilean boost is just a special type of spacetime tuning, we immediately arrive at the following result:

Result 5 (Family of Velocity-Tuned Kernels). *A complete family of velocity-tuned kernels is obtained by applying a galilean boost* $T_\gamma = T_{\underline{\Xi}(\gamma)}$, *with* $\underline{\Xi}(\gamma) \overset{\text{def}}{=} (0; \gamma X_0)$, *to the gaussian spacetime family:*

$$\left\{ \Gamma_{\mu_1 \ldots \mu_n}(\underline{X}; \gamma) \overset{\text{def}}{=} \Gamma_{\mu_1 \ldots \mu_n}(\underline{X}, \underline{\Xi}(\gamma)) \right\}_{n=0}^{\infty} \tag{4}$$

The relevance of using a (seemingly redundant) parametrised ensemble of local operators is best illustrated by means of an example.

Example 1. Consider a point stimulus $L_0(\mathbf{x}, t) = A\delta(\mathbf{x} - \mathbf{c}t)$, moving at fixed velocity **c**. According to (4), the lowest order velocity-tuned filter is given in by:

$$G(\mathbf{x}, t; \sigma, \tau, \mathbf{v}) = \frac{1}{\sqrt{2\pi\sigma^2}^D} \frac{1}{\sqrt{2\pi\tau^2}} \exp\left\{ -\frac{(\mathbf{x} - \mathbf{v}t) \cdot (\mathbf{x} - \mathbf{v}t)}{2\sigma^2} - \frac{t^2}{2\tau^2} \right\} \tag{5}$$

in which the velocity **v** is related to the parameter vector $\boldsymbol{\gamma}$ in (4) by $\mathbf{v} \stackrel{\text{def}}{=} \boldsymbol{\gamma}\sigma/\tau$. Convolving the above input with this kernel yields the following result:

$$L_{\mathbf{v},\sigma,\tau}(\mathbf{x}, t) = \frac{A}{\sqrt{2\pi\sigma_\Delta^2}^D} \exp\left\{ -\frac{(\mathbf{x} - \mathbf{c}t)^2}{2\sigma_\Delta^2} \right\} \tag{6}$$

with $\Delta \stackrel{\text{def}}{=} \frac{\|\mathbf{c} - \mathbf{v}\|}{v_0}$ and $\sigma_\Delta \stackrel{\text{def}}{=} \sigma\sqrt{1 + \Delta^2}$. This shows an ensemble of fuzzy blobs centered at the expected location of the stimulus, the most pronounced member of which is indeed the one for which the tuning velocity coincides with the stimulus velocity ($\Delta = 0$). However, it is the velocity *ensemble* rather than any individual member that carries the information about stimulus velocity and that allows us to extract this in a way that is both robust and independent of contrast A. In order to appreciate the kind of ensemble operations one can think of, consider for example the average $< \mathbf{v} >$. A straightforward symmetry argument shows that this equals:

$$< \mathbf{v} > (\mathbf{x}, t) \stackrel{\text{def}}{=} \frac{\int d\mathbf{v}\, \mathbf{v}\, L_{\mathbf{v},\sigma,\tau}(\mathbf{x}, t)}{\int d\mathbf{v}\, L_{\mathbf{v},\sigma,\tau}(\mathbf{x}, t)} = \mathbf{c} \qquad \forall (\mathbf{x}, t) \tag{7}$$

(just note that $< \mathbf{v} - \mathbf{c} >$ vanishes identically). In our example, this average turns out to be a global constant, which, of course, is due to the uniform motion of the input stimulus. A similar, though non-trivial ensemble integral can be evaluated to obtain the variance $\Delta v = \sqrt{< \|\mathbf{v} - \mathbf{c}\|^2 >}$. It is important to realise, however, that by proceeding in this way we *enforce* a singlue-valued "cross-section" of the observable, i.c. velocity field. It is clear that this is by no means necessary: one could think of segmenting parameter space into subdomains and applying similar operations on each subdomain independently, leading to a multi-valued result (*transparency*). This is a conceivable thing to do especially if such a segmentation is apparent, as in the case of a superposition of two point stimuli for two clearly distinct values of **c**. In that case we can use (7) for each component of transparent motion by restricting the integrations to the respective velocity segments. Although an N-valued representation is certainly plausible in the limiting case when a segmentation of parameter space into N segments is "obvious", it remains an intriguing problem of how to deal with transient regions, occuring e.g. when there is an apparent jump in the velocity field of the input stimulus. Clearly, our representation entails all these cases. The problem as such is one of *pattern extraction* from the output of the tuned local neighbourhood operators in the product space of locations and tuning parameters [18].

