3D Multi-frequency Fully Correlated Causal Random Field Texture Model

Michal Haindl[1,2(✉)] and Vojtěch Havlíček[1]

[1] Institute of Information Theory and Automation,
The Czech Academy of Sciences, Prague, Czechia
{haindl,havlicek}@utia.cz
[2] Faculty of Management, University of Economics, Jindřichův Hradec, Czechia
http://www.utia.cz/

Abstract. We propose a fast novel multispectral texture model with an analytical solution for both parameter estimation as well as unlimited synthesis. This Gaussian random field type of model combines a principal random field containing measured multispectral pixels with an auxiliary random field resulting from a given function whose argument is the principal field data. The model can serve as a stand-alone texture model or a local model for more complex compound random field or bidirectional texture function models. The model can be beneficial not only for texture synthesis, enlargement, editing, or compression but also for high accuracy texture recognition.

1 Introduction

The visual appearance of surface materials and object shapes are crucial for visual scene understanding or interpretation. Visual aspects of surface materials which manifest themselves as visual textures even if there is still missing a rigorous definition of the texture [12]. Thus solid visual scene modeling or interpretation cannot avoid a sound and physically correct texture model quality. The correct material modeling is hindered by the considerable variability of a physical appearance and thus its corresponding textural representation based on changing observation conditions. Several texture modeling methods were published, but most of them do not account for simultaneously variable illumination and viewing conditions.

Real surface material visual appearance is a very complex physical phenomenon which intricately depends on the incident and reflected spherical angles, time, and light spectrum among other at least 16 physical quantities [12]. Although, the general and physically correct material reflectance function should be at least sixteen dimensional [12] which is recently unmeasurable, and even if some simplifying assumptions have to be inevitably accepted, the essential dependencies have to be respected. Among them, these are spectral, illumination, and viewing parameters. Its state-of-the-art approximation, which allows expressing spectral, spatial, illumination angle, and observation angle

S. Palaiahnakote et al. (Eds.): ACPR 2019, LNCS 12047, pp. 423–434, 2020.
https://doi.org/10.1007/978-3-030-41299-9_33

visual dependencies of a measured material is the Bidirectional Texture Function (BTF). BTF significantly improves the visual realism of a modeled object at the expense of non-trivial measurements and mathematical modeling of these vast BTF data spaces. Static random field-based BTF texture modeling demands complex seven-dimensional models. It is far from being a straightforward generalization of any 3D model (required for customary static color textures) with supplementing just four additional dimensions.

Compound Markov random field models (CMRF) consist of several submodels each having different characteristics along with an underlying structure model which controls transitions between these submodels [18]. The exceptional CMRF [7,14] models allow analytical synthesis at the cost of a slightly compromised compression rate due to the non-parametric control field data. Methods based on different Markov random fields [3–5,11,13,15] combine an estimated range-map with synthetic multiscale smooth texture using Markov models. Any of the above CMRF or BTF-CMRF model are build from a set of simpler two or three dimensional random field models. Such a efficient novel three dimensional model is presented in this contribution.

The ideal synthetic texture should be visually indiscernible for any observation or illumination directions from the given measured natural texture sample. A qualitative analysis of modeling results requires a still non-existing reliable mathematical criterion or very impractical and expensive visual psycho-physics. Our results [8] illustrate that neither the standard image quality criteria (MSE [23], VSNR [1], VIF [21], SSIM [22], CW-SSIM [24]) nor the STSIM [25] or ζ [19] texture criteria can be reliably used for texture quality validation.
Our main contributions are the following:

- Introduction of a new efficient multispectral descriptive texture model which can be applied for high-quality material appearance modeling or recognition.
- Analytical Bayesian solution for the model parameters.
- Fast recursive model synthesis.

2 3D Multi-spectral Multi-frequency CAR Texture Model

The seven-dimensional bidirectional texture function (BTF) reflectance model is currently the state-of-the-art general reflectance function model, approximation which can be reliably measured [2,12]. However, to model such a function, we need its factorization to a set of lower two or three-dimensional models because they can be represented with significantly fewer amount of parameters to be estimated. The proposed multispectral model represents such a three-dimensional building factor for a high-quality BTF texture model.

