Linear Image Reconstruction from a Sparse Set of α-Scale Space Features by means of Inner Products of Sobolev Type*

Remco Duits, Bart Janssen, Frans Kanters and Luc Florack

Eindhoven University of Technology, Den Dolech 2, NL-5600 MD Eindhoven, The Netherlands,
R.Duits@tue.nl, B.J.Janssen@tue.nl, F.M.W.Kanters@tue.nl, L.M.J.Florack@tue.nl
http://www.bmi2.bmt.tue.nl/image-analysis/people/rduits

Abstract. Inner products of Sobolev type are extremely useful for image reconstruction of images from a sparse set of α-scale space features. The common (non)-linear reconstruction frameworks, follow an Euler Lagrange minimization. If the Lagrangian (prior) is a norm induced by an inner product of a Hilbert space, this Euler Lagrange minimization boils down to a simple orthogonal projection within the corresponding Hilbert space. This basic observation has been overlooked in image analysis for the cases where the Lagrangian equals a norm of Sobolev type, resulting in iterative (non-linear) numerical methods, where already an exact solution with non-iterative linear algorithm is at hand. Therefore we provide a general theory on linear image reconstructions and metameric classes of images. By applying this theory we obtain visually more attractive reconstructions than the previously proposed linear methods and we find connected curves in the metameric class of images, determined by a fixed set of linear features, with a monotonic increase of smoothness. Although the theory can be applied to any linear feature reconstruction or principle component analysis, we mainly focus on reconstructions from so-called topological features (such as top-points and grey-value flux) in scale space, obtained from geometrical observations in the deep structure of a scale space.

Keywords. Scale Space, Sobolev Spaces, Gelfand Triples, Tikhonov Regularization, Top Point Reconstruction, Deep Structure, Flux Features.

1 Introduction

In linear scale space theory one obtains a so-called α-scale space representation $u_f^\alpha : \mathbb{R}^d \times \mathbb{R}^+ \to \mathbb{R}$ of a grey value image image $f \in \mathbb{L}_2(\mathbb{R}^d)$ by means of a holomorphic semi group generated by $-(-\Delta)^\alpha$, $0 < \alpha \leq 1$. Such a scale space is obtained by means of a convolution

$$u_f^\alpha(\mathbf{x}, s) = (K_s^\alpha * f)(\mathbf{x}), s > 0, \mathbf{x} \in \mathbb{R}^d,$$

where $K_s^\alpha = \mathcal{F}^{-1}[\boldsymbol{\omega} \mapsto e^{-s\|\boldsymbol{\omega}\|^{2\alpha}}]$. These isotropic linear scale space representations follow from a list of fundamental axioms, cf.[1] and the most common cases are $\alpha = 1$ and $\alpha = \frac{1}{2}$ leading to respectively a diffusion system and $\alpha = 1/2$ a potential system on the upper space $s > 0$. In these cases the convolution kernel equals respectively the Gaussian kernel and the Poisson kernel:

$$K_s^1(\mathbf{x}) = \frac{1}{(4\pi s)^{d/2}} e^{-\frac{\|\mathbf{x}\|^2}{4s}} \qquad \text{and} \qquad K_s^{\frac{1}{2}}(\mathbf{x}) = \frac{2}{\sigma_{d+1}} \frac{s}{(s^2 + \|\mathbf{x}\|^2)^{\frac{d+1}{2}}}. \tag{1}$$

In α-scale space one can use all kinds of differential invariants to detect local structure such as corners and lines. These differential invariants $\Phi : C^\infty(\mathbb{R}^d) \to C^\infty(\mathbb{R}^d)$ are algebraic (not necessarily linear) combinations of so-called α-derivatives of $f \in \mathbb{L}_2(\mathbb{R}^d)$, given by

$$D_s^{\mathbf{n}} f := D^{\mathbf{n}}(K_s^\alpha * f) = (D^{\mathbf{n}} K_s^\alpha) * f \ , \qquad s > 0, \alpha \in (0, 1], \mathbf{n} = (n^1, \ldots, n^d) \in \mathbb{N}^d, \tag{2}$$

* The Netherlands Organization for Scientific Research is gratefully acknowledged for financial support.

such that they are Euclidean invariant, i.e. $\Phi \circ \mathcal{U}_g = \mathcal{U}_g \circ \Phi$, for all g within the Euclidean motion group $G = \mathbb{R}^d \rtimes SO(d)$, where $\mathcal{U} : G \mapsto \mathcal{B}(\mathbb{L}_2(\mathbb{R}^d))$ is given by $\mathcal{U}_g\psi(\mathbf{x}) = \psi(R^{-1}(\mathbf{x} - \mathbf{b})), g = (\mathbf{b}, R) \in G$.

Notice that the α-derivatives given by (2) are *bounded* (and thereby well-posed) operators on $\mathbb{L}_2(\mathbb{R}^d)$, which is clearly not the case for the usual derivative operators. This statement directly follows from the Plancherel Theorem and the fact that $\omega \mapsto \omega^p e^{-s|\omega|^{2\alpha}}$ is uniformly bounded on $\mathbb{R}$ for all $p \in \mathbb{N}$. In case $d = 2$ a complete set of functionally independent differential invariants up to second order is given by $\{u, u_x^2 + u_y^2, u_{xx} + u_{yy}, u_{xx}u_y^2 - 2u_{xy}u_xu_y + u_{yy}u_x^2, u_xu_yu_{xx} - u_x^2u_{xy} + u_y^2u_{xy} - u_xu_yu_{yy}\}$. By introducing Gauge coordinates w, v, where w is along the gradient of $u = u_f^\alpha$ and v orthogonal to it, i.e. $\mathbf{e}_w(\mathbf{x}) = \frac{\nabla_\mathbf{x} u}{\|\nabla_\mathbf{x} u\|}$ and $\mathbf{e}_v(\mathbf{x}) = R_{\frac{\pi}{2}}\mathbf{e}_w$, these differential invariants are simply expressed as $\{u, u_w, u_{ww} + u_{vv}, u_{vv}, u_{vw}\}$. Beyond Euclidean invariance, there is affine invariance. A differential invariant $\Phi : C^\infty(\mathbb{R}^d) \to C^\infty(\mathbb{R}^d)$ is called affine invariant if $\mathcal{U}_g \Phi = \Phi \mathcal{U}_g$, for all $g \in \mathbb{R}^d \rtimes A(d)$, where $A(d) = \{M \in GL(d) \ : \ \det M = 1\}$. In the case $d = 2$, we have that every element in $A(2)$ is a composite of a rotation, shearing and an axis-rescaling. The respective matrix representations of these linear mappings are given by

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \qquad S_c = \begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix}, \qquad L_\lambda = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix}, \tag{3}$$

$\theta \in [0, 2\pi), c \in \mathbb{R}$ and $\lambda \in \mathbb{R}$. Consequently, a differential invariant which is also invariant under (the left regular actions of) the subgroups $\{S_c \mid c \in \mathbb{R}\}$ and $\{L_\lambda \mid \lambda \in \mathbb{R}\}$ is affine invariant. The affine invariants up to second order are given by $\{u, \det H_\mathbf{x}(u), (R_{\pi/2}\nabla_\mathbf{x} u)^T H_\mathbf{x}(u) R_{\pi/2}\nabla u\}$ or equivalently in Gauge-coordinates $\{u, u_{vv}u_{ww} - u_{vw}^2, u_w^2 u_{vv}\}$. With respect to affine invariants on $u_f^\alpha(\cdot, s)$ it should be noticed that they do not correspond to affine features on the original image f, since in general for $g = (A, \mathbf{b})$ and $u_{\mathcal{U}_g f}^\alpha(\cdot, s) = \mathcal{U}_g\, u_f^\alpha(\cdot, s)$ holds for all $f \in \mathbb{L}_2(\mathbb{R}^2)$ if and only if A is orthogonal.

In this article we focus on the question:
" *Given a sparse set of linear scale space features can we obtain an approximate reconstruction of the original image f in a fast operational and linear method ?* ".

A Euclidean invariant way of obtaining features from an image f is by means of differential invariants (on u_f^α) as described above, but they are mostly non-linear with respect to f. Therefore we select points where certain differential invariants vanish and compute the n-th order jet at these points, where n equals the order of the differential invariant, which are linear features of the type given by equality (2), so-called α-derivatives. Notice to this end that this does not affect Euclidean invariance of the reconstruction algorithm we will present.

