TEXTURE SPECTRAL SIMILARITY CRITERIA

Havlíček, M., Haindl, M.[1]
[1] Institute of Information Theory and Automation, Czech Academy of Sciences, Prague, CZECH REPUBLIC
{havlimi2,haindl}@utia.cas.cz

Abstract

Two novel criteria capable of assessing spectral similarity and modelling plausibility of synthetic Bidirectional Texture Functions (BTF), static colours textures are presented. The criteria credibly compare their spectral contents. Their primary aim is to support optimal modelling algorithms and measurement setup development by comparing the originally measured target texture with its synthetic simulations. The suggested spectral similarity criteria are extensively tested on measured natural BTF textures and an artificial distinctly coloured texture, and favourably compared with several alternative spectral similarity criteria. The performance quality of the proposed criteria is demonstrated on a long series of specially designed monotonically spectrally degrading experiments.

Keywords: Spectral Criteria, Texture Quality, Bidirectional Texture Functions

1 Introduction

An automatic texture, or more generally image, quality assessment, and mutual-similarity evaluation of two or more of them, presents a very important but still unsolved complex problem. Recent validation of the state-of-the-art image and texture fidelity criteria (Haindl and Kudělka, 2014) on the web-based benchmark (http://tfa.utia.cas.cz) has demonstrated that none of these published criteria can be used for the texture quality validation at all. There is still a pressing need for a reliable criterion for such a validation, e.g., to support BTF texture model development (i.e., a comparison of the original measured texture with a synthesized or reconstructed one, evaluation of optimal parameter settings for such a model) or texture database retrieval.

The Bidirectional Texture Function (BTF) (Haindl and Filip, 2012) is a 7-dimensional function describing surface texture appearance variations due to varying illumination and viewing angles. Because the appearance of real materials dramatically changes with illumination and viewing variations, any reliable representation of material visual properties requires capturing of its reflectance in as wide range of light and camera position combinations as possible. Thus this function is typically represented by thousands of images per material sample, each taken for a specific combination of the illumination and viewing condition. The primary purpose of any synthetic BTF texture approach is to reproduce and enlarge a given measured texture image so that ideally both natural and synthetic texture will be visually indiscernible and simultaneously to achieve a significant compression capability.

The similarity metrics also play an important role in efficient content-based image retrieval (e.g., from digital libraries, or multimedia databases). Surprisingly, many already developed approaches are limited to mono-spectral images, which is clearly a major disadvantage as colour is arguably the most significant visual feature.

The psycho-physical evaluations, i.e., quality assessments performed by humans, currently represent the only trustworthy alternative. Methods of this type require time-demanding experiment design setup, strictly controlled laboratory conditions, and representative sets of human testers, i.e., sufficient numbers of individuals, ideally from the general public, naive with respect to the purpose and design of the experiment. Such experiments are thus extremely impractical, expensive, generally demanding, and hence non-transferable into daily routine practice, operable on demand, and ideally in real time. For hyperspectral textures, such experiments are even impossible because not all spectra can be visualized simultaneously.

In this article we restrict our attention to only the spectral (e.g., colour) composition comparison, which represents only a partial answer to the image quality assessment, and propose a novel solution to this problem. Two introduced spectral criteria together with their previously published alternatives are extensively tested and mutually compared using our suggested test series.

The rest of the paper is organized as follows: Section 2 briefly presents existing possibilities to compare the image colour composition. Section 3 explains in detail our own new criteria, Section 4 describes the performed criteria validation experiments and shows the achieved results. Section 5 summarizes the paper with a discussion and compares our proposed criteria with their existing alternatives.

2 Alternative spectral measures

In this section we first briefly survey some of the existing criteria capable of comparing image colour compositions. The straightforward option for the image colour content comparison is to use a three-dimensional histogram, which approximates the image colour distribution. Let us denote by a_ϱ and b_ϱ the ϱ-th bin of the three-dimensional histogram of the images A and B, where A is the template visual texture or image and similarly B is the texture or image to be compared, and Y_r^A denotes the r-th multi-spectral pixel from the experimental image A where $r = [r_1, r_2, r_3]$ is a multi-index with row, column, and spectral components, respectively. The range of the histogram multi-index $\varrho = [i,j,k]$ depends on a colour space C in which the image is represented (for example in the standard 24-bit RGB colour space, the range of all three components of the multi-index is an integer from < 0; 255>).