The texture region (not necessarily continuous) is represented by a multi-frequency generalization (3DmfCAR) of the adaptive 3D causal auto-regressive random (3DCAR) field model [6,10] which combines the principal random field

Y with the auxiliary field $\tilde{Y}$. This model can be analytically estimated as well as easily synthesized. The 3DmfCAR model is defined:

$$\breve{Y}_r = \sum_{s \in I_r^c} A_s \breve{Y}_{r-s} + \breve{e}_r \qquad \forall r \in I \tag{1}$$

where $\breve{Y}_r = \begin{bmatrix} Y_r \\ \tilde{Y}_r \end{bmatrix}$ is a $\breve{d} \times 1$ vector, $\breve{d} = d + \tilde{d}$, $A_s = \begin{bmatrix} A_{s,1} \vdots A_{s,2} \\ \dots \ . \ \dots \\ A_{s,3} \vdots A_{s,4} \end{bmatrix}$ are $\breve{d} \times \breve{d}$ parametric matrices, Y_r $(d \times 1)$ is a pixel from image Y, I is a finite rectangular lattice, $r = [r_1, r_2, r_3]$ is a multiindex $r \in I$ where r_1, r_2, r_3 are row, column and spectral indices, and d is a number of spectral bands of the image Y. $\tilde{Y}_r$ $(\tilde{d} \times 1)$ is a pixel created by some processing of the measured image pixels

$$\tilde{Y}_r = f(Y_s \, \forall s \in I_r^c), \tag{2}$$

where the only restriction we assume is that the argument of the function $f(\,)$ is limited to pixels from the causal contextual neighborhood I_r^c.

Such an example can be a convolution with some convolution filter kernel h, i.e., $\tilde{Y}_r = h * y_r$ with zero padding for $s \notin I_r^c$ such as a low-pass filtered or upsampled Gaussian pyramid rough level. In this paper, we will further assume the function $f(\,)$ to be a median (3DmdCAR). The model (1) can be expressed in the matrix form:

$$\breve{Y}_r = \gamma \breve{X}_r + \breve{e}_r, \tag{3}$$

where γ is the $\breve{d} \times \breve{d}\,\eta_A$ parameter matrix $\gamma = [A_1, \dots, A_{\eta_A}]$, $\eta_A = card(I_r^c)$, I_r^c is a causal neighborhood, $\breve{e}_r = \begin{bmatrix} e_r \\ \tilde{e}_r \end{bmatrix}$ is a Gaussian white noise vector with zero mean and a constant but unknown $\breve{\Sigma}$ covariance matrix.

$$\breve{X}_r = [\breve{Y}_{r-s}^T : \ \forall s \in I_r^c], \tag{4}$$

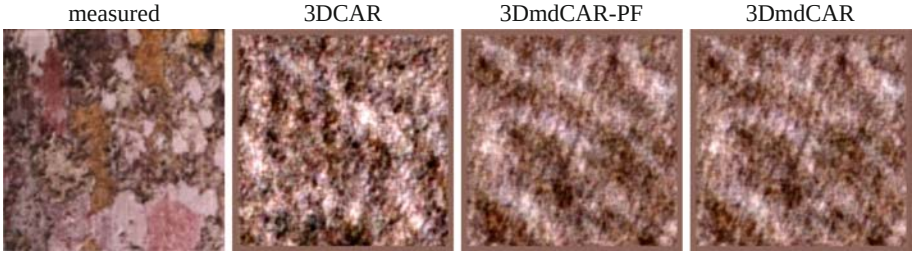

Fig. 1. Measured bark texture and its 3DCAR and 3DmdCAR synthesis, respectively.

2.1 Parameter Estimation

The model parameters $\gamma, \breve{\Sigma}$ are estimated from measured texture Y_{sample} and its modified version $\tilde{Y}_{sample}$ using Bayesian estimations.

These Bayesian parameter estimates (conditional mean values)

$$\hat{\gamma}_{r-1}^T = V_{xx(r-1)}^{-1} V_{xy(r-1)} \tag{5}$$

and

$$\hat{\Sigma}_{r-1} = \frac{\lambda_{(r-1)}}{\beta(r)}. \tag{6}$$

can be accomplished using fast, numerically robust and recursive statistics [10], given the known 3DmfCAR process history

$$\breve{Y}^{(t-1)} = \left\{ \breve{Y}_{t-1}, \breve{Y}_{t-2}, \ldots, \breve{Y}_1, \breve{X}_t, \breve{X}_{t-1}, \ldots, \breve{X}_1 \right\} :$$

$$\hat{\gamma}_{t-1}^T = V_{xx(t-1)}^{-1} V_{xy(t-1)}, \tag{7}$$

$$V_{t-1} = \tilde{V}_{t-1} + V_0, \tag{8}$$

$$\tilde{V}_{t-1} = \begin{pmatrix} \sum_{u=1}^{t-1} \breve{Y}_u \breve{Y}_u^T & \sum_{u=1}^{t-1} \breve{Y}_u \breve{X}_u^T \\ \sum_{u=1}^{t-1} \breve{X}_u \breve{Y}_u^T & \sum_{u=1}^{t-1} \breve{X}_u \breve{X}_u^T \end{pmatrix} = \begin{pmatrix} \tilde{V}_{yy(t-1)} & \tilde{V}_{xy(t-1)}^T \\ \tilde{V}_{xy(t-1)} & \tilde{V}_{xx(t-1)} \end{pmatrix}, \tag{9}$$

where t is the traversing order index of the sequence of multi-indices r and is based on the selected model movement in the lattice I, V_0 is a positive definite initialization matrix (see [10]) which is in our experiment the identity matrix.

