State of the Art on: Functional Maps for Shape Matching and Function Transfer

Michele Colombo, michele11.colombo@mail.polimi.it

1. Introduction to the research topic

An important task in the field of computer graphics and geometry processing is that of non-rigid matching between 3D shapes. As first intuition, this means to find correspondences between points on different 3D shapes when these are related by deformations. The problem is particularly challenging because the space of possible correspondences, represented as mapping of points, is exponential in size. This problem has usually been tackled by establishing a mapping between few, sparse keypoints and then extending it to a dense correspondence (see [23] for a complete survey). These kind of approaches, beyond being computationally hard, have several other downsides such as the difficulty of enforcing map continuity and global consistency.

A breakthrough in the field was represented by [17], which proposed a completely new approach to non-rigid matching: to map functions defined on the shapes, instead of points, thus representing the correspondence as a mapping between function spaces. This framework, besides being more general than the classical notion of point-to-point correspondences, has some really good properties:

- The mapping between function spaces is linear, independently of the arbitrary complexity of the transformation between the shapes, and can be found by solving a linear optimization problem.
- It is compactly represented as a matrix and the usual algebraic operations have a straightforward interpretation in the mapping context.

The main journals on the topic are listed here, together with the h5-index[1] and impact factor[2]:

Journals	h5-index	impact
ACM Transactions on Graphics	85	13.20
IEEE Transactions on Visualization and Computer Graphics	70	10.89
Computer Graphics Forum	55	9.73
Computers & Graphics	30	3.46

The main conferences involve venues from different fields: computer graphics (e.g. ACM SIGGRAPH, ACM SIGGRAPH Asia, Eurographics), computer vision (e.g. IEEE/CVF Conference on Computer Vision and Pattern Recognition, IEEE/CVF International Conference on Computer Vision, European Conference on Computer Vision, International Conference on 3D Vision) and machine learning (e.g. NeurIPS).

1.1. Preliminaries

Let us consider two bi-dimensional manifolds $\mathcal{M}$ and $\mathcal{N}$, and a function $f : \mathcal{M} \to \mathbb{R}$. Let T be a point-to-point mapping $T : \mathcal{M} \to \mathcal{N}$, which maps each point p on $\mathcal{M}$ on a point $T(P)$ on $\mathcal{N}$. If we suppose T to be bijective, we can transfer f from $\mathcal{M}$ to $\mathcal{N}$ by means of T and obtain a function $g : \mathcal{N} \to \mathbb{R}$ such that $g = f \circ T^{-1}$. This induced transformation between spaces of real-valued functions is the functional representation of the mapping T and we denote it as $T_F : \mathcal{F}(\mathcal{M}, \mathbb{R}) \to \mathcal{F}(\mathcal{N}, \mathbb{R})$.

[1] source: https://scholar.google.com
[2] source: https://www.guide2research.com

We note that the original mapping T can be recovered from T_F (intuitively by transferring Dirac's deltas centered in each point of $\mathcal{M}$), thus it is at least as expressive as T. We also note that, for any fixed bijective map T, T_F is a *linear* map between the corresponding function spaces, meaning that $T_F(\alpha_1 f_1 + \alpha_2 f_2) = \alpha_1 T_F(f_1) + \alpha_2 T_F(f_2)$.

Given a basis $\{\phi_i^{\mathcal{M}}\}$ for the functional space $\mathcal{F}(\mathcal{M}, \mathbb{R})$, we can write any function $f : \mathcal{M} \to \mathbb{R}$ as $f = \sum_i a_i \phi_i^{\mathcal{M}}$ and

$$T_F(f) = T_F(\sum_i a_i \phi_i^{\mathcal{M}}) = \sum_i a_i T_F(\phi_i^{\mathcal{M}})$$

