Multiscale Analysis for Images on Riemannian Manifolds∗

Felipe Calderero† and Vicent Caselles‡

Dedicated to the memory of Prof. Vicent Caselles

Abstract. In this paper we study multiscale analyses for images defined on Riemannian manifolds and extend the axiomatic approach proposed by Álvarez, Guichard, Lions, and Morel to this general case. This covers the case of two- and three-dimensional images and video sequences. After obtaining the general classification, we consider the case of morphological scale spaces, which are given in terms of geometric equations, and the linear case given by the Laplace–Beltrami flow. We consider in some detail the case of image metrics given in terms of the structure tensor and compute some cases of such a tensor for video. Then we comment on the connections with variational formulations of image diffusion comparing the anisotropies that appear. Finally, we include numerical experiments illustrating some of the models. Namely, we compare some examples for still images using the Laplace–Beltrami flow and some variational models. We also consider several examples in video: the mean curvature motion and the extension of the morphological and Galilean invariant scale spaces to the video manifold case, and the Laplace–Beltrami flow. We point out that the number of models that appear is huge, and we have restricted ourselves to such cases for the sake of brevity and illustration.

Key words. multiscale analysis, structure tensor, curvatures

AMS subject classifications. 68U10, 35Q80, 46N60

DOI. 10.1137/130923142

1. Introduction. Our purpose in this paper is to study multiscale analyses for images defined on Riemannian manifolds. A multiscale analysis represents a given image at different scales of smoothing, the scale being related to the size of the neighborhood which is used to give an estimate of the brightness of the picture at a given point. It is a basic preprocessing step for shape recognition [46] (see [31, 15] and references therein).

The systematic study of multiscale analyses for images was the purpose of the axiomatic approach proposed in [2]. Based on a series of axioms which define the structure of the multiscale space and a set of geometric and photometric invariants, multiscale analyses were defined in terms of (viscosity) solutions of a parabolic equation. In the case of linear multiscale analysis they obtained the Gaussian scale space (already proposed and studied in [47, 42, 43, 73, 34, 33, 45, 72], etc., using also an axiomatic approach in some of those papers). Besides
the Gaussian scale space, classification covers many of the classical models that were proposed in the literature, like the Perona–Malik equation [57] (see also [20]), the Rudin–Osher–Fatemi model [59], or the mean curvature motion as proposed in [3].

Assuming the invariance under contrast changes (i.e., monotone rearrangements of the gray levels), multiscale analyses were given in terms of geometric equations [27, 26, 56] that diffuse the level sets of the image with functions of their principal curvatures. Following [2, 31] we refer to them as morphological scale spaces since they are related to a PDE formulation of mathematical morphology [62]. The case \( N = 2 \) is of particular interest and leads to the motion of level lines by a function of curvature [51, 28, 37, 38]. Of particular interest is the affine morphological scale space (AMSS) [2, 60, 61, 54], which is affine invariant and corresponds to the motion of level lines by the power \( 1/3 \) of its curvature. The case of scale spaces for three-dimensional (3D) images gives rise to geometric motions that depend on functions of the two principal curvatures of the level surfaces [2, 55]. In particular, the AMSS is related to the power \( 1/4 \) of the positive part of Gaussian curvature (times the sign of mean curvature) and has been studied in [19, 53, 52, 66]. Of particular interest is the case of video sequences. In that case, the Galilean invariant scale spaces are characterized similarly to the 3D case and Gaussian curvature is replaced by acceleration [2, 31, 29, 30].

Scale spaces based on anisotropic diffusion have been the object of systematic study by Weickert [70, 68], and, although they fall into the general set of nonlinear models described in [2], they were not axiomatically studied there.

Variational models also give a different approach to image diffusion. They are also basic ingredients in the regularization of inverse problems. Let us mention here the work of Rudin, Osher, and Fatemi [59], who introduced total variation as an image regularizer by its ability to restore edges. A more general formulation is given in [40, 64, 39, 63], where the authors consider images defined on Riemannian manifolds where the metric depends on the image and reflects the anisotropy of the underlying problem (for edge preservation, for color image restoration, for texture analysis, etc.). Their basic energy functional is the Polyakov action, which is the extension of the Dirichlet integral to maps between Riemannian manifolds [40, 64]. Note that the definition of images on Riemannian manifolds also allows us to work on images and videos that may not necessarily be defined on a flat support. For instance, this framework can be naturally applied to images and videos suffering from severe lens distortion, captured by cameras with a nonplanar geometry (such as a fisheye lens), or imaged on curved mirrors such as those used in art and illusion.

The axiomatic approach used to classify scale spaces was also used in [17] in order to classify interpolation operators according to a set of structural requirements and invariances. Examples are given by the Laplace equation, the AMLE, or the interpolation of level lines by straight lines (related to inpainting/disocclusion [49, 48]). This approach was later extended to image interpolation on surfaces in [16].

Our purpose in this paper is to extend the classification of multiscale analyses to images defined on Riemannian manifolds. From the analytic point of view the two basic ingredients are the papers [2] and [17, 16], and our results are an extension of them. Our motivation stems from several facts. First, we attempt to include anisotropic scale spaces in the general formalism of multiscale analysis. Second, we reflect the underlying basic models that appear in different contexts and share the same ideas. In this context, in a variational setting, it
is interesting to point out the role played by the Polyakov action as a unifying principle. Third, we attempt to define new scale spaces for video that incorporate motion estimates and anisotropies. This will be done by defining a metric on the domain of the video that computes the distance between moving particles and incorporates edge preserving anisotropies. Let us say at this point that more than for novelty (although some will appear), we aim for an analysis that unifies and makes systematic the definition of general scale spaces in different contexts, making explicit the common framework.

This paper is divided in two main parts. The first part (including sections 2, 3, 4, 5, 6, and 7) presents the technical details, while the second (section 8) is devoted to numerical experiments. Particularly, the study of morphological scale spaces and the classification of linear scale spaces are the two main purposes of the first part of the paper. Besides that, let us mention in particular three other points. First, we discuss in some detail the problem of invariances with respect to geometric transformations of the image plane (rotations and affine maps). Note that the nonlinear structure of the manifold makes it much more difficult to impose this set of invariances as axioms. Instead, we define the multiscale analysis as an intrinsic object and thus independent of the manifold parameterization. As a consequence we are imposing some invariances with respect to transformations in the tangent plane around each point, namely rotation invariance. This suffices to characterize morphological scale spaces in terms of geometric equations that depend on the principal curvatures of the level sets of the image. On the other hand we do not consider translation and affine invariance. Translation invariance could be imposed by requiring the invariance with respect to maps induced in the tangent spaces at two points by parallel transport. But we have not considered this in detail since we want to preserve the anisotropy at each point. We have also not included in our general formulation the invariance with respect to affine transformations. It is interesting to note at this point that in the case where the image domain is $\mathbb{R}^N$ the manifold point of view permits us to construct some examples of affine invariant multiscale spaces (even new linear and affine invariant multiscale spaces). This deserves a more detailed study (its performance and applications), which we postpone to a subsequent paper to avoid a longer extension of the present paper [7]. Second, in the context of manifolds, multiscale analyses for video fall into the same context as multiscale analyses of 3D images, the only difference being due to the specific metric for video. This is related to the fact that we have not considered the invariance with respect to Galilean transformations (which is included as an aspect of our intrinsic treatment). Third, we give different examples of metrics and we study the role of the structure tensor as a metric. As is well known, the structure tensor is seen as a metric in the image plane [70, 71, 39, 11, 10, 58]. Here we illustrate how the structure tensor is the natural metric induced in the image plane by considering the image as a manifold of patches. This permits us to compute examples of structure tensors for video sequences.

In the second part of the paper we will describe the experiments that illustrate the models developed in this part. Since the number of models that appear is enormous, we will not give a full numerical experimentation on them. Instead, we will select a few of them in order to illustrate the main concepts, leaving a more deep study for later work.

Although their study goes beyond the scope of this work, the current framework opens a wide range of new and challenging problems. Particularly, we would like to mention the topics related to numerical implementation and discretization of the operators associated with
the multiscale analyses. For instance, those include the discretization of classical multiscale models (such as the scale-space theory in [73]), the multiscale mathematical morphology, or other particular cases, such as the mean curvature motion. The efficient implementation of the multiscale analyses also deserves further attention, for instance, the approximation in some cases by the convolution with spatially varying kernels.

Let us finally summarize the plan of the paper. In section 2 we collect some basic notation and definitions about Riemannian manifolds. In section 3 we define the basic set of axioms satisfied by multiscale analyses for images defined on Riemannian manifolds and express them in terms of solutions of a (eventually degenerate) parabolic equation. We give some examples, namely for video, and study in more detail the case of images defined in two-dimensional (2D) manifolds, where a more precise classification can be given. In section 4 we consider multiscale analyses that commute with contrast changes, leading to morphological scale spaces that operate on the level sets of the image (by functions of their curvatures). They are expressed as (viscosity) solutions of geometric equations defined on Riemannian manifolds. They represent a rich family of scale spaces, and we are far from understanding all possibilities that these models contain. In particular, we note that due to the manifold structure, we are not discriminating between the case of (geometric) 3D images (e.g., medical images) and video besides the different metrics that can be assigned to them. In particular, as the acceleration studied in [30] can be considered as a principal curvature, functions of principal curvatures of the manifold (e.g., the Gaussian curvature operator) contain the basic operators considered in the formalization of [2, 30], at least when the motion of the video is reflected by Galilean motions. Otherwise we take into account the optical flow computed on the given video sequence. In section 5 we consider the case of linear multiscale analyses, naturally obtaining that they are expressed as solutions of the Laplace–Beltrami equation. In section 6 we study the possibility of defining metrics in terms of structure tensors. First, in subsection 6.1, we identify structure tensors as the natural metric obtained by considering the image (assuming it to have a square integrable gradient) as a manifold of patches in $L^2$ of the patch domain. Then in subsection 6.2 we compute some structure tensors for video. We end these theoretical developments by some discussion in section 7 on diffusion operators defined by variational models written in terms of the Dirichlet integral (Polyakov action) on the Riemannian manifold (following the formulation in [40, 64, 39, 63]). This reflects on one hand the unifying underlying principle reflected by the Dirichlet integral and on the other the main differences with multiscale analyses.

Section 8 is devoted to numerical experiments. As we already said, we cannot cover the whole set of possibilities, and some interesting analysis (e.g., morphological cases for still images, the affine invariant ones) have been left for later consideration. We have restricted ourselves to some linear scale spaces for still images (subsection 8.1), to reflect their main difference with respect to diffusions written in terms of the gradient flow of the Dirichlet integral. Then in subsection 8.2 we consider the case of a video with a metric on the domain of the video that computes the distance between moving particles and contains edge preserving anisotropies. In subsection 8.2.1 we have studied the case of mean curvature motion in the video equipped with the above metric (which amounts to considering the case of a scale space based on a motion compensated median filter). Similarly, in subsection 8.2.2 we consider the extension of the Galilean invariant scale space in [2, 30] to the video manifold as defined above.
This amounts again to a numerical scheme similar to that developed by Guichard in [30] with the motion compensated video. Finally, we conclude by considering in subsection 8.2.3 some cases of linear scale space in the video manifold. These experiments illustrate the features of the models, and we have to mention that we have left for later consideration some interesting cases that require a more exhaustive exploration.

2. Preliminaries. We collect in this section some basic notation and definitions about Riemannian manifolds.

Let $(\mathcal{M}, g)$ be a smooth Riemannian manifold. As a particular case we can consider $\mathcal{M} = \mathbb{R}^N$ (or a domain in $\mathbb{R}^N$) endowed with a general metric $g_{ij}$. As usual, given a point $\xi \in \mathcal{M}$, we denote by $T_\xi \mathcal{M}$ the tangent space to $\mathcal{M}$ at the point $\xi$. By $T^*_\xi \mathcal{M}$ we denote its cotangent space.

Let $\xi$ be a point on $\mathcal{M}$, $U \subseteq \mathbb{R}^N$ be an open set containing 0, and $\psi : U \rightarrow \mathcal{M}$ be any coordinate system such that $\psi(0) = \xi$. Let $g_{ij}(\xi)$ and $\Gamma^k_{ij}(\xi)$ (indices $i, j, k$ run from 1 to $N$) denote, respectively, the coefficients of the first fundamental form of $\mathcal{M}$ and the Christoffel symbol computed in the coordinate system $\psi$. For simplicity we shall denote by $G(\xi)$ the (symmetric) matrix $(g_{ij}(\xi))$ and by $\Gamma^k(\xi)$ the matrix formed by the coefficients $(\Gamma^k_{ij}(\xi))$, $i, j, k = 1, \ldots, N$.

The scalar product of two vectors $v, w \in T_\xi \mathcal{M}$ will be denoted by $\langle v, w \rangle$, and the action of a covector $p^* \in T^*_\xi \mathcal{M}$, on a vector $v \in T_\xi \mathcal{M}$, will be denoted by $\langle p^*, v \rangle = p^*_i v^i$, where we use Einstein’s convention that repeated indices are summed. This convention will be used in the rest of the paper. Let $\psi : U \rightarrow \mathcal{M}$ be a coordinate system such that $\psi(0) = \xi$, and $g_{ij}(\xi)$ are the coefficients of the first fundamental form at $\xi \in \mathcal{M}$ in $\psi$. Then, if $v, w \in T_\xi \mathcal{M}$, we have $\langle v, w \rangle_\xi = g_{ij}(\xi) v^i w^j$, where $v^i, w^j$ are the coordinates of $v, w$ in the basis $\frac{\partial}{\partial x^i}|_\xi$ of $T_\xi \mathcal{M}$ (if $\xi \in \mathcal{M}$ is fixed, we will also denote $(v, w)_\xi$ instead of $\langle v, w \rangle_\xi$). Using this basis for $T^*_\xi \mathcal{M}$ and the dual basis on $T^*_\xi \mathcal{M}$ if $p^* \in T^*_\xi \mathcal{M}$ and $v \in T_\xi \mathcal{M}$, we have $\langle p^*, v \rangle = p^*_i v^i$. Notice that we may write $\langle p^*, v \rangle = g^{ij}(\xi) p^i v^j$, where $p^i$ are the coordinates of the vector $p$ associated with the covector $p^*$. The relation between both coordinates is given by

$$p_i = g_{ij}(\xi) p^j \quad \text{or} \quad p^j = g^{ij}(\xi) p_j,$$

where $g^{ij}(\xi)$ denotes the coefficients of the inverse matrix of $g_{ij}(\xi)$. We shall write (2.1) as

$$p^* = Gp \quad \text{or} \quad p = G^{-1} p^*.$$

In this way $G : T_\xi \mathcal{M} \rightarrow T^*_\xi \mathcal{M}$. In the case that $\psi$ is a geodesic coordinate system centered at $\xi$, the matrix $G$ is the identity matrix $I = (\delta_{ij})$, and $I$ maps vectors to covectors, i.e., $I : T^*_\xi \mathcal{M} \rightarrow T^*_\xi \mathcal{M}$ (with the same coordinates in the dual basis). We shall denote by $I^{-1}$ the inverse of $I$, mapping covectors to vectors.

If $U \subseteq \mathbb{R}^N$ and $\psi : U \rightarrow \mathcal{M}$ is a coordinate system with $\psi(0) = \xi$, then $\psi \circ d\psi(0)^{-1} : U' \subseteq T_\xi \mathcal{M} \rightarrow \mathcal{M}$ is a new coordinate system. If we identify $T_0 \mathcal{M}$ with $\mathbb{R}^N$ and $\{e_i\}$ denotes its canonical basis, then $e'_i = d\psi(0)e_i$ satisfy $\langle e'_i, e'_j \rangle = g_{ij}(\xi)$. From now on, we shall use this identification; thus we shall interpret that any coordinate system around a point $\xi \in \mathcal{M}$ is defined on a neighborhood of 0 in the tangent space $T_\xi \mathcal{M}$. 
Maps. Symmetric maps. Quadratic forms. We shall also use this coordinate system to express a bilinear map $A : T_{\xi}M \times T_{\xi}M \rightarrow \mathbb{R}$. Indeed, if $(A_{ij})$ is the matrix of $A$ in this basis and $v, w \in T_{\xi}M$, we may write $A(v, w) = A_{ij}v^jw^i$. If $A^i_j = g^{ik}(\xi)A_{kj}$, then $A^i_j$ determines a map, denoted by $A : T_{\xi}M \rightarrow T_{\xi}M$, such that $A(v, w) = \langle Av, w \rangle = (GAv, w)$. Observe that $G(\xi)A : T_{\xi}M \rightarrow T_{\xi}M$. Observe also that our notation $A^i_j$ already indicates that $A = (A^i_j)$ maps vectors to vectors. In our notation, frequently we shall not distinguish between matrices and maps.

As usual, we say that a linear map $C : T_{\xi}M \rightarrow T_{\xi}M$ is symmetric if $(Cv, w) = (Cw, v)$ for any $v \in T_{\xi}M$, $w \in T_{\xi}M$. From now on, we shall use the notation

\[ SM_{\xi}(N) := \{ A : T_{\xi}M \rightarrow T_{\xi}M, A \text{ is symmetric} \}. \]

We shall also write

\[ S_{\xi}(N) := \{ A : T_{\xi}M \rightarrow T_{\xi}M, G(\xi)A \in SM_{\xi}(N) \}. \]

If we want to stress that $G(\xi)$ is the metric in $T_{\xi}M$, we shall write $(T_{\xi}M, G(\xi))$ and denote $SM_{\xi}(N, G), S_{\xi}(N, G)$, instead of $SM_{\xi}(N), S_{\xi}(N)$, respectively.

If $A \in S_{\xi}(N), v \in T_{\xi}M, c \in \mathbb{R}$, we define the quadratic polynomial

\[ Q(x) = \frac{1}{2} \langle Ax, x \rangle + \langle v, x \rangle + c, \quad x \in T_{\xi}M. \]

Note that

\[ Q(x) = \frac{1}{2} \langle A'x, x \rangle + \langle p, x \rangle + c, \quad x \in T_{\xi}M, \]

where $A' = G(\xi)A \in SM_{\xi}(N), p = G(\xi)v \in T_{\xi}M$.

Notice that if $A : T_{\xi}M \rightarrow T_{\xi}M$, we define $A^t : T_{\xi}^*M \rightarrow T_{\xi}^*M$ by

\[ (A^t_p, v) = \langle p, Av \rangle \quad \forall v \in T_{\xi}M, p \in T_{\xi}^*M. \]

We define $A^{t,g} : T_{\xi}M \rightarrow T_{\xi}M$ by

\[ \langle A^{t,g}v, w \rangle = \langle v, Aw \rangle \quad \forall v, w \in T_{\xi}M. \]

From now on, when the point $\xi \in \mathcal{M}$ is understood we write $G$ instead of $G(\xi)$. Notice that $GA^{t,g} = A^{t}G$.

