Pose-Consistent 3D Shape Segmentation Based on a Quantum Mechanical Feature Descriptor

Mathieu Aubry[1,2], Ulrich Schlickewei[1] and Daniel Cremers[1]

[1] Department of Computer Science, TU München
[2] Adobe Systems Inc. Cambridge, MA

Abstract. We propose a novel method for pose-consistent segmentation of non-rigid $3D$ shapes into visually meaningful parts. The key idea is to study the shape in the framework of quantum mechanics and to group points on the surface which have similar probability of presence for quantum mechanical particles. For each point on an object's surface these probabilities are encoded by a feature vector, the Wave Kernel Signature (WKS). Mathematically, the WKS is an expression in the eigenfunctions of the Laplace–Beltrami operator of the surface. It characterizes the relation of surface points to the remaining surface at various spatial scales. Gaussian mixture clustering in the feature space spanned by the WKS signature for shapes in several poses leads to a grouping of surface points into different and meaningful segments. This enables us to perform consistent and robust segmentation of new versions of the shape.

Experimental results demonstrate that the detected subdivision agrees with the human notion of shape decomposition (separating hands, arms, legs and head from the torso for example). We show that the method is robust to data perturbed by various kinds of noise. Finally we illustrate the usefulness of a pose-consistent segmentation for the purpose of shape retrieval.

1 Introduction

Research in cognitive science suggests that human shape understanding is based on a decomposition of the shape in smaller parts [7]. Inspired by this insight, many algorithms in three-dimensional shape analysis rely on a segmentation of the objects' surface in meaningful parts.

Such a segmentation can be the building block of shape retrieval techniques where an object is recognized as the sum of its parts [23, 24, 12]. Other interesting applications include CAD, reverse engineering and medical image analysis [1], texture mapping [10] and texture superresolution [6].

In this work we propose a method for automatically determining visually meaningful, pose-consistent segmentations of non-rigid $3D$ shapes. Our approach builds upon a quantum mechanical feature descriptor and upon Gaussian mixture clustering in the feature space over several articulations of a shape.

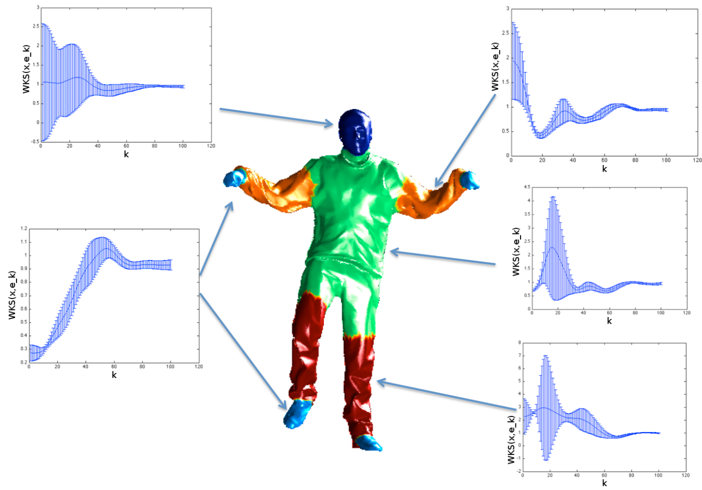

Fig. 1. The basic idea of our segmentation is to group those points in which quantum particles of different energy levels have similar probabilities to be measured. The clustering is achieved with an Expectation-Maximization using a Gaussian mixture distribution. Thus, any cluster is described by a mean descriptor and its variance.

1.1 Related Work

Shape segmentation is a classical problem in shape analysis. For recent surveys on existing methods we refer the reader to [17, 1, 3].

The problem of *pose-consistent segmentation of shapes* has only recently become to the focus of researchers. The task consists in extracting a meaningful partitioning of a shape which identifies the segments consistently over several poses of the shape.

Following the intuition that meaningful shape parts should be rigid, some approaches cluster points whose movement through the different poses is approximately described by the same Euclidean motion. The works [2, 8, 16] fall in this category. Of course, these methods depend on a precomputed correspondence between the articulated shapes which is computationally a very demanding problem.

