Semantic Shape Context for the Registration of Multiple Partial 3D Views

Samir Khoualed[1]
samir.khoualed@lasmea.univ-bpclermont.fr
Umberto Castellani[2]
umberto.castellani@univr.it
Adrien Bartoli[1]
adrien.bartoli@gmail.com

[1] LASMEA
CNRS/Université Blaise Pascal
Clermont-Ferrand, France

[2] VIPS
University of Verona
Verona, Italy

Estimating rigid transformations that align corresponding points of partial 3D views is a critical issue for various tasks in computer vision (e.g., object recognition, object tracking and 3D model reconstruction). The ICP algorithm [2] is the gold standard for pairwise view alignment, but it requires a sufficient overlap among the views and a coarse pre-registration to avoid getting stuck in a local minimum. In particular, according to the taxonomy proposed in [5], when an initial estimate is unknown and more than two views are involved, the problem is called *multiview surface matching*. Three main sub-problems need to be solved [5]: (i) determining which views overlap, (ii) determining the relative pose between each pair of overlapping views, and (iii) determining the absolute pose of the views. Many works address these issues. Focusing on (i) and (ii), two main rough categories of methods have been introduced: *local* methods [6, 8], which are based on point-to-point correspondences, estimated based on a point signature describing local surface properties, and *global* methods which estimate directly the matching of the whole views by comparing global surface characteristics. Combining local and global information is a promising approach, but only few methods use this strategy, to the best of our knowledge (see [7], for an extensive overview of 3D shape matching methods). An interesting and effective approach has been proposed in [4], which is an extension of the so called *shape context* [1] to the 3D domain.

In this paper, we improve the basic idea of the shape context by combining local descriptors with the Bag-of Words (BoW) paradigm. We focus on the problem of multiview surface matching by addressing the early two above-described sub-problems. The proposed point description and matching approach is based on the following steps:

1. **Local points description**. Several local point descriptors are computed in order to capture the local shape variation in the point neighborhood [6]. We focus on the following geometric measures to compute the descriptors for each feature point:

   - **Shape Index** *si* [9]. The Shape Index is defined as:

     $$si = -\frac{2}{\pi}\arctan\left(\frac{k_1+k_2}{k_1-k_2}\right) \qquad k_1 > k_2$$

     where k_1, k_2 are the principal curvatures of a generic vertex. The Shape Index varies in $[-1,1]$ and provides a local categorization of the shape into primitive forms such as spherical cap and cup, rut, ridge, trough, or saddle [9].

   - **Beta Value** *bv*. The Beta Value of vertex p is represented by recording the distance from all nearby vertices v to p, along normal $\vec{n}_p$ at p.

2. **Visual vocabulary construction**. The set of point descriptors collected from all the views of the same model are properly clustered in order to obtain a fixed number of 3D visual *words* (i.e., the set of cluster centroids) [3].

3. **Context definition**. Each local descriptor is assigned to a visual word, and a BoW representation is defined by counting the number of points assigned to each word. In particular, for a fixed point its context is defined as the set of BoWs computed on several regions which are defined by concentric shells centered on the fixed point itself.

4. **Point matching**. The matching between two points is computed by comparing their respective signatures and by taking into account the different kinds of descriptors. Both the *local* and *contextual* contributions are considered. Figure 1 shows an example of this phase.

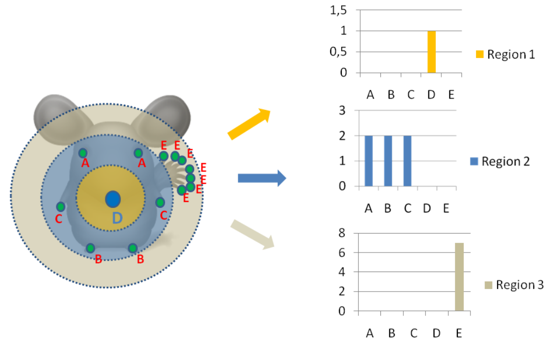

Figure 1: Context definition for a given point (colored in blue). Three sub-regions are identified by three shells. The BoW representation is computed for each shell. The set of three BoWs forms the SSC.

The main idea of this approach consists in the fact that the proposed context encodes not only the spatial relationship between points, but also their ‘class’ w.r.t. each local descriptor (i.e., points associated to the same cluster belong to the same class). We thus call this new representation *Semantic Shape Context* (SSC), where here the semantic is inferred by the point classification. It is worth noting that the choice of local point descriptors is not the focus of this work, since in principle any set of local descriptors can be used and cast in the proposed context. The effectiveness of the SSC is shown by proposing a multiview surface matching framework which implements a fully automatic model registration pipeline.

[1] S. Belongie, J. Malik, and J. Puzicha. Shape matching and object recognition using shape contexts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(4):509–522, 2002.

[2] P.J. Besl and H.D. McKay. A method for registration of 3-D shapes. *IEEE Transactions on pattern analysis and machine intelligence*, 14 (2):239–256, 1992.

[3] G. Csurka, C. Dance, L. Fan, J. Willamowski, and C. Bray. Visual categorization with bags of keypoints. In *Workshop on Statistical Learning in Computer Vision, ECCV*, volume 2004, 2004.

[4] A. Frome, D. Huber, R. Kolluri, T. Bulow, and J. Malik. Recognizing objects in range data using regional point descriptors. *Lecture Notes in Computer Science*, pages 224–237, 2004.

[5] D.F. Huber and M. Hebert. Fully automatic registration of multiple 3D data sets. *Image and Vision Computing*, 21(7):637–650, 2003.

[6] A.E. Johnson and M. Hebert. Using spin images for efficient object recognition in cluttered 3D scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(5):433–449, 1999.

[7] A.S. Mian, M. Bennamoun, and R.A. Owens. Automatic correspondence for 3d modeling: An extensive review. *International Journal of Shape Modeling*, 11(2):253, 2005.

[8] A.S. Mian, M. Bennamoun, and R.A. Owens. A novel representation and feature matching algorithm for automatic pairwise registration of range images. *International Journal of Computer Vision*, 66(1): 19–40, 2006.

[9] S. Petitjean. A survey of methods for recovering quadrics in triangle meshes. *ACM Computing Surveys*, 34(2), 2002.