The intuitive way is to compute the three-dimensional histograms difference:

$$\Delta H(A,B) = \sum_{\varrho \in C} |a_\varrho - b_\varrho| \quad , \tag{1}$$

which is a special case of the Minkowski distance (city block distance, also called Manhattan distance):

$$\Delta_q H(A,B) = \left(\sum_{\varrho \in C} |a_\varrho - b_\varrho|^q \right)^{1/q} \quad .$$

Plausible alternative values of the index q used in practice include the Euclidean distance (q = 2) or the Chebyshev distance (maximum / chessboard distance, q = 1)

$$\Delta_\infty H(A,B) = \sum_{\varrho \in C} \max \{ |a_i - b_i|, |a_j - b_j|, |a_k - b_k| \} \quad ,$$

where a_i; a_j; a_k represents 1st, 2nd and 3rd components of vector a_ϱ and similarly for b_i; b_j; b_k. For $0 < q < 1$ (fractional dissimilarity) the Minkowski distance is not a metric because it violates the triangle inequality (Howarth and Ruger, 2005).

Several other approaches for three-dimensional histogram comparison have been suggested, such as the histogram intersection (Swain and Ballard, 1991):

$$\cap H(A,B) = 1 - \frac{\sum_{\varrho \in C} \min \{a_\varrho, b_\varrho\}}{\sum_{\varrho \in C} b_\varrho} \quad , \tag{2}$$

the squared chord (Kokare et al., 2003):

$$d_{sc}(A,B) = \sum_{\varrho \in C} \left(\sqrt{a_\varrho} - \sqrt{b_\varrho} \right)^2 \tag{3}$$

and the Canberra metric (Kokare et al., 2003):

$$d_{can} = \sum_{C_0} \frac{|a_\varrho - b_\varrho|}{a_\varrho + b_\varrho} \ , \tag{4}$$

where $C_0 = \{\varrho : a_\varrho + b_\varrho \neq 0\} \subset C.$

Another measure, based on χ^2 statistic was suggested in (Zhang and Lu, 2003):

$$\chi^2(A,B) = \sum_{C_0} \frac{2\left(a_\varrho - \frac{a_\varrho + b_\varrho}{2}\right)^2}{a_\varrho + b_\varrho} \ . \tag{5}$$

The information theoretic measures can be also considered for evaluating the colour distribution differences. One possible option is a symmetric modification of the Kullback-Leibler divergence – a variant of the empirical Jeffrey divergence:

$$J(A,B) = \sum_{C^0} a_\varrho \log \frac{2a_\varrho}{a_\varrho + b_\varrho} + b_\varrho \log \frac{2b_\varrho}{a_\varrho + b_\varrho} \ , \tag{6}$$

where $C^0 = \{\varrho : a_\varrho b_\varrho \neq 0\} \subset C.$

The Jeffrey divergence is numerically stable, symmetric and robust with respect to noise and the size of histogram bins (Puzicha et al., 1997).

3 Proposed spectral-similarity criteria

The proposed spectral-similarity criteria allow the comparison of the spectral contents between two textures or arbitrary images.

A possible measure is to use a modified structural similarity metric (SSIM) (Wang et al., 2004) developed for texture comparison, as the texture-spectral-composition comparison might be considered a very special case of this task. The mono-spectral structural similarity metric (SSIM), which compares local statistics in corresponding sliding windows in two images in either the spatial or wavelet domain. This approach consists of three terms that compute and compare luminance, contrast and structure of the images. We have generalized this mono-spectral criterion to multispectral textures and removed the structure term, because in the case of the spectral quality comparison the structure term is irrelevant. It results in a redefined reduced rSSIM:

$$rSSIM(A,B) = \frac{1}{\sharp\{r_3\}} \sum_{\forall r_3} \frac{2\mu_{A,r_3}\mu_{B,r_3}}{\mu^2_{A,r_3} + \mu^2_{B,r_3}} \frac{2\sigma_{A,r_3}\sigma_{B,r_3}}{\sigma^2_{A,r_3} + \sigma^2_{B,r_3}} \ , \tag{7}$$

where $\#_{r3}$ is the spectral index cardinality, i.e., the number of spectral components, A_{r3} is the mean of space A_{r3} and Ar_3 is the standard deviation of space A_{r3}; similarly for B_{r3} and B_{r3}.

