

Fine-grained material classification using micro-geometry and reflectance

Christos Kampouris^{1,2}, Stefanos Zafeiriou¹, Abhijeet Ghosh¹ and Sotiris Malassiotis²

{c.kampouris12, s.zafeiriou, ghosh}@imperial.ac.uk, malasiot@iti.gr

¹Department of Computing, Imperial College London, UK

²Centre for Research and Technology Hellas (CERTH), Thessaloniki, Greece

Abstract. In this paper we focus on an understudied computer vision problem, particularly how the micro-geometry and the reflectance of a surface can be used to infer its material. To this end, we introduce a new, publicly available database for fine-grained material classification, consisting of over 2000 surfaces of fabrics¹. The database has been collected using a custom-made portable but cheap and easy to assemble photometric stereo sensor. We use the normal map and the albedo of each surface to recognize its material via the use of handcrafted and learned features and various feature encodings. We also perform garment classification using the same approach. We show that the fusion of normals and albedo information outperforms standard methods which rely only on the use of texture information. Our methodologies, both for data collection, as well as for material classification can be applied easily to many real-world scenarios including design of new robots able to sense materials and industrial inspection.

Keywords: material classification, micro-geometry, reflectance, photometric stereo

1 Introduction

Materials recognition from their visual appearance is an important problem in computer vision having numerous applications spanning from scene analysis to robotics and industrial inspection. Knowledge of the material of an object or surface can provide useful information about its properties. Metallic objects are usually rigid while fabric is deformable and glass is fragile. Humans use daily this kind of information when they interact with the physical world. For example, they are more careful when they hold a glass bottle than a plastic one and they avoid walking on ice or other slippery surfaces. It would be useful in many applications, such as robotic manipulation or autonomous navigation, if a computer vision system had the same ability.

Due to the limited amount of data available the problem of material recognition/classification was not among the most popular computer vision areas [1–7].

¹ <http://ibug.doc.ic.ac.uk/resources/fabrics>

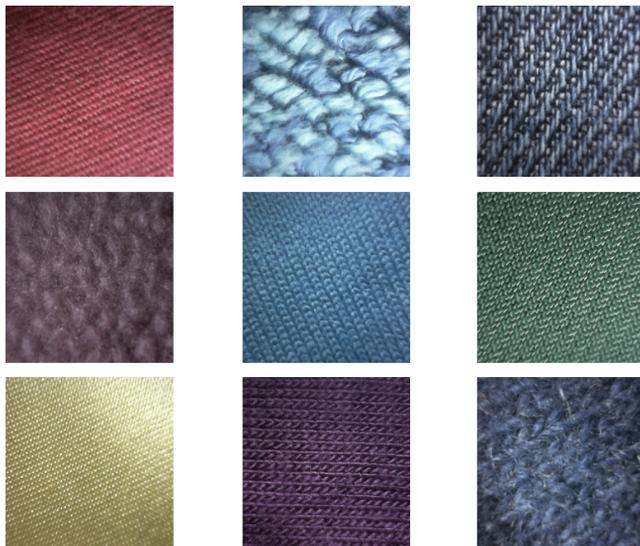


Fig. 1. Samples from our database. From left to right, top to bottom: cotton, terrycloth, denim, fleece, nylon, polyester, silk, viscose, and wool.

But with the advent of the Internet and crowd-source annotation, collection and annotation of large databases of materials and textures was made feasible, which in turn brought renewed attention to the problem of material classification [8–12]. Currently the trend is to use materials captured in unconstrained conditions, also referred to as “in-the-wild”, and mainly in the large scale [10, 11, 9, 8]. In this paper, we take a different direction which has, to the best of our knowledge, has received limited attention. We study the problem of fine-grained material, and particularly fabric, classification in the fine scale introducing the first large database suitable for the task.

Recognizing materials from images is a challenging task because the appearance of a surface depends on a variety of factors such as its shape, its reflectance properties, illumination, viewing direction etc. Material recognition is usually treated as texture classification. Statistical learning techniques have been employed [1–3] on databases of different materials taken under multiple illumination and viewing directions [13–15]. Despite the very high classification rates achieved this way [3], the introduction of data captured “in-the-wild” have recently showed that the problem is extremely challenging in real world conditions.

On the other hand, humans do not rely only on vision to recognize materials. The sense of touch can be really useful to discriminate materials with subtle differences. For example, when someone wants to find out the material a garment is made of, they usually rub it to assess its roughness. This gives us the hint that the micro-geometry of the garment surface provides useful information for its material.