If $A \in S_{\xi}(N)$, then $GA \in SM_{\xi}(N)$ and $(GAv, w) = \langle v, GAw \rangle$, that is, $\langle Av, w \rangle = \langle v, Aw \rangle$.

That is, $A^{t,g} = A$.

Rotations in the tangent space. Let us define a rotation $R : T_{\xi}M \rightarrow T_{\xi}M$ as a linear map that satisfies

\[ \langle Rv, Rw \rangle = \langle v, w \rangle \quad \forall v, w \in T_{\xi}M. \]

Notice that rotations satisfy

\[ R^tG = G. \]
Let $B : T_{\xi}\mathcal{M} \rightarrow T_{\xi}\mathcal{M}$ be a matrix such that $B^{-1}B^t = G^{-1}$. Thus $B^tGB = I$ and $B$ is mapping an orthonormal basis of $(T_{\xi}\mathcal{M}, I)$ to an orthonormal basis of $(T_{\xi}\mathcal{M}, G(\xi))$.

If $R : T_{\xi}\mathcal{M} \rightarrow T_{\xi}\mathcal{M}$ is a rotation, then

$$(B^{-1}RB)^tJ B^{-1}RB = I.$$ 

That is, $B^{-1}RB$ is a classical rotation.

**Gradient and Hessian.** Given a function $u$ on $\mathcal{M}$, let us denote by $D_\mathcal{M}u$ and $D^2_{\mathcal{M}}u$ the gradient and Hessian of $u$, respectively. In a coordinate system, $D_\mathcal{M}u$ is the covector $\frac{\partial u}{\partial x}$, and $D^2_{\mathcal{M}}u$ is the matrix $\frac{\partial^2 u}{\partial x_i \partial x_j} - \Gamma^k_{ij} \frac{\partial u}{\partial x^k}$ which acts on tangent vectors. Thus, with this notation $D^2_{\mathcal{M}}u(\xi) : T_{\xi}\mathcal{M} \times T_{\xi}\mathcal{M} \rightarrow \mathbb{R}$ is a bilinear map, $\xi \in \mathcal{M}$. Let us write $\nabla_{\mathcal{M}}u$ as the vector of coordinates $g^{ij} \frac{\partial u}{\partial u^j}$. Then $\|\nabla_{\mathcal{M}}u(\xi)\|^2 = \langle \nabla_{\mathcal{M}}u(\xi), \nabla_{\mathcal{M}}u(\xi) \rangle_{\xi}$. To simplify our notation we shall write $\nabla u$ and $\Delta u$ instead of $D_\mathcal{M}u$ and $\nabla_{\mathcal{M}}u$. The vector field $\nabla u$ satisfies $\langle \nabla u, v \rangle_{\xi} = du(v)$, $v \in T_{\xi}\mathcal{M}$, $du$ being the differential of $u$.

3. **Multiscale analyses on images defined on $\mathcal{M}$.** Let $C_b(\mathcal{M})$ denote the space of bounded continuous functions in $\mathcal{M}$ with the maximum norm. We think of $C_b(\mathcal{M})$ as the space of images on $\mathcal{M}$. We denote by $C^\infty_b(\mathcal{M})$ the space of infinitely differentiable functions on $\mathcal{M}$.

Following [2], recall that a multiscale analysis is defined as a family of transforms $(T_s)_{s \geq 0}$ which, when applied to the original function $u(x)$, yield a sequence of functions $u(s, x) = T_s u(x)$. In the current context, assume that $T_s : C_b(\mathcal{M}) \rightarrow C_b(\mathcal{M})$ is a nonlinear operator for any $s \geq 0$ and $u(s, x) = T_s u(x)$, $u \in C_b(\mathcal{M})$.

Let $(\kappa) := \kappa_n$ be an increasing sequence of nonnegative constants:

$$Q((\kappa)) := \{ u \in C^\infty_b(\mathcal{M}) : \|D^\alpha u\|_\infty \leq \kappa_n \forall n \geq 0, \forall |\alpha| \leq n \}.$$ 

As usual, $O(f)$ (resp., $o(f)$) will denote any expression which is bounded by $C|f|$ for some constant $C > 0$ (resp., such that $\frac{o(f)}{|f|} \rightarrow 0$ as $f \rightarrow 0$).

Let us summarize here the sets of different axioms that we consider below.

**Architectural axioms:**

- **Recursivity** $T_0(u) = u$, $T_s(T_t u) = T_{s+t} u \forall s, t \geq 0, \forall u$. This axiom is commonly referred as the semigroup property.

- **Infinitesimal generator** $\frac{T_{ah}u-u}{h} \rightarrow A(u)$ as $h \rightarrow 0^+$. This holds for any $u \in C^\infty_b(\mathcal{M})$.

- **Regularity axiom** $\|T_s(u + h\tilde{u}) - (T_s(u) + h\tilde{u})\|_\infty \leq Mh s \forall h, s \in [0, 1], \forall u, \tilde{u} \in Q((\kappa))$, where the constant $M$ depends on $Q((\kappa))$.

- **Locality** $T_s(u)(\xi) - T_s(\tilde{u})(\xi) = o(s)$ as $s \rightarrow 0^+$, $\xi \in \mathcal{M}$, $\forall u, \tilde{u} \in C^\infty_b(\mathcal{M})$ such that $D^\alpha u(\xi) = D^\alpha \tilde{u}(\xi)$ for all multi-indices $\alpha$.

**Comparison principle:**

- **Comparison principle** $T_s u \leq T_s \tilde{u} \forall s \geq 0, \forall u, \tilde{u} \in C^\infty_b(\mathcal{M})$ such that $u \leq \tilde{u}$. 


Morphological axioms:

[Gray level shift invariance] $T_s(0) = 0$, $T_s(u + \kappa) = T_s(u) + \kappa \forall s \geq 0$, $\forall u \in C_b^\infty(M)$, $\forall \kappa \in \mathbb{R}$.

[Gray scale invariance] $T_s(f(u)) = f(T_s(u)) \forall s \geq 0$, $\forall u \in C_b^\infty(M)$, and for any strictly increasing function $f : \mathbb{R} \to \mathbb{R}$.

Remark 1. We shall not explicitly consider geometric invariances. But, as we shall point out below, invariance with respect to rotations in the tangent plane is implicitly contained in the architectural axioms that imply that multiscale analyses are intrinsic, in the sense that they do not depend on a particular parameterization. In [7] we shall also consider $\mathbb{R}^N$ as a manifold with a given metric and analyze the issue of affine invariance.

Theorem 3.1. Let $T_s$ be a multiscale analysis satisfying the recursivity, infinitesimal, and regularity axioms. Then $A(u_r) \to A(u)$ in $C_b(M)$ if $u_r, u \in C_b^\infty(M)$ and $D^\alpha u_r \to D^\alpha u$ in $C_b(M) \forall \alpha$ with $|\alpha| \geq 0$.

Proof. We follow along the lines of the corresponding result in [2]. Let $u_r, u \in C_b^\infty(M)$ and $D^\alpha u_r \to D^\alpha u$ in $C_b(M) \forall \alpha, |\alpha| \geq 0$. Let us write $u_r = u + h_r E_r$, where $E_r = \frac{u - u_r}{h_r}$. We later will make precise the value of $h_r$; let us only say that $h_r \to 0$. We have

$$\frac{T_s u_r - u_r}{s} - \frac{T_s u - u}{s} = \frac{T_s(u + h_r E_r) - (T_s u + h_r E_r)}{s}.$$

Let us consider the distance in the Fréchet space $C_b^\infty(M)$:

$$d(u_1, u_2) := \sum_{k=0}^{\infty} \frac{1}{2^k} \max_{|\alpha| = k} \|D^\alpha u_1 - D^\alpha u_2\|_\infty.$$

Observe that $d(u_r, u) \to 0$ as $r \to \infty$. Let

$$C(k) := \max_r \max_{|\alpha| \leq k} \|D^\alpha u_r - D^\alpha u\|_\infty.$$

We could take $|\alpha| = k$ and then make $C(k)$ increasing. Since

$$\frac{\max_{|\alpha| = k} \|D^\alpha u_r - D^\alpha u\|_\infty}{1 + \max_{|\alpha| = k} \|D^\alpha u_r - D^\alpha u\|_\infty} \leq 2^k d(u_r, u),$$

we have

$$\max_{|\alpha| = k} \|D^\alpha u_r - D^\alpha u\|_\infty \leq 2^k (1 + C(k)) d(u_r, u).$$

Let us choose $h_r = d(u_r, u)$. Then for any $|\alpha| = k$ we have

$$\|D^\alpha E_r\|_\infty = \|D^\alpha u_r - D^\alpha u\|_\infty \leq 2^k (1 + C(k)) \forall r.$$

That is, $\{E_r\}_r \subseteq Q((\kappa))$ for $\kappa_k := 2^k (1 + C(k))$.

Then, by the regularity axiom, we have

$$\frac{\|T_s(u + h_r E_r) - (T_s u + h_r E_r)\|_\infty}{s} \leq M h_r,$$
where the constant $M$ depends on $Q((\kappa))$ and thus does not depend on $r$. Letting $s \to 0+$ we have
\[
\|A u_r - A u\|_\infty \leq M h_r.
\]
Our conclusion follows.

**Theorem 3.2.** Let $T_s$ be a multiscale analysis satisfying all architectural axioms and the comparison principle. Then there exists a function $F : SM_\xi(N) \times T^*_\xi M \times \mathbb{R} \times M \to \mathbb{R}$ increasing with respect to its first argument such that
\[
\frac{T_s u - u}{s} \to F(D^2(u \circ \psi)(0), D(u \circ \psi)(0), u(\xi), \xi, G, \Gamma^k) \quad \text{in } C_b(M) \text{ as } s \to 0+
\]
\[
\forall u \in C^\infty_b(M), \psi \text{ a coordinate system around } \xi \in M. \text{ The function } F \text{ is continuous in its first three arguments. If we assume that } T_s \text{ is gray level shift invariant, then the function } F \text{ does not depend on } u.
\]
Notice that we did not denote explicitly the argument $\xi$ for $G, \Gamma^k$. Notice that the first argument in $F$ is a symmetric map from $T_\xi M$ to $T^*_\xi M$.

We denote by $B_r(\xi)$ the geodesic ball of radius $r > 0$ and center $\xi \in M$.

**Proof.** Again, we follow along the lines of the corresponding result in [2]. Let $u_k \in C^\infty_b(M)$, $k = 1, 2$, be two functions which are $C^2$ at the point $\xi \in M$. Let $\psi : U \to M$ be a coordinate system such that $\psi(0) = \xi$. Let us consider the Taylor expansion of $u_k \circ \psi$ around 0:
\[
u_k \circ \psi(x) = u_k \circ \psi(0) + \frac{\partial(u_k \circ \psi)}{\partial x^i}(0)x^i + \frac{1}{2} \frac{\partial^2(u_k \circ \psi)}{\partial x^i \partial x^j}(0)x^i x^j + O(|x|^3), \quad k = 1, 2.
\]

Assume that both Taylor expansions to second order coincide. Let us prove that
\[
A(u_1)(\xi) = A(u_2)(\xi).
\]
Let $f_\epsilon(\zeta) = u_1(\zeta) + \epsilon |\psi^{-1}(\zeta)|^2$. Clearly, $f_\epsilon(\zeta) \geq u_2(\zeta)$ for $\zeta \in B_r(\xi)$ for some $r > 0$. Let $w$ be a smooth test function with support in $B_r(\xi)$ such that $w(\xi) = 1$ on $B_{r/2}(\xi)$. Let $\tilde{f}_\epsilon = w f_\epsilon + (1-w)u_2$. Then $\tilde{f}_\epsilon$ has the same derivatives as $u_1$ at $\xi$ and $\tilde{f}_\epsilon \geq u_2$. Then, by the comparison principle, $T_s \tilde{f}_\epsilon \geq T_s u_2$. Since $\tilde{f}_\epsilon(\xi) = u_2(\xi)$, then using the locality principle we have
\[
\frac{T_s u_1(\xi) - u_1(\xi)}{s} + o(s) = \frac{T_s \tilde{f}_\epsilon(\xi) - \tilde{f}_\epsilon(\xi)}{s} \geq \frac{T_s u_2(\xi) - u_2(\xi)}{s}.
\]
Letting $s \to 0+$ we have
\[
A(u_1)(\xi) \geq A(u_2)(\xi).
\]
By symmetry, we obtain (3.1).

This implies that for any $u \in C^\infty_b(M)$ which is a function of class $C^2$ at the point $\xi \in M$, we may write
\[
A(u)(\xi) = F(D^2(u \circ \psi)(0), D(u \circ \psi)(0), u(\xi), \xi, G, \Gamma^k),
\]
where we made explicit the dependence of $F$ on the metric and the connection since they appear in the Taylor expansion of $u$ written in intrinsic coordinates. Then Theorem 3.1 proves the continuity of $F$ with respect to its first three arguments.
The monotonicity with respect to its first argument follows as a consequence of Lemma 3.3 below.

Remark 2. We could also have written $F$ as a function $\hat{F} : S_\xi(N) \times T_\xi M \times \mathbb{R} \times M \to \mathbb{R}$ so that $\hat{F}(A, v, c, \xi, G, \Gamma^k) = F(GA, Gv, c, \xi, G, \Gamma^k)$.

Lemma 3.3. Let $\xi \in M$, and let $\psi : U \to M$ be a coordinate system around $\xi$. Let $G, \Gamma^k$ be the metric coefficients and the Christoffel symbols of $M$ in the coordinate system $\psi$ at the point $\xi$. Let $A_1, A_2 : T_\xi M \to T_\xi^* M$ be two matrices such that $A_1, A_2$ are symmetric, $p \in T_\xi M$, $c \in \mathbb{R}$. If the symmetric matrix given by $A_2 - A_1$ is a positive semidefinite matrix (in that case, we will write $A_1 \leq A_2$), then

$$F(A_1, p, c, \xi, G, \Gamma^k) \leq F(A_2, p, c, \xi, G, \Gamma^k).$$

Thus $F$ is elliptic.

Proof. Consider the quadratic polynomials $Q_i : T_\xi M \to \mathbb{R}$ defined in the coordinate system $\psi$ by

$$Q_i(x) = \frac{1}{2}(A_i x, x) + (p, x) + c, \quad i = 1, 2.$$

Observe that $Q_1(x) \leq Q_2(x) \forall x \in T_\xi M$ and $Q_1(0) = Q_2(0)$.

By suitably multiplying by a test function $\varphi$ whose support is contained in $B_r$ and such that $\varphi(\xi) = 1$ for $\xi \in B_r/2$, we may assume that we have functions $u_1, u_2$ with

$$(3.2) \quad u_1(\xi) \leq u_2(\xi) \quad \forall \xi \in M.$$

Moreover, the Taylor expansion of each $u_k$ at $\xi$ is $Q_k$, $k = 1, 2$. Hence, by the comparison principle, we have

$$T_s(u_1)(\xi) - u_1(\xi) \leq T_s(u_2)(\xi) - u_2(\xi).$$

Dividing by $s$, letting $s \to 0$, and using the regularity principle we get

$$F(A_1, p, c, \xi, G, \Gamma^k) \leq F(A_2, p, c, \xi, G, \Gamma^k).$$

Theorem 3.4. Let $T^s$ be a multiscale analysis satisfying all architectural axioms, the comparison principle, and gray level shift invariance. If $u_0$ is bounded and uniformly continuous in $M$ and $u(s, \xi) = T^s u_0(\xi)$, then $u(s, \xi)$ is a viscosity solution of

$$(3.3) \quad u_s = F(D^2_{\xi M} u, Du, \xi, G, \Gamma^k),$$

with $u(0, \xi) = u_0(\xi)$.

The proof that $u(s, \xi) = T_s u_0(\xi)$ is the viscosity solution of (3.3) follows as in [2, 31].

Moreover, the Taylor expansion of each $u_k$ at $\xi$ is $Q_k$, $k = 1, 2$. Hence, by the comparison principle, we have

$$T_s(u_1)(\xi) - u_1(\xi) \leq T_s(u_2)(\xi) - u_2(\xi).$$

The next lemma is crucial in what follows. It relates the matrices and vectors defining a quadratic polynomial in two coordinate systems around a point $\xi \in M$. For a proof, we refer the reader to [16].

Lemma 3.5. Let $U^1, U^2$ be two neighborhoods of 0 in $\mathbb{R}^N$, and let $\psi_i : U^i \to M$ be two coordinate systems around the point $\xi \in M$, i.e., $\psi_i(0) = \xi$. Let us denote by $\Psi$ the diffeomorphism given by the change of coordinates $\Psi = \psi_1^{-1} \circ \psi_2 : U^2 \to U^1$. Let $G, \Gamma$ (resp., $\overline{G}, \overline{\Gamma}$)
be the metric coefficients and the Christoffel symbols of \( M \) in the coordinate system \( \psi_1 \) (resp., \( \psi_2 \)) at the point \( \xi \). Let \( Q : U^1 \to \mathbb{R} \) be the quadratic polynomial

\[
Q(v) = \frac{1}{2}(GAv, v) + (p, v) + c. \tag{3.4}
\]

Let \( \overline{Q}(\bar{v}) := (Q \circ \psi)(\bar{v}) \). Then \( \overline{Q}(\bar{v}) = Q'(\bar{v}) + O(|\bar{v}|^3) \) in a neighborhood of 0, where \( Q' \) is the quadratic polynomial

\[
Q'(\bar{v}) = \frac{1}{2}(GB^{-1}AB\bar{v}, \bar{v}) + \frac{1}{2}(\Gamma(B^t p)(\bar{v}), \bar{v}) - \frac{1}{2}(B^t \Gamma(p)(B\bar{v}), \bar{v}) + (B^t p, \bar{v}) + c, \tag{3.5}
\]

and \( B = D\psi(0) \).

Observe that as a map \( B : T_\xi M \to T_\xi M \), while \( B^t : T_\xi^* M \to T_\xi^* M \). It satisfies

\[
B^t G = \overline{G} B^{-1}. \tag{3.6}
\]

In the next proposition we make explicit the dependence of \( F \) in the metric \( G \) and the connection \( \Gamma^k \). This can be done by expressing \( F \) in a geodesic coordinate system.