The optimal causal functional contextual neighborhood I_r^c can be solved analytically by a straightforward generalization of the Bayesian estimate in [10].

2.2 Model Synthesis

The principal multispectral texture field synthesis can be computed using three possible simple and fast alternatives. The auxiliary random field can be similarly separately synthesized if there is a knowledge of the principal field. To simplify our notation, we will not differentiate measured Y_r (in Sect. 2.1) and synthesized Y_r from the model (this section) because their usage is clear from the context.

Complete Model Synthesis. The complete model $\breve{Y}$ synthesis uses direct application of the 3DmfCAR model Eq. (1)

$$\breve{Y}_r = \hat{\gamma} \breve{X}_r + \mathcal{N}(0, \hat{\Sigma}), \tag{10}$$

for simultaneous synthesis of the principal Y and auxiliary $\tilde{Y}$ random fields. Both random fields are causal thus the required contextual neighbors $Y_{r-s}, \tilde{Y}_{r-s}\ \forall s \in I_r^c$ are known from previous model synthesis steps.

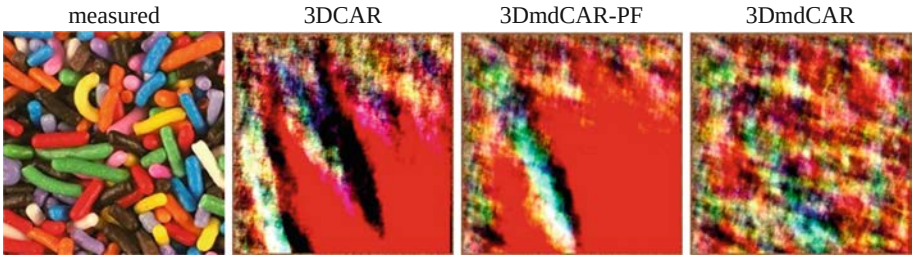

Fig. 2. Measured sweets texture and its unstable 3DCAR and 3DmdCAR-PF synthesis, respectively.

Complete Model Combined Synthesis. The alternative complete model $\breve{Y}$ synthesis (3DmfCAR-PF) combines the principal field Y synthesis using the model equation:

$$Y_r = \sum_{s \in I_r^c} (A_{s,1} Y_{r-s} + A_{s,2} \tilde{Y}_{r-s}) + e_r \qquad \forall r \in I \tag{11}$$

while the auxiliary field $\tilde{Y}$ is separately computed from (2) using just estimated neighbors $Y_s \in I_r^c$ as arguments for the corresponding function $f(\,)$ for computing $\tilde{Y}_r$ pixels.

Principal Model Synthesis. In most applications the auxiliary field $\tilde{Y}$ is not needed, and the model (11) is sufficient for the principal field Y synthesis.

Auxiliary Model Synthesis. If there is a need to synthesize the auxiliary field to a known principal field, it can be easily done using the corresponding part of the 3DmfCAR model equation:

$$\tilde{Y}_r = \sum_{s \in I_r^c} (A_{s,3} Y_{r-s} + A_{s,4} \tilde{Y}_{r-s}) + \tilde{e}_r \qquad \forall r \in I. \tag{12}$$

The auxiliary random field can be alternatively synthesized directly from the measured data ($Y_{measured}$) if there are used in (12) instead of estimated principal pixels Y_r, but then this field cannot be enlarged.

3 Texture Measurement Database

We verified the model on color textures cutouts from our large (more than 2000 high-resolution 7.9 MB (4288×2848) color textures categorized into 14 thematic classes and 20 subclasses) Prague color texture database [9,16]. All these textures are natural textures or man-made material textures. Some tested color textures are also from the VisTex database [20] where all textures have 512×512 size.