In this paper, we study how the micro-geometry and the reflectance of a surface can be used to infer its material. To recover the micro-geometry we use photometric stereo, which gives us the normal map and the albedo of the surface by capturing images under different illumination conditions. An advantage of photometric stereo over other techniques, like binocular stereo, is that there is no correspondence problem between images. All images are captured under the same view point and only the illumination is different. That makes photometric stereo a fast and computational efficient technique that suits well the time constraints of a real-world application.

To be able to apply photometric stereo in unconstrained environments, we built a portable photometric stereo sensor using low cost, off-the-shelf components. The sensor recovers the normal map and the albedo of various surfaces fast and accurately and has the ability to capture fine micro-structure details, as we can see in fig. 1. Using the sensor we collected surface patches of over 2000 garments. Knowing the fabric of a garment can help robots manipulate clothes more robustly and facilitate the automation of household chores like laundry [16] and ironing.

In contrast to most datasets [17, 18] where fabrics are treated as one of the material classes, we investigate the problem of classifying 9 different fabric classes - cotton, denim, fleece, nylon, polyester, silk, terrycloth, viscose, and wool. This fine-grained classification problem is much more challenging due to the large intra-class and small inter-class variance that fabrics have. We perform a thorough evaluation of various feature encoding methodologies such as Bag of Visual Words (BoVW) [1], Pyramid Histogram Of visual Words (PHOW) [19], Fisher Vectors [20, 21], Vector of Locally-Aggregated Descriptors (VLAD) [22] using both SIFT [23], as well as features from Convolutional Neural Networks (CNNs) [24]. We show that using both normals and albedo the classification accuracy is improved. In summary, our main contributions are:

- The first large, publicly available database of garment surfaces captured under different illumination conditions.
- An evaluation of texture features and encodings which shows that the micro-geometry and reflectance of a fabric can be used to classify its material more accurately than by using its texture.

2 Related Work

At first, texture classification was implemented on 2-D patches from image databases such as Brodatz textures [25]. One of the first efforts to collect a large database suitable for studying appearance of real-world surfaces was made in [13]. The so-called CURET database contains image textures from around 60 different materials, each observed with over 200 different combinations of viewing and illumination directions. Leung and Malik [1] used the CURET database for material classification, introducing the concept of 3D textons. They used a filter bank to obtain filter responses for each class of the CURET database and they clustered them to obtain the 3-D textons. Then, they created a model for every

training image by building its texton frequency histogram and they classified each test image by comparing its model with the learned models, introducing what is now known as the BoVW representation. In [1] a set of 48 filters were used for feature extraction (first and second derivatives of Gaussians at 6 orientations and 3 scales making a total of 36, 8 Laplacian of Gaussian filters and 4 Gaussians). A similar approach was followed in [2] using the so-called MR8 filter. The MR8 filter bank consists of 38 filters but only 8 filter responses. To achieve rotation invariance the filter bank contains filters (Gaussian and a Laplacian of Gaussian) at multiple orientations but only the maximum filter response across all orientations is used. Later Varma and Zisserman [3] questioned about the necessity of filter banks by succeeding comparable results using image patch exemplars as small as 3×3 instead of a filter bank. The KTH-TIPS (Textures under varying Illumination, Pose and Scale) image database was created to extend the CURET database by providing variations in scale [4]. Despite the almost perfect classification rate in CURET database the results of the same algorithms in KTH-TIPS showed that the performance drops significantly [5].

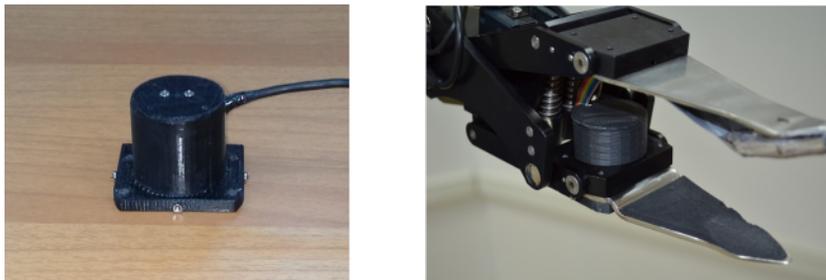


Fig. 2. The photometric stereo sensor on a table (left) and installed in the gripper of an industrial robot (right)

Sharan et al. created the more challenging material database FMD [18] with images taken from the Flickr website and captured under unknown real-world conditions (10 materials with 100 samples per material). Liu et al. [26] used the FMD and a Bayesian framework to achieve 45% classification rate while Hu et al. [27] by using the Kernel Descriptor framework improved the accuracy to 54%. Sharan et al. [7] used perceptually inspired features to further improve the results to 57%.