**Proposition 3.6.** Let \( T_s \) be a multiscale analysis on \( M \) satisfying the architectural axioms and the comparison principle. Let \( \psi_1 : U^1 \to M \) be a coordinate system around \( \xi \in M \). Let \( G, \Gamma \) be the metric coefficients and the Christoffel symbols of \( M \) in the coordinate system \( \psi_1 \) at the point \( \xi \), respectively. For any symmetric matrix \( X = (X_{ij}) : (T_\xi M, I) \to (T_\xi^* M, I) \) in \( SM_\xi(N, I) \), \( q \in (T_\xi^* M, I) \), and \( a \in \mathbb{R} \), let us define the function

\[
H(X, q, a, \xi, I, 0); \tag{3.7}
\]

that is, \( H \) is the function \( F \) obtained when using a geodesic coordinate system. Then

\[
F(A, p, a, \xi, I, \Gamma^k) = H(B^t(A - \Gamma(p))B, B^t p, c, \xi) \tag{3.8}
\]

for any matrix \( A \in SM_\xi(N) \) and any covector \( p \), where \( BB^t = G^{-1} \). Moreover, the function \( H \) satisfies

\[
H(A', p', c, \xi) = H(R^t A'R, R^t p', c, \xi), \tag{3.9}
\]

where \( A' : (T_\xi M, I) \to (T_\xi^* M, I) \) is any matrix in \( SM_\xi(N, I) \), \( p' \in (T_\xi^* M, I) \), and \( R \) is any Euclidean rotation in \( (T_\xi^* M, I) \).

Our notation \( BB^t = G^{-1} \) contains a slight abuse of notation since \( B : T_\xi M \to T_\xi M \) and \( B^t : T_\xi^* M \to T_\xi^* M \). The correct notation should be \( BI^{-1}B^t = G^{-1} \).

Although the proof of Proposition 3.6 is similar to the proof of Proposition 2 in [16], we give it in detail since, besides being crucial for what follows, it contains some basic ideas and notation that are used in the paper. We also make explicit the rotation invariance contained in the above formulas.

**Proof.** We shall use the notation of Lemma 3.5. For convenience, then we use the symmetric map \( GA \). Since \( Q \circ \psi_1^{-1} = Q \circ \psi_2^{-1} \) in \( \psi_1(U^1) \cap \psi_2(U^2) \), with a slight abuse of notation (act \( T_s \) on polynomials)

\[
\lim_{s \to 0} T_s(Q \circ \psi_1^{-1})(\xi) - Q \circ \psi_1^{-1}(\xi) = F(GA, p, \xi, G, \Gamma^k) \tag{3.10}
\]
and
\begin{equation}
\lim_{s \to 0} \frac{T_s(Q \circ \psi^{-1}_2)(\xi) - Q \circ \psi^{-1}_2(\xi)}{s} = F(G B^{-1} A B + \Gamma(B' p) - B' \Gamma(p) B, B' p, c, \xi, G, \Gamma^k),
\end{equation}
we have
\begin{equation}
F(GA, p, c, \xi, G, \Gamma^k) = F(G B^{-1} A B + \Gamma(B' p) - B' \Gamma(p) B, B' p, c, \xi, \overline{G}, \overline{\Gamma}^k)
\end{equation}
or, using (3.6),
\begin{equation}
F(GA, p, c, \xi, G, \Gamma^k) = F(B'(GA - \Gamma(p)) B + \Gamma(B' p), B' p, c, \xi, \overline{G}, \overline{\Gamma}^k).
\end{equation}
Now, for any symmetric matrix $X = (X_{ij}) \in SM_{\xi}(N)$, any $q \in T^*_\xi M$, and $a \in \mathbb{R}$, let us define the function $\tilde{F}$ by the identity
\begin{equation}
\tilde{F}(X, q, a, \xi, G, \Gamma^k) = F(X + \Gamma(q), q, a, \xi, G, \Gamma^k).
\end{equation}
In terms of $\tilde{F}$, (3.13) can be written as
\begin{equation}
\tilde{F}(GA - \Gamma(p), p, c, \xi, G, \Gamma^k) = \tilde{F}(B'(GA - \Gamma(p)) B, B' p, c, \xi, \overline{G}, \overline{\Gamma}^k).
\end{equation}
By varying the quadratic polynomials, the above equation holds for any matrix $A = (A'_{ij})$ such that $GA \in SM_{\xi}(N)$, any invertible matrix $B : T\xi M \to T\xi M$, and any $p \in T\xi M$ considering $\overline{G}$ such that $B' G = \overline{G} B^{-1}$. This holds in particular for any rotation $R$ in $T\xi M$ (so that $R' G R = G$):
\begin{equation}
\tilde{F}(GA - \Gamma(p), p, c, \xi, G, \Gamma^k) = \tilde{F}(R'(GA - \Gamma(p)) R, R' p, c, \xi, \overline{G}, \overline{\Gamma}^k).
\end{equation}
Now, we choose $\psi_1$ as a geodesic coordinate system around $\xi$ for which $G = I$, and $\Gamma^k = 0$. In this case, (3.6) can be written as $\overline{G} = B'C = B'B$. We may write (3.15) as
\begin{equation}
\tilde{F}(IA, p, c, \xi, I, 0) = \tilde{F}(B'C AB, B' p, c, \xi, B'B, \overline{\Gamma}^k),
\end{equation}
and this identity holds for any symmetric matrix $IA \in SM_{\xi}(N, I)$, any vector $p \in T\xi M$, and any invertible matrix $B$ (with the metric $\overline{G} = B'B$). Once again, we change variables and write $A' = B'C AB$, $p' = B' p$, $B' = B^{-1}$. Then we write (3.17) (note that the left- and right-hand sides have been interchanged) as
\begin{equation}
\tilde{F}(A', p', c, \xi, \overline{G}, \overline{\Gamma}^k) = \tilde{F}(B'C A'B', B' p', c, \xi, I, 0),
\end{equation}
and this identity holds for any symmetric matrix $A' : T\xi M \to T\xi M$ in $SM_{\xi}(N, \overline{G})$, any $p' \in T\xi M$, and any invertible matrix $B' : T\xi M \to T\xi M$, where $\overline{G} = (B'C)^{-1} B'^{-1}$. This clearly shows that $\tilde{F}$ does not depend on $G$ and $\Gamma^k$ in the last two arguments. All its dependence is contained in the first argument. Let us introduce the function $H$ to make this explicit.
Now, for any symmetric matrix \( X = (X_{ij}) : (T_{\xi}\mathcal{M}, I) \to (T^*_{\xi}\mathcal{M}, I) \) in \( \text{SM}_{\xi}(N, I) \), any \( q \in (T^*_{\xi}\mathcal{M}, I) \), and scalar \( a \), let us define the function \( H \) by the identity

\[
H(X, q, a, \xi) = \tilde{F}(X, q, a, \xi, I, 0).
\]

(3.19)

Note that by (3.14) and (3.19), we have

\[
H(X, q, a, \xi) = F(X, q, a, \xi, I, 0);
\]

that is, \( H \) is the function \( F \) obtained when using a geodesic coordinate system. Hence, (3.18) can be written as

\[
(3.20) \quad \tilde{F}(A', p', c, \xi, G, \Gamma^k) = H(B^n A'B', B^n p', c, \xi),
\]

and using (3.14), (3.19), we have formula (3.8). Note the role of \( B' \) which makes \( B^n A'B' : (T_{\xi}\mathcal{M}, I) \to (T^*_{\xi}\mathcal{M}, I) \) symmetric. In particular, if we take \( \psi_1 \) as a geodesic coordinate system around \( \xi \) and \( \psi_2 \) as a rotation \( R \) with respect to \( \psi_1 \) so that \( B' = R, \overline{G} = R^t I R = I \), and \( \Gamma^k = 0 \) at the point \( \xi \), then (3.15) can be written as

\[
(3.21) \quad \tilde{F}(A, p, c, \xi, I, 0) = \tilde{F}(R^t A R, R^t p, c, \xi, I, 0),
\]

that is, as

\[
(3.22) \quad H(A', p', c, \xi) = H(R^t A'R, R^t p', c, \xi),
\]

where \( A' : (T_{\xi}\mathcal{M}, I) \to (T^*_{\xi}\mathcal{M}, I) \) is any matrix in \( \text{SM}_{\xi}(N, I) \), \( p' \in (T^*_{\xi}\mathcal{M}, I) \), and \( R \) is any Euclidean rotation in \( (T_{\xi}\mathcal{M}, I) \).

The general expression of rotation invariance is written in terms of \( \tilde{F} \) in (3.15). \( \blacksquare \)

**Remark 3.** Let us write the rotation invariance in the tangent plane in terms of \( F \). Let us consider the quadratic form \( Q \) in the coordinate system \( \psi_1 : B_{T_{\xi}\mathcal{M}}(0, r) \to \mathcal{M} \) given by

\[
Q(v) = \frac{1}{2} (Sv, v) + (p, v) + c,
\]

where \( S \in \text{SM}_{\xi}(N) \). Then we consider the rotation \( R : T_{\xi}\mathcal{M} \to T_{\xi}\mathcal{M} \) and define

\[
\bar{Q}(w) = \frac{1}{2} (SRw, Rw) + (p, Rw) + c = Q(Rw).
\]

Consider the function \( u(\zeta) = Q(\psi_1^{-1}(\zeta)) \). In the coordinate system \( \psi_2 : B_{T_{\xi}\mathcal{M}}(0, r) \to \mathcal{M} \) given by \( \psi_2(v) = \psi_1(Rv) \), \( u(\zeta) = \bar{Q}(R^{-1} \psi_1^{-1}(\zeta)) = \bar{Q}(\psi_2^{-1}(\zeta)) \). That is, \( u \) is expressed by \( \bar{Q} \) in the coordinate system \( \psi_2 \). Then by the regularity axiom

\[
\frac{T_s u(\xi) - u(\xi)}{s} \to F(S, G^{-1} p, \xi, G, \Gamma^k),
\]

\[
\frac{T_s u(\xi) - u(\xi)}{s} \to F(R^t S R, G^{-1} R^t p, \xi, G, \Gamma^k).
\]
Thus

\[(3.23) \quad F(S, G^{-1} p, \xi, G, \Gamma^k) = F(R^t S R, G^{-1} R^t p, \xi, G, \Gamma^k).\]

Notice that the metric does not change, but the connection does.

**Remark 4.** Let us express the rotation invariance given in (3.9) in different ways. Using (3.9), if \( BB^t = G^{-1} \), then

\[(3.24) \quad H(B^t A B, B^t p, c, \xi) = H(R^t B^t A B R, R^t B^t p, c, \xi),\]

where \( A : (T_\xi M, G) \to (T_\xi M, G) \) is any matrix in \( SM_\xi(N, G) \), \( p \in (T_\xi M, G) \), and \( R \) is any Euclidean rotation in \( (T_\xi M, I) \). Note that \( B^t A B : (T_\xi M, I) \to (T_\xi M, I) \) is any matrix in \( SM_\xi(N, I) \), and \( B^t p \in (T_\xi M, I) \). Note that \( BR : (T_\xi M, I) \to (T_\xi M, G) \) satisfies

\[(BRv, BRw)_g = \langle v, w \rangle_I.\]

That is, \( BR \) is an isometry matrix from \( (T_\xi M, I) \) to \( (T_\xi M, G) \).

We can also write the rotation invariance of \( H \) in a different way. By taking \( A' = GA - \Gamma(p) \) in (3.16), we get

\[(3.25) \quad \tilde{F}(A', p, c, \xi, G, \Gamma^k) = \tilde{F}(R^t A' R, R^t p, c, \xi, G, \Gamma^k)\]

for any rotation \( R \) in \( (T_\xi M, G) \) and use (3.20) to obtain

\[(3.26) \quad H(B^t A' B, B^t p, c, \xi) = H(B^t R^t A' R B, B^t R^t p, c, \xi).\]

Note that \( RB \) is an isometry from \( (T_\xi M, I) \) to \( (T_\xi M, G) \).

### 3.1. Further remarks on rotation invariance as a consequence of the invariance by changes of coordinates.

Let us add some remarks to point out how rotation invariance follows as a consequence of invariance with respect to changes of coordinates. Then we will comment on its implication for video. Let us use the notation of Lemma 3.5. Let \( \psi_1 : U_1 \to (M, g) \), \( \psi_2 : U_2 \to (M, \overline{\mathbf{g}}) \) be two coordinate systems, and let \( \Psi := \psi_1^{-1} \circ \psi_2 : U_2 \to U_1 \) be the change of coordinates. Let \( Q \) be a quadratic polynomial in \( U_1 \) given by the data \( (GA, p) = (Gv, c) \) and \( \overline{Q} \) be the corresponding quadratic polynomial in \( U_2 \) given by the data \( (GA, \overline{\mathbf{g}}) = (Gv, c) \). Let \( B = D\Psi(0) \). Recall that

\[(3.27) \quad \overline{Q} = B^t GB.\]

Since by the computations in Lemma 3.5 we have

\[\overline{GA} - \Gamma(B^{-1} G B^{-1} v) = B^t (GA - \Gamma(Gv)) B,\]

we can write (3.13) as

\[F(GA, Gv, c, \xi, G, \Gamma^k) = F(\overline{GA}, \overline{Gv}, c, \xi, \overline{G}, \overline{\Gamma}^k).\]
In terms of \( \tilde{F} \) (note (3.15)) we have
\[
\tilde{F}(GA - \Gamma(Gv), Gv, c, \xi, G, \Gamma^k) = \tilde{F}(B^i(GA - \Gamma(Gv))B, B^iGv, c, \xi, G, \Gamma^k) = \tilde{F}(G\overline{A} - \overline{\Gamma}(G\overline{B}^{-1}v), G\overline{B}^{-1}v, c, \xi, G, \Gamma^k).
\]
Note that if \( \overline{G} = I \) and \( \overline{\Gamma} = 0 \), and \( B_G \) satisfies \( B_G\overline{B}_G^t = G^{-1} \), then the above is written as
\[
\tilde{F}(GA - \Gamma(Gv), Gv, c, \xi, G, \Gamma^k) = \tilde{F}(B_G^t(GA - \Gamma(Gv))B_G, B_G^tGv, c, \xi, I, 0).
\]

Now, if \( B_G \) is such that \( B_G\overline{B}_G^t = G^{-1} \) and \( B \) satisfies (3.27), then from (3.29) and (3.28) we have
\[
\tilde{F}(B_G^t(GA - \Gamma(Gv))B_G, B_G^tGv, c, \xi, I, 0) = \tilde{F}(B_G^tB^i(GA - \Gamma(Gv))B_B, B_G^tB^iGv, c, \xi, I, 0).
\]

In terms of \( H \) we may write
\[
H(B_G^t(GA - \Gamma(Gv))B_G, B_G^tGv, c, \xi) = H(B_G^tB^i(GA - \Gamma(Gv))B_B, B_G^tB^iGv, c, \xi).
\]

Let us analyze (3.30). Note that
\[
B_B B_G^t B^i = B_G^{-1}B^i = G^{-1} = B_G B_G^t.
\]

If we identify \( B_B B_G^t \) with \( B_G \), formula (3.30) becomes trivial. But using the next lemma we deduce that \( B_B B_G^t \) can be identified with \( B_G R \) for a Euclidean rotation \( R \). Then (3.30) is equivalent to saying that
\[
H(B_G^t(GA - \Gamma(Gv))B_G, B_G^tGv, c, \xi) = H(R^t B_G^t(GA - \Gamma(Gv))B_B R, B_G^tB^iGv, c, \xi)
\]
for any Euclidean rotation \( R \).

**Lemma 3.7.** Let \( B_0 \) be a solution of \( B_0 B_0^t = G^{-1} \). Then any other solution has the form \( B_0 R \) for a Euclidean rotation \( R \).

**Proof.** Let \( B \) be another solution of \( B_B^t = G^{-1} \). Then \( B_0 B_0^t = B_B^t \). Thus \( B^{-1}B_0 B_0^t B^{-1} = T \). Thus there is a Euclidean rotation \( R \) such that \( B^{-1}B_0 = R^t \). Thus \( B = B_0 R \). \( \blacksquare \)

**Remark 5.** Let us illustrate this with an example. Let \( \mathcal{M} = \{(x, t) : x \in \mathbb{R}^2, t \in \mathbb{R}\} \). Let us consider the metric \( g(x, t)_{ij} \) so that if \((y, \tau)\) denote the coordinates in \( T_{(x, t)}\mathcal{M}\), then
\[
g(x, t)((y, \tau), (y, \tau)) = (y - v(x, t)\tau)^2 + \tau^2,
\]
where \( v(x, t) \) is the optical flow of the sequence. The meaning of this metric is that being at the point \((x, t)\) of a video a displacement \((y, \tau)\) of a point is not penalized if it follows the law of motion at \((x, \tau)\). For any \( v \in \mathbb{R}^2 \), let us write \( G(v) = \begin{pmatrix} 1 & -v \\ -v & 1 + |v|^2 \end{pmatrix} \).
Then \( \det(G(v)) = 1 \), and

\[
G(v)^{-1} = \begin{pmatrix}
I + v \otimes v & v \\
v & 1
\end{pmatrix} = \begin{pmatrix}
I & 0 \\
0 & 0
\end{pmatrix} + \hat{v} \otimes \hat{v},
\]

where \( \hat{v} = (v,1)^t \), and \((a \otimes b)(x) = (a, x)b, a, b, x \in \mathbb{R}^N \) (here \( N = 2, 3 \)). Let us compute the solutions \( C \) of \( CC^t = G(v)^{-1} \). For that let us compute a matrix \( D \) such that \( D^tD = G(v) \). A solution is

\[
D = \begin{pmatrix}
I & -v \\
0 & 1
\end{pmatrix}.
\]

If \( v = v(x,t) \), we denote \( D \) as \( B(v(x,t)) \), the boost given by the velocity \( v(x,t) \) (i.e., \( B(v)(x,t) = (x-\mathbf{v}t,t) \)). Then \( C = D^{-1} = B(-v(x,t)) \) is a solution of \( CC^t = G(v(x,t))^{-1} \). Any other solution has the form \( B(-v(x,t))R \), where \( R \) is a Euclidean rotation in \( \mathbb{R}^3 \).

