Material Recognition for Efficient Acquisition of Geometry and Reflectance

Michael Weinmann(✉) and Reinhard Klein

Institute of Computer Science II, University of Bonn, Bonn, Germany
mw@cs.uni-bonn.de

Abstract. Typically, 3D geometry acquisition and reflectance acquisition techniques strongly rely on some basic assumptions about the surface reflectance behavior of the sample to be measured. Methods are tailored e.g. to Lambertian reflectance, mirroring reflectance, smooth and homogeneous surfaces or surfaces exhibiting mesoscopic effects. In this paper, we analyze whether multi-view material recognition can be performed robust enough to guide a subsequent acquisition process by reliably recognizing a certain material in a database with its respective annotation regarding the reconstruction methods to be chosen. This allows selecting the appropriate geometry/reflectance reconstruction approaches and, hence, increasing the efficiency of the acquisition process. In particular, we demonstrate that considering only a few view-light configurations is sufficient for obtaining high recognition scores.

Keywords: Material recognition · Reflectance · Set-based classification

1 Introduction

The goal of accurately capturing details in surface geometry and reflectance behavior has led to a huge number of different methods and respective setups. However, current state-of-the-art acquisition procedures are rather designed regarding the expected reflectance behavior. The acquisition process is guided by the user who chooses the acquisition routines based on the impression of the material appearance he obtains when looking at the material sample.

In the domain of reflectance acquisition, it is well-known that smooth, homogeneous materials can be represented well with analytical BRDF models. These typically only depend on the direction of the incoming light and the view direction. In contrast, materials exhibiting mesoscopic effects of light exchange on surface structures imaged to a size of approximately one pixel cannot be modeled by using simple BRDF models. For such materials, current state-of-the-art techniques acquire data-driven BTFs which consider the spatial material variations in addition to the view direction and the direction of the incoming light.

In a similar way, 3D reconstruction techniques typically also depend on some basic assumptions about material reflectance. Many of the methods such as most multi-view stereo techniques, photometric stereo and structured light systems are based on assuming Lambertian reflectance behavior. Some more sophisticated extensions allow considering the wider range of opaque surfaces. In contrast,

© Springer International Publishing Switzerland 2015
L. Agapito et al. (Eds.): ECCV 2014 Workshops, Part III, LNCS 8927, pp. 321–333, 2015.
DOI: 10.1007/978-3-319-16199-0_23

other reconstruction techniques are specialized on mirroring surfaces. All the aforementioned geometry reconstruction techniques consider only a small fraction of the possible surface materials and are not tailored to consider arbitrary surface reflectance.

Without a-priori knowledge about the material properties of the material sample, the naive way would be applying several different techniques and merging their results. However, this is highly inefficient regarding acquisition time, and hardware components are stressed unnecessarily as many of the taken images do not have an influence on the final reconstruction and, thus, have to be neglected. For more efficient geometry and reflectance acquisition procedures in case of missing information about the material properties of the considered material sample or object, it is therefore desirable to automatically select only the appropriate techniques instead of applying several different methods.

In this paper, we focus on this task by investigating a-priori information in form of a database of material measurements to classify a measured material based on a small set of photos (see Fig. 1). Depending on the annotations for the closest match in the database, corresponding methods can easily be determined. For an almost mirroring metal, for instance, a shape-from-specularity approach could be used for geometry reconstruction and a BRDF measurement could be started for measuring the surface reflectance. If the considered material sample is classified as a material with strong mesoscopic effects such as present in e.g. leather, a structured light based geometry acquisition in combination with a BTF acquisition could be proposed. We demonstrate that material recognition can be achieved with a high reliability by looking at the characteristic material appearance under a few viewpoints. At first sight, this problem might seem to be not as interesting any more due to the successful studies on databases such as the CUReT database [7]. However, such databases offer only a small intra-class variance in the appearance of the involved material samples. With recent, more challenging databases with larger intra-class variances of the respective material samples such as the ALOT database [12] and the database in [30], there is a need to obtain further insights into recognizing materials using multiple view-light directions for the reference/query sets.

In summary, the key contributions of our work are

– an approach to classify material instances which can serve as an initialization for an efficient acquisition process, and
– a study for using set-based classifiers to find the closest material in the database from a set of view-light configurations which might not necessarily be contained in the database.