Recently, in [10] the authors presented the first database with textures “in-the-wild” with various attributes (including the material of the texture) and in [9] a very large-scale, open dataset of materials in the wild, the so-called Materials in Context Database (MINC) was proposed [9]. Another database (UBO2014) that appeared recently for synthesising training images for material classification appeared in [28] (7 material categories, each consisting of measurements of 12 different material samples, measured in a darkened lab environment with con-

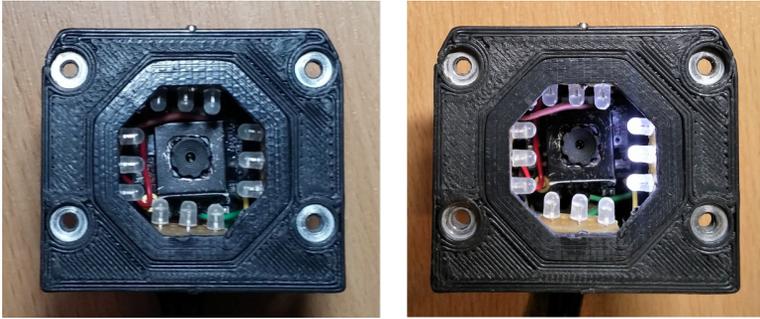


Fig. 3. Bottom view of our photometric stereo sensor having all lights off (left) and with one light on (right)

trolled illumination). The above databases are suitable for training and testing generic material classification methodologies from appearance information.

In this paper we study the problem of fine-grained material classification using both the reflectance properties (albedo) and the micro-geometry (surface normals) of a surface. The most closely related work on fusing albedo and normals for material classification is the work on [29] which proposes a rotational invariant classification method. The authors applied successfully their method on real textures but they used only 30 samples belonging to various material classes (fabric, gravel, wood). In contrast, our database contains around 2000 samples and is suitable for fine-grained classification.

Regarding recognition of materials using micro-geometry the only relevant published technology is the tactile GelSight sensor [30, 31]. In [30] the authors used the Gelsight sensor to collect a database of 40 material classes with 6 samples per class (the database is not publicly available). The Gelsight sensor is able to provide very high resolution height maps but cannot recover albedo [31] (contrary to our sensor which provides both micro-geometry and albedo).

