21^{st} *Computer Vision Winter Workshop*
Luka Čehovin, Rok Mandeljc, Vitomir Štruc (eds.)
Rimske Toplice, Slovenia, February 3–5, 2016

Significance of Colors in Texture Datasets

Milan Šulc, Jiří Matas
Czech Technical University in Prague, Faculty of Electrical Engineering,
Department of Cybernetics, Center for Machine Perception,
Technická 2, 166 27 Praha 6, Czech Republic
{sulcmila,matas}@fel.cvut.cz

Abstract. *This paper studies the significance of color in eight publicly available datasets commonly used for texture recognition through the classification results of "pure-color" and "pure-texture" (colorless) descriptors. The datasets are described using the state-of-the-art color descriptors, Discriminative Color Descriptors (DD) [15] and Color Names (CN) [28]. The descriptors are based on partitioning of the color space into clusters and assigning the image probabilities of belonging to individual clusters. We propose a simple extension of the DD and the CN descriptors, adding the standard deviations of color cluster probabilities into the descriptor. The extension leads to a significant improvement in recognition rates on all datasets. On all datasets the 22-dimensional improved CN^{σ} descriptor outperforms all original 11-, 25- and 50-dimensional descriptors. Linear combination of the state-of-the-art "pure-texture" classifier with the CN^{σ} classifier improves the results on all datasets.*

1. Introduction

Visual recognition based on texture and color are well established computer vision disciplines with several surveys available, e.g. [3, 10, 19, 20, 27, 30]. The state-of-the-art in texture recognition has been recently dominated in terms of accuracy by methods based on deep Convolutional Neural Networks (CNNs) [5, 6], yet the pre-CNN approaches may be preferable in real-time applications for their performance without parallel processing. Although it has been shown that several texture description methods can benefit from adding color information [13], a large number of the pre-CNN texture recognition techniques has been evaluated only on gray-scale images. Since many publicly available datasets used for texture recognition contain color information, we decided to evaluate the accuracy of color-statistics based methods to measure the significance of color information in the datasets.

The first contribution of this paper is a study of the significance of color information in available datasets commonly used for evaluation of texture recognition methods. In total we evaluate 8 texture datasets, namely FMD (Flickr Material Database), ALOT (A Lot Of Textures), KTH-TIPS (Textures under varying Illumination, Pose and Scale), KTH-TIPS2a, KTH-TIPS2b, CUReT (Columbia-Utrecht Reflectance and Texture), VehApp (Vehicle Appearance) and AniTex (Animal Texture).

The second contribution of the paper is an improvement of the state-of-the-art color descriptors, Discriminative Color Descriptors (DD) [15] and Color Names (CN) [28]. DD and CN are based on partitioning of the color space into clusters and assigning each color the probabilities of belonging to individual clusters. Our extension to the DD and the CN descriptors adds the standard deviation for each color cluster to the descriptor. This leads to an improvement in recognition rates on all 8 tested datasets, as shown in the experiments in Section 5.

The third contribution of the paper are experiments combining a state-of-the-art "pure-texture" descriptor with the improved CN^{σ} descriptor, leading to further increase in recognition accuracy.

The rest of the paper is organized as follows: Sections 2.1 and 2.2 review the state of the art in texture and color recognition, respectively. The selected "pure-color" image descriptors and our extension to them are introduced in Section 3. Publicly available color-image databases commonly used for texture classification are described in Section 4. Section 5 describes the experiments and presents the results.

The observations are discussed and conclusions are drawn in Section 6.

2. State of the Art

2.1. Texture-Based Classification

A large number of texture recognition techniques has been proposed, many of them being described in the surveys [3, 19, 20, 30]. In this section we only review the recent development and the state-of-the-art.