If also $\mathcal{N}$ is equipped with a basis $\{\phi_j^{\mathcal{N}}\}$, we can express $T_F(\phi_i^{\mathcal{M}})$ as $\sum_j c_{ji} \phi_j^{\mathcal{N}}$ for some coefficient c_{ji} and the complete mapping of f is then

$$T_F(f) = \sum_i a_i \sum_j c_{ji} \phi_j^{\mathcal{N}} = \sum_j \sum_i a_i c_{ji} \phi_j^{\mathcal{N}}$$

We note that the coefficients c_{ji} depend only on the two bases and not on the function f and they completely specify the mapping T_F, thus we can identify the mapping with the matrix C (made by coefficients c_{ji}), which acts as a change of basis between $\phi_i^{\mathcal{M}}$ and $\phi_j^{\mathcal{N}}$. If f is represented by a vector of coefficients $\mathbf{a}$ then $T_F(\mathbf{a}) = C\mathbf{a}$. In this framework, matching two shapes resorts to the problem of finding such (possibly infinite) matrix C.

This is done by posing some constraints and finding the matrix C that best satisfies them in the least-square sense. The first kind of constraint is function preservation: if we know (or suppose) that two functions f and g should be in correspondence on the two shapes, we can enforce this by minimizing $\|C\hat{f} - \hat{g}\|^2$, where $\hat{f}$ and $\hat{g}$ are their vector coefficients. Such functions may be:

Landmarks Corresponding points on the two shapes, e.g. expressed as distances or gaussians. In the latter case, uncertainty about the position can be encoded as the variance of the distribution.

Descriptors Corresponding probe functions computed on the two shapes, which characterize each point of the object independently of the (isometric) deformation applied. See section 2.2.1 for some examples.

Segments Corresponding parts on the two shapes, e.g. expressed as (soft) indicator or distances.

Actually, functional preservation can be expressed in a more informative way by exploiting product preservation, thus allowing to find the mapping with fewer descriptors, as proposed in [16].

Another possible constraint to apply is the commutativity with respect to some operators R_F and S_F, defined on $\mathcal{F}(\mathcal{M}, \mathbb{R})$ and $\mathcal{F}(\mathcal{N}, \mathbb{R})$. This is expressed by minimizing $\|R_F C - S_F C\|^2$. In particular, the commutativity with the Laplace-Beltrami operator is preserved by isometries and it is used as regularizer for C. Note that all these constraints are linear in the matrix C and can be, thus, efficiently minimized.

Further works have proposed alternative procedures for computing C. For instance, [12] has proposed to refine low-dimensional maps by progressively increasing the size of the map, alternating conversions to and from point-to-point maps and considering more and more elements of the basis.

1.2. Research topic

In the previous description we have assumed the presence of a basis on the manifolds, without directly addressing the problem. This is actually a crucial step in the framework and has a great impact on its result.

In order to represent the mapping through a compact matrix, we need to choose a basis of a suitable subspace of functions, in which the most natural ones have a good approximation. Besides compactness, another important characteristic for a basis is stability, which means that the space of functions spanned is stable under small or nearly-isometric shape deformations.

The most common choice, which was adopted in [17] and in the great majority of the following works, is to use the first eigenfunctions of the Laplace-Beltrami operator [6, 22], which is the manifold equivalent of the Fourier basis. Being them ordered by increasing frequency, this corresponds to a low-pass filter approximation. Their multi-scale behavior allows to choose a suitable number of elements to balance the trade-off between dimension and power of representation. This choice has been proved to be optimal for representing continuous, bounded-variation functions (see [1]). LB eigenbasis is also orthonormal, which allows a simple computation of the representation coefficients, and it can be easily computed for meshes by means of the cotangent scheme [14].

Despite being generally a good choice, it has some downsides in practical usage. The representation of detailed functions requires the use of many elements, but the space of function spanned becomes unstable at higher frequencies, therefore it is not always suited for transferring such functions [3]. As we will see in next section, many alternatives or extension have been proposed to overcome this limitations. They usually target a specific class of functions or specific applications.