If we make a change of coordinates so that \( B = D\Psi(0) \) is given by a boost \( B(w) \), where \( w \in \mathbb{R}^2 \), then the metric \( G \) is

\[
B(w)^tG(v)B(w) = G(v + w).
\]

Then (3.31) is in this case

\[
H(B(-v)^t(GA - \Gamma(Gv))B(-v), B(-v)^tGv, c, \xi)
\]

(3.32)

\[
= H(R^qB(-v)^t(GA - \Gamma(Gv))B(-v)R, R^qB(-v)^tGv, c, \xi)
\]

for any Euclidean rotation \( R \) in \( \mathbb{R}^3 \). Thus the boosting is absorbed by the metric and, again, the only invariance contained in the change of coordinates is with respect to Euclidean rotations in \( \mathbb{R}^3 \).

3.2. Multiscale analyses when \( N = 2 \). Before continuing, let us give some more notation. Let \( e_i, i = 1, \ldots, N \), be an orthonormal basis of \( T_{\xi}M \) and \( e_i^* \) be its dual basis. For each vector \( q \in T_{\xi}M \), let \( \nu_q = \frac{G^{-1}q}{|G^{-1}q|_{\xi}} \), where \( |G^{-1}q|_{\xi} := (G^{-1}q,q) \). Let \( R_q \) be a rotation in \( T_{\xi}M \) so that \( R^q_q = |q|_{\xi}^{-1}e_1^* \).

Following the proof of Proposition 3 in [16] and using (3.26) we have the following proposition.

Proposition 3.8. Assume that the multiscale operator \( T_{\xi} \) satisfies the architectural axioms, the comparison principle, and gray level shift invariance. Then there is a function \( \mathcal{H} : SM(N,I) \times \mathbb{R} \times M \rightarrow \mathbb{R} \) such that \( \mathcal{H}(B^tA\dot{B}, B^tq, c, \xi) = \mathcal{H}(B^tR^q\dot{A}R_qB, |q|_{\xi}^{-1}, \xi) \) for any \( A \in SM(N,G), q \in T_{\xi}M, q \neq 0, c \in \mathbb{R}, \xi \in M \). Moreover, \( \mathcal{H} \) is a continuous and nondecreasing function of \( A \).

Let us now fix \( N = 2 \). Let \( \xi \in M \), and let \( \{e_1, e_2\} \) denote a fixed orthonormal basis of \( T_{\xi}M \). Let \( \{e_1^*, e_2^*\} \) be its dual orthonormal basis in \( T_{\xi}^*M \).

Let us denote \((a \otimes b)(x) = (a, x)b, a \in T_{\xi}M, b, x \in T_{\xi}^*M \). Let \( R_q \) be the rotation in \( T_{\xi}M \) given by

\[
R_q = e_1^* \otimes \nu_q + e_2^* \otimes \nu^q q,
\]
where \((\nu_q, \nu^+_q)\) forms an orthonormal basis of \(T_{\xi}M\). Then \(R^t_q = \nu_q \otimes e^*_1 + \nu^+_q \otimes e^*_2\) and \(R^t_q g = |G^{-1}q|_g e^*_1 = |q|^{-1}_g e^*_1\).

If \(A : T_{\xi}M \to T_{\xi}M\), then

\[
(R^t_q AR_q e_1, e_1) = (A\nu_q, \nu_q), \quad (R^t_q AR_q e_2, e_2) = (A\nu_q^+, \nu_q^+),
\]

\[
(R^t_q AR_q e_1, e_2) = (A\nu_q^+, \nu_q), \quad (R^t_q AR_q e_2, e_2) = (A\nu_q^+, \nu_q^+).
\]

Then, if \(B : T_{\xi}M \to T_{\xi}M\) is such that \(BI^{-1}B^t = g^{-1}\) and \(\hat{\xi} = B^{-1}\xi\), then \(\hat{\xi}\) is a Euclidean basis of \(T_{\hat{\xi}}M\) and

\[
(B^t R^t_q AR_q B \hat{e}_1, \hat{e}_1) = (A\nu_q, \nu_q), \quad (B^t R^t_q AR_q B \hat{e}_1, \hat{e}_2) = (A\nu_q^+, \nu_q^+),
\]

\[
(B^t R^t_q AR_q B \hat{e}_2, \hat{e}_1) = (A\nu_q^+, \nu_q), \quad (B^t R^t_q AR_q B \hat{e}_2, \hat{e}_2) = (A\nu_q^+, \nu_q^+).
\]

Then for any symmetric \(A : T_{\xi}M \to T_{\xi}^*M\) the matrix of \(B^t R^t_q AR_q B\) in the Euclidean basis \(\{\hat{e}_1, \hat{e}_2\}\) is

\[
B^t R^t_q AR_q B = \begin{pmatrix}
A(\nu_q, \nu_q) & A(\nu_q, \nu^+_q) \\
A(\nu^+_q, \nu_q) & A(\nu^+_q, \nu^+_q)
\end{pmatrix}.
\]

Since \(A(\nu_q, \nu^+_q) = A(\nu^+_q, \nu_q)\), we see that \(R^t_q AR_q\) depends only on the three scalars \(A(\nu_q, \nu_q), A(\nu_q, \nu^+_q), A(\nu^+_q, \nu^+_q)\). Note that these quantities are intrinsic; i.e., they do not depend on the coordinate system.

We will now write \(\mathcal{H}(A(\nu_q, \nu_q), A(\nu_q, \nu^+_q), A(\nu^+_q, \nu^+_q), |p|, \xi)\) instead of \(\mathcal{H}(B^t R^t_q AR_q B, |p|, \xi)\).

**Definition 3.9.** Let \(\mathcal{H} : SM(2) \times \mathbb{R} \times M \to \mathbb{R}\). We shall say that \(\mathcal{H}(A, s, \xi)\) is elliptic if \(\mathcal{H}\) is a nondecreasing function of \(A\), in the sense that if \(A_1 \leq A_2\) (as defined in Lemma 3.3), then \(\mathcal{H}(A_1, s, \xi) \leq \mathcal{H}(A_2, s, \xi)\). If \(A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}\) and we define \(\mathcal{H}(a, b, c, s, \xi) = \mathcal{H}(A, s, \xi)\), we shall also say that \(\mathcal{H}(a, b, c, s, \xi)\) is elliptic.

We now come to the main theorem of this subsection. For simplicity of notation let us introduce the terms:

\[
\Lambda_1(u, \xi) = D^2_{\xi,M} u \left( \frac{\nabla u}{|\nabla u|_\xi}, \frac{\nabla u}{|\nabla u|_\xi} \right)(\xi),
\]

\[
\Lambda_2(u, \xi) = D^2_{\xi,M} u \left( \frac{\nabla u}{|\nabla u|_\xi}, \frac{\nabla u^+}{|\nabla u^+|_\xi} \right)(\xi),
\]

\[
\Lambda_3(u, \xi) = D^2_{\xi,M} u \left( \frac{\nabla u^+}{|\nabla u^+|_\xi}, \frac{\nabla u^+}{|\nabla u^+|_\xi} \right)(\xi).
\]

**Theorem 3.10.** Assume that the interpolation operator \(T_s\) satisfies the architectural axioms, the comparison principle, and gray level shift invariance. Let \(\mathcal{H}\) be the elliptic function given in Proposition 3.8. Let \(u(s, x) = T_s u(x), u \in C_b(M)\). Then \(u\) is a viscosity solution of

\[
u_s = \mathcal{H}(\Lambda_1(u, \xi), \Lambda_2(u, \xi), \Lambda_3(u, \xi), |\nabla u|_\xi, \xi) \quad \text{in} \ M.
\]
That is, for any $\varphi \in C^\infty([0, \infty) \times \mathcal{M})$ with bounded derivatives such that $u - \varphi$ has a strict local maximum (minimum) at $(s_0, \xi_0)$ and $\nabla \varphi(s_0, \xi_0) \neq 0$, we get

\begin{equation}
\varphi_s(s_0, x_0) \leq H(\Lambda_1(\varphi, \xi_0), \Lambda_2(\varphi, \xi_0), \Lambda_3(\varphi, \xi_0), |\nabla \varphi|_{\xi_0}, \xi_0)
\end{equation}

(resp., $\geq$).

For a proof, see Proposition 1 in [17] or Theorem 1 in [16].

Let us finish this subsection with some complementary information about $H$. For simplicity, in the next proposition we shall not denote the argument $\xi$ of $H$.

**Proposition 3.11.** (i) If $H$ does not depend upon its first or its third argument, then it also does not depend on its second argument. In other terms, we have the following:

If $H(\alpha, \beta, \gamma, \delta) = \hat{H}(\beta, \gamma, \delta)$, then $H = \hat{H}(\gamma, \delta)$.

If $H(\alpha, \beta, \gamma, \delta) = \hat{H}(\alpha, \beta, \delta)$, then $H = \hat{H}(\alpha, \delta)$,

where $\alpha, \beta, \gamma, \delta \in \mathbb{R}$.

(ii) If $H$ is differentiable at $(0, 0, 0, 0)$, then $H$ may be written as $H(A, \delta) = \text{tr}(BA) + d \delta$, where $B$ is a nonnegative matrix with constant coefficients and $d$ is a real constant.

Thus if we assume that $H$ is differentiable at $(0, 0, 0, 0)$, then we may rewrite (3.33) as

\begin{equation}
u_s = a\Lambda_1(u, \xi) + 2b\Lambda_2(u, \xi) + c\Lambda_3(u, \xi) + d|\nabla u|_{\xi} = 0,
\end{equation}

where $a, c \geq 0$ and $ac - b^2 \geq 0$, which is the same as saying that the matrix $B$ above is nonnegative.

If we take $a = c = 1$, $b = d = 0$, we recover the Laplace–Beltrami operator. If we choose $a = 1$, $b = c = d = 0$, we obtain the extension to 2D manifolds of the so-called infinite Laplacian (used in the context of image interpolation in [17, 13, 1, 18]).

4. The morphological axiom. Throughout this section we assume that $T_s$ is a multiscale analysis satisfying all architectural axioms, the comparison principle, and gray level shift invariance.

Let us recall the axiom:

[Gray scale invariance] $T_s(f(u)) = f(T_s(u)) \forall s \geq 0, \forall u \in C^\infty_b(\mathcal{M})$, and for any strictly increasing function $f : \mathbb{R} \to \mathbb{R}$.

It is also called the morphological axiom. Note that by taking $f(x) = x + C$ for a constant $C$, it implies the gray level shift invariance. It is interesting also to note that the above axiom implies the following:

[Linear gray scale invariance] $T_s(\lambda u) = \lambda T_s(u) \forall s \geq 0, \forall u \in C^\infty_b(\mathbb{R}^N \times \mathbb{R}^N)$, and for any $\lambda \geq 0$.

The next lemma can be proved as in [2].

**Lemma 4.1.** Assume that $T_s$ satisfies all architectural axioms, the comparison principle, and the morphological axiom. Then

\begin{equation}F(\lambda A + \mu p \otimes p, \lambda p, \xi, G, \Gamma^k) = \lambda F(A, p, \xi, G, \Gamma^k) \end{equation}

$\forall A \in \mathcal{SM}_\xi(N), \forall p \in T^*_\xi \mathcal{M}, \forall \xi \in \mathcal{M}, \forall \lambda \geq 0, \mu \in \mathbb{R}$.
In particular, taking $\mu = 0$, we have
\begin{equation}
F(\lambda A, \lambda p, \xi, G, \Gamma^k) = \lambda F(A, p, \xi, G, \Gamma^k) \quad \forall A \in \text{SM}_\xi(N), \forall p \in T^*_\xi \mathcal{M}, \forall \xi \in \mathcal{M}, \forall \lambda \geq 0.
\end{equation}
Hence $F(0, 0, z) = 0$. This property of $F$ can be proved using only the axiom of linear gray scale invariance.

Taking $\lambda = 1$, we have
\begin{equation}
F(A + \mu p \otimes p, p, \xi, G, \Gamma^k) = F(A, p, \xi, G, \Gamma^k)
\end{equation}
\begin{equation}
\forall A \in \text{SM}_\xi(N), \forall p \in T^*_\xi \mathcal{M}, \forall \xi \in \mathcal{M}, \forall \mu \in \mathbb{R}.
\end{equation}
Both properties can be written in terms of $\tilde{F}$ and $H$.

Let $Q_p = I - \frac{\rho G^{-1}_p}{(\rho - 1)p_p}$, $p \in T^*_\xi \mathcal{M} \setminus \{0\}$. Then $Q_p : T^*_\xi \mathcal{M} \to T^*_\xi \mathcal{M}$ and $Q_p^t : T^*_\xi \mathcal{M} \to T^*_\xi \mathcal{M}$.

**Theorem 4.2.** Let $T^*_\xi$ be a multiscale analysis satisfying all architectural axioms, the comparison principle, and the morphological axiom. Then
\begin{equation}
F(A, p, \xi, G, \Gamma^k) = F(Q_p^t A Q_p, p, \xi, G, \Gamma^k) \quad \forall A \in \text{SM}_\xi(N), \forall p \in T^*_\xi \mathcal{M} \setminus \{0\}, \forall \xi \in \mathcal{M}.
\end{equation}
A similar statement holds for $\tilde{F}$. Let $B$ be such that $B^t G B = I$. In terms of $H$ we have
\begin{equation}
H(B^t (A - \Gamma(p))B, B^t p, \xi) = H(B^t Q_p^t (A - \Gamma(p))Q_p B, B^t p, \xi)
\end{equation}
\begin{equation}
\forall A \in \text{SM}_\xi(N), \forall p \in T^*_\xi \mathcal{M} \setminus \{0\}, \forall \xi \in \mathcal{M}.
\end{equation}

**Proof.** We follow the proof of Theorem 4 in [2]. First observe that for any $A \in \text{SM}_\xi(N)$ and any $p \in T^*_\xi \mathcal{M}$ we have
\begin{equation*}
A = Q_p^t A Q_p + (I - Q_p)^t A Q_p + Q_p^t A (I - Q_p) + (I - Q_p)^t A (I - Q_p),
\end{equation*}
and $I - Q_p = \frac{\rho G^{-1}_p}{(\rho - 1)p_p}$.

Thanks to (4.3), by choosing $\mu$ to cancel $(I - Q_p)A(I - Q_p)$ we see that
\begin{equation*}
F(A, p, c, \xi, G, \Gamma^k) = F(Q_p^t A Q_p + (I - Q_p)^t A Q_p + Q_p^t A (I - Q_p), p, c, \xi, G, \Gamma^k).
\end{equation*}
We can also cancel $(I - Q_p)^t A Q_p + Q_p^t A (I - Q_p)$ using the ellipticity of $F$ [2]. Indeed we observe that $(I - Q_p)^t A Q_p, Q_p^t A (I - Q_p) : T^*_\xi \mathcal{M} \to T^*_\xi \mathcal{M}$, $G Q_p, G (I - Q_p) \in T^*_\xi \mathcal{M}$,
\begin{equation*}
(I - Q_p)^t A Q_p, Q_p^t A (I - Q_p) \leq \epsilon G Q_p + \frac{M}{\epsilon} G (I - Q_p)
\end{equation*}
for some $M > 0$. Using (4.3) and the ellipticity of $F$ and letting $\epsilon \to 0^+$, we obtain (4.4).

Working with
\[ B^t(A - \Gamma(p))B = B^tQ_p^t(A - \Gamma(p))Q_pB + B^t(I - Q_p)^t(A - \Gamma(p))Q_pB \]
\[ + B^tQ_p^t(A - \Gamma(p))(I - Q_p)B + B^t(I - Q_p)^t(A - \Gamma(p))(I - Q_p)B \]

and observing that the last three terms can be dominated by \( B^t \rho \otimes B^t \), we obtain (4.5). \( \blacksquare \)

**Theorem 4.3.** Let \( T_s \) be a multiscale analysis satisfying all architectural axioms, the comparison principle, and the morphological axiom. Then there is \( \bar{K} \), which is a symmetric and increasing function of its arguments such that
\[ \tilde{F}(X, p, c, \xi, G, \Gamma^k) = |G^{-1}p|_g\bar{K}(\lambda_1, \ldots, \lambda_{N-1}, \xi), \]
where \( \lambda_1, \ldots, \lambda_{N-1} \) are the eigenvalues of \( G^{-1}Q_p^tXQ_p \). Here \( X \) denotes \( D^2_M u = S - \Gamma(p) \), \( p \) being a covector.

Note that to speak about the eigenvalues of a matrix the matrix acts in the same linear space. Thus we speak about the eigenvalues of \( G^{-1}Q_p^tAQ_p = Q_p^{t,g}G^{-1}AQ_p \) (since we have \( GQ_p^{t,g} = Q_p^tG \)), where \( A : T^*_\xi M \to T^*_\xi M \). The last expression shows that the linear map \( G^{-1}Q_p^tAQ_p \) is symmetric in \( T^*_\xi M \).

If \( N = 2 \), \( e_p = \frac{G^{-1}p}{|G^{-1}p|} \), and \( e_p^\perp \) is a unit vector orthogonal to \( e_p \) in \( T^*_\xi M \), then
\[ \text{Trace}_g(Q_p^tAQ_p) = \text{Trace}(G^{-1}Q_p^tAQ_p) = \langle G^{-1}Ae_p^\perp, e_p^\perp \rangle, \]
and, as usual, we define the curvature by
\[ \text{curv}_g(u) := \frac{\text{Trace}_g(Q_p^tAQ_p)}{|G^{-1}p|_g} = \frac{\lambda_1}{|G^{-1}p|_g}. \]

**Proof.** Recall that (see (3.25))
\[ (4.6) \quad \tilde{F}(X, p, c, \xi, G, \Gamma^k) = \tilde{F}(R^tXR, R^tR, c, \xi, G, \Gamma^k) \]
for any \( X \in \text{SM}_\xi(N), \ p \in T^*_\xi M \setminus \{0\}, \forall \xi \in \mathcal{M}, \) and for any rotation \( R \) in \( T^*_\xi M \). That is, \( \langle Re_i, Re_j \rangle = \delta_{ij} \). Let us prove that \( \tilde{F} \) is a function of \( \lambda_1, \ldots, \lambda_{N-1} \). Let us choose a rotation so that \( Re_p = e_p \). Then one can check that \( R^tR^t|e_p = e_p \) and \( Re\perp e_p \) for any \( x \perp e_p \). Also
\[ R^tR = R^tGG^{-1}p = GR^t|e_p|G^{-1}p|_g = Ge_p|G^{-1}p|_g = p. \]
It is also easy to check that \( Q_pR = RQ_p \). Then, with \( X = GA - \Gamma(p) \),
\[ \tilde{F}(R^tQ_p^tXQ_pR, p, c, \xi, G, \Gamma^k) = \tilde{F}(R^tQ_p^tXQ_pR, R^tR, c, \xi, G, \Gamma^k) = \tilde{F}(Q_p^tXQ_p, p, c, \xi, G, \Gamma^k). \]
Thus, we can eliminate the \( R \) without touching the \( p \).