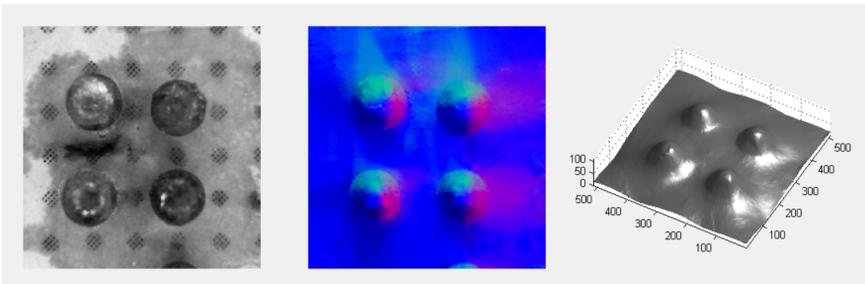


Fig. 4. Albedo, normals, and 3-D shape of our calibration target

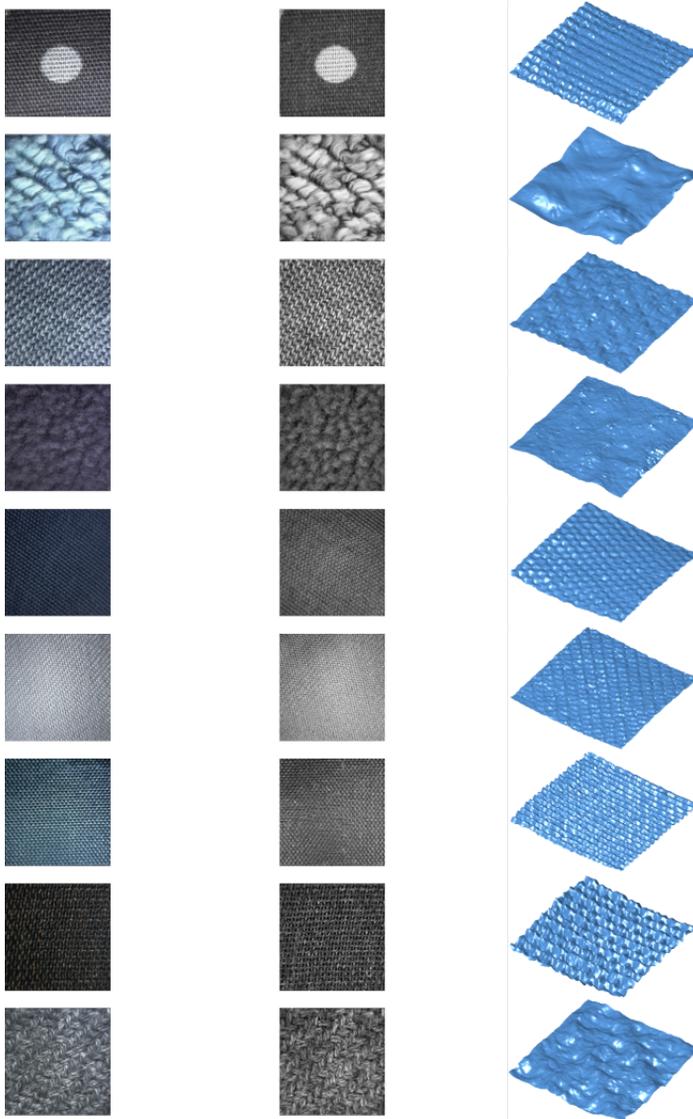


Fig. 5. The 9 main fabric classes with albedo and 3D shape

3 Photometric Stereo Sensor

Photometric stereo [32] is a computer vision method that uses three or more images of a surface from the same viewpoint but under different illumination conditions to estimate the normal vector and albedo at each point. By integrating

the normal map the 3-D shape of the surface can be acquired. Photometric stereo has found application in a wide variety of fields, ranging from medicine [33] to security [34] and from robotics [35] to archaeology [36].

We used this method to acquire the micro-geometry and the reflectance properties of garment surfaces. In order to capture a large number of garments, we built a small-sized, portable photometric stereo sensor using low-cost, off-the-shelf components, fig. 2. The sensor consists of a webcam with adjustable focus, white light emitting diode (LED) arrays and a micro-controller to synchronize the LED arrays with the camera. It features an easy to use USB interface that connects directly to a computer with a standard USB port. All the components are enclosed in a 3-D printed cylindrical case with a diameter of 38mm and height of 40mm . The case has also a rectangular base that allows the sensor to be mounted in the gripper of a robot, fig. 2 (right).

The sensor captures 4 RGB images of a small patch of the inspected surface, one for each light source, in less than a second. The working distance of the camera is 3 cm and to achieve focus we had to modify appropriately the camera optics. The sensor resolution is 640×480 pixels but we crop all images to 400×400 pixels to avoid out-of-focus regions at the corners due to the very shallow depth of field. The corresponding field of view is approximately $10\text{mm} \times 10\text{mm}$, adequate to capture the micro-structure of most fabrics.

We use near-grazing illumination to maximize contrast and eliminate specularities. Because the light sources are close to the inspected surface the lighting field is not uniform. To confront this issue, we have captured with our sensor a flat surface with known albedo and use these images to normalize every new surface we capture.

As can be seen in fig. 3, we use LED arrays to illuminate uniformly the inspected surface. However in such a close distance the illumination is not directional and we have to know the light vectors in each position to apply photometric stereo. To do so, we used a chrome sphere (2 mm diameter) in different locations to calibrate our lights and by interpolation we found the light vectors for each pixel. Next, we run photometric stereo to get the albedo and the normal map using the method of Barsky and Petrou [37] to deal with shadows and highlights. At the end we integrate the normals to obtain the 3D shape of the surface [38].

We evaluated the accuracy of our sensor by using a calibration target, consisting of a 2×2 grid of spheres of diameter of 2 mm. As we can see in figure 4 the sensor can accurately reconstruct the shape of the spheres. Since our goal is to capture the geometry of a surface for classification, we didn't further assess the reconstruction accuracy.

4 Dataset

Using the photometric stereo sensor we collected samples of the surface of over 2000 garments and fabrics. We visited many clothes shops with the sensor and a laptop and captured all our images "in the field". For every garment we kept

information (attributes) about its material composition from the manufacturer label and its type (pants, shirt, skirt etc.). The dataset reflects the distribution of fabrics in real world, hence it is not balanced. The majority of clothes are made of specific fabrics, such as cotton and polyester, while some other fabrics, such as silk and linen, are more rare. Also, a large number of clothes are not composed of a single fabric but two or more fabrics are used to give the garment the desired properties. Fig. 5 presents samples for major classes, along with its albedo and its 3-D micro-geometry.