2. Main related works

2.1. Classification of the main related works

The main related works can be classified according to the specific part of the framework they tackle, we will analyze two main aspects:

- Descriptors: they concentrate on finding new and good probe functions, used to pose the function preservation constraints. Such function should be able to characterize well each portion of the shape, while being pose-invariant. Belong to this category works like [2, 19, 21, 4, 5, 18, 7].
- Basis: they proposes new bases for the representation of functions on the shapes, as alternatives or complements to the standard Laplace-Beltrami eigenfunctions. They may address a particular kind of task or function class. Some relevant works are [13, 10, 11, 9, 8].

Another important distinction is between learning-based approaches, in which there is something learned from a dataset and axiomatic-based approaches, in which every part is set at design time and then applied to the shapes at hand.

2.2. Brief description of the main related works

2.2.1 Descriptors

Many descriptors already existed before the advent of functional maps and they have been naturally employed in this framework as function preservation constraints. The most common are the Heat Kernel Signature and Wave Kernel Signature. HKS [19] defines a set of functions $k_t : \mathcal{M} \rightarrow \mathbb{R}^+$ which represent, for each point $x \in \mathcal{M}$, the amount of heat remaining in x at time t, after placing an initial δ distribution of heat at point x. By discretizing time in h instants we get a set of h function with a different scale behaviour (the lower t and the narrower is the scale characterizing the signature). Being intrinsic, HKS is isometry-invariant but also robust to small non-isometries. WKS [2] exploits a similar idea and associates, to each point $x \in \mathcal{M}$, the probability of a quantum particle to be found in x, for different values of the energy of the particle. With respect to HKS it explicitly models the influence of frequency (energy is a function of frequency) on the scale of the descriptor and it is more informative and robust. Another commonly used descriptor is SHOT [21], which tries to better balance the trade-off between descriptiveness and robustness.

All the previously described methods are axiomatic-based, but learning-based descriptors are gaining increasing attention. In [7] a neural network is proposed to learn an optimal transformation of some predefined input descriptors. It is trained using a supervised loss that minimizes the error with respect to a known pointwise map. Further developments [5, 18] use an unsupervised loss, based purely on desired geometrical properties of the mapping, and thus they do not require annotated datasets. [4] takes again a supervised approach, but it is able to learn the descriptors completely from the raw 3D shape and computes the loss in the spectral domain for an efficient computation.

2.2.2 Bases

As we saw, the choice of the basis is a crucial aspect for both representing functions and find a precise and stable matching between shapes. Several works have proposed alternatives or complements to the common Laplace eigenbasis.

Binary Sparse Frame In [10] a frame has been proposed to represent and transfer step functions with more accuracy than LBO eigenbasis. This is actually a frame and not a basis since it is a set of (possibly) linear dependent functions. The representation of a signal is then not unique and chosen according to sparsity criteria. Being the elements of the basis binary functions, this frame is well suited for representing and transferring step functions, but not well applicable in different scenarios.

Coordinate Manifold Harmonics In [11] the LBO eigenbasis is extended with extrinsic information from the shape, namely the normalized error in the representation of the spatial coordinates of the points with respect to the low-pass filter given by the first k eigenfunctions of LBO. These three functions, for x, y and z axes, are added in sequence and are orthogonal by construction to the previous elements, thus forming a orthonormal basis. This basis, along with its specific pipeline, has been developed for the task of mesh-transferring, performed through matching with a model shape presenting appropriate mesh structure. Due to extrinsic information, this basis is note pose-invariant and has limited applicability in the general context of non-rigid matching.

Localized Manifold Harmonics The Localized Manifold Harmonics, presented in [13], are an extension of the traditional LBO eigenfunctions but localized in a given region of the shape. Each new function ψ_j to be added to the basis is obtained by minimizing the energy:

$$\varepsilon(\psi_j) = \varepsilon_S(\psi_j) + \mu_R \varepsilon_R(\psi_j) + \mu_\perp \varepsilon_\perp(\psi_j)$$

where ε_S is the Diriclet energy and it is minimized by the Laplace-Beltrami eigenfunctions, ε_R is a term penalizing non-zero values outside a region R defined on the shape and $\varepsilon_\perp$ softly enforces the orthogonality to previous elements. As can be noted, functions ψ_j are defined on the global surface, but their non-zero values are localized in a region R, defined by a (possibly soft) indicator. They can be naturally added to the traditional LB basis, while preserving orthonormality. Note that if $\mu_R = 0$, the minimization of this problem gives exactly the ordered sequence of the first LB eigenfunctions.