Other methods employ local feature descriptors and group points with similar signatures. In [24], Toldo et al. cluster convex regions of similar curvature using normalized graph cuts. This approach is inspired by the minima rule in cognitive science. Because the principal curvatures are not isometry-invariant, pose-consistency is not theoretically granted. Indeed, typically the intrinsic distances on a shape do not change significantly from one pose to another, which make isometric deformations a good mathematical model for shape articulations. Shapira et al. [18] use the Shape Diameter Function (SDF) for clustering. The SDF measures at each point the diameter of the shape in inward-normal direction and therefore captures volumetric information on the shape's surface. Again, the SDF is not isometry-invariant whence the segmentation results are not guaranteed to be pose-invariant.

A class of very powerful, isometry-invariant tools for shape analysis rely on the study of the spectrum of the Laplace–Beltrami operator. Our approach belongs to this class. In the geometry processing community, these ideas first appeared in the work [9] of Lévy. Rustamov [15] introduced the Global Point Signature which encodes all local and global information about a point on the shape's surface. Very nice shape segmentation results where shown as an application. However, because the signs and the ordering of the Laplace eigenfunctions can flip from one articulation to another, it is not easy to identify segments over different poses. Reuter [14] proposed a watershed-based segmentation employing a single, user-selected Laplace eigenfunction. Robustness is ensured by persistence-based denoising of the basins. In order to identify labels over different poses, Reuter proposes to align the eigenfunctions of different shapes by comparing persistence diagrams. In [19], Sharma et al. use a constrained spectral clustering approach to segment a single deformable shape. The constraints enforce certain pairs of points to belong to the same segment or to belong to different segments and are given by user input. Label transfer to different shape poses is achieved by registering the shapes. Again, this step involves reordering and sign-flipping of the Laplace eigenfunctions.

To overcome the sign and ordering problem, Sun et al. [22] introduced a very nice, physically motivated feature descriptor, the Heat Kernel Signature (HKS) which encodes the heat dissipation process on the surface. They showed that the HKS contains all information to characterize points uniquely. While the HKS proved to be the current state-of-the-art feature descriptor [5], it has several draw-backs. First of all, the natural parametrization domain for the HKS is time which does not have an intrinsic meaning for a shape. Secondly, due to the exponential decay in diffusion processes, the HKS mixes local and global scales in an intransparent way. In contrast to this, our quantum mechanical feature descriptor, the WKS, is parametrized on the energy domain which has by means of eigenenergies an intrinsic interpretation for a shape. Furthermore, different scales are clearly separated by the WKS.

A persistence-based segmentation technique using the HKS was presented by Skraba et al. [20]. Similarly to Reuter's work, this method is based upon the watershed approach using the HKS function for a user-fixed value of the time t. This value determines whether more local or global features should guide the segmentation.

1.2 Contribution

In this work we present a novel approach for automatically finding pose-consistent segmentations of $3D$ shapes. Our work builds upon a quantum mechanical feature descriptor, the Wave Kernel Signature (WKS). A segmentation is computed in two steps: In a learning step we use several different poses of a shape to build clusters of points for which the probability to find quantum mechanical particles at different energy levels is similar. In the segmentation step, new poses of the shape are partitioned by sorting each point in the most likely cluster. The main contributions can be summarized as follows.

– We show how to incorporate the framework of Quantum Mechanics to pose-consistent shape segmentation. By grouping points in which particles over different energy levels have similar probabilities to be measured, we exploit global as well as local shape information in the segmentation process.
– Our method inherently guarantees consistent transfer of labels to different shape poses, without the need of computationally expensive shape registrations.
– Relying on a clustering in the feature space, our method is easily implemented and fully automatic.

Experimental results show that our segmentation results agree with the human intution, that labels are consistently carried over to new poses and that our method can cope with perturbed data. Finally, we illustrate the usefulness of meaningful shape decompositions with an experiment on shape retrieval.

