Multispectral Texture Fidelity Measure

Miloš Kudělka Jr   Michal Haindl

The Institute of Information Theory and Automation
of the Czech Academy of Sciences
Prague, Czech Republic
Email{kudelka,haindl}@utia.cz

Abstract—**Automatic texture fidelity assessment that would correspond to the human visual perception is an important, but still unsolved computer vision problem with numerous useful applications in the various vision application areas such as image compression and modeling, video streaming or fast image database retrieval. The problem is not satisfyingly solved even for the most simple static monospectral texture representation thus progress in the automatic assessment of texture fidelity is required. We propose improved multiresolution texture fidelity measure based on Markovian random field texture model, which correlates well with human texture fidelity evaluation obtained from texture fidelity benchmark.**

Keywords-**image texture; multispectral texture fidelity; visual quality measure; Markovian texture model;**

I. Introduction

The texture fidelity is an appearance similarity between the original target texture and its synthetic or natural alternative. The pair of high fidelity textures does not need to be pixelwise identical, but both textures should be visually indiscernible from each other, whatever the observation conditions might be. There are a lot of applications of texture fidelity assessment. Be it image compression, inpainting, and restoration, content-based image retrieval, searching for optimal image (texture) model parameters or comparison of different mathematical models. Such an assessment can be done through psycho-physical experiments [1], where a large number of humans need to evaluate image or texture fidelity or quality manually. It requires long experiment design, strictly controlled laboratory conditions and the experiment itself is very time-consuming. It is impractical and not feasible to perform such experiments in real-time situations on a daily basis.

This is a reason for the need of new texture measure that would evaluate quality automatically. There are published attempts to measure subjectively defined texture properties such as regularity [2], roughness, coarseness, directionality [3], etc. Others tries to test general texture quality [4]–[10]. In our previous work, we tested several state-of-the-art image quality and texture fidelity measures [11]. Measures we tested were - the mean-squared error (MSE) [12], the visual signal-to-noise-ratio [13] (VSNR), the structural similarity (SSIM) index [14], the complex wavelet - structural similarity (CW-SSIM) index [15], the visual

information fidelity (VIF) methods [16], and the structural texture similarity measure (STSIM-1, STSIM-2) [5]. All of these measures consider only gray-scale images, but our previously published measure ζ also uses multi-spectral information [10]. We have shown that all of the tested state-of-the-art image quality evaluation methods corresponds poorly to human perception and cannot be used at all for texture fidelity assessment. The only one aimed directly on textures is STSIM [5]. While it performs better than other image quality measures, its results are still not satisfying enough.

Both STSIM and our ζ measure display significantly higher correlation with human perception [10], [11], thus we compare our new proposed fidelity measure to these two measures.

A. Structural Texture Similarity Measure

The STSIM measures are based on a set of statistics computed for each texture subband factor. They are extensions of the CW-SSIM measure [15] in three versions, STSIM-1, STSIM-2, and STSIM-M [5]. STSIM-1 is created from CW-SSIM by replacing the 'structural' term with terms that compare first-order auto-correlations of corresponding subband coefficients $\rho_Y^m(0,1)$ in the horizontal and $\rho_Y^m(1,0)$ in the vertical direction. $Y, \widetilde{Y}$ are the target and a compared textures. Y_r is a pixel at location $r \in I$, where I is a discrete two-dimensional rectangular lattice, the multiindex $r = [r_1, r_2]$ is composed of r_1 row and r_2 column index, respectively. In the equations for a single subband m, the p is typically set to $p = 1$,

$$\text{STSIM-1}^m(Y, \widetilde{Y}) = \left(l_{Y,\widetilde{Y}}^\alpha \right)^{\frac{1}{4}} \left(c_{Y,\widetilde{Y}}^\alpha \right)^{\frac{1}{4}} \left(c_{Y,\widetilde{Y}}^{\alpha,\beta}(0,1) \right)^{\frac{1}{4}}$$
$$\left(c_{Y,\widetilde{Y}}^{\alpha,\beta}(1,0) \right)^{\frac{1}{4}} , \quad (1)$$