3 Conclusion and Discussion

In this paper we have shown how the basic family of local scale-space operators may give rise to a gamut of other families, each of which is characterised, apart from scale, by some physical tuning parameter. We have presented a formalism for generating such families from the underlying gaussian scale-space family in a way that makes the a priori equivalence of all tuning parameter values manifest. We have illustrated the formalism

by an example of velocity tuning, incorporating all possible galilean boosts so as to yield ensembles of velocity-tuned local scale-space operators (cf. Reichardt operators).

We have argued that ensemble averages, or other functional operations in parameter space, rather than the output of individual kernels as such, may provide a robust, operational method for extracting valuable information about the observable of interest. The appealing aspect of this method is that it does not aim for a single-valued "expectation value" for the observable right from the beginning and that single-valuedness is treated on equal foot with multi-valuedness.

The formalism should be readily applicable to other parameters of physical interest, such as frequency, stereo disparity, etc., yielding ensembles of frequency- or disparity-tuned local neighbourhood operators, etc.

References

1. A. Witkin, "Scale space filtering," in *Proc. International Joint Conference on Artificial Intelligence*, (Karlsruhe, W. Germany), pp. 1019–1023, 1983.
2. J. J. Koenderink, "The structure of images," *Biol. Cybern.*, vol. 50, pp. 363–370, 1984.
3. T. Lindeberg, "Scale-space for discrete signals," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 12, no. 3, pp. 234–245, 1990.
4. B. M. ter Haar Romeny, L. M. J. Florack, J. J. Koenderink, and M. A. Viergever, "Scale-space: Its natural operators and differential invariants," in *International Conf. on Information Processing in Medical Imaging*, vol. 511 of *Lecture Notes in Computer Science*, (Berlin), pp. 239–255, Springer-Verlag, July 1991.
5. L. Florack, B. ter Haar Romeny, J. Koenderink, and M. Viergever, "Scale-space." Submitted to IEEE PAMI, November 1991.
6. R. A. Young, "The gaussian derivative model for machine vision: I. retinal mechanisms," *Spatial Vision*, vol. 2, no. 4, pp. 273–293, 1987.
7. P. Bijl, *Aspects of Visual Contrast Detection*. PhD thesis, University of Utrecht, University of Utrecht, Dept. of Med. Phys., Princetonplein 5, Utrecht, the Netherlands, May 1991.
8. J. J. Koenderink and A. J. van Doorn, "Representation of local geometry in the visual system," *Biol. Cybern.*, vol. 55, pp. 367–375, 1987.
9. J. J. Koenderink and A. J. Van Doorn, "Operational significance of receptive field assemblies," *Biol. Cybern.*, vol. 58, pp. 163–171, 1988.
10. J. J. Koenderink and A. J. van Doorn, "Receptive field families," *Biol. Cybern.*, vol. 63, pp. 291–298, 1990.
11. P. Werkhoven, *Visual Perception of Successive Order*. PhD thesis, University of Utrecht, University of Utrecht, Dept. of Med. Phys., Princetonplein 5, Utrecht, the Netherlands, May 1990.
12. D. J. Heeger, "Model for the extraction of image flow," *Journal of the Optical Society of America-A*, vol. 4, no. 8, pp. 1455–1471, 1987.
13. D. Heeger, "Optical flow using spatiotemporal filters," *International Journal of Computer Vision*, vol. 1, pp. 279–302, 1988.
14. E. H. Adelson and J. R. Bergen, "Spatiotemporal energy models for the perception of motion," *Journal of the Optical Society of America-A*, vol. 2, no. 2, pp. 284–299, 1985.
15. W. E. Reichardt and R. W. Schögl, "A two dimensional field theory for motion computation," *Biol. Cybern.*, vol. 60, pp. 23–35, 1988.
16. P. J. Olver, *Applications of Lie Groups to Differential Equations*, vol. 107 of *Graduate Texts in Mathematics*. Springer-Verlag, 1986.
17. J. J. Koenderink, "Scale-time," *Biol. Cybern.*, vol. 58, pp. 159–162, 1988.
18. A. J. Noest and J. J. Koenderink, "Visual coherence despite transparency or partial occlusion," *Perception*, vol. 19, p. 384, 1990. Abstract of poster presented at the ECVP 1990, Paris.