It is a well known problem that the construction of f from $u_f(\cdot, s)$ for any $s > 0$ fixed is extremely ill-posed, as it requires an inverse convolution of the low-pass filter K_s^α. By considering multiple orientations (leading to a so-called orientation score) and proper directed wavelets one can get around this problem, as is shown in [2], but here we will consider multiple scale representations rather than multiple orientation representations of images. In theory an α-scale space is analytic in (x, y, s), but again computing an image reconstruction by means of Taylor expansion of (in)finite order from a single point in scale space is highly ill-posed. Nevertheless, it *is* possible to give an approximate reconstruction of f from a sparse set of linear features obtained by means of α-derivatives, which should be considered as an interpolation rather than an inverse convolution. But first we will study the topological structure of a scale space, also known as *deep structure*, as it seems most reasonable that the features should *at least* capture the topological structure of the scale space.

2 Deep Structure

The topological structure in a scale space and in particular the change of topological structure of $u(\cdot, s)$ over $s > 0$, reflects the hierarchical structure of objects (like blobs) in an image. As the resolution increases extrema disappear until at finite scale $S > 0$ only one extremum is left, cf.[3]. Points in scale space where a saddle and extremum annihilate or points where an extremum and a saddle are created are called top-points. The set of top-points is given by

$$\{(\mathbf{x}, s) \mid (\det H_{\mathbf{x}} u(\cdot, s))(\mathbf{x}) = 0 \text{ and } (\nabla_{\mathbf{x}} u(\cdot, s))(\mathbf{x}) = \mathbf{0}\}.$$

Notice that in a top point all affine differential invariants vanish.[1] At these points the topological structure changes. Other interesting points in scale space are scale space saddles[2], these are exactly those points were $\nabla_{\mathbf{x},s} u(\mathbf{x}, s) = (\mathbf{0}, 0)$, cf. [4].For investigation on the stability of top-points we refer to Balmachnova et al.[5].

Another important geometrical quantity is the *grey-value flow* within an α scale space u_f^α of image f. This multi-scale vector field is given by

$$\mathbf{F}_\alpha[u_f^\alpha](\mathbf{x}, s) = (\mathbf{f}_s^\alpha * f)(\mathbf{x}), \tag{4}$$

where $\mathbf{f}_s^\alpha(\mathbf{x}) = \mathcal{F}^{-1}[\boldsymbol{\omega} \mapsto i \frac{1}{\|\boldsymbol{\omega}\|^{2(1-\alpha)}} \boldsymbol{\omega}\, e^{-s\|\boldsymbol{\omega}\|^{2\alpha}}](\mathbf{x})$. To this end we notice that

$$\frac{\partial}{\partial s}[u_f^\alpha] = -(-\Delta)^\alpha u_f^\alpha = \operatorname{div} \mathbf{F}_\alpha[u_f],$$

which is easily verified in the Fourier domain: $-\|\boldsymbol{\omega}\|^{2\alpha} = i\boldsymbol{\omega} \cdot i \frac{1}{\|\boldsymbol{\omega}\|^{2(1-\alpha)}} \boldsymbol{\omega}$. The grey-value flow tells us how the grey-value particles flow within the scale space representation and reveals the interaction between extremal paths in scale space. We will use this flux vector field to compute so-called flux features. For the special case of a Gaussian scale space $\alpha = 1$ the grey-value flow is obtained by means of the gradient as we have $\mathbf{F}_{\alpha=1}[u_f](\mathbf{x}, s) = \nabla_{\mathbf{x}} u_f(\mathbf{x}, s)$ and $\mathbf{f}_s^{\alpha=1} = \nabla_{\mathbf{x}} K_s^1(\mathbf{x})$. For the special case of a Poisson scale space $\alpha = \frac{1}{2}$ the grey value flow is obtained by means of the Riesz transform $\mathbf{F}_{\alpha=\frac{1}{2}}[u_f](\mathbf{x}, s) = \mathbf{R}_{\mathbf{x}} u_f(\mathbf{x}, s)$ and $\mathbf{f}_s^{\alpha=\frac{1}{2}}$ equals the vector-valued conjugate Poisson kernel: $\mathbf{f}_s^{\alpha=\frac{1}{2}}(\mathbf{x}) = \mathbf{R}_{\mathbf{x}} K_s^{1/2}(\mathbf{x}) = \frac{2}{\sigma_{d+1}} \frac{\mathbf{x}}{(s^2+\|\mathbf{x}\|^2)^{\frac{d+1}{2}}}$. By extending a scale space with its flow, one obtains a vector scale space which equals the first order jet of a Gaussian scale space if $\alpha = 1$ and which equals the monogenic scale space, cf.[6], if $\alpha = 1/2$.

3 Linear Image Reconstruction Schemes

In this section we will generalize the standard linear reconstruction scheme from the usual case where the space of images is modelled as $\mathbb{L}_2(\mathbb{R}^2)$ to the more general case of an arbitrary Hilbert space H. Later we consider the case where H is a Hilbert space of Sobolev-type.

Let $\{\tilde{\psi}_k\}_{k=1}^n$ be a set of n continuous linear functionals (features) on a Hilbert space H, which are linearly independent. Then by the Riesz representation theorem there exist unique $\{\psi_k\}_{k=1}^n$ in H such that $\langle \tilde{\psi}_k, f \rangle = (\psi_k, f)_H$, for all $f \in H$. The values $(\psi_k, f)_H$, $k = 1, \ldots, n$ are called features. Let V be the span of $\{\psi_k\}_{k=1}^n$. Two images $f, g \in H$ have the same features iff $f - g \in V^\perp$. This defines an equivalence relation on H and the equivalence classes are given by

$$[f] = \{g \in H \mid g \sim f\} = f + V^\perp \, .$$

Theorem 1. *Inside the metameric class of images, the unique element with minimal H-norm equals the orthogonal projection of f onto V:*

$$\mathcal{P}_V f = \sum_{k=1}^n (\psi^k, f) \psi_k, \tag{5}$$

[1] The other way around need not be true: $u_{vv} u_w = 0$ and $\det Hu(\cdot, s) = 0$ is equivalent to $u_w = \det Hu(\cdot, s) = 0$ or $u_{vv} = u_{vw} = 0$. In the latter case the isophote curvature and the flowline curvature vanish, i.e. the isophotes and flowlines are straight lines.

[2] As is shown in [1], there do not exist interior extrema (with respect to scale and position) in α-scale spaces.

where the reciprocal base vectors are given by $\psi^k = \sum_{l=1}^{n} G^{kl}\psi_l$*, where* $\sum_{k=1}^{n} G^{ik}G_{kj} = \delta^i_j$*, with Gramm-matrix:* $G = [G_{kj}] = [(\psi_k, \psi_j)_H]$.

Proof. By the Pythagoras theorem we have

$$\min_{g\in[f]} \|g\|^2 = \min_{g\in[f]} \|g - \mathcal{P}_V f + \mathcal{P}_V f\|^2 = \min_{g\in[f]} \|g - \mathcal{P}_V f\|^2 + \|\mathcal{P}_V f\|^2 \tag{6}$$

and this equals $\|\mathcal{P}_V f\|^2$ only in the case $g = \mathcal{P}_V f$. Finally we notice that a closed[3] linear subspace of a Hilbert space is again a Hilbert space and thereby the orthogonal projection is unique. □

The special case $H = \mathbb{L}_2(\mathbb{R}^2)$ in Theorem 1, is a standard linear reconstruction scheme in image analysis, see for example [7] and later [8], [9]. In image analysis this reconstruction theory is usually put in a more indirect Euler-Lagrange framework. For example in the work of Nielsen and Lillholm, cf. [7], [8], where it is already mentioned that the prior need not be an $\mathbb{L}_2$-norm and that there exist much better priors (in the sense that one clearly obtains visually more appealing image reconstructions), such as minimal entropy and a first order Sobolev-norm. But they were unaware of Theorem 1 and thereby they used iterative non-linear schemes to approximate the global minimum. In case of a minimum entropy based prior, which is not[4] a norm induced by an inner product, this is plausible (and probably the best you can get). However, in the case where the prior equals a norm of Sobolev type the exact minimum is given by $P_V f$. Lemma 1 shows that if the prior is a norm on a Hilbert space this boils down to the same result as in Theorem 1 (*as it should*).