$$l_{Y,\widetilde{Y}}^\alpha = \left(\frac{2\mu_Y^\alpha \mu_{\widetilde{Y}}^\alpha + C_0}{(\mu_Y^\alpha)^2 + (\mu_{\widetilde{Y}}^\alpha)^2 + C_0} \right) ,$$

$$c_{Y,\widetilde{Y}}^\alpha = \left(\frac{2\sigma_Y^\alpha \sigma_{\widetilde{Y}}^\alpha + C_1}{(\sigma_Y^\alpha)^2 + (\sigma_{\widetilde{Y}}^\alpha)^2 + C_1} \right) ,$$

$$c_{Y,\widetilde{Y}}^{\alpha,\beta}(a,b) = 1 - 0.5 \left| \rho_Y^{\alpha,\beta}(a,b) - \rho_{\widetilde{Y}}^{\alpha,\beta}(a,b) \right|^p ,$$

$$\rho_Y^{\alpha,\beta}(a,b) = \frac{E\left\{ \left[Y_r^\alpha - \mu_Y^\alpha \right] \left[Y_{r_1+a,r_2+b}^\beta - \mu_Y^\beta \right]^* \right\}}{\sigma_Y^\alpha \sigma_Y^\beta} ,$$

where $l_{Y,\widetilde{Y}}^\alpha$ is a luminance, μ_Y^α is a mean value of the α-th subband of the texture Y, σ_Y^α is its standard deviation, $*$ denotes the complex conjugate, $\alpha = \beta = m$, and C_0, C_1 are small positive constants.

STSIM-2 [5] adds cross-band correlation coefficient $\rho_Y^{m,n}(0,0)$ between subbands m and n and $\alpha_i = m_i$, $\beta_i = n_i$

$$\text{STSIM-2}(Y,\widetilde{Y}) = \frac{\sum_{m=1}^{N_b} \text{STSIM-1}^m(Y,\widetilde{Y})}{N_b + N_c} +$$

$$\frac{\sum_{i=1}^{N_c} c_{Y,\widetilde{Y}}^{\alpha_i,\beta_i}(0,0)}{N_b + N_c} , \quad (2)$$

where N_b is the number of subbands and N_c is the number of possible cross-band correlations. STSIM-1, STSIM-2 $\in \langle 0; 1 \rangle$ with 1 being the best value. STSIM-M (STSIM-Mahalanobis) [5] chooses another approach. Rather than combining aforementioned terms into a single measure, it uses them to create feature vectors f_Y and $f_{\widetilde{Y}}$ and then calculates the *Mahalanobis distance* between the feature vectors. Therefore, to compute the distance between two textures, STSIM-M requires statistics based on the whole set and the results are relative only to this set, which is unfavorable for our cause and therefore the STSIM-M was not included in our tests.

B. ζ Texture Similarity Measure

The previously published [10] measure ζ measures cross-prediction error when using data from the original texture Y and estimated parameters $\widetilde{\gamma}$ (5) from the synthetic texture $\widetilde{Y}$

$$\zeta(Y,\widetilde{Y}) = \frac{1}{|I|} \sum_{\forall r \in I} |Y_r - \widetilde{\gamma}_{r-1} Z_r| . \quad (3)$$

This measure is based on the similar multispectral, but single-scale and unnormalized, texture model (3DCAR, see explanation in section II-A) as the newly proposed measure.

II. TEXTURE FIDELITY MEASURE

The proposed texture fidelity measure is derived from the generative multispectral textural model from the wide-sense Markovian random field family.