Several recent texture recognition algorithms report excellent results on standard datasets while ignoring the available color information. A number of them is based on the popular Local Binary Patterns, such as the Pairwise Rotation Invariant Co-occurrence Local Binary Pattern of Qi et al. [22] or the Fast Features Invariant to Rotation and Scale of Texture of Sulc and Matas [26]. A cascade of invariants computed by scattering transforms was proposed by Sifre and Mallat [24] in order to construct an affine invariant texture representation. Mao et al. [18] use a bag-of-words model with a dictionary of so called active patches: raw intensity patches that undergo further spatial transformations and adjust themselves to best match the image regions. While the Active Patch Model doesn't use color information, the authors claim that adding color will further improve the results. Cimpoi et al. [4], using Improved Fisher Vectors (IFV) for texture description, show further improvement when combined with describable texture attributes learned on the Describable Textures Dataset (DTD) and with color attributes.

Recently, Cimpoi et al. [5, 6] pushed the state-of-the-art in texture recognition using a new encoder denoted as FV-CNN-VD, obtained by Fisher Vector pooling of a very deep Convolutional Neural Network (CNN) filter bank of Simonyan and Zisserman [25]. The CNN filter bank operates on (preprocessed) RGB images. The method achieves state-of-the-art accuracy, yet may not be suitable for real-time applications when evaluated without a high-performance GPU.

2.2. Color Statistics for Classification

Color information is processed by many state-of-the-art descriptors in Computer Vision, including the neurocodes of Deep CNNs or different extensions of SIFT incorporating color. Yet we are interested in simpler color statistics, not making use of spatial information.

Standard approaches to collect color information include color histograms (based on different color representations), color moments and moment invariants. Sande et al. [27] provide an extensive evaluation of such descriptors. The Color Names (CN) descriptor by Weijer et al. [28] is based on models learned from real-world data obtained from Google by searching for 11 color names in English. The Color Names have shown to be a successful color attribute for object detection [12] and recognition [14]. The model assigns each pixel the probability of belonging to one of the 11 color clusters. A similar approach is used by the Discriminative Color Descriptor (DD) of Khan et al. [15], where the color values are clustered together based on their discriminative power in a classification problem with the objective to minimize the drop of mutual information of the final representation.

Khan et al. [13] study the strategies of combining color and texture information. They carried out a comparison of pure color descriptors on the publicly available KTH-TIPS2a, KTH-TIPS2b, and FMD datasets, and on another small dataset denoted as Texture-10. Since the results of Color Names and Discriminative Color Descriptors outperformed other color descriptors in texture classification, we will describe the usage of CN and DD in more detail in Section 3 and use the models in our experiments in Section 5.

3. Selected Color Descriptors

Based on the findings of Khan et al. [13] and on our preliminary results, we consider the Color Names [28] and Discriminative Color Descriptors [15] the best match for our experiments for their superior classification accuracy.

While each of the approaches creates the color models based on a different criteria, the result is a soft assignment of clusters to each RGB value. In both cases the assignment is performed using a lookup table, which creates a mapping from RGB values to probabilities over C clusters c_i, i.e. $p\left(c_i \mid x\right)$. In this work we use the lookup tables provided by the authors of the methods, i.e. the 11-dimensional Color Names representation by [28] and the universal color 11-, 25- and 50-dimensional representations by [15].

The models assume uniform prior over the color names $p(c_i)$. The conditional probabilities for each cluster c_i given an image I are computed as an average over all N pixels x_n in the region:

$$p\left(c_i \mid I\right)=\frac{1}{N} \sum_{x_n \in I} p\left(c_i \mid x_n\right) \tag{1}$$

The standard descriptor D for image I is then a vector containing the probability of each cluster:

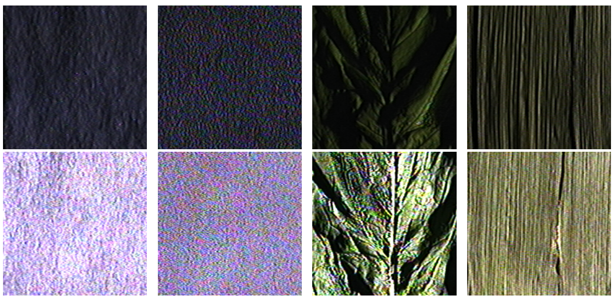
$$D(I)=\left[\begin{array}{c} p\left(c_1 \mid I\right) \\ p\left(c_2 \mid I\right) \\ \vdots \\ p\left(c_C \mid I\right) \end{array}\right] \tag{2}$$