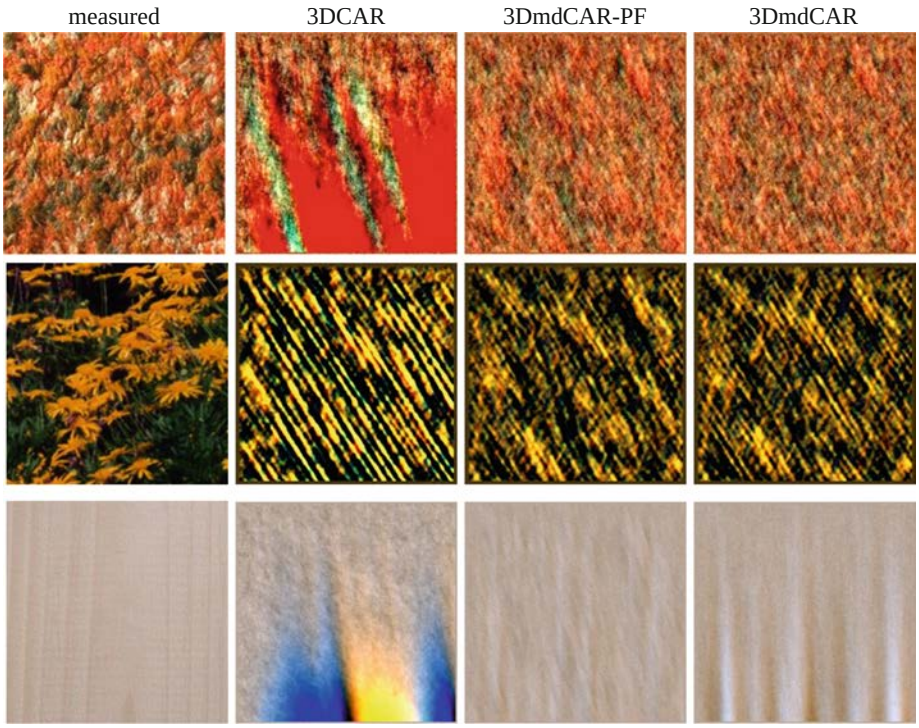

Fig. 3. Measured textile, flowers, and wood textures (first column), their unstable 3DCAR synthesis (second column) and the stabilizing effect of the additional auxiliary field in the 3DmdCAR model application.

4 Results

All our experiments were provided on color texture sets. Some textures were modeled using a single 3DmdCAR model, while others more complicated textures used a combination of several such models using the concept of the random compound field. Figure 1 illustrates the difference between simple 3DCAR model synthesis of a bark texture and both versions of the 3DmdCAR synthesis with all model using the same contextual neighborhood I_r^c. The 3DmdCAR model tends to stabilize some unstable 3DCAR model as is illustrated in three examples in Fig. 3 where all models in each row share the same contextual neighborhood I_r^c with the corresponding unstable 3DCAR models. Figure 2 suggests the stronger stabilizing effect of the complete model synthesis (3DmdCAR) over its combined synthesis (3DmdCAR-PF) alternative. Although the possible instability problem can be easily solved by just increasing the model order, this stabilizing tendency is the advantage of this otherwise more complex proposed model. Figure 4 shows visual quality improvement of the proposed 3DmdCAR model over our previously published 3DCAR model on the plant's texture example. The influence of the median window size on the visual synthetic texture appearance in a range of

Table 1. Spectral modeling per pixel error ζ (13) compared to the measured original.

	3DCAR	3DmdCAR-PF	3DmdCAR
Figure 1 bark	6.49	7.36	7.97
Figure 2 sweets	41.88	64.29	29.07
Figure 3 textile	23.54	8.30	8.89
Figure 3 flowers	16.62	10.91	10.80
Figure 3 wood	51.30	5.81	7.60
Figure 4 plants	16.96	8.93	8.61

median windows 2×2–11×11 is illustrated on the lichen texture in Fig. 5. The larger is the median estimation window; the more enhanced is the low frequencies in the model.