A. Generative Textural Model

Let us assume that multispectral texture image is composed of d spectral planes (e.g., $d = 3$ for simple colour textures). $Y_r = [Y_{r,1}, \ldots, Y_{r,d}]^T$ is the multispectral pixel at location $r \in I$, where I is discrete two dimensional rectangular lattice, the multiindex $r = [r_1, r_2]$ is composed of r_1 row and r_2 column index, respectively. The spectral planes are modeled by 3-dimensional Causal Auto-regressive Random (3DCAR) model. The 3DCAR representation assumes that the multispectral texture pixel Y_r can be modeled as a linear combination of its neighbors:

$$\begin{aligned} Y_r &= \gamma Z_r + \epsilon_r , \\ Z_r &= [Y_{r-s}^T : \forall s \in I_r]^T \end{aligned} \quad (4)$$

where Z_r is the $d\eta \times 1$ data vector with multiindices r, s, $\gamma = [A_1, \ldots, A_\eta]$ is the $d \times d\eta$ unknown parameter matrix with square sub-matrices A_s. Some selected contextual causal or unilateral neighbor index shift set is denoted I_r and $\eta = cardinality(I_r)$. A unilateral neighborhood I_r (the left upper orientation) is defined as $I_r \subset I_r^U = \{s : s_1 < r_1 \ or \ (s_1 = r_1, s_2 < r_2)\}$ and similarly ([17]) its subset - the causal neighborhood. The neighborhood order is based on the Euclidean distance from r. The white noise vector ϵ_r has normal density with zero mean and unknown covariance matrix Σ, same for each pixel. The texture is analyzed in a chosen direction, where multi-index t changes according to the movement on the image lattice. Parameter estimation of a 3DCAR model using either the maximum likelihood, or the least square or Bayesian methods can be found analytically. The Bayesian parameter estimates $\hat{\gamma}$ of the 3DCAR model using the normal-gamma parameter prior, given the known history of the random process $Y^{(t-1)} = \{Y_{t-1}, Y_{t-2}, \ldots, Y_1, Z_t, Z_{t-1}, \ldots, Z_1\}$ for the given pixel position can be computed using the following statistics [17]:

$$\hat{\gamma}_{t-1}^T = V_{zz(t-1)}^{-1} V_{zy(t-1)} , \quad (5)$$

$$\hat{\Sigma}_{r-1} = \frac{\lambda_{(r-1)}}{\beta(r) - d\eta + d + 1} , \quad (6)$$

$$\begin{aligned} V_{t-1} &= \begin{pmatrix} \sum_{u=1}^{t-1} Y_u Y_u^T & \sum_{u=1}^{t-1} Y_u Z_u^T \\ \sum_{u=1}^{t-1} Z_u Y_u^T & \sum_{u=1}^{t-1} Z_u Z_u^T \end{pmatrix} + V_0 \\ &= \begin{pmatrix} V_{yy(t-1)} & V_{zy(t-1)}^T \\ V_{zy(t-1)} & V_{zz(t-1)} \end{pmatrix} , \end{aligned} \quad (7)$$

$$\lambda_{t-1} = V_{yy(t-1)} - V_{zy(t-1)}^T V_{zz(t-1)}^{-1} V_{zy(t-1)} ,$$

where $\beta(r) = \beta(0) + r - 1$, $\beta(0)$ is an initialization constant, and the positive definite matrix V_0 represents a prior knowledge (see [17] for details). Moreover, the parameter estimates (5),(6) can be efficiently computed for all

pixel positions using a numerically robust recursive formula [17], which is advantageous for texture segmentation applications. Finally, the optimal contextual neighborhood I_r can be found analytically by maximizing the corresponding posterior probability [17]. The posterior probability density $p(Y_r \mid Y^{(r-1)}, \hat{\gamma}_{r-1})$ of the model can be easily evaluated as well [17] and used for the optimal model selection. The conditional mean value predictor of the one-step-ahead predictive posterior density for the normal-gamma parameter prior is

$$E\left\{ Y_r \mid Y^{(r-1)} \right\} = \hat{\gamma}_{r-1} Z_r \ . \qquad (8)$$
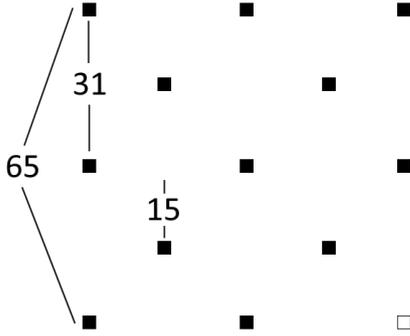
Figure 1. Functional neighborhood I_r. The white square is the central pixel r, black squares are its neighbors $(s \in I_r)$ and the numbers are the pixel distances between neighbors.