We propose to add another statistics to the color descriptor, the standard deviation of the color cluster probabilities in the image:

$$\sigma(c_i \mid I)=\sqrt{\frac{1}{N} \sum_{x_n \in I}\left[p\left(c_i \mid x_n\right)-p\left(c_i \mid I\right)\right]^2} \tag{3}$$

We concatenate the standard deviations to the original descriptor to get the extended representation:

$$D^\sigma(I)=\left[\begin{array}{c} p\left(c_1 \mid I\right) \\ p\left(c_2 \mid I\right) \\ \vdots \\ p\left(c_C \mid I\right) \\ \sigma(c_1|I) \\ \sigma(c_2|I) \\ \vdots \\ \sigma(c_C|I) \end{array}\right] \tag{4}$$



(a) Felt (b) Polyester (c) Lettuce leaf (d) Corn husk

Figure 1: Examples of four texture classes from the CUReT database.

4. Color Texture Datasets

This section reviews publicly available texture datasets that contain color information. Databases available only in the gray-scale version, such as Brodatz, UIUCTex or UMD, are omitted.

4.1. CUReT

The Columbia-Utrecht Reflectance and Texture (CUReT) image database [8] commonly used for texture recognition[1] contains 5612 images of 61 classes. There are 92 images per class, with different combinations of view- and illumination-direction.

The standard experimental protocol divides the dataset into two halves, using 46 training images per class for training and 46 images for testing. Examples of four selected classes from the dataset are displayed in Figure 1.

4.2. KTH-TIPS

The Textures under varying Illumination, Pose and Scale (KTH-TIPS) database [9, 11] was collected by Fritz, Hayman and Caputo with the aim to supplement the CUReT database, concerning texture variations in real-world conditions. The dataset contains 81 images for each of 10 selected materials, taken with different combination of pose, illumination and scale. The dataset contains samples of different color for several materials, each of the samples appears several times. In the experimental protocol the dataset is randomly divided into halves, 40 images per class are used for training and the remaining 41 images are used for testing. It is thus probable, that each of the samples appear in the training data set.

4.3. KTH-TIPS2

The KTH-TIPS2 database [2, 17], gathered by Mallikarjuna, Targhi, Hayman and Caputo, largely followed the procedure used for the previous KTH-TIPS database, with some differences in scale an illumination. The database also contains images from the previous KTH-TIPS dataset. The objective of the database is to provide a better means of evaluation: It contains 4 physical samples for each of 11 materials and images of no physical sample are present in both training and test set. The database contains 108 images of each physical sample. There are two version of the database: KTH-TIPS2a and KTH-TIPS2b. In

[1] http://www.robots.ox.ac.uk/ vgg/research/texclass/setup.html

(a) Corduroy (b) Lettuce (c) Wood (d) Wool

Figure 2: Examples of four texture classes from the KTH-TIPS2 database. Each image belongs to a different physical sample.

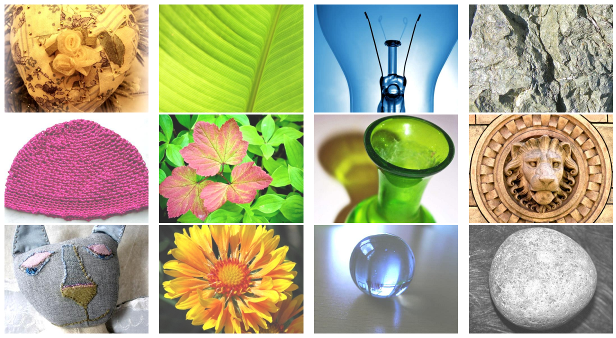
the KTH-TIPS2a dataset, 144 images are are missing (namely there are four samples with only 72 images). In the experimental protocol, three samples from each class form the training set and the remaining sample is used for testing. In the case of the KTH-TIPS2b dataset, one sample forms the training set and the remaining three form the test set. Examples from all four samples of four selected classes from the database are displayed in Figure 2.