The second criterion we propose is the symmetric mean exhaustive minimum distance:

$$\zeta(A,B) = \nu(A,B) + \nu(B,A) \ , \tag{8}$$

$$\nu(A,B) = \frac{1}{M} \sum_{(r_1,r_2)\in\langle A\rangle} \min_{(\acute{r}_1,\acute{r}_2)\in N} \left\{ \rho\left(Y^A_{r_1,r_2,\bullet}, Y^B_{\acute{r}_1,\acute{r}_2,\bullet}\right)\right\} \ , \tag{9}$$

where $Y_{r1,\ r2,\bullet}^{A}$ represents the multi-spectral pixel at location (r_1, r_2) in the image A, • denotes all the corresponding spectral indices, and similarly for $Y_{\acute{r}1,\acute{r}2,\bullet}^{B}$. Further, ρ is an arbitrary vector metric (we used namely Manhattan, Euclidean and maximum metrics), N is the set of not yet used (during the counting, explained below) spatial indices of the image B pixels, M = min {#{A}, #{B}}, #{A} is the number of multi-spectral pixels in A, and similarly for #{B}. We define min {Ø} = 0.

The term v(A, B) is evaluated using the raster scanning. The algorithm scans the pixels in image A, from the left top, and searches for the pixel index in set N (which contains all spatial indices in the image B at the beginning of the process) for which the corresponding pixel is the closest one, in the sense of the used metric ρ. When such a pixel is found, the distance between this pixel and the scanned pixel from the image A, measured by ρ, is added to the sum and the pixel index of ($\acute{r}_1$, $\acute{r}_2$) is removed from the set N. The algorithm proceeds to the right bottom of the image A and stops when either it reaches that corner of the image A or N becomes an empty set. The term v (B, A) is computed similarly. Either v(A, B) or v(B, A) can be non-symmetric while ζ(A, B) is always symmetric.

Modifications of the proposed criterion (9) restricted to colour textures, which take into account colour differences just noticeable by colour psychometric methods in the CIE Lab space are easily possible.

Notice that the proposed spectral-similarity criterion ζ (8) is applicable to any number of spectral bands, not only for the usual three spectral bands of the standard colour images.

4 Comparison and Results

Both suggested spectral criteria together with their previously published alternatives have been extensively tested on the set of ten controllable degradation experiments. The main goal of the performed experiments is to investigate how the individual spectral similarity criteria are affected by the spectral distribution comparing the texture with its modified versions.

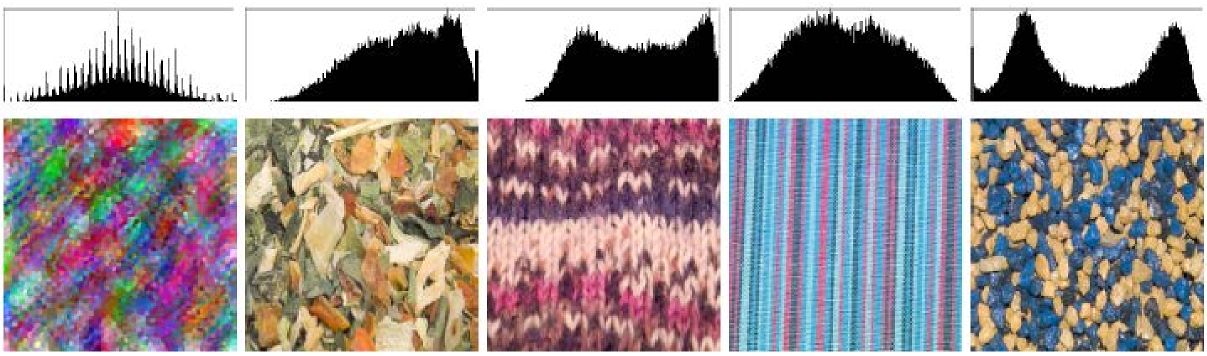

Figure 1 – 2D histograms comparison of our synthetic test colour2 texture (left) and several measured colour textures from the ALOT (Burghouts and Geusebroek, 2009) database