Lemma 1. *In the Euler Lagrange framework, the unique solution of the minimization of the convex positive energy* $E(g) = \frac{1}{2}(g, g)_H$*, under the conditions* $(\psi_i, g)_H = c_i \in \mathbb{C}$ *(fixed) for* $i = 1, \ldots, n$*, where* $\{\psi_i\}$ *are linearly independent in* H*,satisfies*

$$< DE(g), f >= (g, f) = \sum_{i=1}^{n} \lambda^i (\psi_i, f)_H, \quad \textit{for all } f \in H \tag{7}$$

and is given by the orthogonal projection $g = \mathcal{P}_V f$*, given by (5), of the original image* f *on the linear span* V *of the filters* ψ_i*.*

Proof. The uniqueness and last part of the proof is already given by Theorem 1. With respect to the first part we only mention that the Gateaux variation of the energy at g in the direction of f is given by

$$< DE(g), f >= \lim_{\lambda\to 0} \frac{E(g + \lambda f) - E(f)}{\lambda} = (g, f), \text{ for all } f \in H$$

and the Lagrange multipliers must be equal to $\lambda^i = (\psi^i, g)$, where $\{\psi^i\}_{i=1}^n$ denotes the reciprocal basis.□

Instead of iterative schemes we use directly compute the exact solution $P_V f$, for the cases (recall the example in section 4.1) where the Hilbert space H equals a space of Sobolev-type, for example $\mathbb{H}^{k,2}_\gamma(\mathbb{R}^2) = \{f \in \mathbb{L}_2(\mathbb{R}^2) \mid (Rf, Rf) = (R^2 f, f) < \infty\}$, where $R = (I + \gamma^{2k}|\Delta|^k)^{\frac{1}{2}}$, $k = 1, 2, \ldots, \gamma > 0$, as explained in Corollary 1. In the next chapter we will deal with some important theoretical issues that inevitably arise when working with spaces of Sobolev type, but it is not necessary to understand all details to understand the algorithm from a practical point of view. At this point, if the reader is not interested in these more theoretical aspects, we directly refer to Figures 2 and 1 and Corollaries 1, 2.

[3] A finite dimensional subspace is always closed, but this statement reveals how to deal with the case $n \to \infty$.

[4] A norm on a vector space V is induced by an inner product iff the parallelogram law $\|\mathbf{x} + \mathbf{y}\|^2 + \|\mathbf{x} - \mathbf{y}\|^2 = 2(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2)$ holds for all $\mathbf{x}, \mathbf{y} \in V$. In this case the inner product is given by $(\mathbf{x}, \mathbf{y}) = \frac{1}{4}\left\{\|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x} - \mathbf{y}\|^2 + i\|\mathbf{x} - i\mathbf{y}\|^2 - i\|\mathbf{x} + i\mathbf{y}\|^2\right\}$.

4 Gelfand Triples

Let H be a complex Hilbert space and $\mathcal{R}$ an unbounded, positive and selfadjoint operator on H, for which the inverse $\mathcal{R}^{-1}$ is bounded. Note that the boundedness of $\mathcal{R}^{-1}$ implies that the domain $D(\mathcal{R}) = \{f \in H \mid \mathcal{R}f \in H\}$ equals $D(\mathcal{R}) = \mathcal{R}^{-1}(H)$.

Definition 1. *Define the space H^I as the linear space $D(\mathcal{R})$ equipped with the inner product $(f,g)_I = (\mathcal{R}f, \mathcal{R}g)_H$ for all $f, g \in H$.*

Notice that H^I is again a Hilbert space: Let f_n be a Cauchy sequence in H^I. Then $\mathcal{R}f_n$ is a Cauchy sequence in H. H is a Hilbert space, so $\mathcal{R}f_n \to g$ in H, for some $g \in H$. But then, since $\mathcal{R}^{-1}$ is bounded, it follows that f_n is also a Cauchy sequence in H: $\|f_n - f_m\| \leq \|\mathcal{R}^{-1}\|\|\mathcal{R}f_n - \mathcal{R}f_m\|$. So $f_n \to f$ in H, for some $f \in H$. Now $\mathcal{R}$ is self adjoint and therefore closed, so $f \in \mathcal{D}(\mathcal{R})$ and $\mathcal{R}f = g$. Now we have $\mathcal{R}f_n \to \mathcal{R}f$ in H, so $f_n \to f$ in H^I and $f \in H^I$.

Definition 2. *Define the Hilbert space H^{-I} as the completion of H equipped with the inner product $(f,g)_{-I} = (\mathcal{R}^{-1}f, \mathcal{R}^{-1}g)_H$.*

The operator $\mathcal{R}$ on H induces the map $\tilde{\mathcal{R}} : H^I \to H$ by $\tilde{\mathcal{R}}f = \mathcal{R}f$ for all $f \in H^I = D(\mathcal{R})$. Since $\|\tilde{\mathcal{R}}f\|_H = \|f\|_I$ for all $f \in H^I$, the map $\tilde{\mathcal{R}}$ is an isometry. By boundedness of R^{-1}, it follows that $\tilde{R}$ is also surjective and hence a unitary map.

Define $\check{\mathcal{R}} : D(\mathcal{R}) \to H^{-I}$ by $\check{\mathcal{R}}f = \mathcal{R}f$ for all $f \in D(\mathcal{R})$. Since $\|\check{\mathcal{R}}f\|_{-I} = \|f\|_H$ for all $f \in D(\mathcal{R})$, the map $\check{\mathcal{R}}$ is closable and its extension is an isometry. Since $\mathcal{R}(D(\mathcal{R})) = H$ and H is dense in H^{-I} the closure is also surjective, hence a unitary map. Write $\tilde{\check{\mathcal{R}}}$ for the closure of $\check{\mathcal{R}}$. Hence the following triple (known as Gelfand Triple) is obtained

$$H^I \overset{\tilde{\mathcal{R}}}{\hookrightarrow} H \overset{\tilde{\check{\mathcal{R}}}}{\hookrightarrow} H^{-I}. \tag{8}$$

It follows by the Riesz representation theorem and the unitarity of $\tilde{\mathcal{R}}$ and $\tilde{\check{\mathcal{R}}}$ that the space H^{-I} is naturally isomorphic to the anti-dual space of H^I under the pairing $\langle F, f\rangle = (\tilde{\check{\mathcal{R}}}^{-1}F, \tilde{\mathcal{R}}f)_H$ for all $F \in H^{-I}$ and $f \in H^I$. Note that by the selfadjoint-ness of $\mathcal{R}$

$$\langle F, f\rangle = (F, f)_H \tag{9}$$

if $F \in H$ for all $f \in H^I$. In this paper $\mathcal{R}$, $\tilde{\mathcal{R}}$ and $\tilde{\check{\mathcal{R}}}$ are all denoted by the same symbol $\mathcal{R}$. From the context it is clear which operator is meant by this symbol.

4.1 Spaces of Sobolev Type

By a Theorem of John von Neumann, [10]p.200, we have that for every closed densely defined operator $\mathcal{A}$ in a Hilbert space H the operator $\mathcal{A}^*\mathcal{A}$ is self adjoint and $(I + \mathcal{A}^*\mathcal{A})$ has a bounded inverse. So a particular case of a Gelfand triple is obtained by setting $\mathcal{R} = (I + \mathcal{A}^*\mathcal{A})^{1/2}$, where $\mathcal{A}$ is a closed densely defined operator. In that case we have

$$(f,g)_I = (\mathcal{R}f, \mathcal{R}g)_H = (f,g)_H + (\mathcal{A}f, \mathcal{A}g)_H\ . \tag{10}$$

Example 1: Let $k \in \mathbb{N}$. Then it is well-known that the operator $D_k = (1 + \gamma^{2k}|\triangle|^k)^{\frac{1}{2}}$ with domain $\mathbb{H}^{k,2}_\gamma(\mathbb{R}^2)$ is an unbounded, positive and self-adjoint operator on $\mathbb{L}_2(\mathbb{R}^2)$ with bounded inverse. So $\mathcal{A} = \gamma(\sqrt{-\Delta})^k$ and $\mathcal{R} = D_k$ and we obtain the Gelfand triple

$$\mathbb{H}^{k,2}_\gamma(\mathbb{R}^2) \hookrightarrow \mathbb{L}_2(\mathbb{R}^2) \hookrightarrow \mathbb{H}^{-k,2}_\gamma(\mathbb{R}^2). \tag{11}$$

For a detailed discussion on these spaces, in particular $\gamma = 1$, we refer to [10]I.10, pp.56.
Remark: The Gelfand triple structure is interesting for the case that $\mathcal{A}$ is unbounded, since if $\mathcal{A}$ is bounded (and thereby $\mathcal{R}$ is bounded) we have by (10) that :

$$\|f\|_H^2 \leq \|f\|_I^2 \leq (1 + \|\mathcal{A}\|^2)\|f\|_H^2,$$

so the norms (and the thereby induced topologies) are equivalent. Moreover, if $\mathcal{A}$ is bounded the set $\mathcal{D}(\mathcal{R})$ equals H and the Riesz representant of $f \mapsto (\phi, f)$ in H^I is given by $\mathcal{R}^{-2}\phi = (I + \mathcal{A}^*\mathcal{A})^{-1}\phi$:

$$(\mathcal{R}^{-2}\phi, f)_I = (\mathcal{R}^{-1}\phi, \mathcal{R}f)_H = (\mathcal{R}\mathcal{R}^{-1}\phi, f)_H = (\phi, f)_H.$$

Summarizing, if $\mathcal{A}$ is bounded the sets H^I and H are equal and by the Riesz representation theorem H is unitary equivalent with its dual H', so operator $\mathcal{A}$ only tells us how H is identified with H'.