4.4. ALOT

The Amsterdam Library of Textures (ALOT) [1] is similar in spirit to the CUReT dataset, yet the number of materials is much higher: it contains 250 texture classes, 100 images per class. The pictures were taken under various viewing and illumination directions and illumination colors. For evaluation, 20 images per class are used for training and the remaining 80 images per class are used for testing. Examples from the ALOT database are displayed in Figure 3.

(a) Fruit sprinkles (b) Pepper (red) (c) Color calibration checker (d) Macaroni

Figure 3: Examples of four texture classes from the ALOT database.



(a) Fabric (b) Foliage (c) Glass (d) Stone

Figure 4: Examples of four texture classes from the FMD database.

4.5. FMD

The Flickr Material database (FMD) was developed by Sharan et al. [23] with the intention of capturing a range of real world appearances of common materials. The dataset contains 1000 images downloaded manually from Flickr.com (under Creative Commons license), belonging to one of the following materials: Fabric, Foliage, Glass, Leather, Metal, Paper, Plastic, Stone, Water or Wood. There are exactly 100 images for each of the 10 material classes. Unlike the dataset described above, FMD was not primarily created for texture recognition, and it includes images of objects with various textures for each material. The dataset also includes binary masks for background segmentation. The standard evaluation protocol divides the images in each class into two halves, 50 images for training and 50 for testing. Examples from the FMD dataset are displayed in Figure 4.

4.6. AniTex

The Animal Texture dataset (AniTex) constructed by Mao et al. [18] contains 3120 texture patch images cropped randomly from the torso regions inside the silhouettes of different animals in the PASCAL VOC 2012 database. There are only 5 classes (cat, dog, sheep, cow and horse), 624 images each. The authors created the dataset to explore less ho-

(a) Cat (b) Dog (c) Sheep (d) Cow

Figure 5: Examples of four texture classes from the AniTex database.

(a) Plane (b) Bicycle (c) Bus (d) Car

Figure 6: Examples of four texture classes from the VehApp database.

mogeneous texture and appearance than available in standard texture datasets. The patches in the dataset come from images under different conditions such as scaling, rotation, viewing angle variations and lighting condition change. For evaluation, the dataset is randomly divided into 2496 training and 624 testing images. Examples from the AniTex dataset are displayed in Figure 5.

4.7. VehApp

The Vehicle Appearance dataset (VehApp) was created by the same authors as AniTex [18] with the same intentions. It contains 13723 images cropped from PASCAL VOC images containing vehicles of 6 classes (aeroplane, bicycle, car, bus, motorbike, train). The images are evaluated in a way similar to AniTex: 80% images are randomly chosen into the training set, the remaining 20% is used for testing. Examples from the VehApp dataset are displayed in Figure 6.

5. Experiments

We compute 8 descriptors for each image in every database: the standard 11-dimensional Color Name descriptor CN and our extended 22-dimensional version CN$^\sigma$; the 11-, 25- and 50- Discriminative Color Descriptors DD11, DD25, DD50 and the extended versions DD11$^\sigma$, DD25$^\sigma$, DD50$^\sigma$ of double dimensionality.

The multiclass classification is then performed for each descriptor separately by combining binary SVM classifiers in a One-vs-All scheme. Linear SVM classifiers were used together with an approximate feature map of Vedaldi and Zisserman [29]. The χ^2 kernel approximations and the histogram intersection kernel approximations were considered, the latter was chosen based on slightly superior performance in preliminary experiments. The Platt's probabilistic output [16, 21] was used in order to estimate the posterior class probabilities to choose the result in the One-vs-All scenario. To minimize the effect of the random splits into training and testset, each experiment is performed 10 times on a different split, with the exception of the KTH-TIPS2 databases with 4 experiments based on the material samples.