All evaluated criteria were tested on BTF texture space measurements (Figure 2 - wood01) (Muller et al., 2004) and colour textures (colour2, wood). The colour texture (256 x 256) has 11 375 distinct colours and it was deliberately manually created to have a 2D histogram (Figure 1 left) with numerous local extrema, which is expected to be demanding for most tested criteria. This assumption is obviously reflected in the results and confirms the quality of the proposed criteria. For the sake of comparison, the histograms of this image and several real measured textures are shown in Figure 1.

Figure 2 – One illustrative BTF wood01 sample from the 6561 measurements

The BTF wood measurements (Figure 2) are measured in all combinations of 81 different spherical illuminations and viewing angles totalling 6561 measured textures. Obviously it is not possible to run all experiments and to verify spectral quality for all, possibly, infinite number (any combination of the continuous spherical illumination and viewing angles) of synthetic BTF space texture components. The measured BTF data (usually several thousand colour images per material) are analysed for their intrinsic dimensionality (Haindl and Filip, 2012) and then subsequently approximated by a small number of BTF subspaces. The tested BTF wood01 measurement space is represented by twenty BTF subspace measurement clusters, which subsequently serve for building the BTF space's mathematical model for this wooden material.

4.1 Artificially degraded textural test series

The spectral similarity criteria are tested on sequences of gradually degraded textures which are generated from the original test texture. This source texture is the first member $A_1^X = A$ of each sequence, so that each subsequent member (except the original) of this sequence is generated from its predecessor in the sequence: $A_t^X = f\,(A_{t-1}^X)$, $t=1,\ldots,100$. Here $Y_{r,t}^A$ denotes the r-th multi-spectral pixel from the experimental image A_t^X, X is the corresponding experiment's label, and $r = [\, r_1, r_2, r_3 \,]$ is a multi-index with row, column, and spectral components, respectively.

We have created ten degraded textural series A - J. Series A, B, D, I, J are created by adding a constant either to one or all spectral channels in the RGB or CIE Lab colour spaces. In the C series the colour of each pixel is replaced with a colour which has the closest higher probability than the replaced pixel colour. The replacement is done with the candidate colour probability. E uses goniometric function based degradation. F adjusts each channel intensity to approach the average spectral intensity. G and H series use two modifications of a random pixel's substitution. Detailed description of all these degradation schemes is beyond the scope of this article and will published elsewhere.