4.2 Trajectory spaces

In a Hilbert Space H we consider the general evolution equation $\frac{du}{ds} = \mathcal{A}u$, with $\mathcal{A}$ a negative unbounded self-adjoint operator. $\mathcal{A}$ is the infinitesimal generator of a holomorphic semi-group. Solutions $u(\cdot) : (0, \infty) \to H$ of this equation are called trajectories. Such a trajectory may or may not correspond to an "initial condition at $s = 0$" in H. The set of trajectories is considered as a space of generalized functions. The test function space is defined to be $\mathcal{S}_{H,\mathcal{A}} = \bigcup_{s>0} e^{s\mathcal{A}}(H)$.

Theorem 2. *Let H be a Hilbert space. Let Q be a strongly continuous, holomorphic semigroup, with infinitesimal generator $\mathcal{A} < 0$. Then $\mathcal{S}_{H,\mathcal{A}}$ consists exactly of those $f \in \mathcal{D}(\mathcal{A}^\infty)$ such that*

$$\sum_{k=1}^{\infty} \frac{s^k}{k!} \|\mathcal{A}^k f\| < \infty \qquad \textit{for a certain } s > 0\ . \tag{12}$$

This is a well-known result in functional analysis and its proof can for example be found in [11] or [1]Appendix, Thm 11. Now by taking $\mathcal{R} = e^{-s\mathcal{A}}$ on H, with bounded inverse $e^{s\mathcal{A}}$ we obtain the following Gelfand-triples:

$$\mathcal{S}_{H,\mathcal{A}} \overset{\mathcal{R}}{\hookrightarrow} H \overset{\mathcal{R}}{\hookrightarrow} \mathcal{S}'_{H,\mathcal{A}}. \tag{13}$$

Example: $H = \mathbb{L}_2(\mathbb{R}^d)$, $\mathcal{A} = \Delta$, then $(\mathcal{R}^{-1}f) = K_s^1 * f$, where K_s^1 is the Gaussian kernel, recall (1), and we obtain the Gelfand triple

$$\mathcal{S} \overset{\mathcal{R}}{\hookrightarrow} \mathbb{L}_2(\mathbb{R}^d) \overset{\mathcal{R}}{\hookrightarrow} \mathcal{S}'. \tag{14}$$

where $\mathcal{S}_{\mathbb{L}_2(\mathbb{R}^d),\Delta} = \mathcal{S} = \{\phi \in C^\infty(\mathbb{R}^d) \mid \forall_{\mathbf{k},\mathbf{q}} \sup_{\mathbf{x}\in\mathbb{R}^d} |\mathbf{x}^{\mathbf{k}}\phi^{(\mathbf{q})}| < \infty\}$ is the usual Schwarz space.[5]

4.3 A Linear Image Reconstruction Based on Inner Products of Sobolev Type

By taking the space of images H^I (for example $H^I = \mathbb{H}_\gamma^{2k,2}(\mathbb{R}^d)$ or $H^I = \mathcal{S}_{\mathbb{L}_2(\mathbb{R}(d)),-(-\Delta)^\alpha}$) in Theorem 1 we obtain the metameric class

$$[f] = \{g \in H^I \mid (\kappa_k, f)_I = (\kappa_k, g)_I,\ \text{for } k = 1, \ldots, n\}\ ,\quad \text{with } \kappa_k = \mathcal{R}^{-2}\psi_k, k = 1, \ldots n,$$

and the optimal solution within this metameric class $\arg\min_{g\in[f]} \|g\|_{H^I}^2$, is given by $g = P_V f$. Notice that $(\kappa_k, g)_I = (\psi_k, g)_{\mathbb{L}_2(\mathbb{R}^2)}$, so the metameric classes still consist of images (within H^I) with the same features. But initially, the space of images was the modeled by $H = \mathbb{L}_2(\mathbb{R}^2)$ space (and not the smooth space H^I), so we want our metameric classes within $\mathbb{L}_2(\mathbb{R}^2)$ rather than H^I. Theorem 3, takes care of this.

[5] In case of the Poisson semigroup ($\mathcal{A} = -\sqrt{-\Delta}$) leads to the Gelfand-Shilov space $\mathcal{S}_1$ rather than the Schwarz space, see [1].

Theorem 3. *Let $f \in H$. Let $[f]$ denote the metameric class of all elements $g \in H$ such that*

$$g \sim f \Leftrightarrow (\mathcal{R}^{-1}f, \mathcal{R}\kappa_k)_I = (\mathcal{R}^{-1}g, \mathcal{R}\kappa_k)_I \ , \ \textit{for all } k = 1, \ldots, n.$$

with $\kappa_k \in H^{II}$, i.e. $\mathcal{R}\kappa_k \in H^I$. Then the unique solution of the minimization problem $\min\limits_{g \in H^I, g \sim f} \|g\|_I$ is given by $g = P_V^{ext} f = \sum\limits_{i=1}^{n} (\mathcal{R}\kappa^i, \mathcal{R}^{-1}f)_I \kappa_i$, with $\kappa^i = \sum\limits_{j=1}^{n} G^{ij}\kappa_j$, where $\sum\limits_{k=1}^{n} G^{ik}G_{kj} = \delta^i_j$, with $G_{kj} = (\kappa_k, \kappa_j)_I$.

Proof. First notice that P_V^{ext} is the natural extension of P_V, i.e. the restriction of P_V^{ext} to H^I equals P_V: $P_V^{ext}|_{H^I} = P_V$, which directly follows by the fact that $\mathcal{R}$ is self adjoint and therefore

$$(\mathcal{R}\psi, \mathcal{R}^{-1}f)_I = (\psi, f)_I \text{ for all } f \in H^I \text{ and } \psi \in H^{II}. \tag{15}$$

Without loss of generality we may assume that $\{\kappa_k\}$ is an orthonormal base in H_I. Then we have $g - P_V^{ext} f \perp P_V^{ext}$, since by $f \sim g$ and (15) it follows that

$$(g - P_V^{ext} f, P_V^{ext})_I = (\mathcal{R}\kappa^j, \mathcal{R}^{-1}f)(g, \kappa_j) - \overline{(\mathcal{R}\kappa^j, \mathcal{R}^{-1}f)}(\mathcal{R}\kappa^j, \mathcal{R}^{-1}f) = 0 \ .$$

So again Pythagoras: $\min\limits_{g \sim f, g \in H_I} \|g\|_I^2 = \min\limits_{g \sim f, g \in H_I} \|g - \mathcal{P}_V^{ext} f + \mathcal{P}_V^{ext} f\|_I^2 = \min\limits_{g \sim f, g \in H_I} \|g - \mathcal{P}_V^{ext} f\|_I^2 + \|\mathcal{P}_V^{ext} f\|_I^2$ we conclude that this equals $\min\limits_{g \sim f, g \in H_I} \|g\|_I^2 = \|\mathcal{P}_V^{ext} f\|_I^2$ iff $g = \mathcal{P}_V^{ext} f$ □

By taking the respective pairs $H = \mathbb{L}_2(\mathbb{R}^d)$, $\mathcal{R} = (I + \gamma^{2k}|\Delta|^k)^{\frac{1}{2}}$ and $H = \mathbb{L}_2(\mathbb{R}^d)$, $\mathcal{R} = e^{\frac{s}{2}|\Delta|^\alpha}$ in Theorem 3 we obtain the following corollaries:

