

Similarity Analysis of Archaeological Potsherds Using 3D Surfaces

Edgar Roman-Rangel¹ (✉) and Diego Jimenez-Badillo²

¹ CVMLab - University of Geneva, Geneva, Switzerland
edgar.romanrangel@unige.ch

² National Institute of Anthropology and History of Mexico (INAH),
Mexico city, Mexico
diego_jimenez@inah.gob.mx

Abstract. This work presents a new methodology for efficient scanning and analysis of 3D shapes representing archaeological potsherds, which is based on single-view 3D scanning. More precisely, this work presents an analysis of visual diagnostic features that help classifying potsherds, and that can be automatically identified by the appropriate combination of the 3D models resulting from the proposed scanning technique, with robust local descriptors and a bag representation approach, thus providing with enough information to achieve competitive performance in the task of automatic classification of potsherd. Also, this technique allows to perform other types of analysis such as similarity analysis.

Keywords: 3D shapes · Potsherds · Classification · Similarity analysis

1 Introduction

The analysis of archaeological potsherds is one of the most demanding tasks in Archaeology, as it helps understanding both cultural evolution and relations (economic and military) between different cultural groups from the past [1].

One of the trending methods for studying potsherds in the last decade consists on generating 3D models by scanning surviving pieces of ceramics found at excavation sites [2]. This approach enables new capabilities with respect to analyzing the original pieces, e.g., automatic and semi-automatic content analysis [2], virtual reconstructions [3], more efficient archiving and sharing of documents [4, 5], etc. However, it results in a very time-consuming task, as it requires scanning each piece several times from different viewpoints and a later, often manual, registration of the different shots to build the full 3D model [2].

In this work, we present a new approach based on fast scanning of single-view 3D surfaces, which is a much faster process to obtain 3D information. We complemented the proposed approach by evaluating the descriptive performance of the 3D SIFT descriptor [6], which is a relatively new extension of the well known SIFT descriptor [7] used for description of 2D images. This evaluation consisted in computing 3D SIFT descriptors for 3D surfaces of potsherds, create

bag representation, and perform classification experiments. Our results show that the 3D surfaces contain enough information to achieve competitive classification performance. Also, we conducted a similarity analysis to evaluate the potential risk for confusing potsherd coming from different ceramic pieces.

The remaining of this paper is organized as follows. Section 2 gives an introduction to the type of data used in this work. Section 3 explains the method used for description of 3D surfaces, namely the 3D SIFT descriptor [6]. Section 4 explains the two types of experiments we performed, and Sect. 5 presents our results. Finally, Sect. 6 presents our conclusions.

2 3D Potsherds

This section describes the dataset of 282 potsherds used in this work. Namely, it explains the origin of the potsherds, the type of digital data generated from them, and basic statistics of the resulting dataset.

2.1 The Data Set

Origin. The ceramic style of the potsherds used in this work is known as “Aztec III Black on Orange”, and it constitutes one of the most important wares in Mesoamerican archaeology [1]. Geographically speaking, this ceramic style corresponds to the entire Basin of Mexico, i.e., the geographical extension of the former Aztec empire, and it used to be part of the common utilitarian assemblage in households during the late Aztec period (1350–1520 C.E.).

During the 1970’s, a large amount of these ceramics was collected on the surface by the Valley of Mexico Survey Projects [8,9]. Later, stylistic analyses of their materials was performed during the 1990’s, along with an important comparison of their geographic distribution with historically-documented polities [10,11], which helped understand the economic relations between dependent communities and the Aztec capital.

Single View Scanning. As part of our proposed approach, we decided to rely on using 3D meshes generated from scanning the potsherds from a single viewpoint. This is, the type of data we analyzed consists on 3D surfaces rather than the traditional approach of using volumetric models [2]. Figure 1 shows a few examples of these single view 3D surfaces.

Scanning these single view 3D surfaces is a faster data acquisition process, in comparison to the traditional approach of using full 3D models, which can take much longer time as it requires the scanning from different viewpoints and a later, often manual, registration step for generating the full model [2].

As shown in the examples of Fig. 1, there are many section of the potsherds that are common across classes, e.g., central sections of the potsherd, which intuitively suggests that the risk for obtaining poor classification performance is high. However, note that there are also some sections of the potsherd containing visual information that is diagnostic for specific classes, thus increasing the potential for achieving a good classification performance.