2 Related Work

Material classification is a challenging problem due to the significant variations in material appearance under different configurations of viewpoint, illumination and surface geometry and a lot of studies have been conducted. In the following, we briefly discuss model-based and appearance-based approaches.

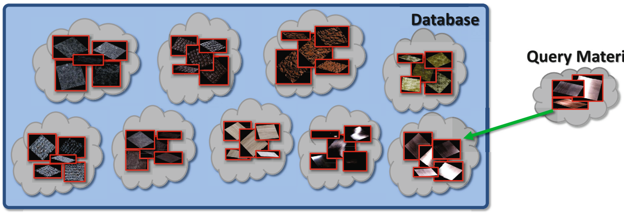

Fig. 1. Material recognition using multi-view information can be formulated as a set-based recognition task. Characteristic material traits observed in the images of a particular material instance form a respective, characteristic material space. The objective is to identify the most similar material instance within the database for an input query material by comparing the material spaces.

Model-based approaches: The inference of knowledge about the considered material surface can be approached by using certain models which capture the variations of material appearance under different view-light configurations. For analyzing physical surface models, histogram models have been used in e.g. [9] and [13] to represent the changes in appearance for materials under varying view-light conditions. In [22], recognition is approached based on a partly Lambertian and partly specular model. The studies in [5] and [10] analyzed the model-based dependency of texture features on illumination. However, such dependencies rely on certain surface characteristics which also applies for the assumed reflectance models. While analytical models might be sufficient to represent the reflectance behavior of locally smooth surfaces with homogeneous reflectance behavior, they do not reflect characteristic material traits that coin many materials with mesoscopic surface reflectance effects. Such effects take place at surface structures imaged to an area of approximately one pixel within an image. More complex material models such as bidirectional texture functions [8] can deal with such mesoscopic effects but have their limitation w.r.t. extremely specular materials. In that case, their data-driven nature requires an ideally continuous angular sampling which would significantly increase the data masses and therefore is rather impractical. Consequently, the selection of such model-based approaches is material-specific, i.e. the fitting procedures are guided by the appropriate model. In addition, the fitting involves the explicit consideration of a multitude of parameters such as the parameters of the reflectance model, the lighting, etc..

Appearance-based approaches: The key components of appearance-based material classification systems are the extraction of discriminative descriptors that reflect characteristic material traits, an efficient and appropriate modeling for the material categories and an appropriate classifier. The probably most widely used descriptors are filterbanks (e.g. [2,3,6,17,18,25,27]), color patches (e.g. [20,24,26,28,30]), denseSIFT (e.g. [20,24,30]), Local Binary Patterns (LBPs) (e.g. [3,19]), kernel descriptors [15] and combinations of multiple of these descriptors (e.g. [1,15,19,20,24,30]). Combining complimentary descriptors for material classification has been demonstrated to lead to superior results. After extracting such descriptors for the images contained in the training set, the typical approach considers representing the appearance of a material in an individual image based on textons as introduced in e.g. [17] and [18] and also followed in e.g. [19,20,24,25,27,28,30]. The resulting per-image representations can then be classified using nearest neighbor classifiers, Bayesian frameworks [20,27], MRF classifiers [26], SVMs [3,14,19,21,30] or other classifiers such as random forests.

While single-image-based material classification represents the most focused task, some acquisition devices also offer to easily acquire several images under several view-light configurations which might significantly facilitate material classification. The expected result of high performance scores is of great importance if further steps of the acquisition procedure depend on the reflectance behavior of the material classified before. In [17] and [18], histograms have been concatenated to form a combined vector for each particular material, which imposes that materials are represented by a consistently handled ordering of the configurations within the combined vector where all the individual image-representations have to be carefully registered. For comparing the combined vector representations, both the number and the IDs of the view-light configurations of both vectors have to coincide. In [6], bidirectional feature histogram manifolds have been introduced. However, having a sparse set of view-light configurations represents a problem to this approach, as the reference manifolds become coarsely sampled. In addition, linear interpolation between neighboring view-light configurations will result in additional sources of inaccuracies which increase with an increasing distance of the neighboring view-light configurations.

We also aim at classifying material instances using only a few images and make use of results from the face recognition domain. For efficiency, we focus on training-free, linear approaches as presented in [4]. In particular, our material recognition approach yields significantly better recognition rates than previous methods when using smaller numbers of view-light configurations.