2 The Wave Kernel Signature – A Quantum Mechanical Feature Descriptor

In this section we describe a quantum mechanical feature descriptor, the WKS, which assigns with each point on an object's surface a vector in $\mathbb{R}^M$. This vector encodes the probability to measure particles of different energy levels in the point. After a brief review of the dynamics of quantum particles on surfaces in 2.1, we give the definition of WKS in 2.2. In 2.3 we outline why WKS is useful for shape analysis, and in particular why it is more convenient than the previously defined HKS. For a more detailed study of the WKS we refer the reader to [4].

2.1 Quantum Particles on Surfaces

A quantum mechanical particle moving on a closed, differentiable surface $X \subset \mathbb{R}^3$ is completely described by its wave function $\psi(x,t) : X \times \mathbb{R}_{>0} \to \mathbb{C}$. This function solves Schrödinger's equation

$$i\frac{\partial \psi}{\partial t}(x,t) = -\Delta_X \psi(x,t), \tag{1}$$

where Δ_X is the Laplace–Beltrami operator of X. While the wave function itself does not have an easy intuitive explanation, for fixed $t > 0$ its squared norm $|\psi(x,t)|^2 : X \to \mathbb{R}$ is the probability density function of the position of the particle at time t.

We now focus on the following physical experiment: Consider a quantum particle on X. Assume that we measure at time $t = 0$ the energy E of this particle and that subsequently we want to determine its position at time $t > 0$.

By the Heisenberg uncertainty relation, we cannot be precise in both, energy and position simultaneously. Thereby, we have to consider a superposition of eigenenergies and eigenstates. Mathematically, the eigenstates and eigenenergies are given by the orthonormal eigenfunctions $\phi_0, \phi_1, \phi_2, \ldots$ and by the corresponding eigenvalues $0 = E_0 > -E_1 \geq -E_2 \geq \ldots$ of the Laplace–Beltrami operator Δ_X. Assume now that the eigenvalues E_k are pairwise distinct which is the case with probability 1. For a particle with energy distribution f_E^2 (hence allowing for uncertainty in the energy), its wave function is given by

$$\psi_E(x,t) = \sum_{k\geq 0} f_E(E_k) \exp\left(-iE_k t\right) \phi_k(x). \tag{2}$$

Using that the functions $\exp\left(-iE_k t\right)_{k\geq 0}$ are orthogonal for the L^2-norm, the average probability that the particle is measured in a point $x \in X$, is computed as

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T |\psi_E(x,t)|^2 dt = \sum_{k\geq 0} f_E(E_k)^2 \phi_k(x)^2. \tag{3}$$

2.2 The Wave Kernel Signature

Now we work out how to use the above insights to design a feature descriptor for shape analysis. For this, it remains to choose the energy distributions f_E^2.

Recall that we aim for a segmentation of shapes undergoing strong pose changes, which correspond mathematically to near-isometric deformations. Therefore we have to optimize our descriptor for robustness to small non-isometric deformations. A perturbation-theoretical analysis which we leave out here due to the lack of space shows that the eigenenergies of a shape under articulation can be modeled as log-normally distributed random variables. More details on this can be found in [4].

This leads us to choose f_E in (3) as a Gaussian distribution in the logarithmic energy $e = \log(E)$ for f_E and we define the Wave Kernel Signature at a point x as

$$\text{WKS}(x,\cdot) : \mathbb{R} \to \mathbb{R}, \quad e \mapsto \frac{1}{C_e} \sum_{k\geq 0} \exp\left(-\frac{\left(e - \log(E_k)\right)^2}{2\sigma^2}\right) \phi_k^2(x), \tag{4}$$

where $C_e = \sum_{k\geq 0} \exp\left(-\frac{(e-\log(E_k))^2}{2\sigma^2}\right)$.

2.3 Comparison of WKS and HKS

The eigenfunctions of the Laplace–Beltrami operator on X can be seen as a generalization of the classical Fourier basis. In this interpretation, eigenvalues play the role of frequencies. Consider a point on a surface as a signal by means of its delta function. Both, the Heat Kernel Signature (HKS) [22] which is defined by

$$\text{HKS}(x,t) = \sum_{k\geq 0} \exp(-E_k t) \phi_k^2(x) \tag{5}$$

and WKS defined by equation (4) are symmetric expressions in the squared Fourier coefficients. Note that the Laplace eigenfunctions depend on the choice of a basis: even in the case of non-repeated Laplace eigenvalues there is a sign ambiguity. Luckily, HKS and WKS are independent of the choice of an orthonormal basis of eigenfunctions. Both descriptors characterize points up to non-rigid motion (cf. [22, 4]).