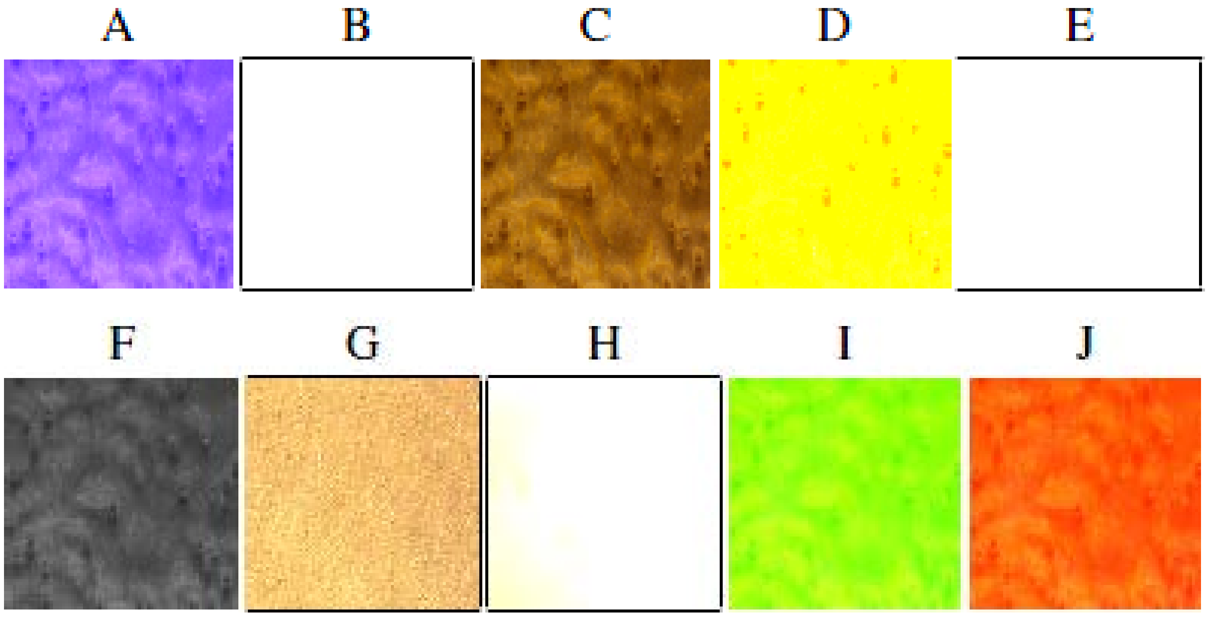

Figure 3 – Final degradations for the wood01 sample for all ten test series

Figure 3 illustrates the effect of single experimental worsening using the last hundredth degraded texture in each test series starting from the original wood BTF measurement on Figure 2. It is possible to see that while some test series (A, C, D, F, I, J) preserve the original spatial arrangement, the others change both spatial and spectral information.

	A	B	C	D	E	F	G	H	I	J	$\oslash$	rank
ΔH	28	17	0	41	73	33	92	4	47	39	37.4	6
$\cap H$	28	34	0	41	31	33	43	27	47	39	32.3	4
d_{sc}	33	16	0	41	74	40	90	5	49	45	39.3	8
d_{can}	15	11	0	58	76	2	83	1	20	24	29.0	3
J	34	36	0	47	46	25	43	24	48	45	34.8	5
χ^2	31	16	0	42	74	39	90	4	50	45	39.1	7
$rSSIM$	2	0	0	20	0	1	0	0	0	0	2.3	2
ζ	0	0	0	0	0	0	0	0	0	0	0.0	1

Table 1 - The strict monotonicity violation percentage of single criteria for the colour2 texture experimental sequences The worst performance in each experiment is highlighted in red.

The tested criteria are applied to quantifying spectral differences between the template texture and the remaining textures in the degradation sequence. For this evaluation all 10 x 100 degraded colour texture images per one tested colour or BTF space component texture are converted to the CIE L*a*b* colour space. As all those sequences are constructed so that monotone spectral composition change is guaranteed, a good criterion should be able to correctly follow this trend. The Table 1 presents the number of monotonicity violations of single criteria in the A-J evaluated experiments. Similar results were achieved also on all other tested textures.

The validation experiments show that it is possible to conduct spectral similarity checking using our criteria on any image from the synthetic BTF space, and this validation performance also holds for the remaining -- and possibly infinite -- number of synthetic images in the corresponding tested BTF space. Unlike many existing approaches, the criterion is not based on three-dimensional histograms, instead representing the estimate of the image spectral distribution, and requiring a sufficiently large data set, which is seldom available. Our criteria neither require the same size of the compared images, nor do they have any limit on the number of spectral bands. The proposed criterion ζ is the only one to rank flawlessly on all deteriorated textures in all controlled degradation experiments. The sSSIM (7) criterion makes sometimes a limited number of errors (Table 1) but ranked always as the second best criterion. The presented criteria propose a reliable fully automatic alternative to psycho-physical experiments, which are, moreover, extremely impractical due to their cost and strict demands on design setup, conditions control, human resources, and time.

5 Conclusions

We present two spectral criteria for comparing spectral similarity of the Bidirectional Texture Functions and colour images. This comparison represents a partial solution for assessing quality of the BTF and multi-spectral textures, as well as colour images. Although the criteria do not consider the spatial distribution of spectral information, they can assist in numerous texture-analytic or synthesis applications. The performance quality of the proposed criteria is demonstrated on a long series of specially designed monotonically spectrally degrading experiments, which also serve for the comparison with the existing alternative criteria. The proposed criterion ζ is the only one to perform faultlessly on all of our extensive validation tests. The validation experiments show that it is possible to conduct spectral similarity checking using our criteria on any image from the synthetic BTF space, and this validation performance also holds for the remaining -- and possibly infinite -- number of synthetic images in the corresponding tested BTF space. The criteria can be used for evaluating image spectral similarity of any images or textures and thus supports the texture-model development. Unlike many existing approaches, the ζ criterion is not based on three-dimensional histograms, instead representing the estimate of the image spectral distribution, and requiring a sufficiently large data set, which is seldom available. Our criteria neither require the same size of the compared images, nor do they have any limit on the number of spectral bands. The ζ criterion is slightly more time-demanding than some alternative criteria.

