STRUCTURAL TEXTURE SIMILARITY FOR MATERIAL RECOGNITION

Jue Lin, Thrasyvoulos N. Pappas

ECE Department, Northwestern University, Evanston, IL 60208

ABSTRACT

We propose a new direct approach for material recognition under diverse illumination and viewing conditions based on visual texture. We apply K-means clustering to feature vectors that consist of steerable filter subband statistics and dominant colors of each texture image in order to obtain a small number of exemplars characterizing each material. We then use structural texture similarity metrics and color composition metrics to compare a query texture to the exemplars for material classification. Experimental results using the CUReT database establish the importance of color and demonstrate that five exemplars per texture provide performance comparable to the state of the art.

Index Terms— Texture analysis, subband statistics, dominant colors, material identification

1. INTRODUCTION

Texture provides valuable information for image analysis for several applications, such as image compression and quality, restoration, biomedical image analysis, computer vision, and content-based retrieval. In particular, texture provides important cues for material identification and characterization [1]. While there has been a considerable amount of work on texture analysis, it remains an open and exciting field. The focus of this paper is on robust material recognition under diverse illumination and viewing conditions. Since material appearance can vary dramatically in such cases, it is important to obtain a compact representation of each material that accounts for the changing appearance. To accomplish this, we use feature vectors that consist of steerable filter subband statistics and dominant colors of a texture image, and rely on K-means clustering to obtain a small number of exemplars that characterize the material under different illumination and viewing conditions. We then rely on texture similarity metrics for comparing a query texture to the exemplars for material classification.

The problem of texture classification and material recognition has been addressed by Leung and Malik [2], Cula and Dana [3], and Varma and Zisserman [4, 5]. They have proposed elaborate approaches that account for variations in texture appearance due to changes in illumination and viewing conditions, and use the CUReT (Columbia-Utrecht reflectance and texture) dataset [6, 7] for algorithm training and testing. Leung and Malik [2] applied a filter bank consisting of 48 filters, performed pixel-wise vector quantization of the filter responses to obtain cluster centers, which they called "textons," and constructed histograms of texton distributions, which they used for image classification. Cula and Dana [3] followed a similar approach. They introduced the notion of *bidirectional texture function (BTF),* which describes the changes in texture appearance as a function of viewing and illumination directions. They proposed a texture recognition method that relies on a bidirectional feature histogram, as well as a method for evaluating the significance of texture images within the BTF. Varma and Zisserman [4] adopted a similar framework, and conducted a comprehensive analysis of the performance of different filter banks, without any priori knowledge of image viewing and illumination conditions as opposed to Leung and Malik's approach. Furthermore, they argued that using filtering is equivalent to linearly transforming raw pixel values into a lower dimensional space [5]. Varma and Zisserman applied both approaches [4, 5] to the problem of material classification under a variety of viewing and illumination conditions using the CUReT dataset [6, 7], and achieved start-of-the-art performance in terms of texture classification accuracy.

In this paper we consider a simpler and more direct alternative to the above elaborate classification approaches [2–5]. In addition, the proposed approach incorporates chrominance information and demonstrates its importance. We propose a new problem formulation that combines K-means clustering with grayscale *structural texture similarity metrics (STSIMs)* [8,9] and *color composition similarity metrics (CCSIMs)* [8, 10]. For each texture image we extract a feature vector that consists of the STSIM subband statistics [8,9] and the dominant colors [11] associated with the CCSIMs [8, 10]. We apply K-means clustering to the feature vectors in order to obtain a small number of exemplars that encompass the changing material appearance. We then use STSIMs and CCSIMs to compare a query texture to the exemplars for material classification.

Our goal is to obtain a better understanding of how the viewing and illumination conditions affect the visual appearance of each material, and to determine the number of exemplars that are sufficient for material discrimination. Our ultimate goal is to be able to determine the illumination and viewing conditions of the exemplars without exhaustive sampling (over multiple illumination and viewing angles) for each material.

Experimental results using the CUReT database establish the importance of color and demonstrate that a small number of exemplars (five) per texture provide performance comparable to the state-of-the art, including the popular deep convolutional neural networks AlexNet and VGGM [12]. Traditional techniques for texture retrieval or segmentation, such as those in [13–18] do not address differences in material appearance due to illumination and viewing conditions, and as such are not applicable to the problem addressed in this paper.