The difference between HKS and WKS lies in the way Fourier coefficients are filtered. HKS can be seen as a collection of low-pass filters parametrized over the time t. The higher t, the more high frequencies are suppressed. In contrast, WKS is a collection of delta filters in the Fourier domain. The precise form of these delta filters is chosen in such a way that robustness to pose changes is granted as outlined in Section 2.2.

Thereby, WKS should allow for more precise localization of features of points in the frequency domain and thus for a higher precision in recognizing corresponding points. For a thorough experimental comparison of HKS and WKS confirming this heuristic we refer the reader to [4].

3 Learning Pose-Invariant Shape Segmentation

Assume now that we are given a shape in several different poses. Our segmentation aims at grouping points in which quantum particles at different energy levels have similar probabilities to be detected. We build clusters in the following way:

- Pick a subset of training poses which are used for learning the clusters. Typically we used 3-5 training poses.
- Compute the WKS for all points of all training shapes, leading to a point cloud in $\mathbb{R}^M$, where M is the number of evaluation energies of the WKS (which is 100 in all our experiments).
- Fit a Gaussian mixture model with K clusters to these training signatures. The computation was done using the EM algorithm initialized by K-means.

Once the learning step is completed, we can segment both the training poses and new poses by assigning with each point the label of the cluster, on which its WKS has the highest score in the Gaussian mixture distribution.

Of course, we could also use other clustering schemes, leading to similar results. In some cases, we found that imposing the same variance to all the Gaussians of the mixture can lead to slightly more robust results. Indeed, this is a simple way to avoid overfitting: if some scale is very consistent in a cluster, the variance for the corresponding Gaussian at this scale will be so small that a slight change at this scale in a test shape will attribute the points to another cluster. A shared variance will avoid that kind of effects.

Fig. 2. Fully unsupervised segmentation of $3D$ shapes: for different shape classes and for different numbers of clusters, our segmentation algorithm is able recognize semantically meaningful parts and to transfer correctly labels through strong pose-deformations. The left and the middle column show segmentations of shapes from the training set, while the right column visualizes the segmentation of new poses. The shapes are courtesy of [25, 5, 21].

4 Experimental Results

4.1 Computational Details

For computing the WKS on triangle meshes, we discretized the Laplacian using the cotan scheme introduced by Pinkall and Polthier [13]. Boundaries were treated with Neumann conditions. We computed the first $N = 300$ eigenvalues and evaluated the WKS at $M = 100$ values of e ranging from $e_{\min} = \log(E_1)$ to $e_{\max} = \frac{\log(E_N)}{1.02}$ with linear increment $\delta = \frac{e_{\max} - e_{\min}}{M}$. The variance was set to $\sigma = 7\delta$. All these values were fixed in all our experiments.

4.2 Segmentation Results

Figure 2 shows results of segmentations of different shapes in several poses for a varying number of clusters. Notice that the labels, visualized by colors, are