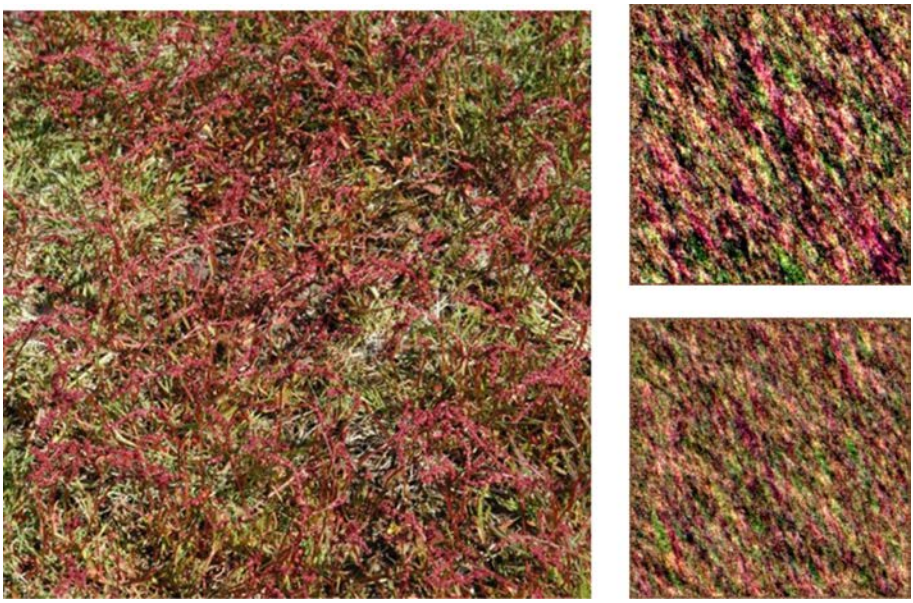

Fig. 4. Measured plants texture (left), its synthesis using 3DCAR (upper right) and 3DmdCAR (lower right) models.

The application of the presented 3DmdCAR model in the more complex compound model is illustrated in Fig. 6. The compound Markov random field models (CMRF) consist of several sub-models each was having different characteristics along with an underlying structure model which controls transitions between these sub-models. The non-parametric control field is estimated using the iterative method [15], and six local Markovian models are the presented 3DmdCAR model.

Table 2. Spectral modeling per pixel error ζ (13) compared to the measured original.

Median	3DCAR	3DmdCAR						
		2×2	3×3	4×4	5×5	7×7	9×9	11×11
Figure 5 lichen	16.46	10.87	10.87	11.19	17.20	10.90	10.75	11.03

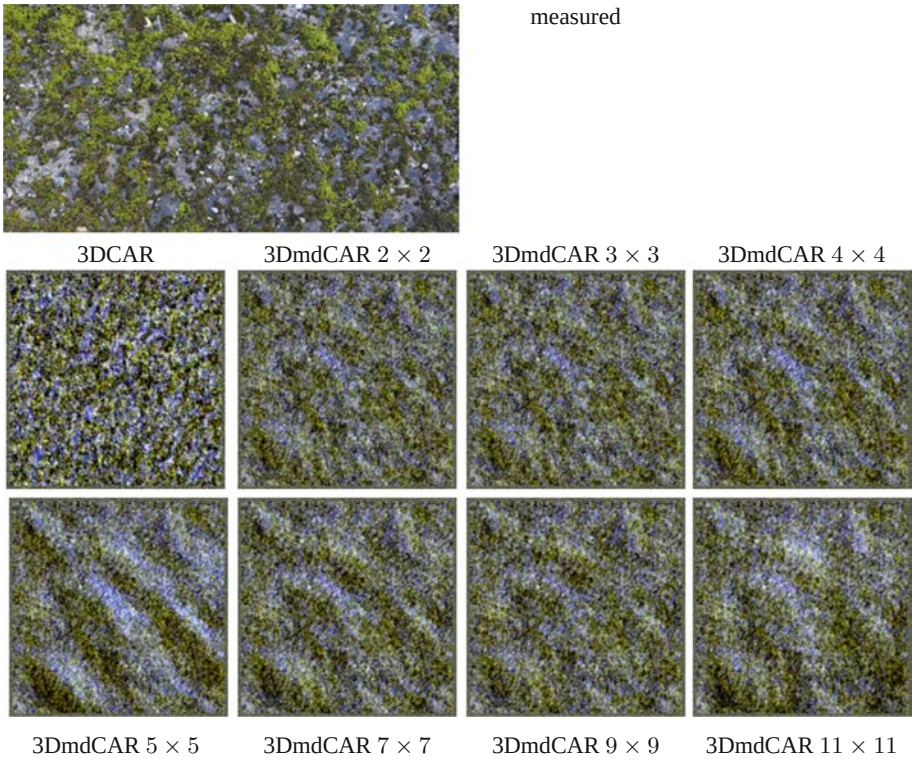

Fig. 5. Measured lichen texture, its 3DCAR synthesis, and several 3DmdCAR syntheses with gradually growing median filter windows.

The 3DmdCAR model can also be beneficially used in texture recognition (supervised or unsupervised) applications. Some preliminary results on the unsupervised bidirectional texture function segmentation can be checked on the Prague texture segmentation data-generator and benchmark [9]. The 3DmdCAR model with 7×7 median there achieves the best average rank over 21 benchmark criteria while the comparable 3DCAR model won only three individual segmentation criteria. Detailed results in these applications will be published elsewhere.