While several state-of-the-art approaches rely on Convolutional Neural Networks (CNN) to gain a marginal performance improvement over the techniques we discussed above, we show that the proposed technique can achieve comparable results with a lot less computation. The proposed technique requires 14 convolution operations per pixel on the grayscale component of the image, CNN-based techniques typically consist of hundreds of layers, each of which applies hundreds of convolutions on tensors with hundreds of channels. We believe that the proposed technique achieves a reasonable balance between performance and speed, and it is preferred for applications for which a GPU is not available.

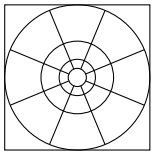

Fig. 1. Steerable filter decomposition with 3 scales, 4 orientations

This paper is organized as follows: In Section 2, we review STSIMs. Section 3 discusses our proposed approach. Experimental results are presented in Section 4 and the conclusions are summarized in Section 5.

2. REVIEW OF TEXTURE SIMILARITY METRICS

In this section we review texture similarity metrics proposed by Zujovic *et al.* [8–10]. These metrics account for the stochastic nature of textures and human perception. The key idea is to replace point-by-point comparisons with comparisons of region statistics computed within each texture. Zujovic *et al.* [8, 10] argued that developing separate metrics for the grayscale component of a texture and its color composition leads to more effective metrics. A comprehensive review of other texture similarity metrics can be found in [9].

2.1. Grayscale Structural Texture Similarity Metrics

The grayscale STSIM metrics [9] consist of a Steerable Filter Decomposition (SFD) [19], a set of subband statistics, formulas for comparing such statistics, and pooling across space and frequency to obtain an objective measure of the similarity of a pair of textures. In the experiments of Section 4 we will assume that the SFD consists of three scales and four orientations for a total of 14 subbands ($4 \times 3 + \text{LP} + \text{HP}$), as shown in Figure 1. The STSIM-2 metric [9] uses the following statistics, which we will adopt for the clustering we present in Section 3:

$$\textbf{means}: \mu_{\mathbf{x}}^m = \mathbb{E}\{x^m(i,j)\} \tag{1}$$

$$\textbf{variances}: (\sigma_{\mathbf{x}}^m)^2 = \mathbb{E}\{[x^m(i,j) - \mu_{\mathbf{x}}^m]^2\} \tag{2}$$

horizontal and vertical autocorrelation coefficients:

$$\rho_{\mathbf{x}}^m(0,1) = \frac{\mathbb{E}\{[x^m(i,j) - \mu_{\mathbf{x}}^m][x^m(i,j+1) - \mu_{\mathbf{x}}^m]\}}{(\sigma_{\mathbf{x}}^m)^2} \tag{3}$$

$$\rho_{\mathbf{x}}^m(1,0) = \frac{\mathbb{E}\{[x^m(i,j) - \mu_{\mathbf{x}}^m][x^m(i+1,j) - \mu_{\mathbf{x}}^m]\}}{(\sigma_{\mathbf{x}}^m)^2} \tag{4}$$

crossband correlation coefficients:

$$\rho_{\mathbf{x}}^{m,n}(0,0) = \frac{\mathbb{E}\{[x^m(i,j) - \mu_{\mathbf{x}}^m][x^n(i,j) - \mu_{\mathbf{x}}^m]\}}{\sigma_{\mathbf{x}}^m \sigma_{\mathbf{x}}^n} \tag{5}$$

where $\mathbf{x}$ denotes a specific texture image, $\mathbf{x}^m$ and $\mathbf{x}^n$ are its m^{th} and n^{th} subbands, and $\mathbb{E}\{x^m(i,j)\}$ is the empirical average of $\mathbf{x}^m$ over all spatial locations (i,j). The crossband statistics only include correlation coefficients between subbands at adjacent scales (excluding the lowpass and highpass bands) for a given orientation, and between orientations at a given scale, for a total of 26 crossband correlations. Overall, the feature vector for each texture includes 82 parameters: 14 subbands $\times$ 4 (mean, variance, horizontal and vertical autocorrelation coefficients) + 26 crossband correlations. In our experiments all of the statistics are computed over the entire texture, i.e., no need for spatial pooling.