Corollary 1. *Let $f \in \mathbb{L}_2(\mathbb{R}^d)$. Let $[f]$ denote the equivalence class of all elements $g \in \mathbb{L}_2(\mathbb{R}^d)$ such that*

$$g \sim f \Leftrightarrow (\psi_k, f)_{\mathbb{L}_2(\mathbb{R}^2)} = (\psi_k, g)_{\mathbb{L}_2(\mathbb{R}^2)}, \ \textit{for all } k = 1, \ldots, n.$$

Then the unique solution of the minimization problem

$$\min_{g \in \mathbb{H}_\gamma^{k,2}(\mathbb{R}^2), g \sim f} \|g\|^2_{\mathbb{H}_\gamma^{k,2}(\mathbb{R}^2)} = \min_{g \in \mathbb{H}_\gamma^{k,2}(\mathbb{R}^2), g \sim f} (g, g)_{\mathbb{L}_2(\mathbb{R}^2)} + (g, \gamma^{2k}|\Delta|^k g)_{\mathbb{L}_2(\mathbb{R}^2)}$$

is given by $g = P_V^{ext} f$, where $P_V^{ext} f = ((I + \gamma^{2k}|\Delta|^k)^{\frac{1}{2}}\kappa^i, (I + \gamma^{2k}|\Delta|^k)^{-\frac{1}{2}} f)_{\mathbb{H}_\gamma^{k,2}(\mathbb{R}^2)}\kappa_i$, where κ^k are the reciproke vectors of $\kappa_k = (I + \gamma^{2k}|\Delta|^k)^{-1}\psi_k$.

Important: Notice that $\kappa_k = R^{k,\gamma}_{d=2} * \psi_k$, where $R^{k,\gamma,\mathbf{0}}_{d=2}(\mathbf{x}) = \mathcal{F}^{-1}(\boldsymbol{\omega} \mapsto \frac{1}{2\pi}(1 + \gamma^{2k}|\boldsymbol{\omega}|^{2k})^{-1})(\mathbf{x})$ is the reproducing kernel (at $\mathbf{0}$) of $\mathbb{H}_\gamma^{k,2}(\mathbb{R}^2)$, for $k > 1$, see Appendix. The reproducing kernels are needed to project $\mathbb{L}_2$-linear onto linear Sobolev-features, which is the basic reason for the succes of our approach to feature reconstruction: The features themselves do not change, but the projection basis is subject to a Tikhonov regularization. In this way " the reconstructed Lena gets rid of her smallpox ", see fig 1.

Corollary 2. *Let $f \in \mathbb{L}_2(\mathbb{R}^d)$. Let $[f]$ denote the equivalence class of all elements $g \in \mathbb{L}_2(\mathbb{R}^d)$ such that*

$$g \sim f \Leftrightarrow (\psi_k, f)_{\mathbb{L}_2(\mathbb{R}^2)} = (\psi_k, g)_{\mathbb{L}_2(\mathbb{R}^2)}, \ \textit{for all } k = 1, \ldots, n.$$

Then the unique solution of the minimization problem

$$\min_{g \in \mathcal{S}_{\mathbb{L}_2(\mathbb{R}^2), -|\Delta|^\alpha}, g \sim f} (g, g)_{\mathbb{L}_2(\mathbb{R}^2)} + \sum_{k=1}^{\infty} (g, s^{2k}|\Delta|^{\alpha k} g)_{\mathbb{L}_2(\mathbb{R}^2)} \tag{16}$$

is given by $g = P_V^{ext} f$, where $P_V^{ext} f = (e^{\frac{s}{2}|\Delta|^\alpha}\kappa^i, e^{-\frac{s}{2}|\Delta|^\alpha} f)_{\mathcal{S}_{\mathbb{L}_2(\mathbb{R}^2), -|\Delta|^\alpha}}\kappa_i$, where κ^i are the reciproke vectors of $\kappa_i = \left(e^{\frac{s}{2}|\Delta|^\alpha}\right)^{-2}\psi_i = K^\alpha_{s/2} * K^\alpha_{s/2} * \psi_i = K^\alpha_s * \psi_i$.

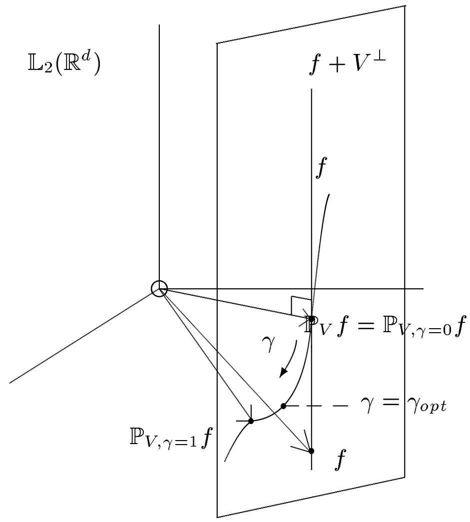


Fig. 1. Illustration of metameric class $f + V^{\perp}$ of images with equal features. For $\gamma = 0$ we have a orthogonal projection. For $\gamma > 0$ this projection is orthogonal in $\mathbb{H}_{\gamma}^{k,2}(\mathbb{R}^2)$ and thereby it is a skew projection in $\mathbb{L}_2(\mathbb{R}^2)$. We obtain a connected curve parametized by γ where smoothness of the projection increases with $\gamma > 0$.

We are mainly interested in the case where the features are obtained by α-derivatives within the scale space image u of the original image f, i.e. $\psi_k(\mathbf{x}) = (\tau_{\mathbf{b}_k} D^{\mathbf{n}_k} K_{s_k}^{\alpha})(\mathbf{x}) = D^{\mathbf{n}_k} K_{s_k}^{\alpha}(\mathbf{x} - \mathbf{b}_k)$ for some multi-index $\mathbf{n}_k$, some position $\mathbf{b}_k \in \mathbb{R}^d$ and some scale s_k. In this case we can analytically compute the Grammian matrix (and thereby analytically compute the solution $g = P_V^{ext} f$):

$$G_{kj} = (\kappa_k, \kappa_j)_{\mathcal{S}_{\mathbb{L}_2(\mathbb{R}^2), -|\Delta|^{\alpha}}} = (\mathcal{R}^{-2}\psi_k, \mathcal{R}^{-2}\psi_j)_{\mathcal{S}_{\mathbb{L}_2(\mathbb{R}^2), -|\Delta|^{\alpha}}} = (\mathcal{R}^{-1}\psi_k, \mathcal{R}^{-1}\psi_j)_{\mathbb{L}_2(\mathbb{R}^2)}$$
$$= (\tau_{\mathbf{b}_k} D^{\mathbf{n}_k} K_{s_k+(s/2)}^{\alpha}, \tau_{\mathbf{b}_j} D^{\mathbf{n}_j} K_{s_j+(s/2)}^{\alpha})_{\mathbb{L}_2(\mathbb{R}^2)} = \left(K_{s_j+s_k+s}^{\alpha}\right)^{(\mathbf{n}_k+\mathbf{n}_j)}(\mathbf{b}_k - \mathbf{b}_j)\ .$$

Important: Here again the features are preserved and the base functions are smoothed. But now the smoothing kernels are given by the α-kernels themselves rather than the reproducing kernels of the isotropic Sobolev-spaces in the previous corollary. So instead of Tikhonov-regularization we smooth according to ordinary α-scale spaces. The relation between these 2 extreme ways of smoothing is given by the Laplace transform with respect to the scale parameter

$$\mathcal{L}[s \mapsto K_s^{\alpha}](\gamma^{-2k}) = \gamma^{2k} R_{d=2}^{k,\gamma,\mathbf{0}} \text{ for all } \gamma > 0, \text{ where } k = \alpha. \tag{17}$$

By interchanging the order of Laplace transformation and integration, it is not difficult to obtain an analytic formula for the Grammian matrix in Corollary 1 (and thereby analytic formulas for the orthogonal projection) for the cases where $\alpha = k$.