This basis can represent functions localized in a specific region in a compact form and with a high level of detail. The main limit of this method is the necessity to define a region on the shape, which is non-trivial to automatize and the limited benefit for functions that present high frequencies on the whole shape.

Product of LB eigenfunctions A recent approach to tackle the trade-off between dimension and expressiveness of the basis, presented in [15] and extended in [8], is that of adding pointwise product of LB eigenfunctions to enrich the traditional basis. This is founded on the fact that functional maps arising from point-to-point correspondences must preserve pointwise product between functions. These eigenproducts carry higher-frequency information and their alignment can be directly derived from the functional map between standard eigenfunctions. Therefore, once a transfer matrix $\tilde{C}$ for the eigenproducts is computed in close form from C, you can use $\tilde{C}$ to transfer functions with higher detail.

Since the computation of $\tilde{C}$ from C is pretty unstable and amplifies possible errors in C, [8] has made some improvements in this sense. They orthogonalized the set of eigenproducts and eigenfunctions, thus removing linear dependency, and proposed an alternative, more stable, way of finding the mapping, which involves computing a point-to-point correspondence and re-computing the functional map in the new basis. Despite providing more robustness, the need to compute point-to-point maps may also constitute a limiting factor of this method.

Embedding Learning In [9] a fully differentiable, and thus learnable, pipeline for estimating correspondence between 3D pointclouds has been proposed. It still relies on the functional map framework and separates the two tasks of learning optimal bases and learning the optimal descriptors to align them. The map is then computed by solving the usual optimization problem.

Here, the basis is considered as a high-dimensional embedding space, linearly-invariant with respect to deformations. To realize such embedding, they propose to train a neural network that takes the coordinates of the points of the shape as input. This is trained by providing pairs of shapes, with known point-to-point correspondence. Given an embedding function, encoded as the weights of the network, it is possible to recover the optimal linear map from the point-to-point correspondence. Therefore, the loss of the model can be assessed by measuring how well this linear map acts on the embedding of the shapes themselves.

[20] tries to do without the need of a known correspondence between shapes: this neural network is trained only with a set of related shapes and some semantic functions for each of them, but no correspondence is required between these functions. Its objective is to learn a coherent set of basis on the shapes. Of course, the learned dictionary is shape dependent and cannot be extended to other shapes.

2.3. Discussion

The selection of a basis poses a trade-off between dimension of the representation and approximation quality. While the Laplace-Beltrami eigenbasis is generally good for low-frequency and smooth functions, it is not well suited for transfer detailed functions. The search for a basis able to represent high level of detail, in a compact and transferable way is still an open problem.

Most of the above mentioned approaches target just a specific problem and are not sufficiently general. Eigenproducts [15, 8] seem to be an interesting direction, even though for the moment they are not used to compute the mapping, but only to represent and transfer functions. Learning-based approaches are also promising, despite the obvious need of data to train.

References

[1] Aflalo, Y., Brezis, H., and Kimmel, R. On the optimality of shape and data representation in the spectral domain. *SIAM J. Imaging Sci. 8*, 2 (2015), 1141–1160.

[2] Aubry, M., Schlickewei, U., and Cremers, D. The wave kernel signature: A quantum mechanical approach to shape analysis. In *Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on* (2011), IEEE, pp. 1626–1633.

[3] Azencot, O., and Lai, R. Shape analysis via functional map construction and bases pursuit, 2019.

[4] Donati, N., Sharma, A., and Ovsjanikov, M. Deep geometric functional maps: Robust feature learning for shape correspondence. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2020), pp. 8592–8601.