B. Contextual Neighborhood

Initially, we tried to use unilateral part of the classical hierarchical or simple square neighborhood. The results were not satisfying. Moreover, the order of such neighborhood had to be small because of the time complexity of the model estimation.

This is the reason we used a functional neighborhood I_r created by manually selecting 12 neighbors from the unilateral part of the square neighborhood (Fig. 1). With such neighborhood, we were able to use significantly higher neighborhood order without compromising the time complexity. The results show that this approach allowed to keep some information from the whole neighborhood. The neighborhood used in our measure is defined on the 65×65 window. Although such an optimal neighborhood can be estimated using, for example, a Bayesian statistics for each texture, we chose a universal sub-optimal neighborhood applicable for a wide range of our benchmark textures.

C. Measure

Before the calculation of the measure, we downsampled the textures twice by the factor of two. Then we upsampled the two images back to the original size and combined these images with the original one together, essentially

creating a 9-spectral $(d = 9)$ image. The idea was to add more information from lower frequencies. This approach proved successful, for results were better than using the original texture resolution only. Our new cross-prediction-based measure (CPM) is, similarly to the measure (3) [10], based on a cross-prediction but using the 3DCAR model. It measures the difference between prediction and cross-prediction

$$\text{CPM}(Y, \tilde{Y}^i) \ = \ \max\left\{ \beta(Y, \gamma, \tilde{\gamma}), \tilde{\beta}(\tilde{Y}^i, \gamma, \tilde{\gamma}) \right\}, \qquad (9)$$

$$\beta(Y, \gamma, \tilde{\gamma}) \ = \ \frac{1}{2^l d} \sum_{i=1}^{d} \alpha_i(Y, \gamma, \tilde{\gamma}) \ ,$$

$$\tilde{\beta}(\tilde{Y}^i, \gamma, \tilde{\gamma}) \ = \ \frac{1}{2^l d} \sum_{i=1}^{d} \tilde{\alpha}_i(\tilde{Y}^i, \gamma, \tilde{\gamma}) \ ,$$

$$\alpha(Y, \gamma, \tilde{\gamma}) \ = \ \frac{\sum_{\forall r \in I_{\{s\}}} (\tilde{\gamma}_{r-1} Z_r - \gamma_{r-1} Z_r)}{|I_{\{s\}}|} \ ,$$

$$\tilde{\alpha}(\tilde{Y}^i, \gamma, \tilde{\gamma}) \ = \ \frac{\sum_{\forall r \in I_{\{s\}}} \left(\gamma_{r-1}{}^i\tilde{Z}_r - \tilde{\gamma}_{r-1}{}^i\tilde{Z}_r \right)}{|I_{\{s\}}|} \ ,$$

where l is a number of bits per spectral band, $I_{\{s\}} \subset I$ is some window identical on both textures $Y, \tilde{Y}^i$, Y, Z_r, γ are data and parameters (5) from the original texture Y, while $\tilde{Y}^i, {}^i\tilde{Z}_r, \tilde{\gamma}$ are data and parameters (5) from the synthetic (or compared) i-th texture. Due to the pixel range normalization $\text{CPM}(Y, \tilde{Y}^i) \in \langle 0; 1 \rangle$ with 0 being the best value as it is a specific kind of an error measure.