5 Flux-features

Besides the regularizing effect of minimizing Sobolev norms there is another advantage of using Sobolev inner products: There exist several interesting features (such as grey-value fluxes) of images which can be constructed by means of Sobolev inner products, but which can not be constructed in the usual framework of $\mathbb{L}_2$-inner products. For example point evaluation $\delta_{\mathbf{a}}(f) = f(\mathbf{a})$ is a continuous linear functional on $\mathbb{H}_{\gamma}^{k,2}(\mathbb{R}^2)$, $k > 1$. As a result (by the Riesz representation Theorem) there exists reproducing kernels, see Appendix A, $R_{k,\gamma}^2 \in \mathbb{H}_{\gamma}^{k,2}(\mathbb{R}^2)$, $k > 1$ such that $\delta_{\mathbf{a}}(f) = f(\mathbf{a}) = (\tau_{\mathbf{a}} R_{k,\gamma}^2, f)$, which makes point evaluations linear features. This can not be done within the more familiar $\mathbb{L}_2(\mathbb{R}^2)$ space. Another example of new possible linear features are so-called flux-features on images. They are

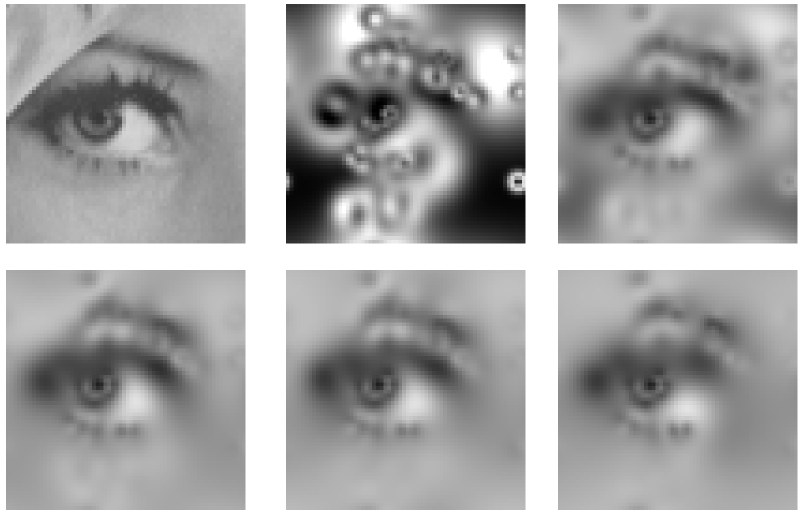


Fig. 2. Reconstruction from 31 top-points of "Lena's eye" with up to second order features only. The upper row shows the original image and reconstructions with $\gamma = 0$ and $\gamma = 5$. The second row shows reconstructions with $\gamma = 22$, $\gamma = 50$ and $\gamma = 250$. The first image in the second row shows the reconstruction with the lowest relative $\mathbb{L}_2$-error. For more details on evaluation we refer to our earlier work,[12],[13].

given by the linear functionals

$$f \mapsto I_\Omega(f) = \int_{\partial\Omega} \frac{\partial f}{\partial n} \, \mathrm{d}\sigma, \qquad f \in \mathbb{H}_\gamma^{2,2}(\mathbb{R}^2).$$

where Ω is a bounded region in $\mathbb{R}^2$, with surface measure $\mu(\Omega) < \infty$ with a orientable piecewise smooth boundary $\partial\Omega$ with outward normal $\mathbf{n}$. These linear functionals are continuous on $\mathbb{H}_\gamma^{k,2}(\mathbb{R}^2)$ if $k \geq 2$, which directly follows by Gauss divergence theorem and Cauchy-Schwarz: $|I_\Omega(f)| \leq \left|\int_\Omega \Delta f \mathrm{d}\mathbf{x}\right| \leq \mu(\Omega)\|f\|_{\mathbb{H}_\gamma^{2,2}(\mathbb{R}^2)}$. As a result there exist flux-kernels $\phi_{\Omega,k,\gamma}$ in $\mathbb{H}_\gamma^{k,2}(\mathbb{R}^2)$, $k \geq 2$, such that

$$I_\Omega(f) = (\phi_{\Omega,k,\gamma}, f)_{\mathbb{H}_\gamma^{k,2}(\mathbb{R}^2)}.$$

The Practical Reason for Flux Features:
In the previous chapters we improved the top-point reconstruction by introducing spaces of Sobolev-type. The idea was to pick a smooth element within the metameric class of images with top-points at the same locations as the locations of the top-points of the scale space of the original image f, where a positive parameter controls the smoothness of this reconstruction. By increasing this parameter ($\gamma > 0$ in Corollary 1 or $s > 0$ in Corollary 2) this representant becomes smoother and thereby the number of extra top points of the scale space of this representant is reduced. Nevertheless, in practice there exist an upper bound to the parameter γ, since if the representant becomes too smooth, the contrasting areas (such as corners/edges) will vanish, which results in so-called "edge-leaking", cf.[12]. Therefore, we propose to reduce the number of extra top-points in the representant by means of extra features. Flux-features obtained by surfaces around critical paths in scale space are highly suitable for this purpose, as they describe how grey-value particles flow between extremal paths.

5.1 Definition and Implementation of Flux Features in Scale Space

Definition 3. *Let $s > 0$, $f \in \mathbb{L}_2(\mathbb{R}^2)$, with α-scale space representation $u_f^\alpha = K_s^\alpha * f$. Let Ω be a bounded region in $\mathbb{R}^2$ with an orientable piecewise smooth boundary $\partial\Omega$ with outward normal $\boldsymbol{n}$. Then we define the flux feature* $I_\Omega^\alpha(f,s) = \left.\frac{\partial}{\partial t} \int_\Omega u_f^\alpha(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}\right|_{t=s}$.

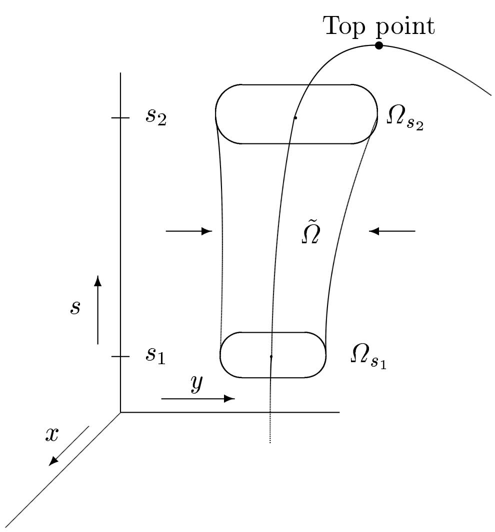
We notice that by means of Gauss' divergence theorem and changing the order of differentiation with respect to $s > 0$ and integration over Ω we can relate flow-vector fields $\mathbf{F}_\alpha u_f^\alpha(\cdot, s)$, recall (4), to flux-features

$$I_\Omega^\alpha(f,s) = \int_\Omega \frac{\partial}{\partial s} u_f^\alpha(\mathbf{x},s)\, \mathrm{d}\mathbf{x} = \int_\Omega -(-\Delta)^\alpha u_f^\alpha(\mathbf{x},s)\, \mathrm{d}\mathbf{x} = \int_\Omega \operatorname{div} \mathbf{F}_\alpha\, u_f^\alpha(\mathbf{x},s)\, \mathrm{d}\mathbf{x} = \int_{\partial\Omega} \mathbf{F}_\alpha\, u_f^\alpha(\mathbf{x},s) \cdot \mathbf{n}(\mathbf{x})\, \mathrm{d}\sigma(\mathbf{x}), \tag{18}$$

Recall that $\mathbf{F}_1$ equals the gradient operator $\mathbf{F}_1 = \nabla$ and $\mathbf{F}_{\frac{1}{2}}$ equals the Riesz transform $\mathbf{F}_{\frac{1}{2}} = \mathbf{R}$. So, in case $\alpha = 1$ and in case $\alpha = 1/2$ we have

$$I_\Omega^1(f,s) = \int_{\partial\Omega} \frac{\partial u_f^1(\cdot,s)}{\partial n}\, \mathrm{d}\sigma(\mathbf{x}) \qquad I_\Omega^{1/2}(f,s) = \int_{\partial\Omega} \mathbf{q}(\mathbf{x},s) \cdot \mathbf{n}(\mathbf{x})\, \mathrm{d}\sigma(\mathbf{x}), \tag{19}$$

where $\mathbf{q} = \mathbf{f}_s^{\frac{1}{2}} * f$ denotes the conjugate Poisson scale space.