4.1 Data collection

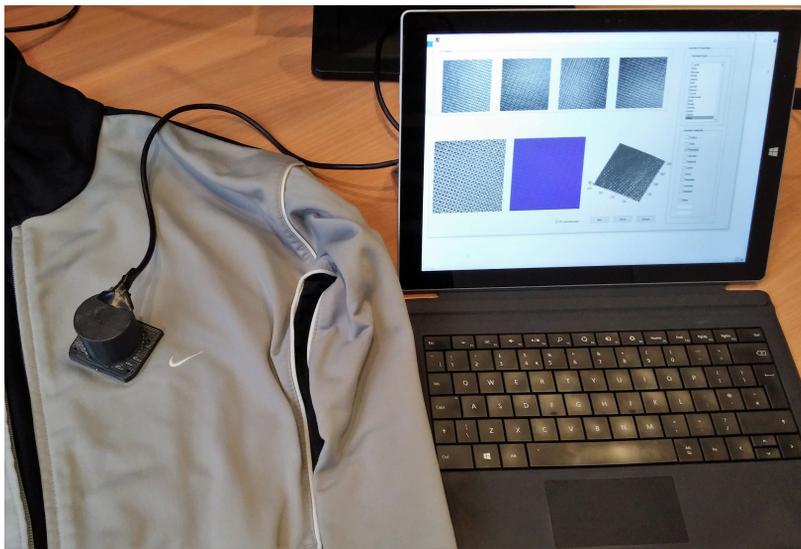


Fig. 6. Capturing setup

The procedure we followed to capture our fabric dataset is depicted in Fig. 6. The photometric stereo sensor is placed on a flat region of the garment, equal to the field of view ($10\text{mm} \times 10\text{mm}$). Initially the sensor turns on one light source to allow the camera to focus on the surface of the garment. Due to the shallow depth of field, this task is very important. Subsequently we capture 4 images, one for each light source. We normalize the captured images using reference images of a flat surface with known albedo.

Fig. 6 shows the graphical user interface we use to collect the data. In the top row the 4 captured images are shown, while in the bottom row we visualize the albedo, the normal map (as an RGB image) and the 3-D shape of the surface. On the right side we can select the fabric of the garment (cotton, wool, etc.) as well as its type (t-shirt, pants, etc.).

5 Experiments

In this section we describe the experiments we run on our fabric database for fine-grained material and garment classification. First of all, a very important aspect of every classification system is the selection of features to use. In our case, we want to use features that distinguish fabric classes from one another. We are mostly interested in features that represent the micro-geometry and reflectance of the surface as they are encoded in normals and albedo respectively. To this end we have followed similar steps as [10]. We have used handcrafted, i.e dense SIFT [23] features ², as well as features coming from Deep Convolutional Neural Networks (CNN) [40, 24]. Deep learning and especially CNNs dominate the field of object and speech recognition. CNNs contain a hierarchical model where we have a succession of convolutions and non-linearities. CNNs can learn very complex relationships, given that exist a large amount of data. In our case we used the VGG-M [40] pre-trained neural network which has 8 layers, 5 convolutional layers plus 3 fully connected. We have applied the non-linear features (dense SIFT and CNNs) to (a) albedo images, (b) the original images under different illuminants, (c) the normals (concatenate the three components, N_x, N_y, N_z , into a single image) and (d) fusing normals and albedo (by creating a larger image consisting of the albedo plus the three components of the normals). The following feature encodings have been applied to the local (dense SIFT) features:

- Feature encoding via the use of BoVW. BoVW uses Vector Quantisation (VQ) [41] of local features (i.e., responses of linear or non-linear filters) by mapping them to the closest visual word in a dictionary. The visual words (vocabulary) are prototypical features, i.e. kind of centroids, and are computed by applying a clustering methodology (e.g., K-means [42]). The BoVW encoding is a vector of occurrence counts the above vocabulary of visual words. BoVW was first proposed in [1] for the purpose of material classification.
- Another popular orderless encoding is the so-called VLAD [22]. VLAD, similar to BoVW, applies VQ to a collection of local features. It differs from the BoVW image descriptor by recording the difference from the cluster center, rather than the number of local features assigned to the cluster. VLAD accumulates first-order descriptor statistics instead of simple occurrences as in BoVW.
- A feature encoding that accumulated first and second order statistics of local features is the FV [21]. FV uses a soft clustering assignment by using a Gaussian Mixture Model (GMM) [43] instead of K-means. The soft-assignments (posterior probability of each GMM component) is used to weight first and second order statistics of local features.

For the case of CNNs features we have used them directly in the classifier or we used FV encoding before performing classification [11]. The classifier we used was

² We have tried other features such as Local Binary Patterns (LBPs) [39] and the MR8 filters in [2] but the results were much poorer than dense SIFT, hence are not reported.

a linear 1 versus "all" Support Vector Machine (SVM) [44]. We have performed four fold cross-validation (i.e., using 75% for training and 25% for testing in each fold).

In all experiments, we used the VLFeat library [45] to extract the SIFT features and to compute the feature encodings (BoVW, VLAD, FV). For the CNNs we used the MatConvNet library [46].