Note that \( G^{-1}R^tG = R^tG^{-1}Q_p^tG = Q_p^{t,g} \),
\[ G^{-1}R^tQ_p^tXQ_pR = R^tG^{-1}Q_p^tXQ_pR = R^tG^{-1}Q_p^{t,g}G^{-1}XQ_pR. \]
Thus \( G^{-1}R^tQ_p^tXQ_pR \) has the same eigenvalues as \( Q_p^{t,g}G^{-1}XQ_p \), which are \( \alpha_1, \ldots, \alpha_{N-1} \). That is, there is a function \( \bar{K} \) such that
\[ \tilde{F}(X, p, c, \xi, G, \Gamma^k) = \bar{K}(\alpha_1, \ldots, \alpha_{N-1}, p, \xi, G). \]
Note that the dependence on $G$ is concentrated in $X = GA - \Gamma(p)$.

Now, let $R$ be a rotation such that $Rg = \epsilon_p$. Then $R^\theta \epsilon_p = \epsilon_q$. Also $RQ_q = Q_p R$. If $p, q$ have the same length, i.e., $|G^{-1}p|_g = |G^{-1}q|_g$, then $R^t p = q$. Then

$$\tilde{F}(Q_p^t XQ_p, p, c, \xi, G, \Gamma^k) = \tilde{F}(X, p, c, \xi, G, \Gamma^k)$$

by the first step we can eliminate the $R$ without touching the $q$)

$$= \tilde{F}(Q_p^t XQ_p, q, c, \xi, G, \Gamma^k).$$

(by the first step we can eliminate the $R$ without touching the $q$)

In other words,

$$\tilde{K}(\alpha_1, \ldots, \alpha_{N-1}, |G^{-1}p|_g, \lambda) = \tilde{K}(\alpha_1, \ldots, \alpha_{N-1}, |G^{-1}p|_g, \lambda).$$

Note that $|G^{-1}p|^2 = (G^{-1}p, p) = (B^t p, B^t p) = 2 |B^t p|^2$. Using (4.1), we know that $\tilde{K}$ is homogeneous of degree one, and we have

$$\tilde{K}(\alpha_1, \ldots, \alpha_{N-1}, |G^{-1}p|_g, \lambda) = \tilde{K}(\alpha_1, \ldots, \alpha_{N-1}, 1, \lambda).$$

In addition, $\tilde{K}$ is a symmetric and increasing function of its arguments. Note that the only dependence on $G$ is made explicit in the formula.

The following remarks help to clarify the proof of Theorem 4.3.

Remark 6. If $R$ is a rotation in the tangent plane $(T_\xi M, G)$, then there exists a Euclidean rotation $\tilde{R}$ of $T_\xi M$ such that $RB = B\tilde{R}$ (where $B$ is such that $B G B^t = I$). Define $\tilde{R}$ so that $RB = B\tilde{R}$ holds. Then

\[
\langle RB v, RB w \rangle = \langle B\tilde{R} v, B\tilde{R} w \rangle \quad \forall v, w \in T_\xi M.
\]

Hence

\[
(B v, B w) = \langle B\tilde{R} v, B\tilde{R} w \rangle \quad \forall v, w \in T_\xi M,
\]

and, thus,

\[
(GB v, B w) = \langle GB\tilde{R} v, B\tilde{R} w \rangle \quad \forall v, w \in T_\xi M.
\]

Since $B^t G B = I$, we deduce that

\[
(v, w) = \langle \tilde{R} v, \tilde{R} w \rangle \quad \forall v, w \in T_\xi M.
\]

That is, $\tilde{R}$ is a Euclidean rotation.

Remark 7. Recall that the dependence of $F$ on $G$ is better expressed through the function $H$ that incorporates a matrix $B$ such that $BB^t = G^{-1}$. Then (following the notation in the proof of Theorem 4.3) by (3.20) we may write

$$\tilde{F}(Q_p^t XQ_p, p, c, \xi, G, \Gamma^k) = H(B^t Q_p^t XQ_p B, B^t p, c, \xi),$$

and using (3.22) we have

(4.7) \quad $H(B^t Q_p^t XQ_p B, B^t p, c, \xi) = H(\tilde{R}^t B^t Q_p^t XQ_p B\tilde{R}, \tilde{R}^t B^t p, c, \xi)$
for any Euclidean rotation $\bar{R}$. Using Remark 6, rotations $R$ in $(T_\xi M, g)$ that fix $e_p$ correspond to Euclidean rotations $\bar{R}$ fixing $B^{-1}e_p$ (which is the kernel of $Q_p B$). Now, observe that then

$$\bar{R}^t B^t p = B^t p.$$  

Indeed, by observing that $B^t = I^{-1} B^{-1} G^{-1}$ ($I$ identifying covectors and vectors with the same coordinates), the above identity is equivalent to

$$I \bar{R}^t I^{-1} B^{-1} G^{-1} p = B^{-1} G^{-1} p.$$  

But $I \bar{R}^t I^{-1} = \bar{R}$ (define $S^i_j = R^i_j$, and let $p$ be any covector; then $(\bar{R} I p)^i = \sum_j \bar{R}^i_j (I p)_j = I(\sum_j S^i_j p_j) = (I S p)^i = (I R^t p)^i$) so that

$$\bar{R} B^{-1} g^{-1} p = B^{-1} G^{-1} p,$$

i.e.,

$$\bar{R} B^{-1} e_p = B^{-1} e_p,$$

which holds true. Thus $\bar{R}^t B^t p = B^t p$ and we can continue the previous equalities in (4.7) for $H$ to obtain

$$H(\bar{R}^t B^t Q^t_p XQ_p B \bar{R}, B^t p, c, \xi).$$

We deduce that $H$ depends only on the eigenvalues of $B^t Q^t_p XQ_p B$ which coincide with those of $BIB^t Q^t_p XQ_p = G^{-1} Q^t_p XQ_p$, which are $\alpha_1, \ldots, \alpha_{N-1}, 0$.

**Remark 8.** Examples of interest are given by motion by mean curvature and by an elliptic function of Gaussian curvature. Motion by mean curvature is like the model in [2] by replacing gradients and Hessians by the corresponding operators on manifolds. The operator based on Gaussian curvature is described below.

**An example for video.** It is interesting to observe that we can apply this to the video case. In that case, $N = 3$ and $M = \{(x, t) : t \in \mathbb{R}, x \in \mathbb{R}^2\}$. Note that when we talk about video, $t$ represents the time coordinate. Let us consider the metric $g(x, t)_{ij}$ so that if $(y, \tau)$ denote the coordinates in $T_{(x,t)} M$, then

$$g(x, t) \begin{pmatrix} (y, \tau), (y, \tau) \end{pmatrix} = A(x, t)(y - v(x, t) \tau)^2 + B(x, t) \tau^2.$$  

As an example, we can take $A(x, t) = \alpha + |\nabla_x I|^p$, $B(x, t) = \beta + (\partial_t I)^p$, $\alpha, \beta > 0$, $p > 0$ (usually $p = 1, 2$), and $\partial_t I = v \cdot \nabla_x I + I_t$. Then det$(G) = A(x, t) B(x, t)$,

$$G^{-1} = \frac{1}{A} \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} + \frac{1}{B} \hat{v} \otimes \hat{v},$$

where $\hat{v} = (v, 1)$. This leads to an anisotropic filter in video. The dependence on the eigenvalues of $Q^t_p XQ_p$ can be expressed as a function of its trace (mean curvature) and determinant (Gaussian curvature) of its submatrix of rank 2. The Gaussian curvature of the level sets of $u$ can be expressed as

$$\det \left( \frac{1}{\sqrt{u}} (I - [e_{Du} \otimes e_{Du}]) D^2_M u + [e_{Du} \otimes e_{Du}] \right),$$
where we wrote \( [a \otimes b](x) = \langle a, x \rangle b = Ga \otimes b \), \( a, b, x \in T_\xi M \). Thus, examples of morphological operators can be given as elliptic functions of the Gaussian and mean curvature. Its well posedness in the context of manifolds using viscosity solution theory has been studied in [6].

**Remark 9.** Examples of first order, like

\[ u_s = |\nabla u|_g, \]

are related to the geodesic mathematical morphology introduced in [24].

**5. The case of linear scale spaces.** To describe the classification of linear scale spaces we need the following lemma.

**Lemma 5.1.** Let \( D \) be a matrix such that

\[ RDR^t = D \]

for all rotations \( R \) in \((T_\xi M, G(\xi))\). Then \( D = \lambda G(\xi)^{-1} \) for some \( \lambda \in \mathbb{R} \).

**Proof.** Let us write \( G \) instead of \( G(\xi) \). Note that \( R^t GR = G \), i.e., \( R^t D = D R \). Thus \( D \) commutes with any Euclidean rotation, and hence there is some \( \lambda \in \mathbb{R} \) such that \( D = \lambda I \). Then \( D = \lambda G^{-1} \). \( \square \)

**Theorem 5.2.** Let \( T_s \) be a multiscale analysis satisfying all architectural axioms, the comparison principle, and gray level shift invariance. Assume that \( T_s \) is linear. Then

\[ u_s = F(D^2_M u, \xi, G), \]

where

\[ F(A, \xi, G) = c(\xi) \text{Tr}(G^{-1} A) \]

for some function \( c \). The ellipticity of \( F \) implies that \( c \geq 0 \).

Thus linear scale spaces on \( M \) are given by the Laplace–Beltrami operator.

**Proof.** Since \( T_s \) is gray level shift invariant, then \( F \) does not depend on \( c \). On the other hand, it does not depend on \( \Gamma^k \). The linearity of \( T_s \) and Theorem 3.2 imply that in terms of the function \( \tilde{F} \),

\[ \tilde{F}(aX_1 + bX_2, ap_1 + bp_2, \xi, G) = a\tilde{F}(X_1, p_1, \xi, G) + b\tilde{F}(X_2, p_2, \xi, G) \]

for any \( X_1, X_2 \in SM(\xi)(N) \), any \( p_1, p_2 \in T_\xi M \), and any \( a, b \in \mathbb{R} \). By taking \( X_1 = X, X_2 = 0, p_1 = 0, p_2 = p, a = 1, b = 1 \), we write

\[ \tilde{F}(X, p, \xi, G) = \tilde{F}(X, 0, \xi, G) + \tilde{F}(0, p, \xi, G) = K'(X, \xi, G) + K''(p, \xi, G), \]
where $K'$ is linear in $X$ and $K''$ is linear in $p$. Moreover, from the rotation invariance of $\tilde{F}$,

$$\tilde{F}(X,p,\xi,G) = \tilde{F}(R^tXR,R^tp,\xi,G) \forall \text{ rotations } R \text{ in } (T_\xi M, G(\xi)),$$

we deduce that

$$K'(X,\xi,G) = K'(R^tXR,\xi,G),$$
$$K''(p,\xi,G) = K''(R^tp,\xi,G).$$

Since $K'$ is linear in $X$, then there exists a matrix $D$ (depending on $\xi, G$) such that

$$K'(X,\xi,G) = \text{Trace}(DX).$$

From the rotation invariance,

$$\text{Trace}(DX) = \text{Trace}(DR^tXR) = \text{Trace}(RDR^tX).$$

Since this is true for all $X$, then $RDR^t = D$. By Lemma 5.1 we have that $D = c(\xi, G)G^{-1}$ for some constant $c(\xi, G)$.

Note that we can write

$$K'(X,\xi,G) = c(\xi, G)\text{Trace}(G^{-1}X) = c(\xi, G)\text{Trace}(BB^tX)$$
$$= c(\xi, G)\text{Trace}(B^tXB) = H(B^tXB, 0, \xi),$$

where $H$ is linear in its first argument (apply the linearity argument to the function $H$ so that there exists a matrix $D'$ depending on $\xi$ such that $H(B^tXB, 0, \xi) = \text{Trace}(D'B^tXB)$; this implies that $D' = c(\xi, G)I$). Then $c$ does not depend on $G$. The ellipticity of $K$ proves that $c \geq 0$.

Now, by (4.6) we have

$$K''(p,\xi,G) = K''(R^tp,\xi,G) \forall \text{ rotations } R \text{ in } T_\xi M.$$ 

Thus $K''$ does not depend on $p$ but only on its modulus; that is,

$$K''(p,\xi,G) = \tilde{K}''(|p|_{g^{-1}},\xi,G)$$

for some function $\tilde{K}''$.

Let us compute the modulus. Observe that

$$\langle R^tp_1, R^tp_2 \rangle = (G^{-1}R^tp_1, R^tp_2) = (RG^{-1}R^tp_1, p_2).$$

From $R^tGR = G$, we have $RG^{-1}R^t = G^{-1}$. Thus

$$(RG^{-1}R^tp_1, p_2) = (G^{-1}p_1, p_2) = \langle p_1, p_2 \rangle.$$ 

Thus $|R^tp|_{g^{-1}} = |p|_{g^{-1}}$ for any covector $p$. Then

$$2\tilde{K}''(|p|_{g^{-1}},\xi,G) = K''(p,\xi,G) + K''(-p,\xi,G) = K''(0,\xi,G) = 0.$$ 

Our claim is proved.
6. The structure tensor as a metric. The structure tensor has been considered as a metric \[10, 58, 70\]. Here we show how to derive it naturally. For that, let us introduce the manifold of patches of a given image.

Let \(\Omega\) be the image domain, and assume that it is a rectangle in \(\mathbb{R}^N\). Let \(u \in L^q(\Omega)\), \(1 \leq q \leq \infty\). We assume that \(u\) has been extended to \(\mathbb{R}^N\) by an even symmetry and then by periodization so that \(u \in L^q_{\text{loc}}(\mathbb{R}^N)\).

Let \(\Omega_p\) be a neighborhood of 0. We take either \(\Omega_p\) as a closed ball or a closed hyperrectangle. Let \(\varphi \in C(\Omega_p)\), \(\varphi \geq 0\), and \(\varphi(x) > 0 \quad \forall x \in \text{int}(\Omega_p)\). We will assume that \(\varphi\) vanishes at the boundary of \(\Omega_p\).

Let \(p_u : \Omega \to L^q(\Omega_p)\), \(1 \leq q \leq \infty\), be the function \(p_u(x) : \Omega_p \to \mathbb{R}\) given by

\[
(6.1) \quad p_u(x)(h) = u(x+h)\varphi(h), \quad h \in \Omega_p.
\]

Let

\[
(6.2) \quad \mathcal{M}P(u, \Omega_p) := \{p_u(x) : x \in \Omega\}
\]

be the manifold of patches determined by \(u\) and \(\Omega_p\).

Notice that we defined the manifold of patches as a map with values in \(L^q(\Omega_p)\) with \(1 \leq q \leq \infty\). The precise value of \(q\) will depend on the supposed regularity of the image \(u\).

This assumption determines the Lipschitz character of \(p_u\) as a map with values in \(L^q(\Omega_p)\). If we assume \(u \in BV(\Omega)\), then we take \(q = 1\); if \(u \in W^{1,2}(\Omega)\), then we take \(q = 2\). This permits us to compute the derivative of the map \(u\). The precise notion of derivative will depend on each case (see \([41, 4]\)). Since we need only the case \(q = 1\) here, we will not discuss the case \(q = 2\), which offers a different perspective and produces some new models of anisotropic diffusion.

Later, in the next subsection we will comment on the convenience of modeling the “manifold of patches” in a continuous framework and its connection to nonlocal means \([12]\).

6.1. The manifold of patches in \(L^2\) of a \(W^{1,2}\)-function. Assume that \(u \in W^{1,2}(\Omega)\), and consider \(p_u : \Omega \to L^2(\Omega_p)\); then \(p_u\) is Lipschitz and the Lipschitz constant is bounded by \(\|u\|_{W^{1,2}}\). Let \(\phi \in C^1(\Omega_p)\) be a test function. Then

\[
(6.3) \quad \left\langle \frac{p_u(y) - p_u(x)}{|y-x|}, \varphi \phi \right\rangle - \left\langle \nabla p_u(x) \left( \frac{y-x}{|y-x|} \right), \varphi \phi \right\rangle \to 0^+ \quad \text{as} \quad y \to x.
\]

Since \(\left\langle \frac{p_u(y) - p_u(x)}{|y-x|}, \nabla p_u(x) \left( \frac{y-x}{|y-x|} \right) \right\rangle\) are uniformly bounded in \(L^2(\Omega_p)\), by a density argument, (6.3) holds for any \(\phi \in C(\Omega_p), \varphi\) of compact support. Then there is a linear map \(p_{\nabla u}(x) \in L^2(\Omega_p)^2\) such that the differential of \(p_u\) at \(x\) in the direction \(v \in \mathbb{R}^2\) is given by \(dp_{\nabla u}(x)(v) = p_{\nabla u,v}(x)\). The \(L^2\) norm of \(dp_{\nabla u}(x)(v)\) is given by

\[
\int_{\Omega_p} |\nabla u(x+h) \cdot v|^2 \varphi(h) \, dh = \int_{\Omega_p} \nabla u(x+h) \otimes \nabla u(x+h) \varphi(h) \, dh(v,v)
\]

\[
=: J_{\varphi}(\nabla u(x+\cdot) \otimes \nabla u(x+\cdot))(v,v).
\]

We see that the structure tensor appears as the metric induced in \(\Omega\) by the map \(p_u : \Omega \to L^2(\Omega_p)\).
Remark 10. From the point of view of modeling, the convenience of modeling the “manifold of patches” using a continuous framework is not clear. If we take a discrete one-dimensional image of an edge \( u : \{1, \ldots, N\} \rightarrow \{0, 1\} \) with \( u(i) = 0 \) for \( i \in \{1, \ldots, p\} \) and \( u(i) = 1 \) for \( i \in \{p + 1, \ldots, N\} \), \( 1 < p < N \), then the discrete patches of size 3 jump from point to point and we do not see the underlying manifold structure. In that case, it is perhaps better to model \( p_u \) as a graph with weights expressing the similarity of patches. But in any case, the modeling being discrete or continuous, the image of \( p_u \) is not an embedded manifold, and it has self-intersections and may also have double coverings depending on the periodicity and structures of the image. Note that the distance between two patches in the image of \( p_u \) is computed in the ambient space (be it \( L^2(\Omega) \) or \( L^1(\Omega) \)) and does not correspond to the geodesic distance computed on \( p_u(\Omega) \). This implies that the nonlocal means filter that takes averages with weights that depend on the distance in the ambient space cannot be interpreted, strictly speaking, as the heat equation in the manifold of patches.