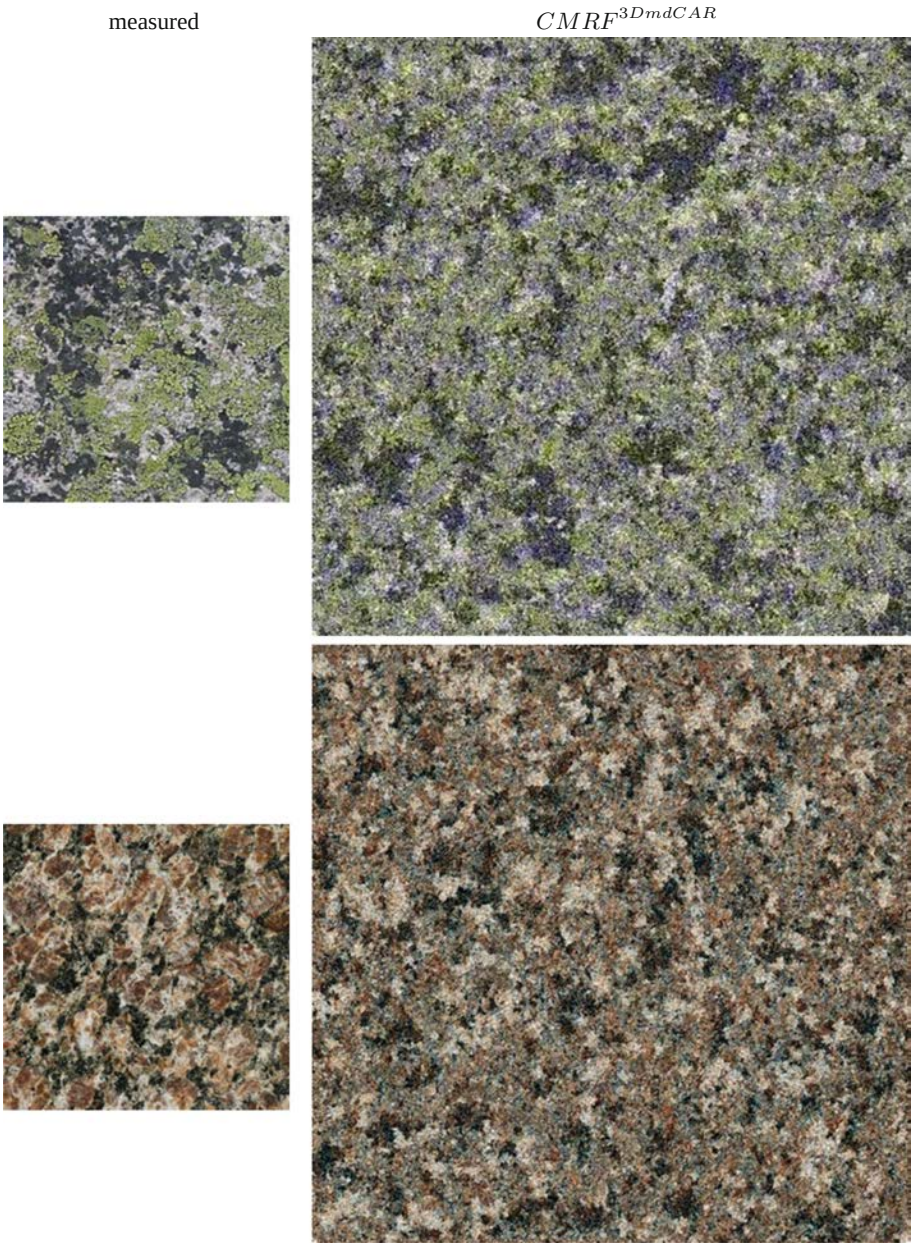

Fig. 6. Measured lichen and stone textures (left), and their enlarged synthesis (lichen upper, stone bottom) using a compound model with several 3DmdCAR submodels.