The most significant difference from the ζ (3) [10] measure is that it calculates the cross-prediction from both "directions," it is normalized and takes the maximum. The idea behind this is that the higher the difference (i.e., the error), the more descriptive the "direction" on the current pair of textures. With this approach, the measure was performing better than the previous similar ζ criterion, regardless of parameters.

III. TEST DATA

The measures are verified on the data gathered from the texture fidelity benchmark [11], which has been created to help the validation of texture fidelity measures. The benchmark contains six real and one handmade artificial ([18]) color texture and their grayscale versions as target textures. Grayscale versions of textures number 2, 4 and 6 (i.e., 9, 11, and 13) were not included in the tests, because there were not enough data for these textures after human testing. The benchmark (test) textures are mathematically synthesized using various random field type models ([1], [19], [20]). We aimed to choose for the benchmark the whole range of synthetic textures with gradually decreasing texture quality as illustrated in Figure 2. Thus the Markovian
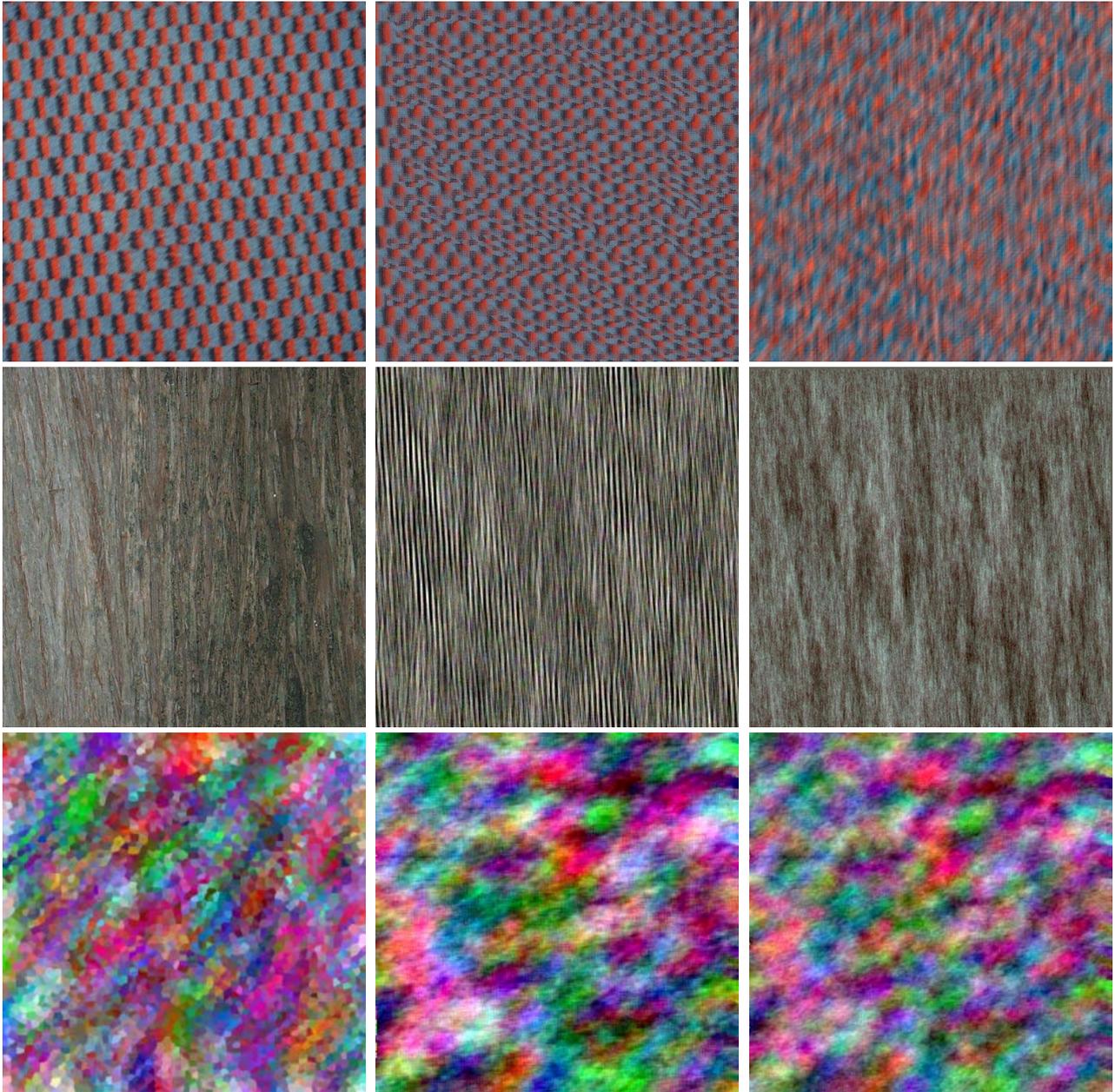
Figure 2. Examples of the original (left column) and synthetic textures (no. 1 upper row - textile, no. 5 middle row - wood, bottom row - the handmade artifical color texture [18]) .