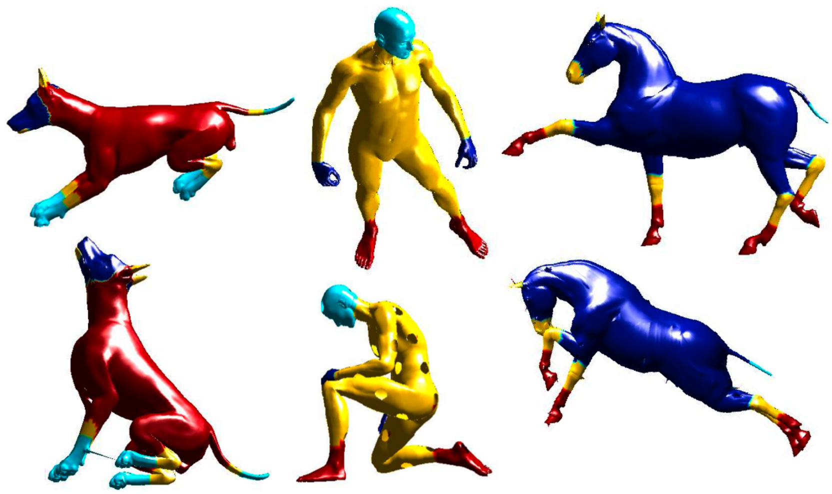

Fig. 3. Robustness of the segmentation results tested on shapes from the SHREC 2010 robustness dataset [5]. On the top, one shape from the training set and at the bottom test shapes with different perturbations. From the left to the right: topology, holes, and shot noise.

automatically transferred correctly to the different articulations and that they are naturally spatially consistent.

4.3 Robustness

To test the robustness of our segmentation, we used the data of the SHREC 2010 benchmark [5]. This dataset contains different shapes undergoing a large variety of poses and of different kinds of perturbations such as topological changes, noise or holes. The method proves stable to such data as can be seen in Figure 3 where some results are visualized.

4.4 Shape Retrieval

As an application of our pose-invariant shape segmentation framework we show an experiment on shape retrieval on the dataset of the SHREC 2010 non-rigid shape retrieval contest [11]. This dataset consists of 10 shape classes each of which contains 20 different shape poses. We choose 5 training shapes from each class and learn a segmentation of these training shapes. As a result of this learning step, we dispose of a Gaussian mixture probability distribution for each shape class. Given a query shape from the database which was not included in the learning process, we compute its WKS at all points and evaluated the negative log-likelihoods of the Gaussian mixtures. The query shape is sorted to the shape class with the maximal log-likelihood. In Figure 4 we visualize the resulting log-likelihood of a query shape for four different shape classes.

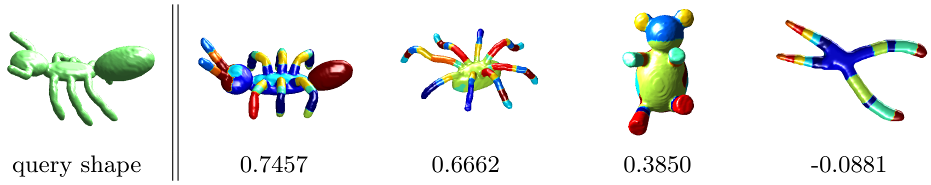

Fig. 4. Shape segmentation applied to shape retrieval on the SHREC 2010 dataset [11]. The four columns on the right show representatives of four shape classes. For each shape class a Gaussian mixture distribution was computed as outlined in Section 3. The resulting segmentations are color encoded. The log-likelihood of the query shape (leftmost column) with respect to these distributions is displayed below each class.

On the 150 query shapes we achieved 72% of correct assignments which is a proof of concept that our part decomposition of shapes is of high informative value for shape recognition.

5 Conclusion

We proposed a novel method for fully unsupervised, pose-consistent $3D$ shape segmentation which arises from a Quantum Mechanical analysis of shapes. By grouping those points in which quantum particles of different energy levels have similar probabilities to be detected, we get an unsupervised partitioning of the shape. Label transfer to different poses is granted by construction without the need of user input or of computationally expensive shape registrations. Interestingly, the computed part decomposition of shapes is consistent with human notions of shape decomposition (torso, head, arms, legs, etc). Finally, we demonstrate that such a segmentation can be efficiently used for shape retrieval.

References

1. Agathos, A., Pratikakis, I., Perantonis, S., Sapidis, N., Azariadis, P.: 3D mesh segmentation methodologies for CAD applications. Computer-Aided Design and Applications 4(6), 827–841 (2007)
2. Anguelov, D., Koller, D., Pang, H.C., Srinivasan, P., Thrun, S.: Recovering articulated object models from 3d range data. In: UAI. pp. 18–26 (2004)
3. Attene, M., Katz, S., Mortara, M., Patané, G., Spagnuolo, M., Tal, A.: Mesh segmentation-a comparative study. In: Shape Modelling International (SMI). IEEE Computer Society (2006)
4. Aubry, M., Schlickewei, U., Cremers, D.: The Wave Kernel Signature - A Quantum Mechanical Approach to Shape Analyis. Tech. rep., TU München, Germany (June 2011)
5. Bronstein, A., Bronstein, M., Bustos, B., Castellani, U., Crisani, M., Falcidieno, B., Guibas, L., Kokkinos, I., Murino, V., Ovsjanikov, M., et al.: SHREC 2010: robust feature detection and description benchmark. Proc. 3DOR (2010)