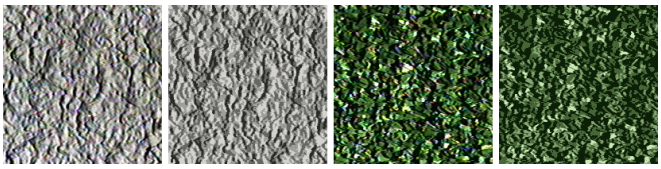

Fig. 2. Original texture and its color composition by ACA

For statistics comparisons and pooling, Zujovic *et al.* [9] proposed two main variations, STSIM-2 and STSIM-M, while another earlier version, STSIM-1, was proposed by Zhao *et al.* [20]. STSIM-1 and STSIM-2 compare statistics in a manner similar to CW-SSIM. The details can be found in [9]. Zujovic *et al.* [9, 21] conducted systematic tests in different application domains and demonstrated the advantage of STSIMs over existing techniques.

An alternative approach is to calculate the dissimilarity between two textures as the Mahalanobis distance between the feature vectors $\mathbf{f}_x$ and $\mathbf{f}_y$ corresponding to the two images

$$Q = (\mathbf{f}_x - \mathbf{f}_y)^T M (\mathbf{f}_x - \mathbf{f}_y) \tag{6}$$

where M is a symmetric positive semi-definite matrix and f is an 82-dimensional vector containing all the image statistics. In STSIM-M, the matrix M is diagonal with the variance of each statistic on the diagonal. Maggioni *et al.* [22] proposed using the intra-class standard deviation instead. Finally, a full matrix can also be estimated using machine learning techniques [23].

2.2. Color Composition Similarity Metrics

The color composition similarity metrics (CCSIMs) [8, 10] are based on a texture representation in terms of spatially adaptive dominant colors (*dominant colors* in short) and the associated percentages. The dominant colors are obtained with the adaptive clustering algorithm (ACA) [24]. The *optimal color composition distance (OCCD)* is then used to compare the color composition of two texture patches. OCCD breaks the histogram of dominant colors into fixed percentage units, finds the optimal mapping between these units, and computes the average distance between them in the *CIE L*a*b** color space [25]. Examples of dominant colors obtained with ACA are shown in Figure 2.

3. PROPOSED APPROACH

The material recognition problem can be formulated as a texture retrieval problem. If we pick a number of exemplars that are representative of the appearance of each material under different illumination and viewing conditions and form a codebook of all exemplars for all materials, then, given a query texture, we can find the exemplar that is most similar to the query in order to determine the material the query belongs to. Thus, the proposed approach consists of two stages, selecting a set of exemplars for each material (learning), and finding the exemplar that is closest to a query (retrieval).

The difference with the texton-based approaches [2–5] is that we use the STSIM subband statistics as feature vectors and K-means clustering for obtaining exemplars for each texture. We then rely on STSIM metrics for retrieval. We thus bypass the intermediate step of obtaining textons by clustering the raw subband filter responses and forming texton histograms characteristic of each texture, which are then compared to a query using the χ^2 statistic. Also, in contrast to the texton-based approaches, we make use of the color information.

3.1. Learning Stage

The grayscale feature vector for each texture patch consists of the STSIM statistics in (1) to (5). In the learning stage, we divide the training textures into groups according to material and apply K-means clustering to the feature vectors in each group. We thus obtain KM exemplars, where M is the number of materials. Since the cluster centers do not necessarily correspond to actual textures, and since texture synthesis is not well-defined by the STSIM parameters, we use the nearest neighbor of the cluster center to visualize each examplar.

Since the feature vectors contain a variety of statistics with different ranges of values, we tried three different normalization schemes: the L^2-norm (no normalization); the Z-norm, whereby the feature vectors normalized by the standard deviation of each component, as in STSIM-M; and the I-norm, whereby the feature vectors are normalized by the intra-class standard deviation, as in STSIM-I. For the Z-norm and the I-norm, the normalization is based only on the training set, which we discuss in the next section.

If we follow a similar approach for the color information, we should use a feature vector that contains the STSIM color composition statistics, that is, the dominant colors and the associated percentages. The OCCD metric could then be used as the distance between two feature vectors. However, Pappas *et al.* [26] found that, for natural images, the majority of segments containing perceptually uniform textures can be characterized by just the first two dominant colors (no percentages) for effective texture classification. Similarly, He and Pappas [27, 28] based their segmentation algorithm on the fact that natural textures consist of intensity variations of a single hue [29]. Hence, the two dominant shades (colors) of that hue should be sufficient to characterize the texture. Thus, a simpler approach is to simply concatenate the two most dominant colors to the vector of grayscale features. In our experiments, described in Section 4, we found that using all four dominant colors without the associated percentages works slightly better than using the two most dominant colors. We used the $L*a*b*$ space for the color features, ordered the colors according to percentage, and normalized all the components by 128 (based on the range of colors in the $L*a*b*$ space).