3 Methodology

We propose an automatic assistance system for guiding the acquisition process where the respective techniques are selected based on a prior material recognition (see Fig. 2). For a query material, we search its best representative within a database of materials with corresponding annotations about how to acquire the respective type of material. Hence, one core component of our system is represented by a material database which contains images of a multitude of material samples taken under different viewing and lighting conditions which are expected to be met during the acquisition with standard devices analyzed in [23].

For a reliable material recognition, we need to consider the spatial variations of a material as well as its change in appearance induced by different viewing and lighting conditions. Therefore, our approach is based on first computing state-of-the-art descriptors to capture the characteristic material traits and the subsequent derivation of a vector-based representation for each of the given images under individual viewing and lighting conditions (see Subsection 3.1). The set

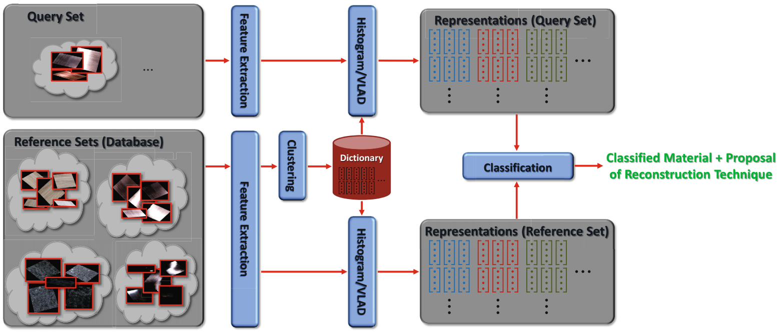

Fig. 2. Overview on the set-based material recognition scheme: After extracting descriptors, we compute a dictionary from the descriptors obtained for the reference data. This dictionary is used to quantize the representation of the content of a particular image into a vector representation. Finally, a set-to-set classification is carried out to find the closest material within the database. The attached annotations for the closest material in the database can be used to guide the subsequent reconstruction process.

of vectors resulting for an individual material sample is then used to obtain its material space. This allows to perform the comparison of different material spaces via set-to-set distances (see Subsection 3.2).

3.1 Representing Materials

In order to be representative for a certain material, material-specific properties have to be included in the set-based representation. Characteristic material traits can be identified in a huge number of different aspects such as color, surface roughness, self-occlusions, interreflections, specularities, and it has been shown beneficial to use several feature descriptors considering different types of attributes (e.g. [20]). We consider the following descriptor types which are densely sampled on a regular grid with a spacing of 5 pixels in our experiments:

- Color: We extract 3×3 color patches as in [20].
- SIFT: For considering the local spatial and directional distribution of image gradients, we extract dense SIFT descriptors as in e.g. [20]. In addition, SIFT descriptors provide robustness to variations in illumination and viewpoint.
- HOG2x2: After computing histograms of oriented gradients, neighboring descriptors are concatenated to a 124-dimensional descriptor as in [31]. As the normalization differs from the scheme used in SIFT, it captures material characteristics in a different way.
- Leung-Malik filters: LM filters [17,18] represent a filterbank with multiple orientations and multiple scales regarding the involved filters. We use 6 orientations and 3 scales.

The extracted descriptors are then used to compute a vector-based representation for each of the masked image regions. In the scope of this paper, we analyze the suitability of the popular bag-of-words representation and the more sophisticated VLAD representation [16]. Based on the descriptors extracted for the images contained in the database, we compute a dictionary of visual words for each descriptor type via k-means clustering. In case of the bag-of-words model, we quantize each descriptor to its closest visual word in the dictionary and form histogram representations. In contrast, the VLAD representation is based on first assigning all the local descriptors $\mathbf{x}_i$ within an image region to their nearest neighbor $\mathbf{c}_j$ with $j = 1, \ldots, k$ in the corresponding dictionary with k visual words for each feature type. Then, the VLAD entries are computed by accumulating the differences of the local descriptors and their assigned visual words. These entries are concatenated to the final VLAD vector, which we normalize to unit length for each of the descriptor types. These vector-based representations for the images of a particular material instance form its corresponding material space. In case of combining several descriptor representations, we simply concatenate the normalized vectors corresponding to the involved descriptor types.

3.2 Set-Based Classification

In contrast to e.g. [6] where a-priori knowledge about the considered viewing and lighting conditions is incorporated for setting up aligned training manifolds, our approach does not rely on the availability of such information. A randomly taken subset of images without knowledge about the imaging parameters should be enough to reliably recognize materials. We use the linear methods presented in [11,32] and [4], where there is no need for parameter learning. Non-linear techniques (e.g. [4]) could be employed as well at the cost of learning hyperparameters such as the kernel width.