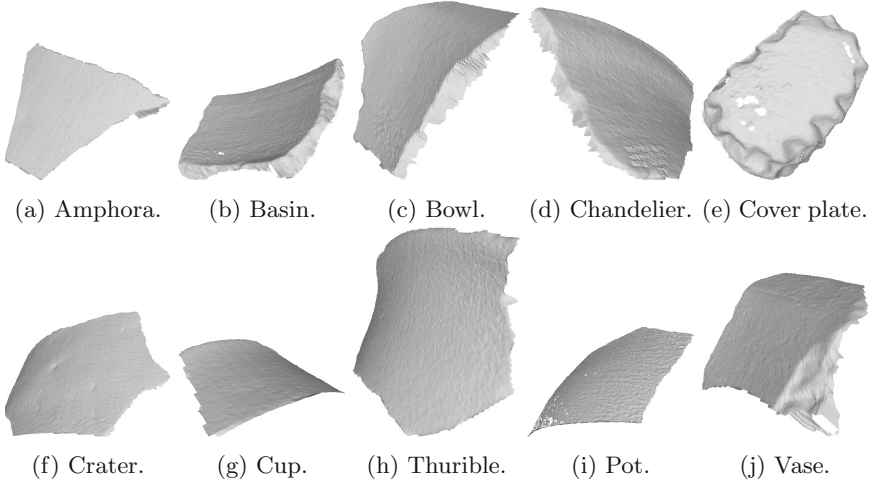


Fig. 1. Examples of single view 3D surfaces scanned from the collection of potsherds. Each example corresponds to one of the classes in our dataset.

Dataset Statistics. The dataset used in this work consists of 282 3D surfaces scanned from potsherds of 10 different ceramic objects, i.e., classes. Figure 2 shows the distribution of 3D surfaces over the 10 ceramic classes. Note that this dataset is not well balanced, i.e., there are much more instances in classes amphora and pot (over 60 in each), while there are only a few instances in classes basin, cover plate, and thurible (only above 10 in each). This behavior results as a consequence of the frequency at which the potsherds are found in the field. Nevertheless, in Sect. 5.2, we conducted an analysis using a balanced subset of this data.

3 Model Description

Given the promising results reported in previous works [6, 12], we use 3D SIFT descriptors and the bag-of-words model (BOW) [13] to represent the 3D surfaces.

3D SIFT. The 3D Scale Invariant Feature Transform (SIFT) [6] is the 3D extension of the well known SIFT descriptor [7].

Given that points on a 3D surface (i.e., mesh vertices) might be non-uniformly spaced, the repeatability of the traditional difference of Gaussian (DoG) [7] detector of points of interest might be negatively affected [6]. Nevertheless, the Gaussian scale space can be constructed (approximated) by using an alternative method which is invariant to distance between vertices but not to their location. Namely, the smoothing of a given vertex \mathbf{v}_i^{s+1} at scale $(s + 1)$ is given by,

$$\mathbf{v}_i^{s+1} = \frac{1}{|V_i^s|} \sum_{\mathbf{v}_j^s \in V_i^s} \mathbf{v}_j^s, \quad (1)$$

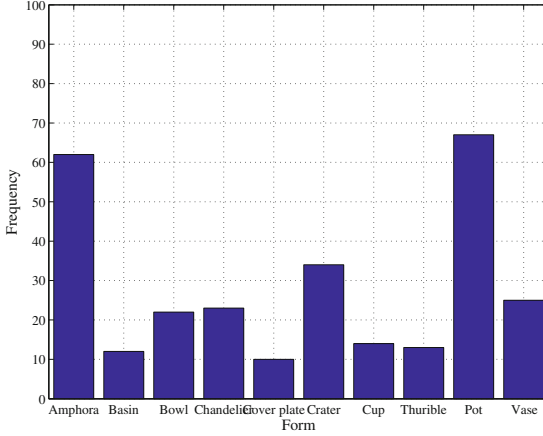


Fig. 2. Frequency of elements over classes in our dataset.

where, V_i^s is the set of first order neighbors of \mathbf{v}_i^s , i.e., the set of vertices sharing an edge with \mathbf{v}_i^s , $|\cdot|$ denotes the cardinality operator, and the summation of the vertices is performed component-wise on the x , y , and z components of \mathbf{v}_i .