Remark 11. As above, we assume that \( \Omega \) is a rectangle in \( \mathbb{R}^N \). Let \( u \in BV(\Omega) \). We assume that \( u \) has been extended to \( \mathbb{R}^N \) so that \( u \in BV_{loc}(\mathbb{R}^N) \). We can consider the map \( p_u : \Omega \rightarrow L^1(\Omega_p) \) as a Lipschitz map and study the metric differential (see [41, 4]). We will not pursue this here.

Let us now concentrate our attention on computing the structure tensor using compensated video.

6.2. The structure tensor on video. Let us consider the manifold
\[ \mathcal{M} = \{(x, t) : t \in \mathbb{R}, x \in \mathbb{R}^2\}. \]
We endow \( \mathcal{M} \) with any of the metrics for video previously mentioned. Let \( u(x, t) \) be a given video, and let \( v(x, t) \) be its optical flow. Fix \( (x, t) \in \mathcal{M} \). Let \( X(x, t + \tau) \) be the trajectory of a point such that
\[ X_\tau(t + \tau) = v(X(x, t + \tau), t + \tau) \]
and \( X(x, t) = x \). Let us use the following notation:
- We denote \( X = (x, t) \in \mathcal{M} \).
- \( \Omega_p(x) \) denotes some patch around \( x \in \mathbb{R}^2 \), like \( x + \Omega_p \), where \( \Omega_p \) is a patch around zero of radius \( R \).
- \( \Omega_p(x, t) = \Omega_p(x) \times (-\eta, \eta) \) for some \( \eta > 0 \). We write \( (h, \tau) \in \Omega_p(x, t) \).
- \( \Omega_p^c(x, t) = \{(X(x + h, t + \tau), t + \tau) : (h, \tau) \in \Omega_p(x, t)\} \). It is a compensated neighborhood, the motion of the points in \( \Omega_p(x) \) by the flow in the time interval \((-\eta, \eta)\).
- \( \Omega_p^c(X) = \{Y = (y, \tau) : g_{ij}(X)Y^iY^j \leq r^2\} \) for some \( r > 0 \).
- Let \( \mu \) be a weight measure on \( \Omega_p^c(x, t) \), either the usual Lebesgue measure or a weight to take into account the deformation induced by the flow map.

The motion compensated video is given by \( \bar{u}(x, \tau) = u(X(x + h, t + \tau), t + \tau) \).

Remark 12. What is the connection between \( \Omega_p^c(x, t) \) and \( (x, t) + \Omega_p^c(X) \)? Are they equivalent infinitesimally? Let us consider a point \((X(x + h, t + \tau), t + \tau) \in \Omega_p^c(x, t)\). It suffices to analyze \( X(x + h, t + \tau) \).

\[ X(x + h, t + \tau) \approx X(x, t) + X_\tau(x, t)\tau + D_x X(x, t)h = X(x, t) + v(x, t)\tau + D_x X(x, t)h, \]
if we define \( y = X(x+h,t+\tau) - X(x,t) \), then (replacing \( \approx \) by \( = \))

\[ y - v(x,t)\tau = D_x X(x,t)h. \]

As in the previous subsection, if we are working with the flow starting at time \( t \), then \( D_x X(x,t) = I \) and (for \( \tau \) small enough)

\[ y - v(x,t)\tau = h. \]

Thus \( \|y - v(x,t)\tau\| = \|h\| \leq R \). We are in a neighborhood \( \Omega_p(x) \) for some radius. We can read this argument to say that given \( \Omega \) with the metric \( G \) given by \((h,\tau)\). This is only approximated, say in a computational perspective.

We are going to compute the structure tensor on a video with two slightly different approaches. In the first one we work with the compensated video. In that case, the patch is cylindrical: a patch in space times a neighborhood in time. In the second computation we work with the original video in the compensated neighborhood; thus the patch is a motion deformed cylinder.

Let us develop the first approach. Let \( \tilde{u}(\bar{x},\tau) \) be the compensated video. Let \( \psi \) be a test function with support in \( \Omega^c_p(x,t) \). Let us compute the differential of the patch manifold map of the compensated video, assuming that the video is in \( W^{1,2} \).

Let us consider the change of variables

\[ \Omega_p(x) \times (-\eta,\eta) \to \Omega^c_p(x,t) \]

given by \((h,\tau) \to Y = (X(x+h,t+\tau),t+\tau)\). We are given in \( \Omega^c_p(x,t) \) a manifold structure with the metric \( G = (g_{ij}(x,t))_{ij} \). This induces a metric \( h_{ij} \) in \( \Omega_p(x) \times (-\eta,\eta) \) given by \( H = D_x X^T G D_x X \). This induces the density \(|H|^{1/2} = |G|^{1/2} |D_x X|\) when integrating in \( \Omega^c_p(x,t) \times (-\eta,\eta) \). We denoted \(|H| = \det(H)\), \(|G| = \det(G)\).

Let \( \Delta = (\Delta x,\Delta t) \) be time and space increments. Thus, the perturbation \((\Delta x,\Delta t)\) enters as the argument in \( \tilde{u}(h + \Delta x,\tau + \Delta t) \). Without loss of generality, assume that \((x,t) = (0,0)\).

Then keeping only terms of first order in \((\Delta x,\Delta t)\), we have

\[ \int_{-\eta}^{\eta} \int_{\Omega_p(x)} (\tilde{u}(h + \Delta x,\tau + \Delta t) - \tilde{u}(h,\tau))\psi(h,\tau)\mu(h,\tau)|H|^{1/2} dhdt \]

\[ = \int_{-\eta}^{\eta} \int_{\Omega_p(x)} (u(X(h + \Delta x,\tau + \Delta t),\tau + \Delta t) - u(X(h,\tau),\tau))\psi(h,\tau)\mu(h,\tau)|H|^{1/2} dhdt \]

\[ = \int_{-\eta}^{\eta} \int_{\Omega_p} [(u_t + v \cdot \nabla u)(X(h,\tau),\tau)]\Delta t + (D_x X(h,\tau))^T \nabla x u(X(h,\tau),\tau)\Delta x] \psi\mu[H|^{1/2} dhdt, \]

where \( \psi = \psi(h,\tau) \), \( \mu = \mu(h,\tau) \). Let us write this integral in \( \Omega^c_p(x,t) \). Since

\[ (D_x X(h,\tau))_\tau = Dv(X(h,\tau),\tau)D_x X(h,\tau) \]

and \( D_x X(h,0) = I \), then

\[ D_x X(h,\tau) = e^{\int_0^\tau Dv(X(h,\bar{\tau}),\bar{\tau}) d\bar{\tau}}. \]
Performing the change of variables \( \tau' = \tau \) (and keeping the notation \( \tau \)), \( y = X(h, \tau) \), we have \( |D_x X(dh d\tau) = dy d\tau, \)

\[
D_x X = e^{\int_0^\tau Dv(X(y,-\tau),\tau)} d\tau,
\]

and writing \( \delta_v u = u_t + v \cdot \nabla_x u, \) \( \bar{\psi}(y, \tau) = \psi(h, \tau), \) \( \bar{\mu}(y, \tau) = \mu(h, \tau), \) we can write the above integral as

\[
= \int_{\Omega_p} [\delta_v u(y, \tau) \Delta t + (D_x X)^t \nabla_x u(y, \tau) \Delta x] \bar{\psi}(y, \tau) \bar{\mu}(y, \tau)|G|^{1/2} d\tau dy.
\]

If we write

\[
\nabla u(y, \tau) = \begin{pmatrix} (D_x X)^t \nabla_x u(y, \tau) \
\delta_v u(y, \tau) \end{pmatrix},
\]

we can write the above integral as

\[
= \int_{\Omega_p(x,t)} \nabla u(y, \tau) \cdot \nabla \bar{\psi}(y, \tau) \bar{\mu}(y, \tau)|G|^{1/2} d\tau dy.
\]

Then the structure tensor is

\[
T_1(X)(\Delta, \bar{\Delta}) = \int_{\Omega_p(x,t)} \nabla u(y, \tau) \otimes \nabla u(y, \tau) \bar{\mu}(y, \tau)|G|^{1/2} d\tau dy(\Delta, \bar{\Delta}).
\]

**Remark 13.** What is the meaning of the scaling \( |G|^{1/2}(Y) \)? Recall that \( |G| = AB \), where \( A = 1 + |
abla_x I|^2, \) \( B = 1 + (\delta_v I)^2. \) Thus saying \( (AB)^{1/2} \bar{\Delta}^2 \leq \alpha \) is like saying that \( \bar{\Delta}^2 \leq \frac{\alpha}{(1+|\nabla_x I|^2)^{1/2}(1+(\delta_v I)^2)^{1/2}}. \) Thus the neighborhood is adaptive (anisotropic) and takes into account space and time edges (anisotropies).

**Remark 14.** What is the meaning of the scaling factor \( |D_x X|? \) Recall that in an occlusion \( |D_x X| \) is big. Thus saying \( |D_x X| \bar{\Delta}^2 \leq \alpha \) is like saying that \( \bar{\Delta}^2 \leq \frac{\alpha}{|D_x X|}. \) Thus the neighborhood is reduced due to occlusions. What is the intuition? When there is a compression like \( X(x) = \frac{x}{\lambda}, \lambda \gg 1, \) then \( |D_x X| = \frac{1}{\lambda n} \) (\( n \) is the space dimension) and \( \bar{\Delta}^2 \leq \frac{\alpha}{|D_x X|} \) means an expansion of the neighborhood.

**Remark 15.** As a possible approximation, we can take \( (D_x X)^t \approx (I + \tau Dv(y, \tau))^t. \) Also \( |D_x X|^{-1} \approx 1 - \int_0^\tau \text{div} v(X(y, -\tau), \tau) d\tau \approx 1 - \tau \text{div} v(h, \tau). \)

In the second computation we work with the original video in the compensated neighborhood; thus the patch is a motion deformed cylinder. In that case the perturbation \( \Delta = (\Delta x, \Delta t) \) enters directly into \( u: u(X + \Delta + Y). \) This leads to a different structure tensor \( T_2. \) In this case, we naturally included the density \( |G|^{1/2} \) and the factor \( |D_x X| \) does not appear. What is the difference with \( T_1? \) The gradients are different.

Let \( p_u: \mathcal{M} \to L^2(\Omega_p(X)). \) This is an abuse of notation to say that \( p_u(X) \in L^2(\Omega_p(X)) \) for \( X \in \mathcal{M}. \) We endow \( L^2(\Omega_p(X)) \) with the density \( |G| \) where \( g \) is the anisotropic metric defined in (4.8). We also use a weight \( \mu \) in case this would be convenient after our understanding of the situation. Let \( \psi \) be a test function with support in \( \Omega_p(X). \) Note that if \( \Delta = (\Delta t, \Delta x) \) is a small perturbation, then \( \psi \) is also a test function with support in \( \Omega_p(X + \Delta). \)
Using the first order Taylor expansion $u(X + \Delta x + Y) - u(X + Y) \approx D_xtu(X + Y)\Delta$ ($D_xt$ denotes the time-space gradient), we write

$$
\langle p_u(X + \Delta) - p_u(X), \psi \rangle_{H^1(\mathbb{R}^d, \mu)}
$$

$$
= \int_{\mathbb{R}^d} \left( u(X + \Delta x + Y) - u(X + Y) \right) \psi(Y) \mu(Y) |G|^{1/2}(Y) dY.
$$

Then the structure tensor is

$$
T_2(X)(\Delta, \Delta) = \int_{\mathbb{R}^d} D_xtu(X + Y) \otimes D_xtu(X + Y) \psi(Y) \mu(Y) |G|^{1/2}(Y) dY(\Delta, \Delta).
$$

**Remark 16.** Should one work with the compensated video in a cylindrical patch or with the original video in a moved patch? The main difference between both computations is that the perturbations ($\Delta x, \Delta t$) appear differently, and this leads to different gradients in the structure tensor: with convective derivative in $T_1$ and using $D_xtu$ in $T_2$.

7. **Variational models for image diffusion.** Finally, we review the variational formulation of image diffusion given in [40, 64, 39, 63], where the authors consider images defined on Riemannian manifolds with a metric that depends on the image and reflects the anisotropy of the underlying problem (designed for edge preservation, for color image restoration, for texture analysis, etc.). The basic energy functional is the Polyakov action, which is the extension of the Dirichlet integral to maps between Riemannian manifolds [40, 64]. We can directly see what the differences are between both formulations (that unify when considering weighted Polyakov actions).

Let us consider two manifolds $(\Sigma, g)$ and $(\mathcal{N}, h)$, where $g, h$ denote the respective metrics. Recall that $\sqrt{\det(G)} dx$ is the volume element in $\Sigma$. Let $\phi : \Sigma \to \mathcal{N}$. The Polyakov action is

$$
P(\phi) = \int_{\Sigma} g^{\mu \nu} \partial_{\mu} \phi^i \partial_{\nu} \phi^j h_{ij} |g|^{1/2} dx,
$$

where $|g|^{1/2} := \sqrt{\det(G)}$ and, as above, $g^{\mu \nu}$ is the inverse of the metric $g$. In the case where $\phi = u$ is a scalar, we have

$$
P(u) = \int_{\Sigma} g^{\mu \nu} \partial_{\mu} u \partial_{\nu} u |g|^{1/2} dx.
$$
Let $L^2(|g|^{1/2})$ denote the space of square integrable functions with respect to the density $|g|^{1/2}$. Let us consider some examples:

(i) Assume that $\Sigma$ is a surface and $\mathcal{N} = \mathbb{R}$, and consider

$$P(u) = \int_\Sigma g^{\mu\nu} \partial_\mu u \partial_\nu u |g|^{1/2} dx.$$  \hfill (7.3)

Then

$$\nabla_{L^2(|g|^{1/2})} P(u) = - \frac{1}{\sqrt{|g|}} \partial_\nu (\sqrt{|g|} g^{\mu\nu} \partial_\mu u).$$

The gradient descent in $L^2(|g|^{1/2})$ is

$$\min \frac{1}{2\Delta} \int_\Sigma (u - u^k)^2 |g|^{1/2} dx + P(u)$$

and leads to

$$u_s = \frac{1}{\sqrt{|g|}} \partial_\nu (\sqrt{|g|} g^{\mu\nu} \partial_\mu u).$$

This is an anisotropic diffusion equation introduced in [64, 39, 40, 63]. If $g$ is a diagonal matrix, then

$$P(u) = \int_\Sigma |\nabla u|^2 dx$$

and

$$\nabla_{L^2(|g|^{1/2})} P(u) = - \frac{1}{\sqrt{|g|}} \Delta u.$$

The gradient descent in $L^2(|g|^{1/2})$ leads to

$$u_s = \frac{1}{\sqrt{|g|}} \Delta u.$$  \hfill (7.5)

(ii) Weighted Polyakov action and descent. Let $\Lambda, \bar{\Lambda}$ be two given densities. Let us consider

$$\min \frac{1}{2\Delta} \int_\Sigma (u - u^k)^2 |g|^{1/2} \Lambda dx + P_\Lambda(u),$$

where

$$P_\Lambda(u) = \int_\Sigma g^{\mu\nu} \partial_\mu u \partial_\nu u |g|^{1/2} \Lambda dx.$$  \hfill (7.7)

Then

$$\nabla_{L^2(|g|^{1/2})} P_\Lambda(u) = - \frac{1}{\sqrt{|g|}} \partial_\nu (\bar{\Lambda} \sqrt{|g|} g^{\mu\nu} \partial_\mu u)$$

and the gradient descent in $L^2(|g|^{1/2})$ leads to

$$u_s \Lambda = \frac{1}{\sqrt{|g|}} \partial_\nu (\bar{\Lambda} \sqrt{|g|} g^{\mu\nu} \partial_\mu u).$$
If $\Lambda = \frac{1}{\sqrt{|g|}}$, we may write

\[(7.8)\quad u_s = \partial_\nu (\bar{\Lambda} \sqrt{|g|} g^{\mu\nu} \partial_\mu u).\]

If $\Lambda = 1$, we may write

\[(7.9)\quad u_s = \frac{1}{\sqrt{|g|}} \partial_\nu (\bar{\Lambda} \sqrt{|g|} g^{\mu\nu} \partial_\mu u).\]

This permits us to obtain anisotropic diffusion equations.

(iii) The TV case is given by

\[(7.10)\quad \mathcal{P}_{TV}(u) = \int_{\Sigma} \langle \bar{G}^{-1} \nabla u, \nabla u \rangle^{1/2} \sqrt{|g|} dx.\]

If $\Sigma$ is a surface and $g = \text{diag}(a, a)$, then $|g|^{1/2} = a$, and

\[(7.11)\quad \mathcal{P}_{TV}(u) = \int_{\Sigma} a^{1/2} |\nabla u| dx.\]

The gradient descent in $L^2$ is

\[u_s = \text{div} \left( a^{1/2} \frac{\nabla u}{|\nabla u|} \right),\]

while in $L^2(|g|^{1/2})$ it is

\[u_s = \frac{1}{a} \text{div} \left( a^{1/2} \frac{\nabla u}{|\nabla u|} \right).\]

(iv) An example in video smoothing. Let us consider the metric given in (4.8). If $A = \alpha + |\nabla_x I|^2$, $B = \beta + |\delta_t I|^2$ ($\delta_t I = v \cdot \nabla_x I + \partial_t I$), then

\[\langle G^{-1} \nabla_{xt} u, \nabla_{xt} u \rangle = \frac{1}{\alpha + |\nabla_x I|^2} |\nabla_x u|^2 + \frac{1}{\beta + |\delta_t I|^2} (v \cdot \nabla_x u + \partial_t u)^2.\]

This leads to an anisotropic filter in video.

Remark 17. We have considered weighted and unweighted versions of the Polyakov action. The weights also appear in [8, 9] and in [21, 22].

Remark 18. Variational models for image diffusion: the nonlocal case. Let us turn to the nonlocal versions of the Dirichlet integral. Let $J$ be a convex function and $w(x, y)$ be a kernel. Let $\bar{\Lambda}$ be a weight function. Let

\[(7.11)\quad \mathcal{E}_{\bar{\Lambda}}(u) = \int_{\Sigma} w(x, y) J(u(x) - u(y)) \bar{\Lambda} dx dy.\]
Although we shall not discuss this here, under some assumptions, after suitable localization, it converges to a weighted Polyakov action [44, 35]; see [12, 69, 71, 11, 10, 50].