models ([1], [19], [20]) were modified to have an inferior parameters setting and the observers would be able to rank their quality in the psychophysical evaluation. Each texture has up to 17 synthesized variants. Four examples of the benchmark synthetic textures can be seen in Figure 2. For more information about the benchmark and human ranking see [11].

IV. RESULTS

We observed the consistency between human and measure-based ranking for each texture fidelity measure. The Spearman's rank correlation coefficient (Tab.I) is used to compare human quality ranks obtained from the benchmark with the ranks stemming from the STSIM-1, STSIM-2, ζ, and our proposed measure. We only compare here results from the leading STSIM and ζ measures because they are

Table I
SPEARMAN'S RANK CORRELATION ρ BETWEEN THE HUMAN RANK AND THE MEASURES RESULTS. TEXTURE SETS 1–7 CONTAIN COLOR IMAGES, WHILE THE REMAINING HAVE GRAYSCALE TEXTURES.

texture no.	CPM	ζ	STSIM-1	STSIM-2
1	0.765	0.866	0.483	-0.539
2	0.706	0.797	0.505	0.400
3	0.723	0.801	0.860	0.571
4	0.797	0.363	0.857	0.813
5	0.782	0.721	0.636	0.370
6	0.859	0.600	0.767	0.767
7	0.534	0.670	0.420	0.174
8	0.711	0.566	0.608	-0.385
10	0.797	0.909	0.892	0.654
12	0.745	0.794	0.782	0.321
14	0.560	0.456	0.662	0.530
$\bar{\rho}$	**0.725**	0.686	0.679	0.334
$\min\{\rho\}$	**0.534**	0.363	0.420	-0.539

Table II
THE MEASURES VALUES, MEAN MEASURE VALUES ($\oslash$), THE DISTANCES OF THE MEAN MEASURE VALUE FROM THE OPTIMUM (δ^*), AND MEASURE VARIANCES, FOR THE GRAYSCALE TEXTILE TEXTURE NO. 10. THE FIRST COLUMN IS A LIST OF VARIOUS QUALITY MODELING EXPERIMENTS.

texture no. 10				
synthesis	CPM	ζ	STSIM-1	STSIM-2
1	0.052	4.5820	0.871	0.918
2	0.055	5.7472	0.873	0.892
3	0.035	5.5038	0.834	0.913
4	0.040	4.3384	0.901	0.930
5	0.062	4.8765	0.881	0.912
6	0.057	4.8806	0.902	0.926
7	0.031	4.5872	0.875	0.930
8	0.075	4.0194	0.922	0.934
9	0.063	4.8738	0.928	0.898
10	0.022	3.8918	0.937	0.943
11	0.092	4.5845	0.924	0.919
12	0.036	4.2380	0.927	0.899
13	0.044	4.5893	0.921	0.910
14	0.124	3.7684	0.951	0.940
15	0.032	3.8319	0.966	0.960
16	0.035	3.6978	0.958	0.952
17	0.109	3.7283	0.927	0.927
$\oslash$	**0.057**	4.455	0.912	0.924
δ^*	**0.057**	-	0.088	0.076
var	**0.001**	0.347	0.001	0.0004