The presented criteria propose an automatic alternative to psycho-physical experiments, which are costly and impractical.

Acknowledgments

This research was supported by the Czech Science Foundation project GAČR 14-10911S.

References

BURGHOUTS, G. J. and GEUSEBROEK, J.-M. 2009. Material-specific adaptation of color invariant features, *Pattern Recognition Letters* 30(3), 306–313.

DICE, L. R. 1945. ‘Measures of the amount of ecologic association between species’, *Ecology* 26(3), 297–302.

HAINDL, M. and FILIP, J. 2012. Visual Texture, Advances in Computer Vision and Pattern Recognition, Springer-Verlag London, London.

HAINDL, M. and KUDˇELKA, M. 2014. Texture fidelity benchmark, in ‘Computational Intelligence for Multimedia Understanding (IWCIM), 2014 International Workshop on’, IEEE Computer Society CPS, Los Alamitos, pp. 1 – 5.
URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7008812&

HOWARTH, P. and RUGER, S. M. 2005. Fractional distance measures for content-based image retrieval., in D. E. Losada and J. M. Fernndez-Luna, eds, ‘ECIR’, Vol. 3408 of Lecture Notes in Computer Science, Springer, pp. 447–456.

JACCARD, P. 1901. Etude comparative de la distribution florale dans une portion des Alpes et du Jura, Impr. Corbaz.

KOKARE, M., CHATTERJI, B. and BISWAS, P. 2003. Comparison of similarity metrics for texture image retrieval, in ‘TENCON 2003. Conference on Convergent Technologies for the Asia-Pacific Region’, Vol. 2, IEEE, pp. 571–575.

MAHY, M., EYCKEN, L. and OOSTERLINCK, A. 1994. Evaluation of uniform color spaces developed after the adoption of cielab and cieluv, Color Research & Application 19(2), 105–121.

MINDRU, F., MOONS, T. and GOOL, L. V. 1998. Color-based moment invariants for viewpoint and illumination independent recognition of planar color patterns, in 'Illumination Independent Recognition of Planar Color Patterns, Proceedings ICAPR98', pp. 113–122.

MULLER, G., MESETH, J., SATTLER, M., SARLETTE, R. and KLEIN, R. 2004. Acquisition, synthesis and rendering of bidirectional texture functions, in Eurographics 2004, STAR - State of The Art Report, Eurographics Association, Eurographics Association, pp. 69–94.

PUZICHA, J., HOFMANN, T. and BUHMANN, J. M. 1997. Nonparametric similarity measures for unsupervised texture segmentation and image retrieval, in Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conf. on', pp. 267–272.

RUBNER, Y., TOMASI, C. and GUIBAS, L. J. 2000. The earth mover's distance as a metric for image retrieval, *International Journal of Computer Vision* 40(2), 99–121.
URL: http://dx.doi.org/10.1023/A:1026543900054

SWAIN, M. J. and BALLARD, D. H. 1991. Color indexing, *International Journal of Computer Vision* 7(1), 11–32. URL: http://dx.doi.org/10.1007/BF00130487

WANG, Z., BOVIK, A. C., SHEIKH, H. R. and SIMONCELLI, E. P. 2004. Image quality assessment: from error visibility to structural similarity, *IEEE Transactions on Image Processing* 13(4), 600–612.
URL: http://dx.doi.org/10.1109/TIP.2003.819861

WYSZECKI, G. and STILES,W. S. 1982. Color science, Vol. 8,Wiley New York.

ZHANG, D. and LU, G. 2003. Evaluation of similarity measurement for image retrieval, in Neural Networks and Signal Processing, 2003. Proceedings of the 2003 International Conference on, Vol. 2, IEEE, pp. 928–931.