All 8 color descriptors are compared in terms of class recognition accuracy in Table 1. The best published results of "pure-texture" (color-less) methods and the results of the state-of-the-art FV-CNN [5] method are attached to the table for comparison. The comparison of the best "pure-color" and "pure-texture" results on all 8 datasets is illustrated in Figure 7.

An experiment on combining efficient classifiers of "pure-texture" and "pure-color" was performed as follows: Each image was described using the CN$^\sigma$ color descriptor (using the same method as above) and the Ffirst [26] texture descriptor (with $n_{\text{conc}} = 3$

Table 1: Recognition accuracy of selected color descriptors on publicly available databases commonly used for texture recognition.

	CUReT	TIPS	TIPS2a	TIPS2b	ALOT	FMD	AniTex	VehApp
# classes	61	10	11	11	250	10	5	6
CN	85.9$\pm$0.6	99.3$\pm$0.9	46.7$\pm$2.0	39.0$\pm$2.5	51.0$\pm$0.5	26.3$\pm$2.4	38.0$\pm$2.0	34.7$\pm$1.0
DD11	68.7$\pm$0.9	95.5$\pm$1.3	43.5$\pm$6.5	36.1$\pm$1.0	38.2$\pm$0.4	24.0$\pm$1.1	32.4$\pm$1.6	33.2$\pm$1.0
DD25	83.4$\pm$0.8	96.8$\pm$0.9	44.0$\pm$7.6	36.0$\pm$2.3	60.9$\pm$0.5	23.9$\pm$1.4	36.0$\pm$1.7	36.9$\pm$0.6
DD50	87.7$\pm$1.0	99.0$\pm$0.7	46.9$\pm$4.8	38.5$\pm$1.5	65.5$\pm$0.4	22.6$\pm$1.4	37.4$\pm$1.1	39.1$\pm$1.0
CN^{σ}	**94.2$\pm$0.6**	**99.8$\pm$0.3**	51.7$\pm$5.7	**42.6$\pm$1.4**	73.9$\pm$0.5	**28.0$\pm$2.2**	**41.7$\pm$1.8**	39.1$\pm$0.7
$\mathrm{DD11}^{\sigma}$	81.9$\pm$0.8	97.6$\pm$1.0	48.5$\pm$3.8	38.3$\pm$1.9	60.1$\pm$0.5	22.7$\pm$1.6	35.9$\pm$2.1	35.8$\pm$0.5
$\mathrm{DD25}^{\sigma}$	88.9$\pm$0.7	99.4$\pm$0.3	49.1$\pm$3.7	39.9$\pm$4.5	75.0$\pm$0.5	23.9$\pm$1.1	39.9$\pm$1.6	39.3$\pm$0.7
$\mathrm{DD50}^{\sigma}$	91.0$\pm$0.7	99.6$\pm$0.2	**53.2$\pm$4.6**	42.0$\pm$2.8	**78.0$\pm$0.5**	25.3$\pm$1.7	38.9$\pm$0.8	**41.2$\pm$0.9**
FV-CNN[5]	99.0$\pm$0.2	–	–	81.8$\pm$2.5	98.5$\pm$0.1	79.8$\pm$1.8	–	–
Pure-texture	99.8$\pm$0.1[24]	99.7$\pm$0.1[4]	88.2$\pm$6.7[26]	76.0$\pm$2.9 [26]	95.9$\pm$0.5 [26]	57.4$\pm$1.7[22]	50.8[18]	63.4[18]

descriptors per image, each describing $c = 7$ consecutive scales). An approximate intersection kernel map is applied to both color and texture descriptors, which are then classified using the One-vs-All Support Vector Machines with Platt's probabilistic outputs. The final scores in Table 2 were then combined using 3 axiomatic approaches, denoted as:

1. PROD: The dot product of both of the scores is used for final decision.
2. SUM: The sum of both of the scores is used for final decision.
3. $\mathrm{SUM}_{0.3}$: The weighted sum of both of the scores is used for final decision, where the weight of color is only 30% of the weight of texture, taking into account the lower performance of the color descriptors on most datasets.