Table III
MEASURES VARIANCES FOR ALL 14 INDIVIDUAL BENCHMARK TEXTURE SETS.

texture no.	CPM	ζ	STSIM-1	STSIM-2
1	0.0016	132.5659	0.0006	0.0003
2	0.0005	30.7192	0.0017	0.0006
3	0.0006	4.8728	0.0012	0.0003
4	0.0008	2.6779	0.0011	0.0003
5	0.0003	1.5416	0.0002	0.0002
6	0.0013	1.9117	0.0015	0.0001
7	0.0010	12.2320	0.0007	0.0001
8	0.0007	0.6327	0.0006	0.0003
10	0.0008	0.3471	0.0012	0.0004
12	0.0002	0.1025	0.0002	0.0001
14	0.0002	0.7608	0.0007	0.0002
median var	**0.0007**	1.9117	0.0007	0.0003

and STSIM for the texture no. 10 can be seen in Tab. II. The texture set no. 10. was chosen because the STSIM-1 results for this set have the highest variance of all benchmark texture sets and the set is grayscale. Hence our measure loses its multispectral advantage. Both CPM and STSIM-1 have similar variance. Tab. III presents the variances for all fourteen texture sets. The CPM median variance is the same with STSIM-1, even though STSIM-1 has slightly higher average variance value overall benchmark sets.

V. CONCLUSIONS

The presented CPM texture fidelity measure is the only genuine multispectral and multiresolution texture qualitative measure. CPM is based on the robust and recursively estimated generative Markovian texture model statistics. The measure is computed analytically and outperforms the best alternatives - STSIM or the previously published ζ fidelity measures. The Spearman's rank correlation illustrates the good correlation between human texture quality ranking and the proposed measure which is verified on the texture fidelity benchmark. The CPM measure has the highest minimum and average correlation from all four tested texture fidelity measures. The criterion values for individual textures also show a higher variance than STSIM-2 which we consider being an advantage for quality ranking between subtly distinguishable synthetic texture realizations. The measure can be favorably used also in various analytical applications such as texture classification or content-based image retrieval. Finally, most authors, including [5], separate color from the structural information when testing texture fidelity. We believe that the color cannot be omitted in a universal texture fidelity measure without compromising its performance and reliability.

the best existing texture fidelity measures known to us. Comparison with other, inferior, alternative measures (MSE, PSNR, VSNR, VIF, SSIM, CW-SSIM) can be consulted in [11]. Tab.I shows correlations of tested measures on all 14 texture sets. The ζ results, presented in [10], show that it performs equally well or better than STSIM. It can be seen, that CPM outperforms ζ and STSIM measures, as it has higher both average and minimum correlations. CPM values for a single pair of textures are normalized to $\langle 0;1 \rangle$, which is another advantage over ζ. CPM is the error measures. Thus its value for two identical textures is zero. The measures variance comparison between CPM

REFERENCES

[1] M. Haindl and J. Filip, *Visual Texture*, ser. Advances in Computer Vision and Pattern Recognition. London: Springer-Verlag London, January 2013.

[2] W.-C. Lin, J. Hays, C. Wu, Y. Liu, and V. Kwatra, "Quantitative evaluation of near regular texture synthesis algorithms," in *CVPR*. IEEE Computer Society, 2006, pp. 427–434.

[3] H. Tamura, S. Mori, and T. Yamawaki, "Textural features corresponding to visual perception," *IEEE Trans. on Sys., Man and Cyb.*, vol. 8, no. 6, pp. 460–472, June 1978.

[4] Y. Rubner, J. Puzicha, C. Tomasi, and J. M. Buhmann, "Empirical evaluation of dissimilarity measures for color and texture," *Computer vision and image understanding*, vol. 84, no. 1, pp. 25–43, 2001.

