

Feature Encoding of Spectral Signatures for 3D Non-Rigid Shape Retrieval

Frederico A. Limberger
<http://www.cs.york.ac.uk/~fal504>
 Richard C. Wilson
<http://www.cs.york.ac.uk/~wilson>

University of York
 York, UK

A large number of local signatures have been created to represent local characteristics of geometric 3D shapes for the purpose of many computer vision, geometry processing and shape analysis tasks. A local signature is a compact representation that characterizes a small region of a shape. They usually capture information about the neighbourhood of a vertex and so they can be directly applied to some important tasks like point correspondence and shape segmentation. For this purpose, it is desirable to compute signatures that are invariant under rigid, non-rigid and isometric deformations, the typical deformations that 3D models undergo. However, local descriptors cannot be immediately applied to the problem of shape retrieval, because this task is not addressed by comparing local signatures but by comparing global descriptors (signatures that represent the shape as a whole). Creating a global descriptor is not a simple task since shapes can have arbitrary number of vertices, edges and faces. To create a generic representation of a shape all important characteristics should be preserved during the encoding process. Even so, the global representation must compress local characteristics using the same basis for all shapes in order to facilitate comparisons.

In this paper, we propose an efficient and discriminative encoding framework to address the problem of creating global signatures for 3D models from local descriptors based on the spectrum of the shape, for the purpose of shape retrieval and classification. In this way, we propose the use of Fisher Vector (FV) to describe the entire representation of a shape. Differently from [1, 2], our approach uses a Gaussian Mixture Model (GMM) as a dictionary of probabilistic visual words, and encodes the global signature using three orders statistics (0-th, 1-st, 2-nd) rather than using only the first order. Further, while the classical Bag-of-Features (BoF) generates a K -dimensional histogram, where K is the vocabulary size, Fisher Vector encoding generates a high-dimensional vector with $2KD$ dimensions, where D is the size of each local descriptor, being more discriminative but still simple to compare, as all shapes are encoded in the same basis. We also propose some improvements to the Wave Kernel Signature (WKS), since it overtakes the Heat Kernel Signature (HKS) and Scale Invariant Heat Kernel Signature (SI-HKS) in retrieval performance. A better scaling is proposed for the eigenvalues of the Laplace-Beltrami operator which captures more information about the shape and we also propose the use of principal curvatures to increase the efficiency of the encoding method.

The Fisher Vector encoding characterizes a large set of vectors by their three-order deviation from a vocabulary, creating a high-dimensional gradient vector representation. The gradient of the log-likelihood, also called *Fisher score*, describes the contribution of each parameter to the generation process. To compute the FV encoding, we write the local shape descriptors wrt. the probabilistic model (Gaussian Mixture Model). To properly compare shape signatures we have also applied L2 Normalization and Power Normalization to the Fisher Vector, also known as the Improved FV [3].

When it comes to improving the WKS (1), we have analysed the influence of the logarithmic scaling to the computation of the signature and we concluded that it loses high frequency information about the spectrum of the shape (even important high frequencies that do not originate from noise are blurred). Then, we analysed the distributions of the differences between eigenvalues of 24 articulated woman shapes (Figure 1) and we propose to use the cubic root scaling (2) rather than the classical logarithmic scale (1) since it fits much better a normal distribution, which is used to handle the differences between same-class shapes.

$$\text{WKS}(x, e) = C_e \sum_{k=1}^{\infty} \phi_k(x)^2 f_E(\Lambda_k)^2 \quad f_E(\Lambda_k)^2 = e^{-\frac{(e - \log(\Lambda_k))^2}{2\sigma^2}} \quad (1)$$

$$f_C(\Lambda_k)^2 = e^{-\frac{(e - \sqrt[3]{\Lambda_k})^2}{2\sigma^2}} \quad e \in [\sqrt[3]{\lambda_{\min}}, \sqrt[3]{\lambda_{\max}}] \quad (2)$$

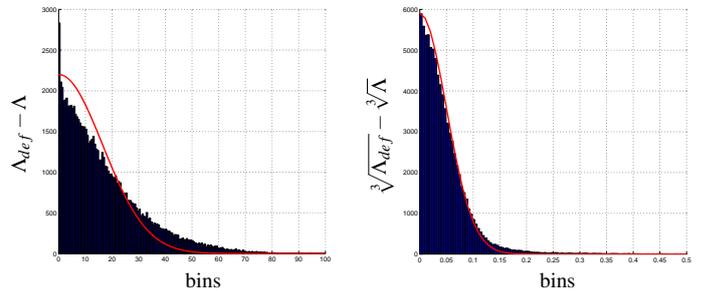


Figure 1: Histograms of the differences between the eigenvalues Λ of the Laplace-Beltrami operator of 24 articulated woman shapes. (left) Histogram of the differences not scaled. (right) Histogram of the differences scaled by the cubic root. On each graph, the red line is a reference normal distribution fitted to the respective histogram.

We also analyze the WKS when combined with encoding methods. These methods are based on the differences between descriptor histograms and a probabilistic model. The more discriminative these histograms are the more the encoding process will be able to distinguish between shapes of different classes. We have noticed that when shapes undergo pose changes, the maximum principal curvature of the vertices remains stable. This is a local feature of the surface, which is isometry invariant and stable at most points under object articulation but not well coded in the WKS. Therefore, as another improvement to the shape retrieval task, we propose the use of principal curvatures to increase the separation of features leading to a more discriminative histogram. Practically, we shift the WKS individually by the maximum principal curvature c for each vertex x as stated in Equation (3). We smooth the maximum principal curvature by taking the mean of the respective neighbour vertices to diminish the influence of noise. In (3), α is a weight that normalizes c accordingly to the signature values. We use $\alpha = 0.015$ in all our experiments.

$$\text{IWKS}(x, e) = C_e \sum_{k=1}^{\infty} \phi_k(x)^2 f_C(\Lambda_k)^2 + c_x \alpha \quad (3)$$

We compare our method against the competitors of the non-rigid tracks of SHREC'11 and SHREC'15. By analyzing our retrieval accuracy in the SHREC'11 track it is clear that our approach improves spectral descriptors to tackle shape retrieval problems since it achieves excellent retrieval measures (DCG very close to 1) and comparable results with best retrieval methods (very similar e-Measure to the top tier methods). Furthermore, our method presents a much better performance when compared to other spectral descriptors, showing its potential.

Our conclusion is that although our method does not beat all other groups in the SHREC'11 and SHREC'15 benchmarks it proves that spectral methods are a good choice to retrieve shapes when combined with informative encoding methods, since our method is close to the best techniques in performance and achieves excellent results in most benchmark classes. We observed that the our worst retrieval performances happen when shapes undergo huge topology changes. Specially in these cases, spectral signatures still need to be improved. In future works, we plan to create a spectral signature that is less variant to major topology changes.

- [1] Alexander M. Bronstein, Michael M. Bronstein, Leonidas J. Guibas, and Maks Ovsjanikov. Shape google: Geometric words and expressions for invariant shape retrieval. *ACM Trans. Graph.*, 30(1):1:1–1:20, February 2011.
- [2] M. Ovsjanikov, A.M. Bronstein, M.M. Bronstein, and L.J. Guibas. Shape google: a computer vision approach to isometry invariant shape retrieval. In *ICCV Workshops*, pages 320–327, Sept 2009.
- [3] Florent Perronnin, Jorge Sánchez, and Thomas Mensink. Improving the fisher kernel for large-scale image classification. In *Proc. ECCV*, pages 143–156, Berlin, 2010.