8. Experimental results.

8.1. Image scale spaces. Let $\Omega \subset \mathbb{Z}^2$ be the discrete image domain. We compare two linear scale spaces for 2D images where the image manifold is given by a diagonal metric $g_{ij}$, in order to see the effect of the anisotropy inside or outside the divergence operator. The first one corresponds to the gradient descent given in

$$
\min \frac{1}{2\Lambda} \sum_{x \in \Omega} (u - u_k)^2 |g|^{1/2} \Lambda + \mathcal{P}_\Lambda(u),
$$

with the weighted Polyakov function

$$
\mathcal{P}_\Lambda(u) = \sum_{x \in \Omega} g^{\mu \nu} \partial_\mu u \partial_\nu u |g|^{1/2} \bar{\Lambda} = \sum_{x \in \Omega} c(x) |\nabla^+ u(x)|^2,
$$

where $\bar{\Lambda} = c : \Omega \rightarrow [0,1]$ is a scalar weighting function and $\nabla^+$ is a forward difference discretization of the gradient. Let us denote by $\nabla_i^+ u(x)$ the $i$th component of $\nabla^+ u(x) \in \mathbb{R}^2$ and define $e_1 = (1,0), e_2 = (0,1)$. Then

$$
\nabla_i^+ u(x) = \begin{cases} 
    u(x + e_i) - u(x) & \text{if } x + e_i \in \Omega, \\
    0 & \text{otherwise}.
\end{cases}
$$

Considering the weight $\Lambda = \frac{1}{\sqrt{|g|}}$, the discretized gradient flow of energy $\mathcal{P}_\Lambda$ is given by defining a scale space with the following evolution equation:

$$
\begin{aligned}
    u^{k+1}(x) &= u^k(x) - \delta s \nabla \mathcal{P}_\Lambda(u^k)(x) = u^k(x) + \delta s \text{div}^-(c(x)\nabla^+ u^k(x))
\end{aligned}
$$

(which in a continuous formulation corresponds to $u_s = \text{div}(c \nabla u)$). The discrete backward divergence operator is defined as the negative adjoint of $\nabla^+$. For a vector-valued image $g : x \in \Omega \mapsto (g_1(x), g_2(x)) \in \mathbb{R}^2$ we have

$$
\text{div}^- g(x) = \sum_{i=1}^2 \begin{cases} 
    g_i(x) & \text{if } x - e_i \notin \Omega, \\
    -g_i(x - e_i) & \text{if } x + e_i \notin \Omega, \\
    g_i(x) - g_i(x - e_i) & \text{otherwise}.
\end{cases}
$$

The scalar function $c$ modulates the diffusion. As in [57], we define it as a decreasing function of the modulus of the discrete gradient:
Figure 1. Edge maps $c(x)$ used for the scale space results shown in Figures 2 and 3.

$$c(x) = \exp \left( -\frac{1}{k^2} |\nabla^+ (G_\sigma * u)(x)|^2 \right),$$

where $G_\sigma$ is a Gaussian smoothing filter. For our experiments we set the filter standard deviation to 1 and $k = 20$. The resulting edge maps are shown in Figure 1 for three images.

For comparison, we also consider the linear scale space given by the Laplace–Beltrami flow when using a diagonal metric given by $g_{ij} = c^{-1} \delta_{ij}$. Note that it corresponds to the gradient descent of an unweighted Polyakov functional. Thus, we consider the gradient descent in $L^2(|g|^{1/2})$ of

$$(8.4) \quad \mathcal{P}(u) = \sum_{x \in \Omega} |\nabla^+ u|^2$$

given by

$$(8.5) \quad u^{k+1}(x) = u^k(x) + \delta s(x) \text{div}^{-} \nabla^+ u^k(x)$$

(which in a continuous formulation corresponds to $u_s = c \Delta u$). Note that $c$ appears outside the divergence operator in the resulting evolution equation.

Figures 2 and 3 show a comparison of both scale spaces for the images Barbara and Boat, respectively.

Results obtained with model (8.3) maintain the edges much better. At a strong edge for which $c \approx 0$ there will be no diffusion across the edge. On the other hand, if $c$ is used as a metric, its impact on the evolution is a local modulation of the time step: $\delta s(x) = \delta s c(x)$. Thus, diffusion is slowed down when $c(x)$ is small, but it does not prevent diffusion across edges.
8.2. Video scale spaces.

8.2.1. A mean curvature motion scale space. Let $u(x,t)$ be a given video sequence, and let $u(s,x,t)$ be the corresponding multiscale analysis, where $s \geq 0$ denotes the scale parameter ($t$ represents the time coordinate). In this subsection we study the scale space for video based on the mean curvature motion in the 3D video manifold described in (4.8). This model is a particular instance of the set of models described in Theorem 4.3 and could be written as
Figure 4. Example of motion-compensated spatio-temporal window of $3 \times 3 \times 3$ pixels. The blue lines correspond to the backward (left) and forward (right) optical flow of the central pixel of the spatial window at Frame $t$.

$$u_s = |\nabla u|_g H_g,$$

where $H_g(s, x, t) = \frac{\lambda_1(s, x, t) + \lambda_2(s, x, t)}{2}$, $\lambda_1$ and $\lambda_2$ being the principal curvatures of the level surface of $\{(s, x, t) \in \mathcal{M} : u = u(s, x, t)\}$ at the point $(x, t)$ (and scale $s$) when the manifold $\mathcal{M}$ is endowed with the metric $g$. Following (4.8) we shall consider the case where $v(x, t)$ is the optical flow of the original sequence and, after compensating the video, we can reduce it to the static case. In any case, following [31], the numerical scheme is reduced to an implementation of the iterated median filter using neighborhoods that are computed with the metric $g$. Although this may be a rough approximation, it will serve as an illustration of the models. Thus, we can just say that we present some experimental results for the geodesic and motion-compensated median scale space for video sequences. At each iteration, this filter computes for each pixel of the video sequence its median value over a geodesic ball of radius $h$ so that the final scale $s = nh^2$ ($n$ being the number of iterations). The geodesic distance is defined on the video manifold in terms of the spatial, temporal, and gray level distances, as described next.

First, a motion-compensated window is defined for each pixel in the following way. Given the video pixel $(x, t) \in \mathbb{R}^3$, i.e., pixel $x = (x_1, x_2) \in \mathbb{R}^2$ at frame $t$, we define its spatial neighborhood in the current frame by a squared window of $(2 \cdot W + 1) \times (2 \cdot W + 1)$ pixels, centered at the pixel being processed (that is, $W$ is the maximum distance from the center to the boundary of the spatial window). Then, for each one of the $T$ previous and the next $T$ frames, a spatial window of the same size is defined, centered at the position of the central pixel in that frame given by the optical flow (see the top part of Figure 4). In other words, the square window in the current frame is temporally translated following the motion trajectory of the pixel being processed. Once the motion-compensated spatial window in each frame has been extracted, they can be aligned to form a $(2 \cdot W + 1) \times (2 \cdot W + 1) \times (2 \cdot T + 1)$ 3D pixel block that can be normally processed (see the bottom part of Figure 4).
Note that the previous choice of the motion-compensated window represents a simplification since we could use any suitable representative of the vector flow at the central pixel or the more sophisticated mapping of each pixel in the central window. In any case, we choose this approach for reasons of simplicity.

In our next step we compute the geodesic distance from the central pixel (i.e., the pixel being processed) to the rest of the pixels in the motion-compensated spatio-temporal window. For that purpose, we define a metric and use it to compute the distance between adjacent pixels in the motion-compensated window. Then, the Dijkstra algorithm [25] is applied on the resulting weighted graph to determine the minimum distance path from the central pixel to the rest of the window pixels. We consider the total cost of the minimum paths provided by the Dijkstra algorithm as the geodesic distance in terms of the specified metric.

Particularly, in all experiments we consider 6-connectivity on the 3D motion-compensated window (for each pixel, the left, right, top, and bottom spatial neighbors; and the previous and next temporal neighbors). Given the coordinates of two pixels in the motion-compensated 3D window \((x, t), (\overline{x}, \overline{t}) \in \mathbb{R}^3\) and their corresponding gray level value in the video sequence, \(u(x, t), u(\overline{x}, \overline{t})\), the metric used to define the weights of the graph is given by

\[
(8.6) \quad w((x, t), (\overline{x}, \overline{t})) = k_{x_1} \cdot (x_1 - \overline{x}_1)^2 + k_{x_2} \cdot (x_2 - \overline{x}_2)^2 + k_t \cdot (t - \overline{t})^2 + k_c \cdot [u(x, t) - u(\overline{x}, \overline{t})]^2
\]

for some parameters \(k_{x_1}, k_{x_2}, k_t, k_c \geq 0\) that scale the magnitude of the geodesic distance in each dimension, respectively. In all experiments shown, we consider \(k_{x_1} = k_{x_2}\). Note that this follows from the expression (4.8) with a particular choice of parameters. Experimentally, we have observed that the scale space obtained is not particularly sensitive to parameter selection.

After computing the geodesic distance from the central pixel to the rest of the window, the pixels belonging to a geodesic ball of radius \(h\) are determined. This is done simply by considering only those pixels whose geodesic distance to the central pixel is smaller than or equal to the ball radius, \(h\). Finally, at each iteration, the median value of this subset (pixels belonging to the geodesic ball of radius \(h\)) is computed and set as the corresponding value of the pixel being processed in the median scale-space output video. After \(n\) iterations, we arrived at the scale \(s = nh^2\) [31].

An example of the median scale space obtained for the Foreman video sequence is shown in Figures 5 and 6 for geodesic balls of radii \(h = 1\) and \(h = 2\), respectively. A motion-compensated window of size \(11 \times 11 \times 5\) pixels was used (that is, a spatial neighborhood of \(11 \times 11\) pixels in the two previous frames, the current frame, and next two frames). To prevent the motion-compensated window from going outside the video domain in any case, video frames were properly extended by mirroring. This simple and fast approximative solution was enough to prevent the appearance of border effects in our results without significantly increasing the computational time involved in processing each frame. Nevertheless, note that this is just a practical choice for the sake of efficiency and that more accurate image extension methods can be used, for instance, image inpainting [23, 14, 5].

The motion trajectory of the pixel being processed, needed to build the motion-compensated window, was estimated by the dense optical flow provided by the Horn and Schunck algorithm [32]. Particularly, the color-based multilevel implementation provided by [65] was used.
The evolution of the median scale space as the number of iteration increases can be observed in both Figures 5 and 6 (each row corresponds to an iteration, increasing from top to bottom). It can be seen that larger values of $h$ and larger numbers of iterations lead to larger simplification of the video sequence but preserving the most relevant information, such as edges and region homogeneity.

Figures 7 and 8 present similar results for a noise corrupted version of the previous sequence (zero-mean Gaussian noise with standard deviation $\sigma = 20$) for geodesic balls of radii $h = 1$ and $h = 2$, respectively. In this situation it can be seen that denoising requires a minimum geodesic ball radius in order to be effective. On one hand, it can be observed that a radius of $h = 1$ (Figure 7) is not large enough to reach the neighboring pixels when the geodesic ball is centered at a noisy pixel. For that reason, the generated scales are similar to the original noisy frames. On the other hand, a radius of $h = 2$ is able to include a larger set of neighboring pixels and the median acts as usual, leading to a significant noise reduction.

To better illustrate the denoising ability of the motion-compensated median scale space, Figure 9 shows scales generated by fixing the number of iterations (particularly, $n = 1$) and varying only the radius of the geodesic ball, $h$. It can be seen that the median scale space leads to a significant noise reduction for scales larger than $s = 3^2$. Figure 10 presents similar results for the Claire sequence, where the original sequence has been corrupted not only by Gaussian noise (zero mean, $\sigma = 20$) but also by flickering artifacts in the form of black rectangles of random size and position that appear at some frames.

As before, it can be observed that the median scale space is robust in the presence of independent and identically distributed spatio-temporal noise, such as the Gaussian noise. Nevertheless, it does not help in removing or shrinking the flickering artifacts. The reason is that, from a random process perspective, flickering artifacts are far from being spatially independent noise.
Figure 5. Median scale space for Foreman sequence (frame size: 176 × 144 pixels). Frames 1, 2, 3, 4, and 5 (columns from left to right) are processed at scales $s = 1, 2, 3, 4, 5,$ and $6$ (rows from top to bottom). The scale is determined by $s = n \cdot h^2$, where $n$ is the number of iterations (the first iteration corresponds to the second row, the second iteration to the third row, and so on), and $h = 1$ is the radius of the geodesic ball. Metric parameters: $k_{x1} = k_{x2} = k_t = k_c = 0.1$. Motion-compensated window of $11 \times 11 \times 5$ pixels.
Figure 6. Median scale space for Foreman sequence (frame size: 176 × 144 pixels). Frames 1, 2, 3, 4, and 5 (columns from left to right) are processed at scales $s = 4, 8, 12, 16, 20, 24$ (rows from top to bottom). The scale is determined by $s = n \cdot h^2$, where $n$ is the number of iterations (the first iteration corresponds to the second row, the second iteration to the third row, and so on), and $h = 2$ is the radius of the geodesic ball. Metric parameters: $k_{r1} = k_{r2} = k_t = k_e = 0.1$. Motion-compensated window of $11 \times 11 \times 5$ pixels.
Figure 7. Median scale space for Foreman sequence (frame size: 176 × 144 pixels) corrupted by Gaussian noise (zero mean, \( \sigma = 20 \)). Frames 1, 2, 3, 4, and 5 (columns from left to right) are processed at scales \( s = 1, 2, 3, 4, 5, \) and 6 (rows from top to bottom). The scale is determined by \( s = n \cdot h^2 \), where \( n \) is the number of iterations (the first iteration corresponds to the second row, the second iteration to the third row, and so on), and \( h = 1 \) is the radius of the geodesic ball. Metric parameters: \( k_{x_1} = k_{x_2} = k_i = k_c = 0.1 \). Motion-compensated window of 11 × 11 × 5 pixels.
Figure 8. Median scale space for Foreman sequence (frame size: 176 × 144 pixels) corrupted by Gaussian noise (zero mean, σ = 20). Frames 1, 2, 3, 4, and 5 (columns from left to right) are processed at scales s = 4, 8, 12, 16, 20, and 24 (rows from top to bottom). The scale is determined by s = n·h², where n is the number of iterations (the first iteration corresponds to the second row, the second iteration to the third row, and so on), and h = 2 is the radius of the geodesic ball. Metric parameters: k_{x1} = k_{x2} = k_{t} = k_{c} = 0.1. Motion-compensated window of 11 × 11 × 5 pixels.
Figure 9. Median scale space for Foreman sequence (frame size: 176 × 144 pixels) corrupted by Gaussian noise (zero mean, $\sigma = 20$). Frames 1, 3, 5, 7, and 9 (columns from left to right) are processed at different scales ($s = n \cdot h^2$) varying the radius of the geodesic ball and fixing the number of iterations ($n = 1$). Rows from top to bottom: $s = 1^2$, $2^2$, $3^2$, $4^2$, $5^2$, and $6^2$. Metric parameters: $k_{x_1} = k_{x_2} = k_t = k_c = 0.1$. Motion-compensated window of $11 \times 11 \times 5$ pixels.
Figure 10. Median scale space for Claire sequence (frame size: 176 × 144 pixels) corrupted by Gaussian noise (zero mean, σ = 20) and flickering artifacts (black rectangles of random size and position). Frames 1, 3, 5, 7, and 9 (columns from left to right) are processed at different scales (s = n · h^2) varying the radius of the geodesic ball and fixing the number of iterations (n = 1). Rows from top to bottom: s = 1^2, 2^2, 3^2, 4^2, 5^2, 6^2, and 7^2. Metric parameters: k_{x_1} = k_{x_2} = k_t = k_c = 0.1. Motion-compensated window of 11 × 11 × 5 pixels.
Particularly, the central assumption in a median filter is that the most frequent value in the neighborhood of a pixel is its correct (noise-free) value. It cannot be ensured that the previous assumption will hold when a spatio-temporal neighborhood is considered in a video corrupted by such a type of artifacts. But it is even worse when the neighborhood is defined in terms of the geodesic distance given by (8.6). For instance, consider that we are centered at a pixel affected by a flickering artifact. In terms of the geodesic distance given by (8.6), the pixel is relatively close to the rest of pixels of the same frame belonging to the black rectangle. On the contrary, the distance to uncorrupted pixels outside the rectangle (from the same or different frame) is in general significantly larger (since they may have a large difference in gray level). In other words, the distance function centered at a corrupted pixel suffers a sudden increase for pixels outside the black rectangle. The consequence of this discontinuity is that, unless the radius of the geodesic ball used to select the neighborhood is large enough, the geodesic ball will include in its interior only the black region. In that situation the median filter will output a black value for the pixel being processed and, in general, it will leave the artifact unaltered. Hence, the random rectangles can be removed only at very large scales (large radius of the geodesic balls). Unfortunately, it is likely that at such scales the most relevant video information has already disappeared.

8.2.2. AMG scale space. In [30], a morphological, affine, and Galilean invariant scale space for video sequences was presented. In its original formulation, the video sequence was processed as a 3D pixel array without considering in its formulation the motion trajectory of the video pixels. In this subsection, we present its motion-compensated extension, that is, considering temporal adjacency not directly defined by the 3D video array but defined by the motion trajectories described by the pixels in the video sequence, provided by the optical flow estimation. Thus, the goal of this subsection is to compare the original formulation to its motion-compensated version, and, for that purpose, both scale spaces have been implemented and the experiments in [30] have been replicated.

From an axiomatic approach, it is shown in [30] that the scale space with the previous properties (morphological, affine, and Galilean invariant) is given by a family of nonlinear parabolic equations, referred to as AMG (affine, morphological, and Galilean). If \( u_0(x,t) \) for \((x,t) \in \mathbb{R}^3\) denotes the video sequence, the general model that generates the scale space \( u(s,x,t) \) can be written as

\[
\frac{\partial u}{\partial s} = |\nabla u| F(\text{curv}(u), \text{accel}(u)),
\]

where the initial condition is \( u(0,x,t) \), \( \nabla u \) denotes the spatial gradient, \( \text{curv}(u) \) represents the (spatial) curvature of the (2D) level lines of \( u \) (on a frame), and \( \text{accel}(u) \) is the acceleration in the direction of the spatial gradient (referred to as apparent acceleration). We note that \( \text{curv}(u) \text{accel}(u) \) is proportional to the Gaussian curvature of the level surfaces of \( u \) [30, 31]. We single out a particular instance of (8.7), the scale space that considers spatial and temporal dependency of the video sequence given by the equation

\[
\frac{\partial u}{\partial s} = \text{curv}(u)^{1/3}|\nabla u| \{\text{sign[\text{curv}(u)\text{accel}(u)]}\}^+, \]

where \( r^+ = \max(r, 0) \). Further details on each term can be found in [30]. We explore the extension of this model to the video manifold \((M, g)\) described in (4.8) (again a particular
instance of the set of models described in Theorem 4.3). Thus this model can be written
in terms of the curvature of the level lines of 𝑢 (spatial information on each frame) and the
Gaussian curvature of the level surfaces of 𝑢 in the manifold (𝑀, 𝑔). Since the numerical
approach can be based on the corresponding one for (8.8), after motion compensation we
recall in detail Guichard’s scheme for (8.8) [30].