5.1 Results

For our classification experiments we didn't use garments with blended fabrics. We kept only those whose composition is 95% of one material. We also discarded fabric classes with too few samples. So, for our experiments we have chosen a subset of 1266 samples which belong to one of the following fabric classes: cotton, terrycloth, denim, fleece, nylon, polyester, silk, viscose, and wool. The number of samples in each class is shown in table 1.

| Cotton | Terrycloth | Denim | Fleece | Nylon | Polyester | Silk | Viscose | Wool |
|--------|------------|-------|--------|-------|-----------|------|---------|------|
| 588 | 30 | 162 | 33 | 50 | 226 | 50 | 37 | 90 |

Table 1. Number of samples per class for fabric classification

Table 2 summarizes the results of our experiments for fabric classification. As we can see, the combination of albedo and normals gives constantly slightly better accuracy. This is true not only for the hand-made features but also for the learned features of CNNs. This provides evidence that using both geometry and texture is useful for fine grained material classification problems.

| Modality | SIFT | | | CNN - VGG-M | | |
|----------|------|------|------|-------------|------|-------|
| | FV | VLAD | BoVW | FV | FC | FC+FV |
| Images | 65.3 | 58.9 | 58.4 | 66.5 | 44.0 | 71.7 |
| Albedo | 63.5 | 56.5 | 56.6 | 67.9 | 47.6 | 71.2 |
| Normals | 72.5 | 57.2 | 54.3 | 73.9 | 41.9 | 74.1 |
| Alb+Norm | 74.7 | 61.7 | 60.1 | 76.1 | 50.5 | 79.6 |

Table 2. Results for fabric classification

We also run experiments for garment recognition to investigate whether the garment type can be estimated by the microgeometry or albedo. Here we split our dataset in 6 classes: blouses, jackets, jeans, pants, shirts, and t-shirts. Table 3 shows the number of samples per garment class.

We used the same approach with fabric classification testing SIFT features in different encodings as well as CNN features. The results are presented in table

| Blouses | Jackets | Jeans | Pants | Shirts | T-shirts |
|---------|---------|-------|-------|--------|----------|
| 294 | 146 | 139 | 205 | 159 | 261 |

Table 3. Number of samples per class for garment classification

4. It is obvious that this problem is a more difficult and requires more research but also in this case the fusion of albedo and normals gives the best results.

| Modality | SIFT | | | CNN - VGG-M | | |
|----------|------|------|------|-------------|------|-------|
| | FV | VLAD | BoVW | FV | FC | FC+FV |
| Images | 50.1 | 48.3 | 44.7 | 55.7 | 40.0 | 61.2 |
| Albedo | 49.5 | 49.7 | 46.6 | 57.3 | 40.6 | 63.4 |
| Normals | 54.9 | 51.3 | 53.3 | 54.2 | 41.9 | 64.3 |
| Alb+Norm | 55.6 | 51.6 | 52.8 | 57.1 | 50.5 | 64.6 |

Table 4. Results for garment classification

Fig. 7 (left) presents the confusion matrix for the combination of albedo and normals using fully-connected deep convolutional features from the VGG-M model and FV pooling. This model achieved 79.6% average accuracy, surpassing both dense SIFT features with FV pooling and CNN with FV pooling (74.7% and 76.1% respectively). As we can see cotton, terrycloth, denim and fleece can be recognized successfully with almost no errors. On the other hand, classes like nylon and viscose are more difficult to classify due to their glossy appearance and the small number of samples in our dataset.

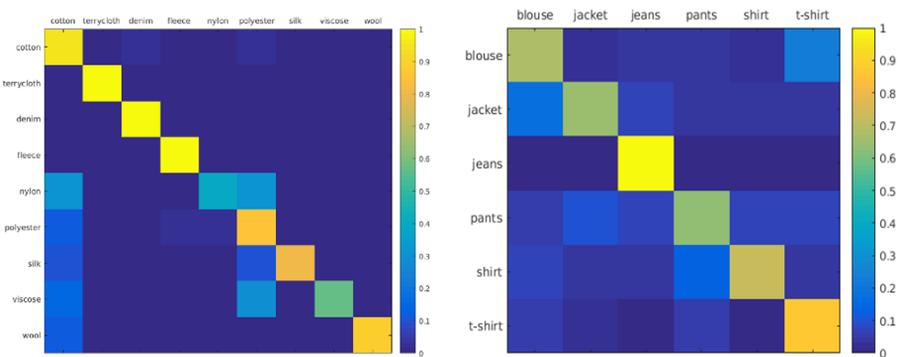


Fig. 7. Confusion matrix of the best fabric (left) and the best garment (right) classifier

The same combination of local features and pooling encoder (FC-CNN + FV) gives also the best result for the problem of garment classification. Here the accuracy is 64.6%, while using dense SIFT we get 55.6% and using CNN + FV, 57.1%. As we can see by observing the confusion matrix in fig. 7 (right) only the classification rate of jeans is high. This is expected since jeans are made by denim fabric and our system can recognize denim accurately (fig. 7 (left)). On the contrary, all other garments can be made of various fabrics and there is no one-to-one correspondence between fabric and garment class.