3.2. Retrieval Stage

Given the codebook of exemplars and the metric parameters (for the Z-norm and the I-norm) from the training stage, we use the STSIM/CCSIM metrics to compare the query texture to all of the exemplars in the codebook in order to find the closest one (most similar). For the exemplars obtained with the L^2-norm, we use STSIM-1 and STSIM-2 for measuring the similarity between the query and the exemplars. In all cases, the OCCD distance of the color features is computed separately, and combined with the grayscale distance in a multiplicative fashion, that is, STSIM*OCCD for STSIM-M and STSIM-I, and (1-STSIM)*OCCD for STSIM-1 and STSIM-2.

4. EXPERIMENTAL RESULTS

In our experiments we used textures from the CUReT database, which contains images of 61 samples of materials photographed under different illumination and viewing conditions, at a distance of two meters from the sample [6, 7]. We used the same subset of CUReT textures under 92 conditions that Varma and Zisserman [4,5] used. However, following the work of Nguyen [30], we eliminated the samples of three materials, "peacock" (57), "orange peel" (55), and "leaf" (23), because they were clearly not textures. We also restored the textures of three materials (33, 44, 46) that had aliasing in the chrominance component, and eliminated several severely

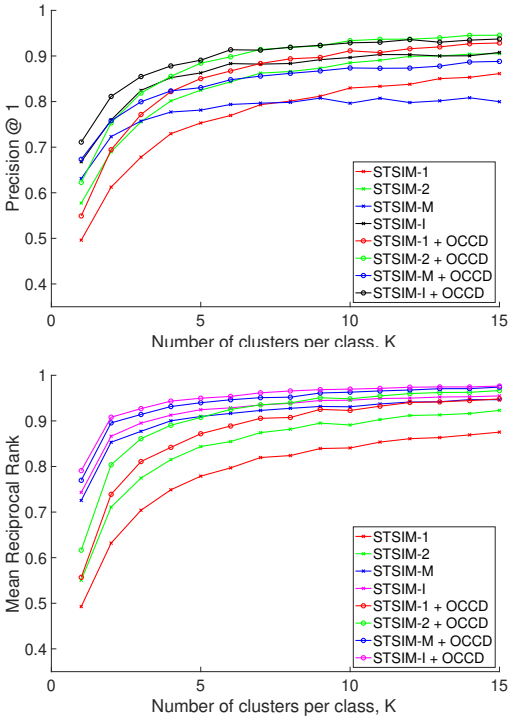

Fig. 3. Performance statistics

aliased samples of one material (05) as well as isolated samples of a few other materials that could not be restored. Finally, the texture samples of materials 02, 06, and 29 appear to be aliased but look natural, so we included them in our experiments. In addition, following [4, 31], we found that applying a normalization based on Weber's law to the pixel values yields slightly better performance, even though Weber's law has been already accounted for in the gamma-corrected image values.

To assess the performance of the proposed techniques, we used five-fold cross validation, whereby the samples for each material were partitioned into a training set (80%) and a testing set (20%). For each of the five runs of the training stage we used K-means to obtain K exemplars for each material. We then evaluated the retrieval performance of the different techniques on the testing set, using precision at one (P@1), the percentage of correct first matches, and mean reciprocal rank (MRR), the average inverse rank of the first correct match [9]. Note that it does not make sense to use mean average precision (MAP) [9,32] because the exemplars span a broad range of appearances and all we need to do is find the closest one. For each technique, the P@1 and MRR were averaged over the five runs. The results are plotted in Figure 3 for different number of exemplars K for the following techniques:

- L^2-norm for clustering and STSIM-1 for classification
- L^2-norm for clustering and STSIM-2 for classification
- Z-norm for clustering and classification (STSIM-M)
- I-norm for clustering and classification (STSIM-I)

Moreover, each technique was tested with and without color. As we discussed in Section 3, the addition of color was done by concatenating four dominant colors to the feature vector for clustering, and OCCD in multiplicative combination with the grayscale metric for classification. Note that performance initially increases sharply