Linear convex hull based classifier. Representing the material instances via vector representations for the respective images, we make use of the convex hull classifier presented in [4]. Here, we assume that the vector representations under the available view-light configurations chosen to represent one of the individual material samples can be represented via convex hulls. The distance between convex hulls can be calculated by using quadratic programming and is abbreviated via CHISD (Convex Hull based Image Set Distance) as in [4].

Linear affine hull based classifier. Similar to [4], we consider affine hulls for representing the material spaces. We calculate the linear affine hull parameters by computing an orthonormal basis for the affine subspace spanned by vectors representing a particular material. The distance between two linear affine hulls abbreviated via AHISD (Affine Hull based Image Set Distance) can be computed using the hyperplane which optimally separates the affine hulls.

Mutual subspace method (MSM). This type of method used in [11,32] represents each class with a subspace formed by the respective vectors, and the similarity between subspaces is determined by comparing the angles between the subspaces.

4 Experimental Results

For computing the histogram and VLAD representations respectively, we used dictionary sizes of 150 for color, 250 for SIFT, 250 for HOG2x2 and 200 for the LM filters throughout all of our experiments as used in [20, 24].

In the scope of our experiments, we aim at analyzing the capabilities of the different set-based recognition techniques. We therefore perform experiments on different datasets for varying numbers of view-light configurations in the reference and query sets. We always take disjoint sets of view-light configurations for the reference/query sets of the material samples.

CUReT Database. For obtaining an intuition of the recognition performance, we use the well-established LM filters and denseSIFT for recognizing the 61 CUReT material samples (see Fig. 3, left). Using 5 randomly chosen view-light configurations for representing both reference and query materials, we already obtain high accuracies of around 95.5% for both LM filters and denseSIFT when using AHISD and CHISD with VLAD representations. MSM methods perform worse by about 5%. The benefit of the high-dimensional VLAD representation is obvious in the fact, that histograms perform significantly worse by $4\% - 11\%$. Using more view-light configurations to span the space for the different material samples, we observe that the accuracy obtained when using the individual descriptors closely approaches the 100% already for about 10 view-light configurations in reference and query sets. In general, there is a tendency that the high-dimensional VLAD description gives better accuracies than using histograms. We also combined the descriptors which additionally increases the performance.

In [6], a selection of 20 material instances of the CUReT database has been analyzed. Using 56 images per material instance for their reference manifolds, a performance of about 98% has been reached for classifying individual textures and a bit more than 70% for using 10 configurations per material. For a fair comparison, we only use LM filters as descriptors. Representing the reference sets with 10 randomly drawn view-light configurations and having a single configuration for the query material, we obtain performances of around 95% for the combination of CHISD and VLAD representations which is only slightly worse. In a direct comparison to using 10 configurations for the reference sets, this combination gives an improvement of about 20%. Using more configurations in the query sets, we already reach more than 99% starting from three configurations. DenseSIFT shows a similar performance.

The high performances reached on this database indicate that the individual material samples appear rather distinctive and that the database is not highly challenging. Additionally, as a consequence of the high performances, a real analysis of the different set-based methods w.r.t. each other is hardly possible and the need for set-based recognition is not yet clearly visible. For this reason, more insights can be obtained by using more challenging datasets with higher intra-class variances in material appearance under different view-light configurations.

Fig. 3. Material samples in the CUReT database [8] (taken from [29]) (left) and the ALOT database [12] (right). For illustration purposes, only a subset of the 250 material samples of the ALOT database is shown.

ALOT Database. This database [12] (see Fig. 3, right) offers significantly more and also a wider range of different material types, which have additionally been observed under different illumination colors. In our experiments, we consider color patches, denseSIFT, HOG2x2, LM filters and their combination. Taking 5 view-light configurations for the reference and the query sets results in an accuracy of about 60% for color, 89% for denseSIFT, 83% for LM filters and 83% for HOG2x2 for VLADs with CHISD or AHISD and 4% less for MSM methods. Using histograms also leads to lower performances. The combination of the descriptors, however, leads to about 94% for AHISD and CHISD with VLAD and little lower accuracies for the MSM methods. Taking 10 configurations per reference/query set, the accuracies of the individual descriptors increase, and for the combination of descriptors we reach slightly above 99% using all the methods. This indicates a trend that the reliability of material recognition increases with increasing numbers of view-light configurations for reference and query sets.