[5] J. Zujovic, T. Pappas, and D. Neuhoff, "Structural texture similarity metrics for image analysis and retrieval," *Image Processing, IEEE Transactions on*, vol. 22, no. 7, pp. 2545–2558, 2013.

[6] T. N. Pappas, D. L. Neuhoff, H. de Ridder, and J. Zujovic, "Image analysis: Focus on texture similarity," *Proceedings of the IEEE*, vol. 101, no. 9, pp. 2044–2057, 2013.

[7] Y. Zhai, D. L. Neuhoff, and T. N. Pappas, "Local radius index - a new texture similarity feature," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, May 2013, pp. 1434–1438.

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

[9] M. Alfarraj, Y. Alaudah, and G. AlRegib, "Content-adaptive non-parametric texture similarity measure," in *2016 IEEE 18th International Workshop on Multimedia Signal Processing (MMSP)*, September 2016, pp. 1–6.

[10] M. Kudělka and M. Haindl, "Texture fidelity criterion," in *2016 IEEE International Conference on Image Processing (ICIP)*. IEEE, September 2016, pp. 2062 – 2066. [Online]. Available: http://2016.ieeeicip.org/

[11] M. Haindl and M. Kudělka, "Texture fidelity benchmark," in *Computational Intelligence for Multimedia Understanding (IWCIM), 2014 International Workshop on*. Los Alamitos: IEEE Computer Society CPS, November 2014, pp. 1 – 5. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=\&arnumber=7008812\&isnumber=7008791

[12] Z. Wang and A. Bovik, "Mean squared error: Lot it or leave it? a new look at signal fidelity measures," *IEEE Signal Processing Magazine*, vol. 26, no. 1, pp. 98 – 117, 2009.

[13] D. M. Chandler and S. S. Hemami, "Vsnr: A wavelet-based visual signal-to-noise ratio for natural images," *Image Processing, IEEE Transactions on*, vol. 16, no. 9, pp. 2284–2298, 2007.

[14] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Processing*, vol. 13, no. 4, pp. 600–612, apr 2004. [Online]. Available: http://dx.doi.org/10.1109/TIP.2003.819861

[15] Z. Wang and E. P. Simoncelli, "Translation insensitive image similarity in complex wavelet domain," in *In Acoustics, Speech, and Signal Processing, 2005. Proceedings. (ICASSP 05). IEEE International Conference on*, 2005, pp. 573–576.

[16] H. Sheikh and A. Bovik, "Image information and visual quality," *Image Processing, IEEE Transactions on*, vol. 15, no. 2, pp. 430–444, 2006.

[17] M. Haindl, "Visual data recognition and modeling based on local markovian models," in *Mathematical Methods for Signal and Image Analysis and Representation*, ser. Computational Imaging and Vision, L. Florack, R. Duits, G. Jongbloed, M.-C. Lieshout, and L. Davies, Eds. Springer London, 2012, vol. 41, ch. 14, pp. 241–259, 10.1007/978-1-4471-2353-8_14. [Online]. Available: http://dx.doi.org/10.1007/978-1-4471-2353-8_14

[18] M. Havlíček and M. Haindl, "Texture spectral similarity criteria," in *Proceedings of the 4th CIE Expert Symposium on Colour and Visual Appearance*. Commission Internationale de L'Eclairage CIE Central Bureau, September 2016, pp. 147 – 154. [Online]. Available: http://div2.cie.co.at/?i_ca_id=985

[19] M. Haindl, "Visual data recognition and modeling based on local markovian models," in *Mathematical Methods for Signal and Image Analysis and Representation*. Springer, 2012, pp. 241–259.

[20] M. Haindl, J. Grim, P. Somol, P. Pudil, and M. Kudo, "A gaussian mixture-based colour texture model," in *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*, vol. 3. IEEE, 2004, pp. 177–180.