4.1 Texture Spectral Similarity

Although there is not any reliable texture similarity criterion, we can quantitatively measure spectral texture similarity. We have recently proposed [17] new reliable criterion for image spectral composition comparison. This ζ criterion simultaneously considers texture spectral similarity as well as the mutual ratios of similar pixels based on the mean exhaustive minimum distance:

$$\downarrow \zeta(A,B) = \frac{1}{M} \sum_{(r_1,r_2)\in\langle A\rangle} \min_{(s_1,s_2)\in U} \left\{\rho\left(Y^A_{r_1,r_2,\bullet}, Y^B_{s_1,s_2,\bullet}\right)\right\} \qquad \geq 0, \qquad (13)$$

where $Y^A_{r_1,r_2,\bullet}$ represents the pixel at location (r_1, r_2) in the image A, $\bullet$ denotes all the corresponding spectral indices, and similarly for $Y^B_{s_1,s_2,\bullet}$. Further, ρ is the maximum vector metric. U is the set of unprocessed pixel indices of B, $M = \min\{\sharp\{A\}, \sharp\{B\}\}$, $\sharp\{A\}$ is the number of pixels in A, and similarly for $\sharp\{B\}$. We define $\min\{\emptyset\} = 0$.

Table 1 illustrates color modeling quality presented in 3DmdCAR model examples. While the bark synthesis from the simpler 3DCAR model has comparable quality, their 3DmdCAR counterparts, textile, flowers, wood, and plants achieved much better appearance from the novel model. This can be expected for the unstable 3DCAR models (Fig. 3 - textile, wood) or unstable 3DmdCAR-PF (Fig. 2 sweets) model, but in the case of stable plants texture (Fig. 4) it is a less obvious conclusion. The ζ criterion value shows that the best spectral modeling quality was achieved for the 3DmdCAR-PF wood texture synthesis (Table 1) moreover, bark, textile, and plants results are slightly worse. The sweets texture, on the contrary, is the most complicated for this model. Similarly, Table 2 indicates the 9×9 median size to be optimal for the lichen texture although faster and smaller $2 \times 2, 3 \times 3$ medians can be used with slight compromise as well.

5 Conclusions

The presented 3DmfCAR model in its median version 3DmdCAR exhibits outstanding modeling capability on a wide range of natural or artificial color textures representing visual properties of surface materials. The model is inherently multispectral, and thus it can be used to model any number of spectral or hyperspectral bands. The major advantage of the model is that it can be analytically estimated as well as easily synthesized and used for seamless texture enlargement to fill any required size. The model tends to stabilize the simpler 3DCAR model of the same order. Visual properties of the model can be easily changed using a modification of the auxiliary field properties. The drawback of the proposed model is more parameters that have to be estimated, and thus also a large learning set required. However, even this more extensive learning set is negligible with the learning set required for convolutional neural networks.

The proposed model can be used either in its stand-alone version or as a primary factor for a more complex BTF or compound BTF models.

Acknowledgments. The Czech Science Foundation project GAČR 19-12340S supported this research.

References

1. Chandler, D.M., Hemami, S.S.: VSNR: a wavelet-based visual signal-to-noise ratio for natural images. IEEE Trans. Image Process. **16**(9), 2284–2298 (2007)
2. Dana, K.J., Nayar, S.K., van Ginneken, B., Koenderink, J.J.: Reflectance and texture of real-world surfaces. In: CVPR, pp. 151–157. IEEE Computer Society (1997)
3. Haindl, M., Filip, J.: Fast BTF texture modelling. In: Chantler, M. (ed.) Texture 2003. Proceedings, pp. 47–52. IEEE Press, Edinburgh, October 2003
4. Haindl, M., Filip, J.: A fast probabilistic bidirectional texture function model. In: Campilho, A., Kamel, M. (eds.) ICIAR 2004. LNCS, vol. 3212, pp. 298–305. Springer, Heidelberg (2004). https://doi.org/10.1007/978-3-540-30126-4_37
5. Haindl, M., Filip, J., Arnold, M.: BTF image space utmost compression and modelling method. In: Kittler, J., Petrou, M., Nixon, M. (eds.) Proceedings of the 17th IAPR International Conference on Pattern Recognition, vol. III, pp. 194–197. IEEE Press, Los Alamitos, August 2004. http://dx.doi.org/10.1109/ICPR.2004.1334501
6. Haindl, M., Havlíček, V.: A multiscale colour texture model. In: Kasturi, R., Laurendeau, D., Suen, C. (eds.) Proceedings of the 16th International Conference on Pattern Recognition, pp. 255–258. IEEE Computer Society, Los Alamitos, August 2002. http://dx.doi.org/10.1109/ICPR.2002.1044676
7. Haindl, M., Havlíček, V.: A compound MRF texture model. In: Proceedings of the 20th International Conference on Pattern Recognition, ICPR 2010, pp. 1792–1795. IEEE Computer Society CPS, Los Alamitos, August 2010. https://doi.org/10.1109/ICPR.2010.442. http://doi.ieeecomputersociety.org/10.1109/ICPR.2010.442
8. Haindl, M., Kudělka, M.: Texture fidelity benchmark. In: 2014 International Workshop on Computational Intelligence for Multimedia Understanding (IWCIM), pp. 1–5. IEEE Computer Society CPS, Los Alamitos, November 2014. https://doi.org/10.1109/IWCIM.2014.7008812. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7008812&isnumber=7008791
9. Haindl, M., Mikeš, S.: Texture segmentation benchmark. In: Lovell, B., Laurendeau, D., Duin, R. (eds.) Proceedings of the 19th International Conference on Pattern Recognition, ICPR 2008, pp. 1–4. IEEE Computer Society, Los Alamitos, December 2008. https://doi.org/10.1109/ICPR.2008.4761118. http://doi.ieeecomputersociety.org/10.1109/ICPR.2008.4761118
10. Haindl, M.: Visual data recognition and modeling based on local markovian models. In: Florack, L., Duits, R., Jongbloed, G., Lieshout, M.C., Davies, L. (eds.) Mathematical Methods for Signal and Image Analysis and Representation. Computational Imaging and Vision, vol. 41, pp. 241–259. Springer, London (2012). https://doi.org/10.1007/978-1-4471-2353-8_14
11. Haindl, M., Filip, J.: Extreme compression and modeling of bidirectional texture function. IEEE Trans. Pattern Anal. Mach. Intell. **29**(10), 1859–1865 (2007). https://doi.org/10.1109/TPAMI.2007.1139. http://doi.ieeecomputersociety.org/10.1109/TPAMI.2007.1139
12. Haindl, M., Filip, J.: Visual Texture. Advances in Computer Vision and Pattern Recognition. Springer, London (2013). https://doi.org/10.1007/978-1-4471-4902-6