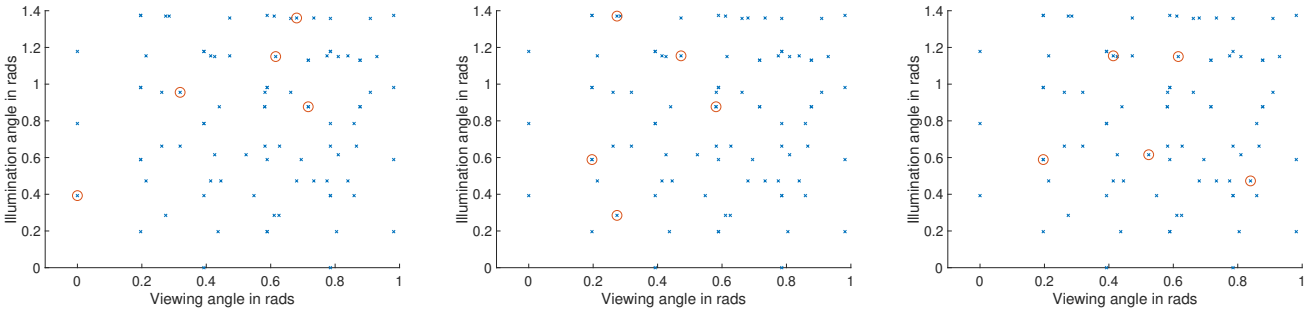

Fig. 4. Viewing and illumination polar angles for Concrete (50), Rug a (18), and Quarry Tile (25). Exemplars using I-norm encircled with red

K=5	K=29	VZ K=5	VZ, $K\approx$70	AlexNet	VGGM
92.1	97.6	80.7	97.2	98.5	98.7

Table 1. Classification accuracy comparison (proposed approach in first two columns)

with K and then plateaus after $K = 5$. The best overall performance is provided by STSIM-I, followed by STSIM-M, STSIM-2, and STSIM-1. Thus, accounting for variance within class (STSIM-I) outperforms accounting for variance across all textures (STSIM-M). As expected, introducing color information gives rise to a significant boost in retrieval performance for all of the techniques.

Table 1 compares our best result, using STSIM-I, with our implementation of the Varma-Zisserman [4] algorithm for $K = 5$ exemplars per class. The table also compares our results with the popular deep convolutional neural networks AlexNet and VGGM [12], which were fine-tuned by the authors. For a fair comparison, we used $K = 29$ for our approach and all the training data ($K \approx 70$) for Varma-Zisserman. While $K = 29$ provides the best results, it is significant and important that high performance is obtained for much lower $K = 5$. We conclude that the performance of the proposed approach is comparable to or better than state-of-the-art methods, while maintaining relatively low computational cost. The performance of the proposed approach is also comparable to more recent results reported in [33, 34].

Finally, we should also mention that the performance of traditional texture retrieval techniques, such as the one in [13], is not competitive in the context of the CUReT database: even when the nearest neighbor, leave-one-out decision rule is used (very large K), the classification accuracy is 52%, most probably due to the simplicity of the texture features.

4.1. Codebook Visualization

Visualizing the exemplars is important for understanding how they capture the material appearance under different viewing and illumination conditions. As we discussed, since the cluster centers do not necessarily correspond to actual textures, we use the nearest neighbor of the cluster center to visualize each examplar. Figure 5 shows $K = 5$ exemplars for three materials, "concrete" (50), "rug a" (18) and "quarry tile" (25). Note that for materials with a strong 3-D structure, the appearance changes in both color and pattern. The changes in the perceived flatness of the materials are particularly interesting. In contrast, for materials with weaker 3-D structure the variations are primarily in darkness.

Figure 4 shows the placement of the exemplars as a function of polar illumination and viewing angles. The samples nearest each

Fig. 5. Exemplars for Concrete (50); Rug a (18); Quarry Tile (25)

cluster center are enclosed in red circles. Observe that the exemplars are well separated in the viewing and illumination angle space. No other particular trends were observed in the exemplar placement. Note that including the azimuthal angles would require 4-D plots; however, most materials in the database have no clear directionality, and as a result, the azimuthal angle does not have a significant effect on the appearance. However, for completeness, we also plotted the exemplar placement as a function of the difference of illumination and viewing angles in the polar and azimuthal directions, and found no particular trends, either.