After computing the Gaussian smoothing of the 3D surface using Eq. (1), the difference of Gaussians is approximated by,

$$d_i^s = \frac{1}{(s \cdot \sigma_{i,0}^2)} \left(\mathbf{v}_i^s - \mathbf{v}_i^{s+1} \right), \quad (2)$$

where, σ_0 denotes the initial variance of the integration parameter, which is independently estimated for each vertex as,

$$\sigma_{i,0} = \frac{1}{|V_i^s|} \sum_{\mathbf{v}_j^s \in V_i^s} \text{abs} \left(\mathbf{v}_i^s - \mathbf{v}_j^s \right), \quad (3)$$

where, $\text{abs}(\cdot)$ indicates absolute.

Implementing this methodology, allows to select those 3D vertices maximizing Eq. (2) as points of interest for which a 3D SIFT descriptor will be computed.

In turn, the computation of the 3D SIFT descriptor requires the computation of a depth map [6], which is estimated by projecting the 3D vertices onto a 2D surface, thus generating a kind of image on which the traditional SIFT descriptor [7] is computed. More precisely, a dominant plane P_i is estimated for each vertex \mathbf{v}_i . Namely, P_i corresponds to the tangent plane of \mathbf{v}_i , which can be found by its normal \mathbf{n}_i and the point \mathbf{v}_i itself.

Once the dominant plane P_i is computed, all the neighboring points to \mathbf{v}_i are mapped onto a 2D array, for which the SIFT descriptor is computed. This is done by filling in the 2D array the distance that exists from the 3D surface to the

dominant plane P_i for each neighboring point. In turn, the neighboring points of \mathbf{v}_i correspond to those vertices within a distance D_i from \mathbf{v}_i , where D_i is defined as,

$$D_i = C\sqrt{s_i}\sigma_{i,0}, \quad (4)$$

where, $\sigma_{i,0}$ is computed as defined in Eq. (3), s_i is the scale at which the point of interest attains its maximum value according to Eq. (2), and C is a parameter to control the level of locality [6].

After computing the sets of 3D SIFT descriptors for each 3D surface, we quantized them to construct bag-of-words representation (BOW), using dictionaries of different sizes. Later, we used the bag representations to compute distances between 3D surfaces and to conduct our similarity analysis of potsherds.

4 Analysis of Potsherds

We performed two types of analysis of the Aztec potsheds using their 3D surfaces. Namely, experiments of automatic classification, and the analysis of similarity.

Classification Performance. By relying on the k -NN classification approach [14], with $k = 1$, we evaluated the classification performance achieved by different BOW representations [13] of 3D-SIFT descriptors [6].

To this end, we estimated vocabularies of different sizes using the k -means clustering algorithm [15]. More precisely, vocabularies of 100, 250, 500, 1000, 2500, and 5000 words, and then compared the resulting bag representations using the Euclidean distance. In Sect. 5.1, we present the results of this evaluation along with the confusion matrix generated by the dictionary that achieved the best classification rate.

Similarity Analysis. To better understand the numeric results of the previously described classification experiments, we estimated the intra-class and inter-class distance that can be expected when using our approach, i.e., BOW [13] of 3D-SIFT descriptors [6] computed over surfaces. We computed these values as:

- Intra-class distance: The average of the pairwise distance between all instances within the class of interest.
- Inter-class distance: The average of the pairwise distance from each instance of the class of interest to all instances on the remaining classes.

Section 5.2 presents the intra-class and inter-class average distance of our dataset computed using the dictionary that achieved the best classification performance. Also in that section, we present a graph-based relation of the, on average, most similar classes given a reference class.

5 Results

This section presents the results of both of our analysis: classification performance and similarity analysis.

5.1 Classification Performance

The first result from the classification performance test indicates that small vocabularies are better to achieve good classification rates. Table 1 shows that 58.45% of the potsherds are correctly classified by using only 100 and 250 words, and that this rate drops quickly with vocabularies of 1000 words or more. Although these results confirm previous observations regarding the impact of the vocabulary size [12], they also contradict some results of recent works [16].

Table 1. Classification performance achieved with vocabularies of different sizes.

vocabulary size	100	250	500	1000	2500	5000
classification rate (%)	58.45	58.45	54.26	47.52	39.36	36.17

By visual inspection, we realized the reason for which small vocabularies work well for our particular 3D structures, i.e., surfaces. Namely, besides of having instances with large amount of variations, these variations can be captured by very local descriptors, i.e., descriptors that capture information only within the closest neighborhood of the point of interest. This particularity of our data allows to represent all important local variations with only a small set of prototype descriptors, i.e., words. See Fig. 1 for visual examples of the dataset.