Database measured for [30]. While the ALOT database [12] gives a more visible impression on the power of set-based recognition, the samples in this database still do not seem to show too extreme intra-class variations under different view-light configurations in comparison to the inter-class variances. In contrast, the material samples of the database in [30] are used to model the variance in different semantic categories. We used the measurements of the 84 material samples used in [30] and further 76 material samples in the database extension (see Fig. 4). For each of these material samples, photos have been taken under 151 different viewing directions and 151 lighting directions leading to 22,801 images per material sample. For some of the categories, several of the samples only exhibit rather subtle differences (e.g. tiles or metals). This makes the dataset challenging. Instead of grouping these samples into semantic categories as in [30], we consider the measurements per material sample individually and focus on recognizing the material samples. As illustrated in Fig. 5 and Fig. 6, the accuracy again increases for an increasing number of configurations considered in the reference/query sets. As before, we observe the trend of VLADs being more discriminative than histograms. Furthermore, the descriptors have been evaluated separately, where denseSIFTs tend to perform best. The difference to the performance of other descriptors is more visible for the histogram representations. AHISD and CHISD almost consistently outperform the MSM methods. Using the combination of different descriptors results in improvements over the accuracies obtained for the individual descriptors. These improvements are larger, if only a few configurations are available for the reference/query sets.

Fig. 4. Some of the materials measured for the database of [30] and its extension

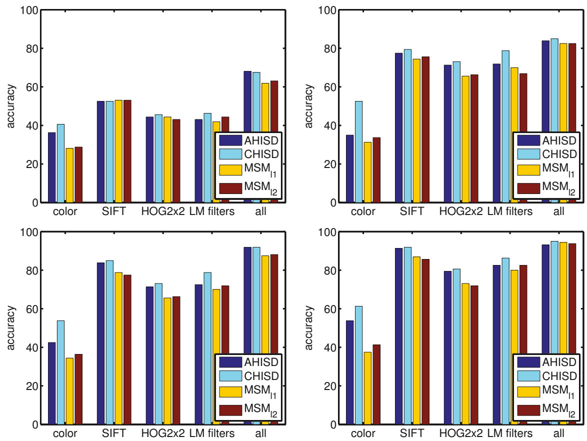

Fig. 5. Accuracies obtained for the data measured in [30] when using histograms and disjoint reference and query sets of 5 (upper left), 10 (upper right), 15 (lower left), 20 (lower right) randomly drawn images of different view-light configurations

In Fig. 7, we illustrate the dependency of the obtained accuracy on the number of view-light configurations in the query sets. We only depict this information for CHISD, which outperformed the other classifiers in the previous experiments. However, we additionally show the performances of using individual descriptors and some combinations of the descriptors. In general, the obtained accuracies increase for taking more images in the query sets and using the VLAD representation leads to accuracies superior to the ones obtained when using histograms. Furthermore, we also analyzed the accuracies obtained for using different combinations of the descriptors. The difference in the obtained accuracies indicates

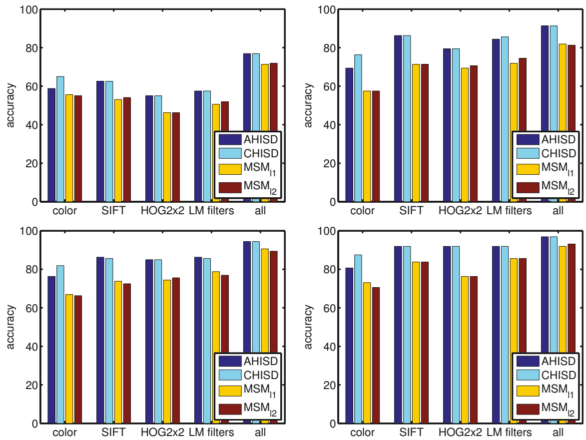

Fig. 6. Accuracies obtained for the data measured in [30] when using VLADs and disjoint reference and query sets of 5 (upper left), 10 (upper right), 15 (lower left), 20 (lower right) randomly drawn images of different view-light configurations