In terms of combining probability distributions [7], the SUM and $\mathrm{SUM}_{0.3}$ schemes represent a *linear opinion pool* and the PROD scheme represents a *logarithmic opinion pool*.

6. Observations and Conclusions

A set of experiments with color-based image descriptors was performed on 8 datasets commonly used for texture classification, leading to interesting insights in color-based classification and in the understanding of available texture-recognition datasets.

One can see that using the simple color descriptors is sufficient for excellent results in specific cases, such as the KTH-TIPS dataset, where materials of the same color appear in both training and test data. Satisfying results can also be obtained on the CUReT and ALOT datasets. The KTH-TIPS2a and KTH-TIPS2b datasets are more difficult for "pure-color" classification, since testing data may come from samples of different colors than training data, as illustrated in Figure 2. The FMD, AniTex and VehApp

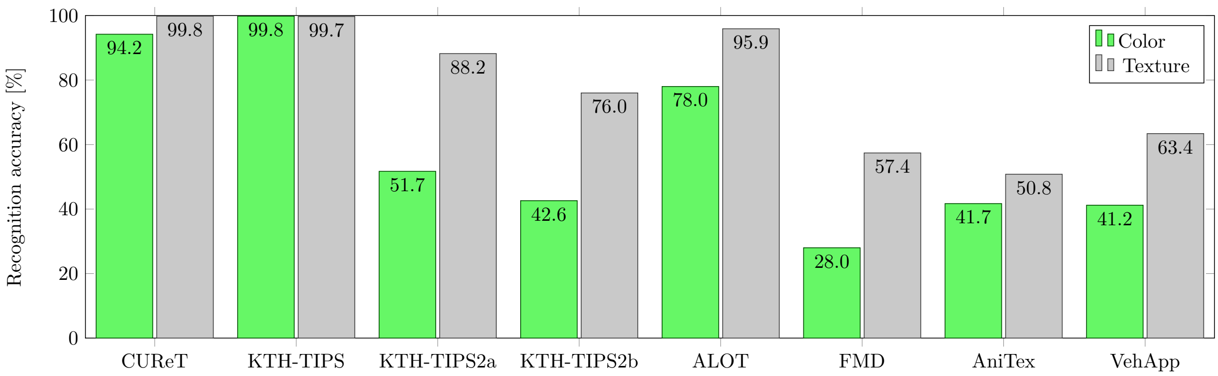


Figure 7: Comparison of the best published results of "pure-texture" descriptors and the best results obtained using "pure-color" descriptors.

Table 2: Recognition accuracy for combinations of "pure-texture" (Ffirst) and "pure-color" (CN^{σ}) descriptors.

	CUReT	TIPS	TIPS2a	TIPS2b	ALOT	FMD	AniTex	VehApp
# classes	61	10	11	11	250	10	5	6
CN^{σ}	94.24±0.60	99.83±0.31	51.73±5.71	42.64±1.43	73.86±0.46	27.98±2.20	41.67±1.77	39.07±0.67
Ffirst	99.65±0.09	99.51±0.53	88.29±6.77	76.60±4.29	96.43±0.23	50.22±1.90	45.72±1.78	54.41±0.66
PROD	99.41±0.15	99.98±0.08	68.13±5.06	60.12±4.06	94.65±0.20	46.58±2.37	49.97±1.50	56.47±0.76
SUM	99.04±0.20	**100.00±0.00**	77.59±5.87	60.35±5.13	92.06±0.29	45.70±2.47	**50.08±1.56**	56.56±0.98
$SUM_{0.3}$	**99.68±0.12**	99.85±0.26	**88.76±6.40**	**77.17±4.23**	**97.05±0.14**	**52.24±1.68**	48.99±1.83	**56.62±0.92**

datasets are quite difficult for their heterogeneous nature, both in terms of texture and color. Yet the color statistics might still provide useful information when combined with other descriptors.