5. CONCLUSIONS

We presented a new direct approach for material recognition based on visual texture. We used a steerable filter decomposition and relied on subband statistics and dominant colors for the characterization of each texture. We then applied K-means clustering to obtain a small number of exemplars that represent each material under different illumination and viewing conditions. To recognize the material that corresponds to a given texture, we used grayscale structural texture similarity metrics and color composition similarity metrics to compare it to the exemplars. We conducted comprehensive experiments with the CUReT database, and showed that five exemplars per texture provide performance that is comparable to the state-of-the art. We also demonstrated the importance of color. In future work we will consider machine learning techniques to further reduce the number of exemplars. Our ultimate goal is to obtain a better understanding of how the viewing and illumination conditions affect the visual appearance of each material.

6. REFERENCES

[1] E. H. Adelson, "On seeing stuff: The perception of materials by humans and machines," in *Human Vision and Electronic Imaging VI*, B. E. Rogowitz and T. N. Pappas, Eds., San Jose, CA, Jan. 2001, vol. 4299 of *Proc. SPIE*, pp. 1–12.

[2] T. Leung and J. Malik, "Representing and recognizing the visual appearance of materials using thee-dimensional textons," *Int. J. Computer Vision*, vol. 43, no. 1, pp. 29–44, 2001.

[3] O. G. Cula and K. J. Dana, "3D texture recognition using bidirectional feature histograms," *Int. J. Computer Vision*, vol. 59, no. 1, pp. 33–60, 2004.

[4] M. Varma and A. Zisserman, "A statistical approach to texture classification from single images," *Int. J. Computer Vision*, vol. 62, no. 1/2, pp. 61–81, 2005.

[5] M. Varma and A. Zisserman, "A statistical approach to material classification using image patch exemplars," *IEEE Tr. Pattern Anal. Machine Intell.*, vol. 31, no. 11, pp. 2032–2047, 2009.

[6] "CUReT: Columbia-Utrecht Reflectance and Texture Database," www1.cs.columbia.edu/CAVE/software/curet/.

[7] K. J. Dana, B. van Ginneken, S. K. Nayar, and J. J. Koenderink, "Reflectance and texture of real-world surfaces," *ACM Tr. Graphics*, vol. 18, no. 1, pp. 1–34, Jan. 1999.

[8] J. Zujovic, *Perceptual Texture Similarity Metrics*, Ph.D. thesis, Northwestern Univ., Evanston, IL, Aug. 2011.

[9] J. Zujovic, T. N. Pappas, and D. L. Neuhoff, "Structural texture similarity metrics for image analysis and retrieval," *IEEE Tr. Image Proc.*, vol. 22, no. 7, pp. 2545–2558, July 2013.

[10] J. Zujovic, T. N. Pappas, and D. L. Neuhoff, "Structural similarity metrics for texture analysis and retrieval," in *Proc. Int. Conf. Image Proc.*, Cairo, Egypt, Nov. 2009, pp. 2225–2228.

[11] J. Chen, T. N. Pappas, Aleksandra Mojsilovic, and Bernice E. Rogowitz, "Adaptive perceptual color-texture image segmentation," *IEEE Tr. Image Proc.*, vol. 14, no. 10, pp. 1524–1536, Oct. 2005.

[12] M. Cimpoi, S. Maji, I. Kokkinos, and A. Vedaldi, "Deep filter banks for texture recognition, description, and segmentation," *Int. J. Computer Vision*, vol. 118, no. 1, pp. 65–94, May 2016.

[13] J. M. Coggins and A. K. Jain, "A spatial filtering approach to texture analysis," *Pattern Recognition Letters*, vol. 3, pp. 195–203, 1985.

[14] A. K. Jain and F. Farrokhnia, "Unsupervised texture segmentation using Gabor filters," *Pattern Recognition*, vol. 24, pp. 1167–1186, 1991.

[15] M. Unser, "Texture classification and segmentation using wavelet frames," *IEEE Tr. Image Proc.*, vol. 4, no. 11, pp. 1549–1560, Nov. 1995.

[16] B. S. Manjunath and W. Y. Ma, "Texture features for browsing and retrieval of image data," *IEEE Tr. Pattern Anal. Machine Intell.*, vol. 18, no. 8, pp. 837–842, Aug. 1996.