Figure 3 shows examples of the two most different words in the 100 words vocabulary. Namely, words number 67 and 95, where the distance between words is estimated as the distance between centroids of their corresponding clusters. More precisely, Fig. 3 shows 3D surfaces, each with highlighted sections that correspond to the 3D SIFT descriptors assigned as instances of those particular words, i.e., closest to the centroid computed by k-means.

More precisely, instances of word 95 (orange circles in Fig. 3) describe a slightly curved section, while instances of word 67 (blue circles) describe roughly flat sections near an abrupt change on the surface, like a hole (Fig. 3a and b) or the starting point of the base of the ceramic (Fig. 3c). However, note that given the nature of the 3D data, a small set of local descriptors suffices for accurate description, as well as for competitive classification performance.

Figure 4 shows the confusion matrix obtained with the vocabulary of 100 words. As shown by its main diagonal, our approach is able to achieve competitive classification performance for most of the classes, and specially for classes Amphora, Chandelier, and Pot, i.e., 81%, 74%, and 64% of the times, a potsherd from those classes is properly assigned to the correct class label. On the other hand, potsherd of classes Basin, Thurible, and Vase seem to be more challenging.

5.2 Similarity Analysis

As explained in Sect. 4, we evaluated the intra-class and inter-class distance of the 3D surfaces, to acquire a better understanding of the classification potential that

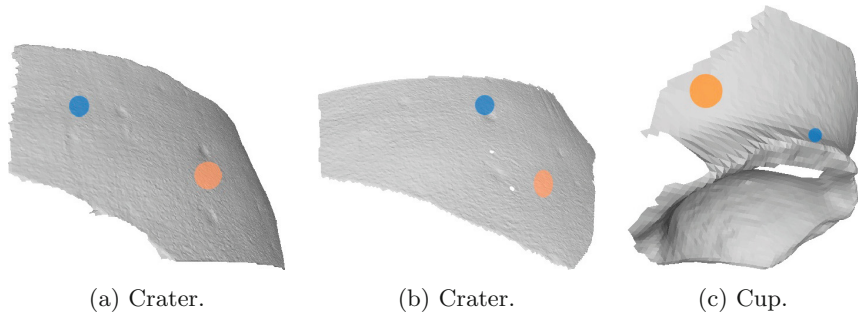


Fig. 3. 3D surfaces of potsherds from crater and cup classes. Highlighted sections correspond to the spatial locality of 3D SIFT descriptors. Blue sections indicate instances of word 67, and orange sections indicate instances of word 95, which are the two most different words in the 100 words vocabulary (Color figure online).

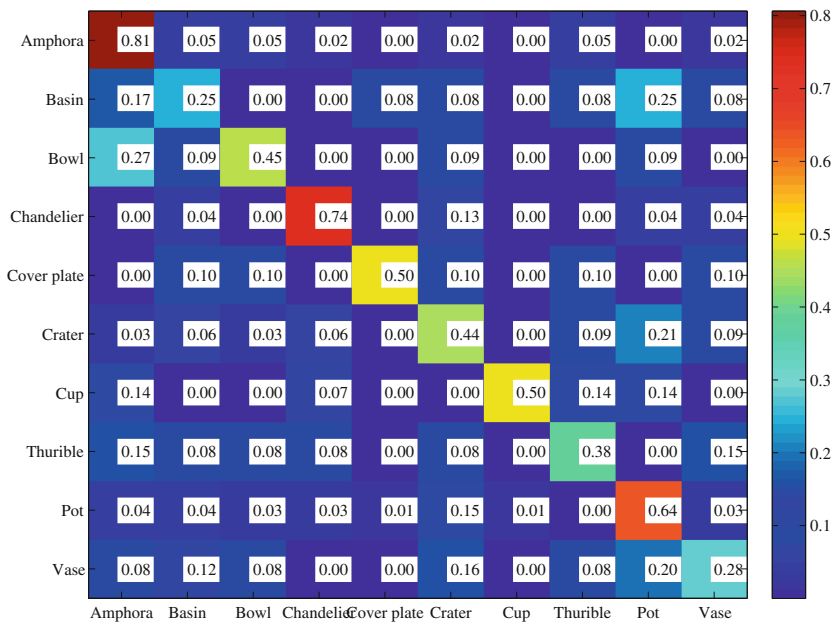


Fig. 4. Confusion matrix obtained from the automatic classification of 3D surfaces scanned from Potshards of the ancient Aztec culture.

the 3D SIFT descriptor has to deal with potsherds. For this analysis we used the vocabulary of 100 words, which achieved the highest classification performance, as shown in Sect. 5.1.