An extension to the Color Names (CN) and Discriminative Color Descriptors (DD) has been proposed (denoted as CN^{σ}, DD^{σ}), significantly improving the recognition accuracy on all 8 tested datasets. The comparison of Color Names (CN) and Discriminative Color Descriptors (DD) descriptors brings a surprising observation: on 6 out of the 8 texture datasets, Color Names outperform even the higher-dimensional Discriminative Color Descriptors DD25, although the opposite may be expected from the findings on different tasks [15]. The improved CN^{σ} outperforms other "pure-color" descriptors on 5 out of 8 datasets, the best results on the remaining 3 datasets are achieved by the improved $DD50^{\sigma}$ descriptor.

Combining a state-of-the-art "pure-texture" classifier [26] with the "pure-color" classifier of CN^{σ} leads to an improvement on all 8 tested datasets. The weights of the classifiers in the combination should be set according to the classifiers performance. Note that by combining the classifiers a 100% accuracy was achieved on the KTH-TIPS. Significant improvements are also achieved on the AniTex and VehApp databases, where [26] performs rather poorly.

The state-of-the-art "pure-texture" and "pure-color" classifiers and their combinations obtain excellent results on simpler texture-recognition problems. They are outperformed by the recent FV-CNN model [5] in the more difficult tasks. Yet the low computational complexity of some "pure-texture" and "pure-color" descriptors is beneficial and their performance may be still interesting for future works, e.g. when used in a cascade classification scheme and followed by FV-CNN in case of ambiguity.

Acknowledgements

Milan Šulc was supported by CTU student grant SGS15/155/OHK3/2T/13, Jiří Matas by The Czech Science Foundation Project GACR P103/12/G084.

References

[1] G. J. Burghouts and J.-M. Geusebroek. Material-specific adaptation of color invariant features. *Pattern Recognition Letters*, 30(3):306–313, 2009. 4

[2] B. Caputo, E. Hayman, and P. Mallikarjuna. Class-specific material categorisation. In *Proc. of IEEE International Conference on Computer Vision (ICCV)*, volume 2, pages 1597–1604. IEEE, 2005. 3

[3] C.-h. Chen, L.-F. Pau, and P. S.-p. Wang. *Handbook of pattern recognition and computer vision*. World Scientific, 2010. 1, 2

[4] M. Cimpoi, S. Maji, I. Kokkinos, S. Mohamed, and A. Vedaldi. Describing textures in the wild. In *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*, pages 3606–3613. IEEE, 2014. 2, 6

[5] M. Cimpoi, S. Maji, I. Kokkinos, and A. Vedaldi. Deep filter banks for texture recognition, description, and segmentation. *arXiv preprint arXiv:1507.02620*, 2015. 1, 2, 5, 6, 7

[6] M. Cimpoi, S. Maji, and A. Vedaldi. Deep filter banks for texture recognition and segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3828–3836, 2015. 1, 2

[7] R. T. Clemen and R. L. Winkler. Combining probability distributions from experts in risk analysis. *Risk analysis*, 19(2):187–203, 1999. 6

[8] K. J. Dana, B. Van Ginneken, S. K. Nayar, and J. J. Koenderink. Reflectance and texture of real-world surfaces. *ACM Transactions on Graphics (TOG)*, 18(1):1–34, 1999. 3

[9] M. Fritz, E. Hayman, B. Caputo, and J.-O. Eklundh. The kth-tips database, 2004. 3

[10] T. Gevers, A. Gijsenij, J. Van de Weijer, and J.-M. Geusebroek. *Color in computer vision: fundamentals and applications*, volume 23. John Wiley & Sons, 2012. 1