[5] Halimi, O., Litany, O., Rodola, E., Bronstein, A. M., and Kimmel, R. Unsupervised learning of dense shape correspondence. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2019), pp. 4370–4379.

[6] Levy, B. Laplace-beltrami eigenfunctions towards an algorithm that "understands" geometry. In *IEEE International Conference on Shape Modeling and Applications 2006 (SMI'06)* (2006), pp. 13–13.

[7] Litany, O., Remez, T., Rodolà, E., Bronstein, A., and Bronstein, M. Deep functional maps: Structured prediction for dense shape correspondence. pp. 5660–5668.

[8] Maggioli, F., Melzi, S., Ovsjanikov, M., Bronstein, M., and Rodolà, E. Orthogonalized fourier polynomials for signal approximation and transfer. In *Proceedings of Eurographics 2021* (2021).

[9] Marin, R., Rakotosaona, M.-J., Melzi, S., and Ovsjanikov, M. Correspondence learning via linearly-invariant embedding, 2020.

[10] Melzi, S. Sparse representation of step functions on manifolds. *Computers & Graphics 82* (2019), 117–128.

[11] Melzi, S., Marin, R., Musoni, P., Bardon, F., Tarini, M., and Castellani, U. Intrinsic/extrinsic embedding for functional remeshing of 3d shapes. *Computers & Graphics 88* (2020), 1–12.

[12] Melzi, S., Ren, J., Rodolà, E., Sharma, A., Wonka, P., and Ovsjanikov, M. Zoomout: Spectral upsampling for efficient shape correspondence. *ACM Transactions on Graphics (TOG) 38*, 6 (Nov. 2019), 155:1–155:14.

[13] Melzi, S., Rodolà, E., Castellani, U., and Bronstein, M. Localized manifold harmonics for spectral shape analysis. *Computer Graphics Forum 37*, 6 (2018), 20–34.

[14] Meyer, M., Desbrun, M., Schröder, P., and Barr, A. H. Discrete differential-geometry operators for triangulated 2-manifolds. In *Visualization and mathematics III*. Springer, 2003, pp. 35–57.

[15] Nogneng, D., Melzi, S., Rodolà, E., Castellani, U., Bronstein, M., and Ovsjanikov, M. Improved functional mappings via product preservation. *Computer Graphics Forum 37*, 2 (2018), 179–190.

[16] Nogneng, D., and Ovsjanikov, M. Informative descriptor preservation via commutativity for shape matching. *Computer Graphics Forum 36*, 2 (2017), 259–267.

[17] Ovsjanikov, M., Ben-Chen, M., Solomon, J., Butscher, A., and Guibas, L. Functional maps: a flexible representation of maps between shapes. *ACM Transactions on Graphics (TOG) 31*, 4 (2012), 30:1–30:11.

[18] Roufosse, J.-M., Sharma, A., and Ovsjanikov, M. Unsupervised deep learning for structured shape matching. In *Proceedings of the IEEE International Conference on Computer Vision* (2019), pp. 1617–1627.

[19] Sun, J., Ovsjanikov, M., and Guibas, L. A concise and provably informative multi-scale signature based on heat diffusion. *Computer graphics forum 28*, 5 (2009), 1383–1392.

[20] Sung, M., Su, H., Yu, R., and Guibas, L. Deep functional dictionaries: Learning consistent semantic structures on 3d models from functions. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems* (Red Hook, NY, USA, 2018), NIPS'18, Curran Associates Inc., p. 483–493.

[21] Tombari, F., Salti, S., and Di Stefano, L. Unique signatures of histograms for local surface description. In *Proc. ECCV* (2010), Springer, pp. 356–369.

[22] Vallet, B., and Lévy, B. Spectral geometry processing with manifold harmonics. *Computer Graphics Forum 27*, 2, 251–260.

[23] Van Kaick, O., Zhang, H., Hamarneh, G., and Cohen-Or, D. A survey on shape correspondence. *Computer Graphics Forum 30*, 6 (2011), 1681–1707.