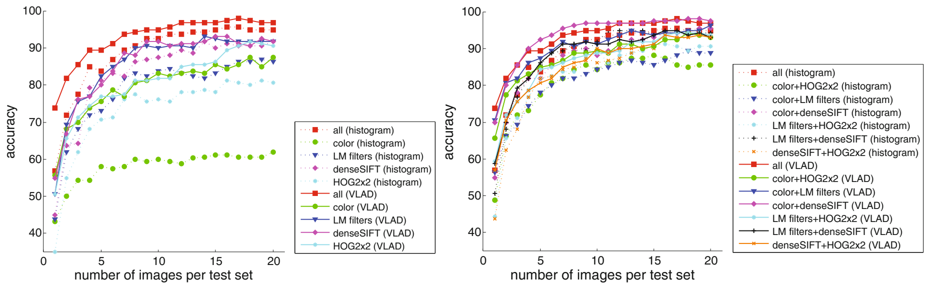

Fig. 7. Accuracies obtained for using sets of 20 view-light combinations for the reference sets and an increasing number of view-light combinations for the query materials (for the data measured in [30]). As expected, the accuracy increases with larger query sets.

that the descriptors carry different amounts of complementary information. In particular, the combination of color and denseSIFT clearly outperforms the remaining combinations of two descriptor types and even slightly outperforms the combination of all four descriptors types. Additionally, it becomes apparent that considering multiple view-light configurations leads to significant performance gains of almost 20% for using 10 configurations for the query sets in comparison to using a single configuration for the query sets when considering the combinations of descriptors. For more view-light configurations in the query set, we observe rather marginal improvements in the accuracies. When analyzing the few misclassified material samples (e.g. two of the tiles and two of the metals have not been properly distinguished) and the respective estimated material labels, we observed that the estimated material and the ground truth material indeed look rather similar and it is even hard to distinguish them as a human. In turn, this means that we still can take the stored parameters for a subsequent acquisition or reconstruction respectively due to the similarity of the materials. As a result, we also obtain a highly reliable recommendation regarding adequate acquisition and reconstruction methods.

5 Conclusions

In this paper, we have presented a study on using set-based recognition schemes in combination with standard descriptors and encodings for material recognition. Our studies demonstrate the benefit of making use of several images of a material sample for different view-light conditions regarding material recognition. There are only little performance gains possible for databases with smaller intra-class variance before reaching the saturation close to 100%, which might have lead to less interest in investigations on material recognition based on several view-light configurations in recent years. However, when considering more challenging databases with larger intra-class variances under different view-light configurations, it is still essential to provide a reliable material recognition with regard to an efficient acquisition relying on a correct recommendation of the acquisition procedure to be used. We have shown that such a material recognition can be achieved with a high reliability by looking at the characteristic material appearance under a few view-light configurations which emphasizes the significant benefit of set-based material recognition in the presence of larger variations in appearance of the individual samples.

Acknowledgments. The research leading to these results was funded by the European Commission's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 323567 (Harvest4D), 2013-2016.

References

1. Burghouts, G.J., Geusebroek, J.M.: Color textons for texture recognition. In: BMVC, pp. 1099–1108 (2006)
2. Caputo, B., Hayman, E., Fritz, M., Eklundh, J.O.: Classifying materials in the real world. Image Vision Comput. **28**(1), 150–163 (2010)
3. Caputo, B., Hayman, E., Mallikarjuna, P.: Class-specific material categorisation. In: ICCV, vol. 2, pp. 1597–1604 (2005)
4. Cevikalp, H., Triggs, B.: Face recognition based on image sets. In: CVPR, pp. 2567–2573 (2010)
5. Chantler, M., McGunnigle, G., Penirschke, A., Petrou, M.: Estimating lighting direction and classifying textures. In: BMVC, pp. 72.1–72.10 (2002)
6. Cula, O.G., Dana, K.J.: 3d texture recognition using bidirectional feature histograms. IJCV **59**(1), 33–60 (2004)
7. Dana, K.J., van Ginneken, B., Nayar, S.K., Koenderink, J.J.: Reflectance and texture of real world surfaces. Tech. rep. (1996)
8. Dana, K.J., Nayar, S.K., Ginneken, B.V., Koenderink, J.J.: Reflectance and texture of real-world surfaces. In: CVPR, pp. 151–157 (1997)