[11] E. Hayman, B. Caputo, M. Fritz, and J.-O. Eklundh. On the significance of real-world conditions for material classification. In *Computer Vision–ECCV 2004*, pages 253–266. Springer, 2004. 3

[12] F. S. Khan, R. M. Anwer, J. van de Weijer, A. D. Bagdanov, M. Vanrell, and A. M. Lopez. Color attributes for object detection. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, pages 3306–3313. IEEE, 2012. 2

[13] F. S. Khan, R. M. Anwer, J. van de Weijer, M. Felsberg, and J. Laaksonen. Compact color–texture description for texture classification. *Pattern Recognition Letters*, 51:16–22, 2015. 1, 2

[14] F. S. Khan, J. Van de Weijer, and M. Vanrell. Modulating shape features by color attention for object recognition. *International Journal of Computer Vision*, 98(1):49–64, 2012. 2

[15] R. Khan, J. Van de Weijer, F. Shahbaz Khan, D. Muselet, C. Ducottet, and C. Barat. Discriminative color descriptors. In *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pages 2866–2873. IEEE, 2013. 1, 2, 7

[16] H.-T. Lin, C.-J. Lin, and R. C. Weng. A note on platts probabilistic outputs for support vector machines. *Machine learning*, 68(3), 2007. 5

[17] P. Mallikarjuna, M. Fritz, A. Targhi, E. Hayman, B. Caputo, and J. Eklundh. The kth-tips and kth-tips2 databases. `http://www.nada.kth.se/cvap/databases/kth-tips`, 2006. 3

[18] J. Mao, J. Zhu, and A. L. Yuille. An active patch model for real world texture and appearance classification. In *Computer Vision–ECCV 2014*, pages 140–155. Springer, 2014. 2, 4, 5, 6

[19] M. Mirmehdi, X. Xie, and J. Suri. *Handbook of texture analysis*. Imperial College Press, 2009. 1, 2

[20] M. Pietikäinen. Texture recognition. *Computer Vision: A Reference Guide*, pages 789–793, 2014. 1, 2

[21] J. Platt. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. *Advances in large margin classifiers*, 10(3), 1999. 5

[22] X. Qi, R. Xiao, C.-G. Li, Y. Qiao, J. Guo, and X. Tang. Pairwise rotation invariant co-occurrence local binary pattern. *PAMI*, 36(11):2199–2213, 2014. 2, 6

[23] L. Sharan, R. Rosenholtz, and E. Adelson. Material perception: What can you see in a brief glance? *Journal of Vision*, 9(8):784–784, 2009. 4

[24] L. Sifre and S. Mallat. Rotation, scaling and deformation invariant scattering for texture discrimination. In *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pages 1233–1240. IEEE, 2013. 2, 6

[25] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014. 2

[26] M. Šulc and J. Matas. Fast features invariant to rotation and scale of texture. In L. Agapito, M. M. Bronstein, and C. Rother, editors, *Computer Vision–ECCV 2014 Workshops, Part II*, volume 8926 of *LNCS*, pages 47–62, Gewerbestrasse 11, CH-6330 Cham (ZG), Switzerland, September 2015. Springer International Publishing AG. 2, 5, 6, 7

[27] K. E. Van De Sande, T. Gevers, and C. G. Snoek. Evaluating color descriptors for object and scene recognition. *PAMI*, 32(9):1582–1596, 2010. 1, 2

[28] J. Van De Weijer, C. Schmid, J. Verbeek, and D. Larlus. Learning color names for real-world applications. *Image Processing, IEEE Transactions on*, 18(7):1512–1523, 2009. 1, 2

[29] A. Vedaldi and A. Zisserman. Efficient additive kernels via explicit feature maps. *PAMI*, 34(3), 2011. 5

[30] J. Zhang and T. Tan. Brief review of invariant texture analysis methods. *Pattern recognition*, 35(3):735–747, 2002. 1, 2