Grey-value flow of grey-value particles through surface $\partial\tilde{\Omega}$ around an extremal path in scale space

Definition 4. *Let $\tilde{\Omega} = \{(\boldsymbol{x}, s) \in \mathbb{R}^2 \rtimes [0,\infty) \mid 0 \leq s_1 \leq s \leq s_2, \boldsymbol{x} \in \Omega_s \subset \mathbb{R}^2\}$, where for all $s > 0$, Ω_s is a bounded region in $\mathbb{R}^2$, with an orientable piecewise smooth boundary $\partial\Omega$ with outward normal $\boldsymbol{n}$. Suppose the lifetime of a grey-value particle through scale space is negatively exponentially distributed with expectation μ^{-1}. Then the grey-value flow trough $\tilde{\Omega}$ is given by*[6]

$$I_{\tilde{\Omega}}^{\alpha,\mu}(f) = \frac{1}{\mu} \int_{s_1}^{s_2} I_\Omega^\alpha(f,s)\, e^{-\mu s}\, \mathrm{d}s. \tag{20}$$

Notice that $\tilde{\Omega}$ need not be equal to $\Omega \times (s_1, s_2)$ as Ω_s may vary in $s > 0$. By straightforward computation and the Plancherel Theorem we have

$$\begin{aligned} I_\Omega^\alpha(f,s) &= -((-\Delta)^\alpha u_f^\alpha(\cdot,s), 1_\Omega)_{\mathbb{L}_2(\mathbb{R}^2)} = -(\boldsymbol{\omega} \mapsto \|\boldsymbol{\omega}\|^{2\alpha} \hat{u}_f^\alpha(\boldsymbol{\omega},s), \widehat{1_\Omega})_{\mathbb{L}_2(\mathbb{R}^2)} \\ &= \left(\boldsymbol{\omega} \mapsto \frac{1}{\sqrt{1+(\gamma\|\boldsymbol{\omega}\|)^{2k}}} \hat{f}(\boldsymbol{\omega}), \boldsymbol{\omega} \mapsto \frac{-\|\boldsymbol{\omega}\|^{2\alpha} e^{-s\|\boldsymbol{\omega}\|^{2\alpha}}}{\sqrt{1+(\gamma\|\boldsymbol{\omega}\|)^{2k}}} \widehat{1_\Omega}(\boldsymbol{\omega})\right)_{\mathbb{L}_2(\mathbb{R}^2\,;\,(1+\gamma\|\boldsymbol{\omega}\|^{2k})\,\mathrm{d}\boldsymbol{\omega})} \\ &= (\phi_{\Omega,k,\gamma}^{\alpha,s}, f)_{\mathbb{H}_\gamma^{k,2}(\mathbb{R}^2)} \end{aligned} \tag{21}$$

with $f \in \mathbb{L}_2(\mathbb{R}^2)$, $\alpha \in (0,1]$ and flux-kernel $\phi_{\Omega,k,\gamma}^{\alpha,s} = \mathcal{F}^{-1}[\hat{\phi}_{\Omega,k,\gamma}^{\alpha,s}] = \mathcal{F}^{-1}[\boldsymbol{\omega} \mapsto \frac{-\|\boldsymbol{\omega}\|^{2\alpha} e^{-s\|\boldsymbol{\omega}\|^{2\alpha}}}{\sqrt{1+(\gamma\|\boldsymbol{\omega}\|)^{2k}}} \widehat{1_\Omega}(\boldsymbol{\omega})]$. Notice that the substitution $k = 0$ in equality (21) yields $I_\Omega^\alpha(f,s) = (\phi_{\Omega,0}^{\alpha,s}, f)_{\mathbb{L}_2(\mathbb{R}^2)}$ only for $s > 0$.

[6] In case $\Omega_s = \Omega$, i.e. Ω_s does not change over scale, and $s_0 = 0$ and $s_1 \to \infty$ we obtain $\psi_{\tilde{\Omega},k}^{\alpha,\mu} = \frac{1}{\mu}\mathcal{L}[s \mapsto \phi_{\Omega_s,k}^{\alpha,s}](\mu)$, where $\mathcal{L}$ denotes the well-known Laplace transform.

In case $s=0$, where $u_f^\alpha(\cdot,0)=f$, we may write $I_\Omega^\alpha(f,0)=(\phi^{\alpha,0}_{\Omega,k,\gamma},f)_{\mathbb{H}^{k,2}_\gamma(\mathbb{R}^2)}$ for $k\geq 2$. For $k=0$ the inner product is not defined. It now follows that $I^{\alpha,\mu}_{\tilde{\Omega}}(f)=(\psi^{\alpha,\mu}_{\tilde{\Omega},k},f)_{\mathbb{H}^{k,2}_\gamma(\mathbb{R}^2)}$, where the netto flux kernel needed to compute $I^{\alpha,\mu}_{\tilde{\Omega}}(f)$ is simply given by $\psi^{\alpha,\mu}_{\tilde{\Omega},k}=\frac{1}{\mu}\int\limits_{s_0}^{s_1}\phi^{\alpha,s}_{\Omega_s,k}e^{-\mu\, s}\mathrm{ds}$. Finally we notice that $\phi^{\alpha,0}_{\Omega_s,k,\gamma}=\kappa_k*\phi^{\alpha,s}_{\Omega,0}$, where the flux-kernel $\phi^{\alpha,s}_{\Omega_s,0}$ is given by

$$\phi^{\alpha,s}_{\Omega_s,0}=\frac{\partial}{\partial s}K^\alpha_s*1_\Omega=(\mathrm{div}\mathbf{F}_\alpha K^\alpha_s)*1_{\Omega_s}. \tag{22}$$

The case where Ω_s equals an ellips: In case of the disk $\Omega_s=B_{\mathbf{0},a}$ there are two reasonable options, either the flux kernel is evaluated via the Fourier domain, which yields:
$\phi^{\alpha,s}_{B_{\mathbf{0},a},0}(\mathbf{x})=-a\int\limits_0^\infty\rho^{2\alpha}J_1\left(\frac{\rho}{a}\right)J_0(\rho\, r)e^{-s\rho^{2\alpha}}\,\mathrm{d}\rho,\qquad r=\|\mathbf{x}\|,$
where we notice that $\mathcal{F}1_\Omega(\boldsymbol{\omega})=\frac{a\,J_1\left(\frac{\rho}{a}\right)}{\rho}$, $\|\boldsymbol{\omega}\|=\rho$, or by means of (18), (19). In case $\alpha=1$ we get

$$\phi^{1,s}_{\Omega,0}(\mathbf{y})=\int\limits_{\partial\Omega}\frac{\partial K^1_s(\mathbf{x}-\mathbf{y})}{\partial r}\mathrm{d}\sigma_{\mathbf{x}}=a\int\limits_0^{2\pi}\left\{\cos\phi\left(\frac{y_1-a\cos\phi}{8\pi s^2}\right)+\sin\phi\left(\frac{y_2-a\sin\phi}{8\pi s^2}\right)\right\}e^{\frac{-(a\cos\phi-y_1)^2-(a\sin\phi-y_2)^2}{4s}}\mathrm{d}\phi,$$

which is easily numerically approximated. Finally, we notice that the ellips-case $\Omega_s=E_{\theta(s),\lambda(s),a(s)}=\{\mathbf{x}\in\mathbb{R}^2\mid\|L^{-1}_{\lambda(s)}R^{-1}_{\theta(s)}\mathbf{x}\|^2=(a(s))^2\}$, $\theta(s)\in[0,2\pi),\lambda(s)>0,a(s)>0$, recall (3), directly follows from the disk case by means of

$$\phi^{\alpha,s}_{\Omega_s=E_{\theta(s),\lambda(s),a(s)},k}(\mathbf{x})=\phi^{\alpha,s}_{B_{\mathbf{0},a},k}(L^{-1}_{\lambda(s)}R^{-1}_{\theta(s)}\mathbf{x}),\qquad\mathbf{x}\in\mathbb{R}^d,s>0.$$

6 Conclusion

The common (non)-linear reconstruction frameworks follow an Euler Lagrange minimization. If the Lagrangian (prior) is a norm induced by an inner product of a Hilbert space this Euler Lagrange minimization boils down to a simple orthogonal projection within the corresponding Hilbert space. This basic observation has been overlooked in image analysis for the cases where the Lagrangian equals a norm of Sobolev type resulting in iterative (non-linear) numerical methods, where we provide exact solutions. By means of Gelfand triples we consider two extreme cases (standard isotropic Sobolev spaces) and trajectory spaces. Although the reconstruction algorithm is slightly simpler for the trajectory space case, best results are obtained by minimizing isotropic Sobolev norms of low order, taking into account that in both cases smoothness of the reconstruction is tuned by a positive parameter. In our opinion a linear reconstruction should at least capture the topology (deep structure) of the scale space of the original image. Therefore we only consider linear features inspired by the deep structure of images. Besides top-points, the well-known scale space singularities, we introduce the concept of flux-features. They describe the transport of grey-value particles between extremal paths that lead to singular points in scale space. By including these features we provide another tool to minimize spurious top-points in the reconstruction image.