Let us first say that we approximate the curvature term \(\text{curv}(𝑢)^{1/3}\) by central finite differ-
ences. The discretization of the term \(|∇𝑢|\{\text{sign}\{\text{curv}(𝑢)\}\text{accel}(𝑢)\}\}^+\) is performed by the same
numerical scheme proposed in [30]:

\[
\begin{align*}
\text{(8.9)} \quad |∇𝑢| \{\text{sign}\{\text{curv}(𝑢)\}\text{accel}(𝑢)\}\}^+ & \approx \min_{𝑣_𝑏, 𝑣_𝑓 \in \mathbb{W}} \frac{1}{\Delta 𝑡^2} \left[ |u(s, x - 𝑣_𝑏, 𝑡 - \Delta 𝑡) − u(s, x, 𝑡)| \\
+ |u(s, x + 𝑣_𝑓, 𝑡 + \Delta 𝑡) − u(s, x, 𝑡)| + \Delta 𝑡| < ∇𝑢, 𝑣_𝑏 − 𝑣_𝑓 > | \right],
\end{align*}
\]

i.e., the minimum with respect to all possible backward (resp., forward) displacement vectors,
\(𝑣_𝑏 \in \mathbb{R}^3\) (resp., \(𝑣_𝑓 \in \mathbb{R}^3\), defined in a 3D spatio-temporal window \(\mathbb{W}\), going from the pixel
being processed in the current frame to the previous (resp., next) frames. Hence, \(x + 𝑣_𝑓 \Delta 𝑡\)
describes a pixel of the next frame and \(x − 𝑣_𝑏 \Delta 𝑡\) a pixel of the previous frame. If only a single
frame is considered before and after the current frame, then the frame interval will be the
unit, that is, \(\Delta 𝑡 = 1\). This is the case for all experiments in this subsection.

Summarizing, the left term in (8.9) is given by the minimum difference in gray level
provided by any of the possible pixel displacements in the spatio-temporal window defined in
the previous and next frames, which simultaneously leads to a small projection in the spatial
gradient direction.

As in the original implementation in [30], we maintain the boundary pixels (in time and
space) to their initial values. For that reason, in all experiments shown there will be no
evolution in the first and last frames of the video sequences.

The only difference between the scale space proposed by [30] and our extension is that, in
the first case, the spatio-temporal window is directly considered in the 3D video array without
motion compensation, while, in our approach, a motion-compensated spatio-temporal window
is used, computed considering motion trajectories of the pixels as described in subsection 8.2.1.

One of the limitations of the original scale space in [30], discussed in the paper by the
author, is the tendency of the original AMG model to remove any object whose trajectory is
not a translational motion. This is illustrated by the following experiment, performed on the
test video sequence in Figure 11. It is formed by 24 frames of 64 × 64 pixels where two circles
are shown: a black circle moving with a constant velocity around a still white circle. Although
the modulus of the speed of the black circle is constant, its acceleration is not null, and this
causes the original AMG model to shrink it until removing it. The effect of the original AMG
model is shown in Figure 12. On the other hand, our motion-compensated version does not
significantly modify the original video, as can be seen in Figure 13. The motion-compensated
model does not shrink or remove any image object as long as a motion trajectory can be found
for its pixels all along the video sequence. Hence, the scale space obtained does not depend
on the nature of the motion of the objects but on the temporal consistency of the motion in
the video.
Figure 11. Original test sequence (24 frames, 64 × 64 pixels per frame). Video frames are shown in increasing temporal order from left to right, from top to bottom.

Figure 12. Video sequence in Figure 11 processed by the AMG equation proposed in [30]. A spatio-temporal window of 3 × 3 × 3 (5 × 3 spatial window in the previous, current, and next frames) was used. Time increment in the discretization of the equation: δs = 0.05. Results at t = 6.

Next, we replicate one of the experiments shown in [30]. Figure 14 shows a synthetic video sequence formed by 24 frames of 64 × 64 pixels. This video represents four black circles in motion. From top to bottom, the first one is in horizontal translation motion, with the speed of two pixels by frame. The second one also has a constant horizontal speed of two pixels by frame, but it has a vertical acceleration since its vertical speed is equal to 1, then 0, then −1, then 1, and so on. The third one moves from left to right with acceleration equal to 1 or −1. The fourth one moves with constant speed on a circumference.

For the scale and window size selected in the experiment in [30] it was shown that the original AMG model did not modify the first circle (since it was a translational motion), it partially shrink the second circle, and it completely removed the last two. When we replicate the experiment we have observed that the final result is particularly sensitive to the relation between speed modulus and window size. For instance, for a spatio-temporal window of size 3 × 3 × 3 pixels, the original AMG model removes the first two circles much faster than the last two, not behaving as theoretically expected. This example is shown in Figure 15. On the contrary, the motion-compensated AMG model does not show this strong sensitivity. As can be observed in Figure 16, for the same scale and window size as the results in Figure 15, the circles have not been modified and look identical to the original test sequence. In any case, the original AMG model behaves as expected when a sufficiently large spatio-temporal window is considered.
The next experiment, also originally in [30], illustrates the behavior of the AMG model in the presence of noise and flickering artifacts. For that purpose, the test sequence in Figure 17 is considered. The top square is moving with a uniform horizontal speed of one pixel per frame. The bottom square moves in random directions but always with a speed of modulus one. The video has been corrupted by noise uniformly distributed in the interval $[-125, 125]$ (the gray levels of the black and white pixels being 0 and 255, respectively). Some flickering artifacts in the form of rectangles of random size and position were introduced in some frames.

In this case, the results are similar to those reported in [30] (see Figures 18 and 19 for results of the original and motion-compensated AMG models, respectively). Both models have the ability to significantly denoise the severely corrupted video sequence, although the original structure of the square is partially shrinked. It can be observed that, for the same scale, the motion-compensated AMG model is more evolved than the original model, in the sense that the level of noise is slightly lower but also the square structure is slightly more smoothed. In other words, denoising by the motion-compensated AMG model is slightly faster than by the original model. In both cases, the flickering artifacts have been significantly removed. This is an interesting property of the generated scale space that may be valuable to some applications, such as old movie restoration [36] or correction of transmission errors in video decoding [67].

Finally, we compare both models for real video sequences. Figures 20 and 21 present the results at different scales of some frames of the Foreman sequence for the original and motion-compensated models, respectively. It can be observed that, in the presence of some motion, the original model generates artifacts (for instance, see the first two columns of Figure 20 that correspond to the scale space of Frames 14 and 15). On the contrary, the motion-compensated model presents no artifacts and leads to a correct video simplification as the scale increases (observe the scale space generated in this case for Frames 14 and 15 in the first two columns of Figure 21).

Similarly, Figures 22 and 23 show the scale space generated for the same sequence but this time corrupted by Gaussian noise and flickering artifacts in the form of black rectangles of arbitrary size and position for the original and motion-compensated models, respectively. Again, it can be seen that the original model leads to severe artifacts on the generated scale space caused by the presence of motion in the video sequence. On the other hand, the motion-
compensated model provides a correct scale space. It can be also observed how the flickering artifacts are properly shrunk and finally removed by the motion-compensated model as the scale increases. Particularly, the first two columns of Figure 23 present two adjacent frames (Frames 14 and 15) and it can be seen how the flickering artifact in Frame 15 introduces some artifacts also at the first scales of Frame 14. Nevertheless, as the scale increases, the information from adjacent frames is propagated into Frame 15 and the flickering artifact is completed with temporally coherent information. Similar results are shown in Figure 24 for the Claire sequence corrupted by noise and flickering artifacts.

8.2.3. Linear scale space. This subsection presents some experiments on the linear scale space for video sequences proposed in section 5 in the video manifold \((\mathcal{M}, g)\) described in (4.8). For the sake of completeness and clarity, here we review the equation involved and present the implementation details.

Given the video sequence \(u(x_1, x_2, t), (x_1, x_2, t) \in \mathbb{R}^3\), the backward and forward optical flows are estimated as in the previous subsection, using the Horn and Schunck algorithm [32] (particularly, the implementation provided by [65]). Then, we compute the following temporal term:

\[
B(x_1, x_2, t) = \beta + |\partial_t u|,
\]
Figure 16. Video sequence in Figure 11 processed by the AMG equation using motion-compensated scale-space. A spatio-temporal motion-compensated window of $3 \times 3 \times 3$ ($3 \times 3$ pixels and previous, current, and next frames) was used. Time increment in the discretization of the equation: $\delta s = 0.05$. Results at $t = 10$.

Figure 17. Original test sequence (24 frames, 64 \times 64 pixels per frame). Video frames are shown in increasing temporal order from left to right, from top to bottom.

where $\beta > 0$ is a parameter weighting the contribution of the temporal dimension to the scale space, and $\partial_t u$ is the convective derivative of the video sequence, that is, the first derivative along the motion trajectory of each pixel. This last term is approximated by finite differences using the backward and forward optical flows at each pixel.

The next term we compute is the tensor product of the optical flow at each pixel (for instance, combining the backward and forward optical flow motion vectors). Given the optical flow vector at each video pixel,

$v(x_1, x_2, t) = (v_{x_1}(x_1, x_2, t), v_{x_2}(x_1, x_2, t))$

and letting $v(x_1, x_2, t) = (v(x_1, x_2, t), 1)$, the tensor product term is given by

$$(v \otimes v)(x_1, x_2, t) = v(x_1, x_2, t) \otimes v(x_1, x_2, t)$$

$$= \begin{pmatrix}
  v_{x_1}^2(x_1, x_2, t) & v_{x_1}(x_1, x_2, t) \cdot v_{x_2}(x_1, x_2, t) & v_{x_1}(x_1, x_2, t) \\
  v_{x_1}(x_1, x_2, t) \cdot v_{x_2}(x_1, x_2, t) & v_{x_2}^2(x_1, x_2, t) & v_{x_2}(x_1, x_2, t) \\
  v_{x_1}(x_1, x_2, t) & v_{x_2}(x_1, x_2, t) & 1
\end{pmatrix}.$$
Next, the spatial terms are computed:

\[(8.11) \quad A(x_1, x_2, t) = \alpha + \| \nabla_{x_1,x_2} u(x_1, x_2, t) \|^2 = \alpha + \left[ \left( \frac{\partial u(x_1,x_2,t)}{\partial x_1} \right)^2 + \left( \frac{\partial u(x_1,x_2,t)}{\partial x_2} \right)^2 \right], \]

where \(\alpha > 0\) is a parameter weighting the contribution of the spatial dimension to the scale space, and \(\| \nabla_{x_1,x_2} u(x_1, x_2, t) \|\) is the norm of the spatial gradient of the video sequence at each pixel.

Then, we determine

\[(8.12) \quad \det(G)(x_1, x_2, t) = A(x_1, x_2, t) \cdot B(x_1, x_2, t), \]

\[(8.13) \quad G^{-1}(x_1, x_2, t) = \frac{1}{A(x_1, x_2, t)} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \frac{1}{B(x_1, x_2, t)} \cdot (v \otimes v)(x_1, x_2, t), \]

where \(\det(G)(x_1, x_2, t) \in \mathbb{R}\) and \(G^{-1}(x_1, x_2, t) \in \mathbb{R}^{3 \times 3}\).
Figure 20. Original AMG scale space proposed in [30] for Foreman sequence (frame size: 176 × 144 pixels). Frames 14, 15, 17, 22, and 24 (columns from left to right) are processed at scales $s = 0.5, 1.5, 2.5, 3.5, 4.5, \text{ and } 5.5$ (rows from top to bottom). Time increment in the discretization of the equation: $\delta s = 0.5$. 
Figure 21. Motion-compensated AMG scale space for Foreman sequence (frame size: 176 × 144 pixels). Frames 14, 15, 17, 22, and 24 (columns from left to right) are processed at scales $s = 0.5, 1.5, 2.5, 3.5, 4.5,$ and $5.5$ (rows from top to bottom). Time increment in the discretization of the equation: $\delta s = 0.5$. 
Figure 22. Original AMG scale space proposed in [30] for Foreman sequence (frame size: 176 × 144 pixels) corrupted by Gaussian noise (zero mean, σ = 20) and flickering artifacts in the form of black rectangles of random size and position. Frames 14, 15, 17, 22, and 24 (columns from left to right) are processed at scales s = 0.5, 1.5, 2.5, 3.5, 4.5, and 5.5 (rows from top to bottom). Time increment in the discretization of the equation: δs = 0.5.
Figure 23. Motion-compensated AMG scale space for Foreman sequence (frame size: 176 × 144 pixels) corrupted by Gaussian noise (zero mean, \( \sigma = 20 \)) and flickering artifacts in the form of black rectangles of random size and position. Frames 14, 15, 17, 22, and 24 (columns from left to right) are processed at scales \( s = 0.5, 1.5, 2.5, 3.5, 4.5, \) and 5.5 (rows from top to bottom). Time increment in the discretization of the equation: \( \delta s = 0.5 \).
Figure 24. Motion-compensated AMG scale space for Claire sequence (frame size: 176 × 144 pixels) corrupted by Gaussian noise (zero mean, $\sigma = 20$) and flickering artifacts in the form of black rectangles of random size and position. Frames 7, 9, 10, 23, and 24 (columns from left to right) are processed at scales $s = 0.5, 1.5, 2.5, 3.5, 4.5$, and 5.5 (rows from top to bottom). Time increment in the discretization of the equation: $\delta s = 0.5$. 
Next, we compute

\[
T(x_1, x_2, t) = G^{-1}(x_1, x_2, t) \cdot \left( \begin{array}{c} \frac{\partial u}{\partial x_1} \\ \frac{\partial u}{\partial x_2} \\ \frac{\partial u}{\partial t} \end{array} \right) = (T_1(x_1, x_2, t) \ T_2(x_1, x_2, t) \ T_3(x_1, x_2, t))
\]

and, finally, the linear scale space is iteratively generated at different scales using the explicit scheme provided by finite differences:

\[
u^{s+\delta s}(x_1, x_2, t) = u^s(x_1, x_2, t) + \delta s \cdot \frac{1}{\sqrt{\det(G)(x_1, x_2, t)}} \cdot \text{div}(T(x_1, x_2, t)),
\]

where the divergence term (with respect to the three variables), \(\text{div}(T(x_1, x_2, t))\), is approximated by central finite differences.

Figure 25 presents the linear scale space generated for the Foreman sequence by the described scheme. It can be seen that the scale space is very conservative, slightly smoothing the sequence and keeping the most relevant video information. The effects of the linear scale space are more obvious when applied to noisy sequences. Figure 26 shows the scale space generated for the same noise corrupted version of the Foreman sequence used in subsection 8.2.1. In this situation, its denoising ability can be seen, particularly for the three largest scales shown (last three rows).

To provide a better understanding of the linear scale space for video sequences, we study the spatial and temporal terms independently. Removing the temporal contribution from the scale space (that is, removing \(B(x_1, x_2, t)\) from (8.12) and (8.13)) leads to applying the well-known spatial linear scale space \[72\] frame by frame, without using any temporal information. Nevertheless, the effect of removing the spatial term and using only temporal information in the generation of the linear scale space (that is, removing \(A(x_1, x_2, t)\) from (8.12) and (8.13)) is not so clear. Figure 27 presents the linear scale space generated exclusively using temporal information. As can be seen, the video sequence remains mainly unaltered for any scale. The reason is that as long as a pixel can have a good correspondence at each frame, that is, a motion trajectory all along the video sequence, the temporal information is consistent and the temporal scale space has no effect.

When the linear scale space using only temporal information is applied to the noisy version of the Foreman sequence (shown in Figure 28), a mild denoising is observed. In this case, the formation of consistent pixel trajectories along the video sequence provides some averaging along the motion trajectories generated by the optical flow. However, since no spatial information is considered in the scale space, the spatial consistency of neighboring pixels with similar gray levels is not assured. That means that even for large scales adjacent pixels may have significant gray level differences as a consequence of noise, and that the video sequence is not further denoised as long as a consistent temporal motion trajectory is found for each noisy pixel.

**Acknowledgments.** We thank P. Arias, J. Calvo, and C. Ballester for interesting discussions on this subject. Particularly, we thank P. Arias for providing us with the experiments in subsection 8.1.
Figure 25. Linear scale space for Foreman sequence (frame size: 176 × 144 pixels). Frames 1, 3, 5, 7, and 9 (columns from left to right) are processed at increasing scales (rows from top to bottom). The exact scale value is shown at the beginning of each row. Parameter setting: $\alpha = 0.01$, $\beta = 0.1$, $\delta s = 50$. 
Figure 26. Linear scale space for Foreman sequence (frame size: 176 × 44 pixels) corrupted by Gaussian noise (zero mean, $\sigma = 20$). Frames 1, 3, 5, 7, and 9 (columns from left to right) are processed at increasing scales (rows from top to bottom). The exact scale value is shown at the beginning of each row. Parameter setting: $\alpha = 0.01$, $\beta = 0.1$, $\delta s = 150$. 
Figure 27. Linear scale space using only temporal information for Foreman sequence (frame size: 176×144 pixels). Frames 1, 3, 5, 7, and 9 (columns from left to right) are processed at increasing scales (rows from top to bottom). The exact scale value is shown at the beginning of each row. Parameter setting: $\beta = 0.1$, $\delta s = 0.5$. 
Figure 28. Linear scale space using only temporal information for Foreman sequence (frame size: 176 × 144 pixels) corrupted by Gaussian noise (zero mean, $\sigma = 20$). Frames 1, 3, 5, 7, and 9 (columns from left to right) are processed at increasing scales (rows from top to bottom). The exact scale value is shown at the beginning of each row. Parameter setting: $\beta = 0.1$, $\delta s = 0.5$. 
REFERENCES

MULTISCALE ANALYSIS OF IMAGES ON MANIFOLDS