6 Conclusion

Material classification is important but challenging problem for computer vision. This work focuses on recognizing the material of clothes. Inspired by the fact that humans use their sense of touch to infer the fabric of a garment, we investigated the use of the micro-geometry of garments to recognize their materials. We built a portable photometric stereo sensor and introduced the first large scale dataset of garment surfaces, belonging to 9 different fabric classes. We tested a number of different features for material classification and showed that using both the micro-geometry and the reflectance properties of a fabric can improve the classification rate. Although there is room for improvement, our system is efficient and practical solution for material classification that can be applied in many different scenarios in robotics and industrial inspection.

7 Acknowledgments

This work has been inspired and initiated by the late Maria Petrou, professor at the department of Electrical and Electronic Engineering, Imperial College London and director of the Informatics Institute of the Centre for Research and Technology Hellas.

The work of Sotiris Malassiotis has been funded by the European Community 7th Framework Programme [FP7/2007-2013] under grant agreement FP7-288553 (CloPeMa). The database collection has taken place while Christos Kampouris was with CERTH funded by CloPeMA. Experiments were conducted and the paper was written while Christos Kampouris was with Imperial College partially funded by the EPSRC project EP/N006259/1 (Computational Imaging and Analysis of Scene Appearance). The work of Abhijeet Ghosh has been funded by the EPSRC Early Career Fellowship EP/N006259/1 and the Royal Society Wolfson Research Merit Award.

References

1. Leung, T., Malik, J.: Representing and recognizing the visual appearance of materials using three-dimensional textons. *International Journal of Computer Vision* **43**(1) (2001) 29–44

2. Varma, M., Zisserman, A.: A Statistical Approach to Texture Classification from Single Images. *International Journal of Computer Vision* **62**(1) (2005) 61–81
3. Varma, M., Zisserman, A.: A statistical approach to material classification using image patch exemplars. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **31**(11) (2009) 2032–47
4. Caputo, B., Hayman, E., Mallikarjuna, P.: Class-specific material categorisation. In: *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on. Volume 2., IEEE* (2005) 1597–1604
5. Caputo, B., Hayman, E., Fritz, M., Eklundh, J.O.: Classifying materials in the real world. *Image and Vision Computing* **28**(1) (2010) 150–163
6. Varma, M., Zisserman, A.: Classifying images of materials: Achieving viewpoint and illumination independence. In: *Computer Vision/ECCV 2002. Springer* (2002) 255–271
7. Sharan, L., Liu, C., Rosenholtz, R., Adelson, E.: Recognizing materials using perceptually inspired features. *International Journal of Computer Vision* **103**(3) (2013) 348–371
8. Cimpoi, M., Maji, S., Kokkinos, I., Mohamed, S., Vedaldi, A.: Describing textures in the wild. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.* (2014) 3606–3613
9. Bell, S., Upchurch, P., Snavely, N., Bala, K.: Material recognition in the wild with the materials in context database. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.* (2015) 3479–3487
10. Cimpoi, M., Maji, S., Kokkinos, I., Vedaldi, A.: Deep filter banks for texture recognition, description, and segmentation. *International Journal of Computer Vision* (2015) 1–30
11. Cimpoi, M., Maji, S., Vedaldi, A.: Deep filter banks for texture recognition and segmentation. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.* (2015) 3828–3836
12. Bell, S., Upchurch, P., Snavely, N., Bala, K.: Opensurfaces: A richly annotated catalog of surface appearance. *ACM Transactions on Graphics (TOG)* **32**(4) (2013) 111
13. Dana, K., Van-Ginneken, B., Nayar, S., Koenderink, J.: Reflectance and Texture of Real World Surfaces. *ACM Transactions on Graphics (TOG)* **18**(1) (Jan 1999) 1–34
14. ALOT: www.science.uva.nl/~mark/alot
15. KTH-TIPS2: www.nada.kth.se/cvap/databases/kth-tips
16. Miller, S., van den Berg, J., Fritz, M., Darrell, T., Goldberg, K., Abbeel, P.: A geometric approach to robotic laundry folding. *The International Journal of Robotics Research* **31**(2) (2012) 249–267
17. Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Fei-Fei, L.: Imagenet: A large-scale hierarchical image database. In: *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on.* (June 2009) 248–255
18. Sharan, L., Rosenholtz, R., Adelson, E.: Material perception: What can you see in a brief glance? *Journal of Vision* **9**(8) (2009) 784
19. Bosch, A., Zisserman, A., Munoz, X.: Image classification using random forests and ferns. In: *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on, IEEE* (2007) 1–8
20. Jaakkola, T.S., Haussler, D., et al.: Exploiting generative models in discriminative classifiers. *Advances in neural information processing systems* (1999) 487–493