For each class in our dataset, Fig. 5 shows the intra-class distance alongside its inter-class distance counterpart. As one could expect for a well behaved scenario, the distance between elements of the same class is shorter than the distance

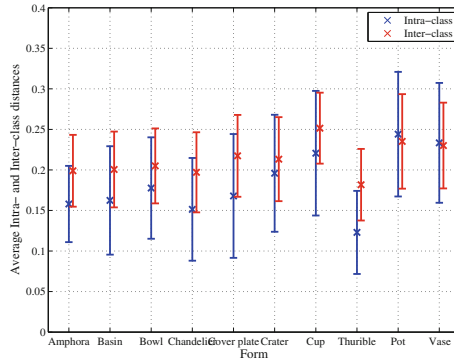
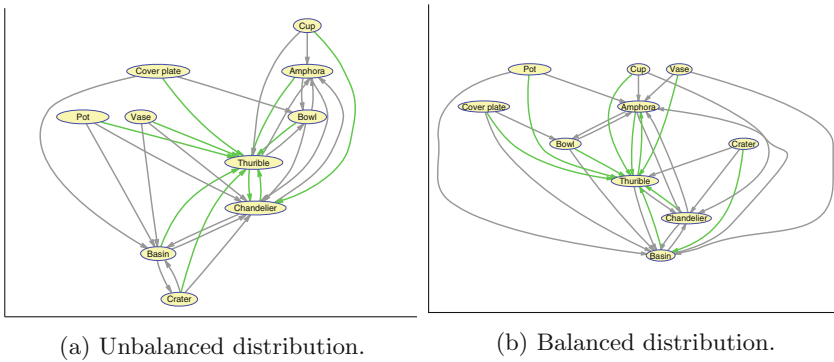


Fig. 5. Intra-class and Inter-class similarity of Potsherds.



(a) Unbalanced distribution.

(b) Balanced distribution.

Fig. 6. Inter-class similarity in our dataset. Green arrows indicate the most similar reference class to each other class (class of interest). Gray arrows point to the second and thirds most similar reference classes to each other class (Color figure online).

between elements of different classes, i.e., the average intra-class distance is lower than the average inter-class distance. This observation is true for 8 of the 10 classes of potsherd, with classes Pot and Vase having slightly larger intra-class distance, as well as larger standard deviation.

A more detailed explanation of the average distance is shown in Fig. 6, where more precisely, the average similarity between pairs of classes is presented. For this purpose, a class of interest c_i is considered similar to a reference class c_r if the average inter-class distance, as explained in Sect. 4, between c_i and c_r is smaller than the average inter-class distance from c_i to any other class.

For each class in our dataset, Fig. 6a shows its three most similar classes. Note that the class Thurable is indicated as the most similar for 8 out of the 9 remaining classes, which might explain the low classification rate shown in the confusion matrix of Fig. 4. As previously shown in Fig. 2, the distribution of instances is not well balanced across classes of potsherds. To verify whether or

not this unbalanced characteristic of the dataset has an impact on our analysis, we randomly selected a subset of 10 instances per class, thus generating a balanced dataset. Figure 6b shows the three most similar classes for each reference class using this balanced dataset. Although a few arrows changed their end node, not important differences were observed when using a balanced dataset. This suggests that this is the natural behavior of the potsherds, and it does not depend on the amount of instances available for each class.

6 Conclusions

We presented a novel approach for automatic analysis of archaeological potsherds, based on the use of 3D surfaces scanned from a single viewpoint. More precisely, potsherds of ceramic artifacts from the ancient Aztec culture. The proposed scanning approach is much faster than the traditional method that takes several shots from different viewpoints and then registers them to construct a complete 3D model. Yet, our methodology produces data with enough information that can be described using 3D SIFT descriptors and bag representations.

