Local Signature Quantization by Sparse Coding

D. BOSCAINI AND U. CASTELLANI

VISION, IMAGE PROCESSING AND SOUNDS LABORATORY
DEPARTMENT OF COMPUTER SCIENCE, UNIVERSITY OF VERONA, ITALY

Abstract We briefly present the research activities in which Computer Vision group of the Department of Computer Science of the University of Verona is involved. Then we focus on one of the latest research direction that directly involve us.

1 Overview of VIPS lab

Shape Analysis has recently become one of the major research topic in the Vision, Image Processing and Sound Laboratory (VIPS) in Verona. The meeting of different research experiences in Computer Vision, Computer Graphics and Pattern recognition of the group members has lead to the development of several project related to local and global shape description, shape retrieval, medial representations and real world applications.

In recent works, for example, novel methods to create local generative shape descriptors have been presented and tested for different applications [1, 2], new kinds of medial representations and related global shape descriptors have been proposed [3, 4], and frameworks to perform shape retrieval have been presented as well [5].

Real world applications of this kind of techniques demonstrated the potential benefits of applying state of the art geometry processing algorithms to measure and classify digital acquisition of real 3D objects. We focused in particular on medical applications, developing classifiers able to discriminate healthy and pathological brains [6, 7] and on digital anthropometry, working on techniques for the measurements of human body characteristics [8].

In our lab undergraduated students are also working on the development of demo applications for capturing and processing 3D information with low cost devices and on the use of geometrical processing for applications related to Human Computer Interaction.

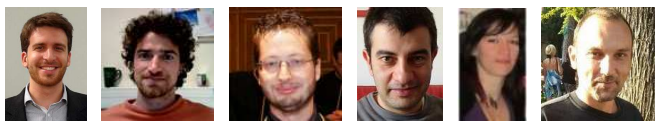


FIGURE 1. *From left to right: D. Boscaini, U. Castellani, D. Peruzzo, A. Giachetti, V. Garro, C. Lovato, VIPS Lab.*

2 Introduction

In 3D object retrieval is very important to define reliable shape descriptors, which compactly characterize geometric properties of the underlying surface. To this aim two main approaches are considered:

- global approaches are effective in describing the whole object but they are very sensitive to missing parts, so they are not suitable for the description of partial models.
- local approaches are more suitable to characterize small parts of the shape but is necessary a strategy for compare local descriptors belonging to different shapes, e.g. compute the correspondences between the two shapes in exam.

With this work we address the problem of the quantization procedure: from a set of local descriptors how can one describe the whole shape? The overall aim is to obtain the advantage of both the approaches. From one side we are able to compare global shapes rather than a set of single points. From the other side, we exploit local informations which, in general, is more robust to noise and missing parts and more suitable to deal with partial objects.

A popular method is to introduce a sort of counting approach by collecting local informations into a histogram which leads to a local-to-global signature. Such examples are signature distance distributions or the bag of words approach. In [9] we propose to go beyond of these approaches by exploiting sparse coding techniques.

3 Sparse coding framework

Suppose we have a sentence s and a dictionary D. We want to explain the sentence s with words contained in D. This problem could be formalized as

$$\min_{\alpha} \|s - \alpha D\|_2^2.$$

The idea is that vector α picks up only the words that describe the sentence s. In general a dictionary is over complete: there are a lot of words with the same or similar meaning. For this reason we might be interested to consider the minimum number of words as possible.

In this case Tychonoff regularization theory could help. Reformulating the previous problem as:

$$\min_{\alpha} \|s - \alpha D\|_2^2 + \lambda R(\alpha). \tag{1}$$

and choosing $R(\alpha) = \|\alpha\|_1$, we are promoting the sparsity of the solution. This problem is known as Lasso formulation.
In general s could be though as a generic signal generated by αD.

Often the dictionary D is not available. We are therefore interested in inferring both the vector α and the dictionary D from the signal s. The problem (1) becomes:

$$\min_{D}\Big(\min_{\alpha} \|s - \alpha D\|_2^2 + \lambda\|\alpha\|_1\Big). \tag{2}$$

SPAMS tool solve problem (2) employing an alternating minimization method between the variables D and α.

As a further step we should consider the more general case, in which, instead of a single signal s, we have a collection of signals $\mathbf{s} = \{s^i\}_{i=1,\dots,N}$. Therefore, equation (2) can be generalized as:

$$\min_{D} \frac{1}{N}\sum_{i=1}^{N} \min_{\alpha^i} \|s^i - \alpha^i D\|_2^2 + \lambda\|\alpha^i\|_1, \tag{3}$$

where $\alpha = \{\alpha^i\}_{i=1,\dots,N}$ is though as a collection of vectors.

4 Local-to-global shape descriptor

Since we want to employ sparse coding techniques for shape analysis purposes we need a signal s related to the geometrical properties of the shape examined and possibly robust to its non-rigid deformations.