9. Dana, K.J., Nayar, S.K.: Histogram model for 3d textures. In: CVPR, pp. 618–624 (1998)
10. Drbohlav, O., Chantler, M.: Illumination-invariant texture classification using single training images. In: Texture 2005: Proceedings of the International Workshop on Texture Analysis and Synthesis, pp. 31–36 (2005)
11. Fukui, K., Yamaguchi, O.: Face recognition using multi-viewpoint patterns for robot vision. In: ISRR, pp. 192–201 (2003)
12. Geusebroek, J.M., Smeulders, A.W.M.: Amsterdam library of textures (ALOT) (June 2014). http://aloi.science.uva.nl/public_alot/
13. van Ginneken, B., Koenderink, J.J., Dana, K.J.: Texture histograms as a function of irradiation and viewing direction. IJCV **31**(2–3), 169–184 (1999)
14. Hayman, E., Caputo, B., Fritz, M., Eklundh, J.-O.: On the significance of real-world conditions for material classification. In: Pajdla, T., Matas, J.G. (eds.) ECCV 2004. LNCS, vol. 3024, pp. 253–266. Springer, Heidelberg (2004)
15. Hu, D., Bo, L., Ren, X.: Toward robust material recognition for everyday objects. In: BMVC, pp. 1–11 (2011)
16. Jegou, H., Douze, M., Schmid, C., Pérez, P.: Aggregating local descriptors into a compact image representation. In: CVPR, pp. 3304–3311 (2010)
17. Leung, T., Malik, J.: Recognizing surfaces using three-dimensional textons. In: ICCV, vol. 2, pp. 1010–1017 (1999)
18. Leung, T., Malik, J.: Representing and recognizing the visual appearance of materials using three-dimensional textons. IJCV **43**(1), 29–44 (2001)
19. Li, W., Fritz, M.: Recognizing materials from virtual examples. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (eds.) ECCV 2012, Part IV. LNCS, vol. 7575, pp. 345–358. Springer, Heidelberg (2012)
20. Liu, C., Sharan, L., Adelson, E.H., Rosenholtz, R.: Exploring features in a bayesian framework for material recognition. In: CVPR, pp. 239–246 (2010)
21. Liu, C., Yang, G., Gu, J.: Learning discriminative illumination and filters for raw material classification with optimal projections of bidirectional texture functions. In: CVPR, pp. 1430–1437 (2013)
22. Osadchy, M., Jacobs, D.W., Ramamoorthi, R.: Using specularities for recognition. In: ICCV, pp. 1512–1519 (2003)
23. Schwartz, C., Sarlette, R., Weinmann, M., Rump, M., Klein, R.: Design and implementation of practical bidirectional texture function measurement devices focusing on the developments at the University of Bonn. Sensors **14**(5), 7753–7819 (2014)
24. Sharan, L., Liu, C., Rosenholtz, R., Adelson, E.H.: Recognizing materials using perceptually inspired features. IJCV **103**(3), 348–371 (2013)
25. Varma, M., Zisserman, A.: Classifying images of materials: achieving viewpoint and illumination independence. In: Heyden, A., Sparr, G., Nielsen, M., Johansen, P. (eds.) ECCV 2002, Part III. LNCS, vol. 2352, pp. 255–271. Springer, Heidelberg (2002)
26. Varma, M., Zisserman, A.: Texture classification: are filter banks necessary?. In: CVPR, vol. 2, pp. 691–698 (2003)
27. Varma, M., Zisserman, A.: Unifying statistical texture classification frameworks. Image and Vision Computing **22**(14), 1175–1183 (2004)
28. Varma, M., Zisserman, A.: A statistical approach to material classification using image patch exemplars. PAMI **31**(11), 2032–2047 (2009)
29. Visual Geometry Group (University of Oxford): Texture classification (June 2014). http://www.robots.ox.ac.uk/~vgg/research/texclass/setup.html

30. Weinmann, M., Gall, J., Klein, R.: Material classification based on training data synthesized using a BTF database. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds.) ECCV 2014, Part III. LNCS, vol. 8691, pp. 156–171. Springer, Heidelberg (2014)
31. Xiao, J., Hays, J., Ehinger, K.A., Oliva, A., Torralba, A.: Sun database: Large-scale scene recognition from abbey to zoo. In: CVPR, pp. 3485–3492 (2010)
32. Yamaguchi, O., Fukui, K., Maeda, K.: Face recognition using temporal image sequence. In: Int. Conf. on Automatic Face and Gesture Recognition, pp. 318–323 (1998)