A The Reproducing kernels $R^2_{k,\gamma}$ of the Sobolev spaces $\mathbb{H}^{k,2}_\gamma(\mathbb{R}^2)$, $k>1$.

In this section we compute the reproducing kernels $R^2_{k,\gamma}$ of the Sobolev spaces $\mathbb{H}^{k,2}_\gamma(\mathbb{R}^2)$, $k>1$. They are needed in our reconstruction algorithms to relate $\mathbb{L}_2$-features to the Sobolev features on $\mathbb{H}^{k,2}_\gamma(\mathbb{R}^2)$, $k>1$. Due to numerical limitations (sampling) the cases which are most practically relevant are $k=0,1,2,3,4$. The (isotropic) functions $\mathbf{x}\mapsto R^2_{k,\gamma}(\mathbf{x})$, which are illustrated in figure 3 , are bounded iff $k>1$, which coincides with the fact that $\mathbb{H}^{k,2}_\gamma(\mathbb{R}^2)$ is a functional Hilbert space iff $k>1$.

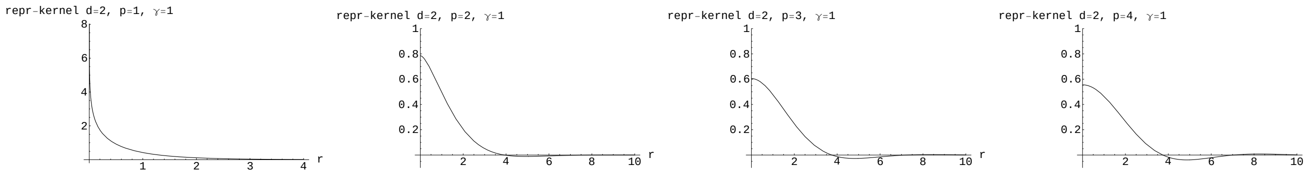


Fig. 3. Plots of the graphs of the functions $\|\mathbf{x}\| \mapsto R^2_{k,\gamma}(\mathbf{x})$, for $k = 1, 2, 3, 4$.

Theorem 4. *The space $\mathbb{H}^{k,2}_{\gamma}(\mathbb{R}^d)$ is a reproducing kernel space iff $k > d/2$. The reproducing kernel (i.e. the Riesz-representant of the continuous point evaluation, $F(\boldsymbol{a}) = (R^d_{\gamma,k,\boldsymbol{a}}, F)_{\mathbb{H}^{k,2}_{\gamma}(\mathbb{R}^d)}$) is then given by*

$$R^d_{\gamma,k,\boldsymbol{a}}(\boldsymbol{x}) = \frac{1}{2\pi\gamma^{2k}}\mathcal{F}^{-1}(\boldsymbol{\omega} \mapsto e^{i\boldsymbol{\omega}\cdot\boldsymbol{a}}\frac{1}{\gamma^{-2k} + \|\boldsymbol{\omega}\|^{2k}})(\boldsymbol{x}) = r^{-\frac{d-2}{2}}\frac{1}{2\pi}\int_0^\infty \rho^{\frac{d}{2}} J_{\frac{d-2}{2}}(\rho r)\frac{1}{1+\gamma^{2k}\rho^{2k}}\mathrm{d}\rho \ , \quad (23)$$

where $r = \|\boldsymbol{x} - \boldsymbol{a}\| > 0$, $\rho = \|\boldsymbol{\omega}\|$ and satisfies the following recursion: $R^d_{\gamma,2k,\boldsymbol{a}} = \frac{1}{2}\left(R^d_{e^{\frac{i\pi}{4k}}\gamma,k,\boldsymbol{a}} + R^d_{e^{\frac{-i\pi}{4k}}\gamma,k,\boldsymbol{a}}\right)$.
In case [7] $d = 2$ we have $R^{d=2}_{\gamma,k,\boldsymbol{a}}(\boldsymbol{x}) = \frac{1}{2\pi\gamma^2} G^{k+1,0}_{0,2k}\left(\frac{\rho^{2k}}{\gamma^{2k}k^k} \mid \boldsymbol{a}_k\right)$, where $\boldsymbol{a}_k = \{a_{kj}\}_{j=1}^{2k} \in \mathbb{Q}^{2k}$ is given by $a_{kj} = \frac{j-1}{k}$ for $1 \leq j \leq k$, $a_{k(k+1)} = \frac{k-1}{k}$ and $a_{kj} = a_{k(j-k-1)}$ for $k+2 \leq j \leq 2k$.
In particular we have $R^{d=2}_{\gamma,k=1,\boldsymbol{a}}(\boldsymbol{x}) = \frac{1}{2\pi\gamma^2}K_0\left(\frac{r}{\gamma}\right)$ and $R^{d=2}_{\gamma,k=2,\boldsymbol{a}}(\boldsymbol{x}) = -\frac{1}{2\pi\gamma^2}kei_0(r/\gamma)$, with $r = \|\boldsymbol{x} - \boldsymbol{a}\|$

References

1. Duits, R., Florack, L., de Graaf, J., ter Haar Romeny, B.: On the axioms of scale space theory. Journal of Mathematical Imaging and Vision **20** (2004) 267–298
2. Duits, R., Duits, M., van Almsick, M.: Invertible orientation scores as an application of generalized wavelet theory. Technical report, TUE, Eindhoven (2004) Technical Report 04-04, Biomedical Image and Analysis, Department of Biomedical Engineering, Eindhoven University of Technology.
3. Loog, M., Duistermaat, J.J., Florack, L.M.J.: On the behavior of spatial critical points under Gaussian blurring. a folklore theorem and scale-space constraints. [14] 183–192
4. Kuijper, A., Florack, L.M.J.: Hierarchical pre-segmentation without prior knowledge. In: Proceedings of the 8th International Conference on Computer Vision (Vancouver, Canada, July 9–12, 2001), IEEE Computer Society Press (2001) 487–493
5. Balmachnova, E., Florack, L., Platel, B., Kanters, F., ter Haar Romeny, B.: Stability of top-points in scale space. (Proceedings 5th Scale Space conference 2005) 62–72
6. Felsberg, M., Duits, R., Florack, L.: The monogenic scale space on a bounded domain and its applications. Proceedings Scale Space Conference, Isle of Skye, UK. (2003) 209–224
7. Nielsen, M., Lillholm, M.: What do features tell about images? [14] 39–50
8. Lillholm, M., Nielsen, M., Griffin, L.D.: Feature-based image analysis. International Journal of Computer Vision **52** (2003) 73–95
9. Kanters, B. Platel, L.M.J. Florack and B.M. ter Haar Romeny, F.: Image reconstruction from multiscale critical points. In Griffin, L., Lillholm, M., eds.: Scale Space Methods in Computer Vision, 4th International Conference, Scale Space 2003, Isle of Skye, UK, Springer (2003) 464–478
10. Yosida, K.: Functional Analysis. Springer Verlag, Berlin, Heidelberg, New York (1980)
11. Robinson, D.: Elliptic Operators and Lie groups. Clarendon Press, Oxford, New York, Tokyo (1991)
12. Janssen, B., Kanters, F., Duits, R., Florack, L.M.J., ter Haar Romeny, B.M.: A linear image reconstruction framework based on sobolev type inner products. (To Appear in Proc. of the Scale Space Conference 2005)
13. Kanters, F., Lillholm, M., Duits, R., Janssen, B., Platel, B., Florack, L.M.J.: Image reconstruction from multiscale top points. (To appear at Proceedings 5th Scale Space Conference 2005)
14. Kerckhove, M., ed.: Scale-Space and Morphology in Computer Vision: Proceedings of the Third International Conference, Scale-Space 2001, Vancouver, Canada. Volume 2106 of Lecture Notes in Computer Science. Springer-Verlag, Berlin (2001)

[7] Where we notice that J_μ is a Bessel function (first kind) of order μ and K_0 equals the well-known BesselK-function of order 0 and the Kelvin function $\mathrm{kei}_0(v) = \frac{1}{2i}\left(K_0\left(e^{\frac{+i\pi}{4}}v\right) - K_0\left(e^{\frac{-i\pi}{4}}v\right)\right)$ and $G^{k+1,0}_{0,2k}$ denotes a Meyer-G function.