For this reason we chose to take advantage of Laplace-Beltrami operator properties and consider as local signatures diffusion geometry descriptors:

- Heat Kernel Signature (HKS), it could be thought as a collection of low-pass filters. This damages the ability of the descriptor to precisely localize shape features,
- Wave Kernel Signature (WKS), it could be thought as a collection of band-pass filters and exhibits superior feature localization.

Once HKS or WKS are computed for each vertex of the shape, we consider them as the collection of signals $\boldsymbol{s} = \{s^i\}_{i=1,\dots,n}$. Then, generative learning techniques are employed for solving problem (3). We consider the resulting dictionary D as the global signature for each shape. The query phase could be done employing the classical leave-one-out approach: comparing the dictionary of the query shape with the dictionary of all the other shapes in the database.

FIGURE 2. *Null shapes of our database taken from TOSCA and Sumner databases. From left to right we find man, dog, cat, man, woman, horse, camel, cat, elephant, flamingo, horse and lioness. In all our experiment we consider correct the matching between two men, cats or horses.*

FIGURE 3. *Examples of deformations types considered in our database, taken from SHREC 11 robust benchmark. From left to right we find the null shape, affine, holes, micro holes, scale, sampling, noise, shot noise, topology and view deformation.*

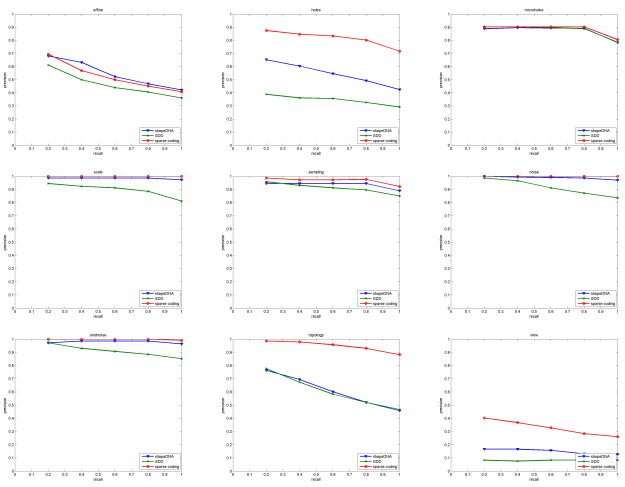
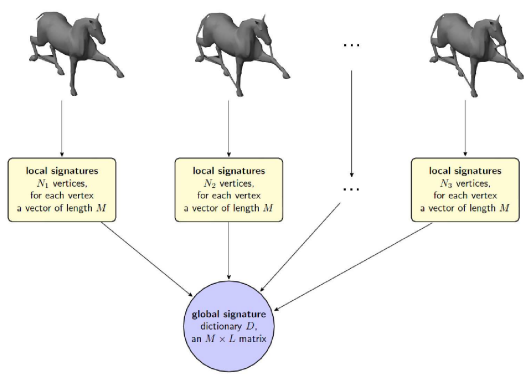


FIGURE 4. *Comparison between Precision-Recall curves of Shape DNA (blue), SDD (green) and sparse coding approach (red) on the SHREC 2011 Robustness benchmark.*

5 Local-to-global class descriptor

If we consider as $\boldsymbol{s}$ the collection of local signatures of each instance of an entire class of shapes, and then repeat the procedure illustrated above, with very little effort we have a descriptor that takes into account the behavior of the entire class of shapes.
In this fashion several deformation of the same object (e.g. noise or topological deformations) can contribute in learning the dictionary.



In order to deal with multiple classes of shapes, several dictionaries $\{D^{c_1}, D^{c_2}, \dots, D^{c_n}\}$ can be trained: one for each available class. In the query phase, given a shape O and its collection of local signatures $\boldsymbol{s}$ we have to solve problem (1) for each dictionary D^{c_i} and we obtain vectors α^{c_i}. At the shape O is assigned the class c such that

$$\|\boldsymbol{s} - \alpha^c D^c\| = \min_i \|\boldsymbol{s} - \alpha^{c_i} D^{c_i}\|.$$

Principal advantages:

- our method is that it allow to compare the signature of a query shape only with the dictionary of the classes of the shapes present in the database considered,
- the dictionary encodes more instances of the same non-rigid deformation, making the proposed signature a descriptor of the entire class of deformations rather than a descriptor of a single shape.

TABLE 1. *Comparison between Nearest Neighbor performances of SDD, Shape DNA and sparse coding approach.*

Deformation	SDD		Shape DNA		sparse coding	
	corrects	%	corrects	%	corrects	%
Affine	45/72	0.67	49/72	0.68	**59/72**	**0.82**
Holes	36/72	0.50	58/72	0.81	**65/72**	**0.90**
Micro holes	**65/72**	**0.90**	64/72	0.89	**65/72**	**0.90**
Scale	**72/72**	**1.00**	71/72	0.99	**72/72**	**1.00**
Sampling	67/72	0.93	69/72	0.96	**72/72**	**1.00**
Noise	71/72	0.99	**72/72**	**1.00**	**72/72**	**1.00**
Shot noise	70/72	0.97	**72/72**	**1.00**	**72/72**	**1.00**
Topology	58/72	0.80	55/72	0.76	**72/72**	**1.00**
View	11/72	0.15	16/72	0.22	**49/72**	**0.64**
Average	495/648	0.76	526/648	0.81	**598/648**	**0.92**

6 Shape-to-shape comparison

The dictionary obtained in this way is the descriptor of an entire class of objects and this cause problems with well known techniques for the evaluation of the retrieval performances of the proposed method, e.g. Precision-Recall curves or dissimilarity matrices. In fact, in order to obtain these evaluation criteria, we need a shape-to-shape comparison rather than a shape-to-class one.