13. Haindl, M., Havlíček, M.: Bidirectional texture function simultaneous autoregressive model. In: Salerno, E., Çetin, A.E., Salvetti, O. (eds.) MUSCLE 2011. LNCS, vol. 7252, pp. 149–159. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-32436-9_13. http://www.springerlink.com/content/hj32551334g61647/
14. Haindl, M., Havlíček, V.: A plausible texture enlargement and editing compound markovian model. In: Salerno, E., Çetin, A.E., Salvetti, O. (eds.) MUSCLE 2011. LNCS, vol. 7252, pp. 138–148. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-32436-9_12. http://www.springerlink.com/content/047124j43073m202/
15. Haindl, M., Havlíček, V.: BTF compound texture model with non-parametric control field. In: The 24th International Conference on Pattern Recognition (ICPR 2018), pp. 1151–1156. IEEE, August 2018. http://www.icpr2018.org/
16. Haindl, M., Mikeš, S.: A competition in unsupervised color image segmentation. Pattern Recogn. **57**(9), 136–151 (2016). https://doi.org/10.1016/j.patcog.2016.03.003. http://www.sciencedirect.com/science/article/pii/S0031320316000984
17. Havlíček, M., Haindl, M.: Texture spectral similarity criteria. IET Image Process. **13**(6), 1998–2007 (2019). https://doi.org/10.1049/iet-ipr.2019.0250
18. Jeng, F.C., Woods, J.W.: Compound Gauss-Markov random fields for image estimation. IEEE Trans. Signal Process. **39**(3), 683–697 (1991)
19. Kudělka, M., Haindl, M.: Texture fidelity criterion. In: 2016 IEEE International Conference on Image Processing (ICIP), pp. 2062–2066. IEEE, September 2016. https://doi.org/10.1109/ICIP.2016.7532721. http://2016.ieeeicip.org/
20. Pickard, R., Graszyk, C., Mann, S., Wachman, J., Pickard, L., Campbell, L.: Vistex database. Technical report, MIT Media Laboratory, Cambridge (1995)
21. Sheikh, H., Bovik, A.: Image information and visual quality. IEEE Trans. Image Process. **15**(2), 430–444 (2006)
22. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Process. **13**(4), 600–612 (2004). https://doi.org/10.1109/TIP.2003.819861
23. Wang, Z., Bovik, A.: Mean squared error: love it or leave it? A new look at signal fidelity measures. IEEE Signal Process. Mag. **26**(1), 98–117 (2009)
24. Wang, Z., Simoncelli, E.P.: Translation insensitive image similarity in complex wavelet domain. In: Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2005, pp. 573–576 (2005)
25. Zujovic, J., Pappas, T., Neuhoff, D.: Structural texture similarity metrics for image analysis and retrieval. IEEE Trans. Image Process. **22**(7), 2545–2558 (2013). https://doi.org/10.1109/TIP.2013.2251645