[17] V. Wouwer, G. Scheunders, P. Livens, and S. van Dyck, "Wavelet correlation signatures for color texture characterization," *Pattern Recognition*, vol. 32, pp. 443–451, 1999.

[18] M. N. Do and M. Vetterli, "Texture similarity measurement using Kullback-Leibler distance on wavelet subbands," in *Proc. Int. Conf. Image Proc.*, Vancouver, BC, Canada, Sept. 2000, vol. 3, pp. 730–733.

[19] W. T. Freeman and E. H. Adelson, "The design and use of steerable filters," *IEEE Tr. Pattern Anal. Machine Intell.*, vol. 13, no. 9, pp. 891–906, Sept. 1991.

[20] X. Zhao, M. G. Reyes, T. N. Pappas, and D. L. Neuhoff, "Structural texture similarity metrics for retrieval applications," in *Proc. Int. Conf. Image Proc. (ICIP)*, San Diego, CA, Oct. 2008, pp. 1196–1199.

[21] J. Zujovic, T. N. Pappas, D. L. Neuhoff, R. van Egmond, and H. de Ridder, "Effective and efficient subjective testing of texture similarity metrics," *Journal of the Optical Society of America A*, vol. 32, no. 2, pp. 329–342, Feb. 2015.

[22] M. Maggioni, G. Jin, A. Foi, and T. N. Pappas, "Structural texture similarity metric based on intra-class statistics," in *Proc. Int. Conf. Image Proc. (ICIP)*, Paris, France, Oct. 2014, pp. 1992–1996.

[23] G. Jin and T. N. Pappas, "Building structural similarity databases for metric learning," in *Human Vision and Electronic Imaging XX*, B. E. Rogowitz, T. N. Pappas, and Huib de Ridder, Eds., San Francisco, CA, Feb. 2015, vol. 9394 of *Proc. SPIE*, pp. 93940N–1–12.

[24] T. N. Pappas, "An adaptive clustering algorithm for image segmentation," *IEEE Tr. Signal Proc.*, vol. SP-40, no. 4, pp. 901–914, Apr. 1992.

[25] A. Mojsilović, J. Hu, and E. Soljanin, "Extraction of perceptually important colors and similarity measurement for image matching, retrieval, and analysis," *IEEE Tr. Image Proc.*, vol. 11, no. 11, pp. 1238–1248, Nov. 2002.

[26] T. N. Pappas, J. Chen, and D. Depalov, "Perceptually based techniques for image segmentation and semantic classification," *IEEE Commun. Mag.*, vol. 45, no. 1, pp. 44–51, Jan. 2007.

[27] L. He and T. N. Pappas, "An adaptive clustering and chrominance-based merging approach for image segmentation and abstraction," in *Proc. Int. Conf. Image Proc.*, Hong Kong, Sept. 2010, pp. 241–244.

[28] L. He, *A Clustering Approach for Color Texture Segmentation*, Ph.D. thesis, Northwestern Univ., Evanston, IL, Aug. 2012.

[29] I. Omer and M. Werman, "Color lines: Image specific color representation," in *IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2004, vol. 2, pp. 946–953.

[30] D. Nguyen, *Visual Texture Analysis for Material Understanding*, Ph.D. thesis, Northwestern Univ., Evanston, IL, Dec. 2017.

[31] J. Malik, S. Belongie, T. Leung, and J. Shi, "Contour and texture analysis for image segmentation," *Int. J. Comput. Vis.*, vol. 43, no. 1, pp. 7–27, 2001.

[32] E. M. Voorhees, "Variations in relevance judgments and the measurement of retrieval effectiveness," *Information Proc. & Management*, vol. 36, no. 5, pp. 697–716, Sept. 2000.

[33] Li Liu, P. Fieguth, Y. Guo, X. Wang, and M. Pietikäinen, "Local binary features for texture classification: Taxonomy and experimental study," *Pattern Recognition*, vol. 62, pp. 135–160, Feb. 2017.

[34] Li Liu, J. Chen, P. Fieguth, G. Zhao, R. Chellappa, and M. Pietikäinen, "From BoW to CNN: Two decades of texture representation for texture classification," *Int. J. Computer Vision*, vol. 127, no. 1, pp. 74–109, Jan. 2019.