6. Goldluecke, B., Cremers, D.: Superresolution texture maps for multiview reconstruction. In: IEEE International Conference on Computer Vision (ICCV). Kyoto, Japan (2009)
7. Hoffman, D.: Visual intelligence: How we create what we see. WW Norton and Company (2000)
8. James, D., Twigg, C.: Skinning mesh animations. ACM Transactions on Graphics (SIGGRAPH 2005) 24(3) (August 2005)
9. Lévy, B.: Laplace-Beltrami Eigenfunctions Towards an Algorithm That ÜnderstandsĞeometry. In: Proc. Int. Conf. on Shape Modeling and Applications. p. 13. IEEE (2006)
10. Lévy, B., Petitjean, S., Ray, N., Maillot, J.: Least squares conformal maps for automatic texture atlas generation. ACM Transactions on Graphics 21(3), 362–371 (2002)
11. Lian, Z., Godil, A., Fabry, T., Furuya, T., Hermans, J., Ohbuchi, R., Shu, C., Smeets, D., Suetens, P., Vandermeulen, D., et al.: SHREC'10 Track: Non-rigid 3D Shape Retrieval. In: Eurographics 3DOR (2010)
12. Ovsjanikov, M., Bronstein, A., Bronstein, M., Guibas, L.: Shape Google: a computer vision approach to invariant shape retrieval. Proc. NORDIA 1(2) (2009)
13. Pinkall, U., Polthier, K.: Computing discrete minimal surfaces and their conjugates. Experimental Mathematics 2(1), 15–36 (1993)
14. Reuter, M.: Hierarchical shape segmentation and registration via topological features of laplace-beltrami eigenfunctions. International Journal of Computer Vision
15. Rustamov, R.: Laplace-Beltrami eigenfunctions for deformation invariant shape representation. In: SGP. pp. 225–233. Eurographics (2007)
16. Schaefer, S., Yuksel, C.: Example-based skeleton extraction. In: Symposium on Geometry Processing. pp. 153–162 (2007)
17. Shamir, A.: A survey on mesh segmentation techniques. In: Computer Graphics Forum. vol. 27, pp. 1539–1556 (2008)
18. Shapira, L., Shalom, S., Shamir, A., Cohen-Or, D., Zhang, H.: Contextual part analogies in 3d objects. International Journal of Computer Vision 89(2-3), 309 – 326 (2010)
19. Sharma, A., von Lavante, E., Horaud, R.P.: Learning shape segmentation using constrained spectral clustering and probabilistic label transfer. In: Proc. European Conference on Computer Vision. LNCS (2010)
20. Skraba, P., Ovsjanikov, M., Chazal, F., Guibas, L.: Persistence-based segmentation of deformable shapes. In: NORDIA. pp. 45–52. IEEE (2010)
21. Sumner, R., Popović, J.: Deformation transfer for triangle meshes. ACM Transactions on Graphics (TOG) 23(3), 399–405 (2004)
22. Sun, J., Ovsjanikov, M., Guibas, L.: A Concise and Provably Informative Multi-Scale Signature Based on Heat Diffusion. Computer Graphics Forum 28(5), 1383–1392 (2009)
23. Tal, A., Zuckerberger, E.: Mesh retrieval by components. In: GRAPP. pp. 142–149 (2006)
24. Toldo, R., Castellani, U., Fusiello, A.: Visual vocabulary signature for 3D object retrieval and partial matching. Proc. 3dOR 8, 21–28 (2009)
25. Vlasic, D., Baran, I., Matusik, W., Popović, J.: Articulated mesh animation from multi-view silhouettes. In: ACM SIGGRAPH 2008 papers. pp. 1–9. ACM (2008)