Using the proposed methodology, we evaluated the impact that the size of different visual dictionaries has on classification tasks. Our results show that small vocabularies suffice to obtain competitive classification performance, which becomes harmed as the size of the dictionary increases. In particular, we obtained the best classification performance by using dictionaries of 100 and 250 words. Namely, our methodology is able to attain competitive classification performance in most cases, with the few exceptions of certain classes of potsherd that are more difficult to categorize, i.e., basin, thurible, and vase. More over, thurible and basin are the two classes that were ranked the most common classes, i.e., the classes whose instances are most similar to instances of other classes, as shown by the quantitative similarity analysis that we conducted.

Overall, this type of analysis is of interest for archaeologists, as it shows the potential that patten recognition techniques have to deal with some of the most challenging problems in Archaeology.

Acknowledgments. This work was founded by the Swiss-NSF through the Tepalcatl project (P2ELP2-152166).

References

1. Cervantes Rosado, J., Fournier, P., Carballal, M.: La cerámica del Posclásico en la Cuenca de México. In: La producción Alfarera en el México Antiguo, vol. V, Colección Científica. Insituto Nacional de Antropología e Historia (2007)
2. Karasik, A., Smilanski, U.: 3D scanning technology as a standard archaeological tool for pottery analysis: practice and theory. *J. Archaeol. Sci.* **35**, 1148–1168 (2008)
3. Kampel, M., Sablatnig, R.: Virtual reconstruction of broken and unbroken pottery. In: Proceedings of the International Conference on 3-D Digital Imaging and Modeling (2003)

4. Razdan, A., Liu, D., Bae, M., Zhu, M., Farin, G.: Using geometric modeling for archiving and searching 3D archaeological vessels. In: Proceedings of the International Conference on Imaging Science, Systems, and Technology, CISST (2001)
5. Larue, F., Di-Benedetto, M., Dellepiane, M., Scopigno, R.: From the digitization of cultural artifacts to the web publishing of digital 3D collections: an automatic pipeline for knowledge sharing. *J. Multimedia* **7**, 132–144 (2012)
6. Darom, T., Keller, Y.: Scale-invariant features for 3-D mesh models. *IEEE Trans. Image Process.* **21**, 2758–2769 (2012)
7. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* **60**, 91–110 (2004)
8. Sanders, W.T., Parsons, J., Santley, R.S.: *The Basin of Mexico: Ecological Processes in the Evolution of a Civilization*. Academic Press, New York (1979)
9. Parsons, J.R., Brumfiel, E.S., Parsons, M.H., Wilson, D.J.: Prehistoric settlement patterns in the southern valley of Mexico: the Chalco-Xochimilco regions. In: *Memoirs of the Museum of Anthropology*, vol. 14. University of Michigan (1982)
10. Hodge, M.G., Minc, L.D.: The spatial patterning of Aztec ceramics: implications for prehispanic exchange systems in the valley of Mexico. *J. Field Archaeol.* **17**, 415–437 (1990)
11. Hodge, M.G., Neff, H., Blackman, M.J., Minc, L.D.: Black-on-orange ceramic production in the Aztec empire’s heartland. *Lat. Am. Antiq.* **4**, 130–157 (1993)
12. Roman-Rangel, E., Jimenez-Badillo, D., Aguayo-Ortiz, E.: Categorization of Aztec Potsherds using 3D local descriptors. In: Jawahar, C.V., Shan, S. (eds.) *ACCV 2014 Workshops*. LNCS, vol. 9009, pp. 567–582. Springer, Heidelberg (2015)
13. Sivic, J., Zisserman, A.: Video Google: a text retrieval approach to object matching in videos. In: Proceedings of the IEEE International Conference on Computer Vision (2003)
14. Cover, T., Hart, P.: Nearest neighbor pattern classification. *IEEE Trans. Inf. Theor.* **13**(1), 21–27 (1967)
15. Lloyd, S.: Least squares quantization in PCM. *IEEE Trans. Inf. Theor.* **28**, 129–137 (1982)
16. Boyer, E., Bronstein, A.M., Bronstein, M.M., Bustos, B., Darom, T., Horaud, R., Hotz, I., Keller, Y., Keustermans, J., Kovnatsky, A., Litman, R., Reininghaus, J., Sipiran, I., Smeets, D., Suetens, P., Vandermeulen, D., Zaharescu, A., Zobel, V.: SHREC 2011: robust feature detection and description benchmark. In: Proceedings of the 4th Eurographics Conference on 3D Object Retrieval (2011)