To overcome this difficulty we define a class-based similarity function between shapes: given two shapes A and B, and the associated class dictionaries D^A and D^B, we consider as the error function between the shapes the function

$$e(A, B) = \|\boldsymbol{s}_A - \alpha_A D^B\|_2^2 + \|\boldsymbol{s}_B - \alpha_B D^A\|_2^2.$$

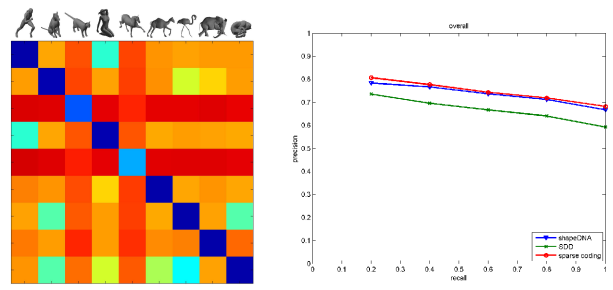


FIGURE 5. *Left: class-signature dissimilarity matrix for noise deformations. Blue colors represent lower values, red colors represent higher values. Right: comparison between Precision-Recall curves of Shape DNA (blue), SDD (green) and sparse coding approach (red) on the SHREC 2011 Robustness benchmark.*

FIGURE 6. *Some instances of SHREC 2007 Partial Matching queryset.*

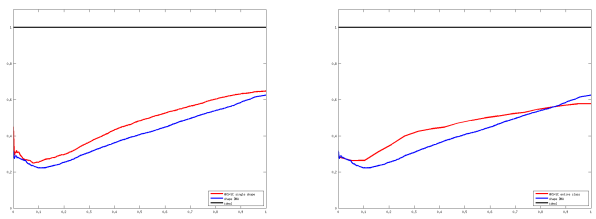


FIGURE 7. *Comparison between NDCG curves of sparse coding approach (red) and shape DNA (blue) on the SHREC 2007 Partial Matching track. Left: single shape descriptor, right: entire class descriptor.*

Conclusions We propose a new approach to deal with the issue of extending local signatures to global shape descriptor based on sparse coding techniques. This seem very promising, in fact sparsity enforces the encoding of essential information.
There is still a lot to do, some open question are:

- is there a criterion for choosing free parameters (e.g. the size of dictionary)?,
- how to adopt a discriminative approach into the generative framework?

References

[1] U. Castellani, M. Cristani, S. Fantoni, V. Murino, Sparse points matching by combining 3D mesh saliency with statistical descriptors, Computer Graphics Forum vol. 27 no. 2 pp. 643–652 (2008)

[2] U. Castellani, M. Cristani, V. Murino, Statistical 3D Shape Analysis by Local Generative Descriptors, IEEE Transactions on Pattern Analysis and Machine Intelligence vol. 33 no. 12 pp. 2555–2560 (2011)

[3] A. Giachetti, C. Lovato, Radial symmetry detection and shape characterization with the multiscale area projection transform, Computer Graphics Forum vol. 31 no. 5 pp. 1669–1678

[4] A. Giachetti, C. Lovato, Centers of Approximate Spherical Symmetry and Radial Symmetry Graphs, Eurographics 2013 - Short Papers pp. 13–16 (2013)

[5] R. Toldo, U. Castellani, A. Fusiello, The Bag of Words approach for retrieval and categorization of 3D objects, The Visual Computer vol. 26 no. 10 pp. 1257–1268 (2010)

[6] U. Castellani, A. Perina, V. Murino, M. Bellani, G. Rambaldelli, M. Tansella, P. Brambilla, Brain morphometry by probabilistic latent semantic analysis Medical Image Computing and Computer-Assisted Intervention (MICCAI) pp. 177-184 (2010)

[7] U. Castellani, A. Perina, V. Murino, M. Bellani, G. Rambaldelli, M. Tansella, P. Brambilla, A new shape diffusion descriptor for brain classification Medical Image Computing and Computer-Assisted Intervention (MICCAI) pp. 426-433 (2011)

[8] C. Lovato, U. Castellani, A. Giachetti, Computer Vision/Computer Graphics Collaboration Techniques, Automatic segmentation of scanned human body using curve skeleton analysis, Lecture Notes in Computer Science vol. 5496 pp. 34–45 (2009)

[9] D. Boscaini, U. Castellani, Local Signature Quantization by Sparse Coding, Proceedings of Eurographics Workshop on 3D Object Retrieval (2013)