21. Perronnin, F., Dance, C.: Fisher kernels on visual vocabularies for image categorization. In: *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, IEEE (2007) 1–8
22. Jégou, H., Douze, M., Schmid, C., Pérez, P.: Aggregating local descriptors into a compact image representation. In: *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, IEEE (2010) 3304–3311
23. Lowe, D.: Object recognition from local scale-invariant features. In: *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*. Volume 2. (1999) 1150–1157 vol.2
24. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556* (2014)
25. Brodatz, P.: *Textures*. Dover Publications (1966)
26. Liu, C., Sharan, L., Adelson, E., Rosenholtz, R.: Exploring features in a bayesian framework for material recognition. In: *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*. (2010) 239–246
27. Hu, D., Bo, L., Ren, X.: Toward Robust Material Recognition for Everyday Objects. *Proceedings of the British Machine Vision Conference 2011* (2011) 48.1–48.11
28. Weinmann, M., Gall, J., Klein, R.: Material classification based on training data synthesized using a btf database. In: *Computer Vision–ECCV 2014*. Springer (2014) 156–171
29. Wu, J., Chantler, M.J.: Combining gradient and albedo data for rotation invariant classification of 3d surface texture. In: *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*, IEEE (2003) 848–855
30. Li, R., Adelson, E.: Sensing and recognizing surface textures using a gelsight sensor. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. (2013) 1241–1247
31. Johnson, M.K., Cole, F., Raj, A., Adelson, E.H.: Microgeometry capture using an elastomeric sensor. In: *ACM Transactions on Graphics (TOG)*. Volume 30., ACM (2011) 46
32. Woodham, R.: Photometric method for determining surface orientation from multiple images. *Optical Engineering* (1980)
33. Smith, L., Smith, M., a.R. Farooq, Sun, J., Ding, Y., Warr, R.: Machine vision 3D skin texture analysis for detection of melanoma. *Sensor Review* **31**(2) (2011) 111–119
34. Zafeiriou, S., Atkinson, G., Hansen, M., Smith, W., Argyriou, V., Petrou, M., Smith, M., Smith, L.: Face recognition and verification using photometric stereo: The photoface database and a comprehensive evaluation. *Information Forensics and Security, IEEE Transactions on* **8**(1) (2013) 121–135
35. Ikeuchi, K., Nishihara, H.K., Horn, B.K., Sobalvarro, P., Nagata, S.: Determining grasp configurations using photometric stereo and the prism binocular stereo system. *The International Journal of Robotics Research* **5**(1) (1986) 46 – 65
36. Einarsson, P., Hawkins, T., Debevec, P.: Photometric stereo for archeological inscriptions. In: *ACM SIGGRAPH 2004 Sketches. SIGGRAPH '04*, New York, NY, USA, ACM (2004) 81–
37. Barsky, S., Petrou, M.: The 4-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **25**(10) (2003) 1239–1252
38. Frankot, R., Chellappa, R.: A method for enforcing integrability in shape from shading algorithms. *IEEE T. Pattern Anal.* **10**(4) (1988) 439–451

39. Ojala, T., Pietikäinen, M., Mäenpää, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **24**(7) (2002) 971–987
40. Chatfield, K., Simonyan, K., Vedaldi, A., Zisserman, A.: Return of the devil in the details: Delving deep into convolutional nets. In: *British Machine Vision Conference*. (2014)
41. Gray, R.M.: Vector quantization. *ASSP Magazine, IEEE* **1**(2) (1984) 4–29
42. Jain, A.K., Dubes, R.C.: *Algorithms for clustering data*. Prentice-Hall, Inc. (1988)
43. Fukunaga, K.: *Introduction to statistical pattern recognition*. Academic press (2013)
44. Hsu, C.W., Lin, C.J.: A comparison of methods for multiclass support vector machines. *Neural Networks, IEEE Transactions on* **13**(2) (2002) 415–425
45. Vedaldi, A., Fulkerson, B.: VLFeat: An open and portable library of computer vision algorithms. <http://www.vlfeat.org/> (2008)
46. Vedaldi, A., Lenc, K.: Matconvnet - convolutional neural networks for MATLAB. *CoRR abs/1412.4564* (2014)