Scale Sensitivity of Textural Features

Michal Haindl(✉) and Pavel Vácha

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

Abstract. Prevailing surface material recognition methods are based on textural features but most of these features are very sensitive to scale variations and the recognition accuracy significantly declines with scale incompatibility between visual material measurements used for learning and unknown materials to be recognized. This effect of mutual incompatibility between training and testing visual material measurements scale on the recognition accuracy is investigated for leading textural features and verified on a wood database, which contains veneers from sixty-six varied European and exotic wood species. The results show that the presented textural features, which are illumination invariants extracted from a generative multispectral Markovian texture representation, outperform the most common alternatives, such as Local Binary Patterns, Gabor features, or histogram-based approaches.

Keywords: Textural features · Texture scale recognition sensitivity · Surface material recognition · Markovian illumination invariant features

1 Introduction

Visual scene understanding is based on shapes and materials. While the shape is stable visual attribute, the surface material appearance vastly changes under variable observation conditions [3], which significantly and negatively affect its recognition as well as its realistic synthesis. Reliable computer-based interpretation of visual information which would approach human cognitive capabilities is very challenging and impossible without significant improvement of the corresponding sophisticated visual information models capable to handle huge variations of possible observation conditions. The appropriate paradigm for such a surface reflectance function models is a multidimensional visual texture. Generative visual texture models are useful not only for modelling physically correct virtual objects material surfaces in virtual or augmented reality environments or restoring images but also for contextual recognition applications such as segmentation, classification or image retrieval.

2 Textural Features

Numerous textural features have been published with miscellaneous recognition successfulness. Only the leading and commonly used textural features are

C. Beltrán-Castañón et al. (Eds.): CIARP 2016, LNCS 10125, pp. 84–92, 2017.
DOI: 10.1007/978-3-319-52277-7_11

selected for this texture scale sensitivity study. These are the two-dimensional causal auto-regressive (2DCAR), local binary patterns (LBP), Gabor, and colour histogram features, respectively.

2.1 2DCAR Illumination Invariant Features

The texture is factorised into K levels of the Gaussian down-sampled pyramid and subsequently each pyramid level is modelled by a wide-sense Markovian type of model - the Causal Auto-regressive Random (CAR) model. The model parameters are estimated and illumination invariants are computed from them. Finally, the illumination invariants from all the pyramid levels are concatenated into one feature vector. Let us assume that each multispectral (colour) texture is composed of C spectral planes (usually $C = 3$), $Y_r = [Y_{r,1}, \ldots, Y_{r,C}]^T$ is the multispectral pixel at location r. The multiindex $r = (r_1, r_2)$ is composed of row index r_1 and column index r_2. The spectral planes are mutually decorrelated by the Karhunen-Loeve transformation and subsequently modelled using either a three-dimensional model or a set of C two-dimensional models. The two-dimensional models assumes that the j-th spectral plane of pixel at position r can be modelled as:

$$Y_{r,j} = \gamma_j Z_{r,j} + \epsilon_r, \qquad Z_{r,j} = [Y_{r-s,j} : \forall s \in I_r]^T$$

where $Z_{r,j}$ is the $\eta \times 1$ data vector, $\gamma_j = [a_1, \ldots, a_\eta]$ is the $1 \times \eta$ unknown parameter vector. Some selected contextual neighbour index shift set is denoted I_r and $\eta = cardinality(I_r)$. The texture is analysed in a chosen direction, where multiindex t changes according to the movement on the image lattice I. Given the known CAR process history $Y^{(t-1)} = \{Y_{t-1}, Y_{t-2}, \ldots, Y_1, Z_t, Z_{t-1}, \ldots, Z_1\}$, $\hat{\gamma}$ can be estimated using fast, numerically robust recursive statistics [2]:

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

where the positive definite matrix V_0 represents prior knowledge. The following features are proved to be colour invariant [10,11]:

$$a_s\ ,\ \forall s \in I_r, \qquad \alpha_1 = 1 + Z_t^T\, V_{zz(t)}^{-1}\, Z_t,$$
$$\alpha_2 = \sqrt{\sum_{\forall r \in I} (Y_r - \hat{\gamma}_t Z_r)^T\, \lambda_t^{-1}\, (Y_r - \hat{\gamma}_t Z_r)},$$
$$\alpha_3 = \sqrt{\sum_{\forall r \in I} (Y_r - \mu)^T\, \lambda_t^{-1}\, (Y_r - \mu)}\ \},$$
$$\beta_1 = \ln\left(\tfrac{\psi(r)^C}{\psi(t)^C}\, |\lambda_t|\, |\lambda_r|^{-1}\right), \qquad \beta_8 = \left(\tfrac{\psi(r)^C}{\psi(t)^C}\, |\lambda_t|\, |\lambda_r|^{-1}\right)^{\frac{1}{2C}},$$
$$\beta_2 = \ln\left(\tfrac{\psi(r)^C}{\psi(t)^C}\, |V_{zz(t)}|\, |V_{zz(r)}|^{-1}\right), \quad \beta_9 = \left(\tfrac{\psi(r)^C}{\psi(t)^C}\, |V_{zz(t)}|\, |V_{zz(r)}|^{-1}\right)^{\frac{1}{2C\eta}},$$
$$\beta_3 = \ln\left(|V_{zz(t)}|\, |\lambda_t|^{-\eta}\right), \qquad \beta_{10} = \left(|V_{zz(t)}|\, |\lambda_t|^{-\eta}\right)^{\frac{1}{2C}},$$
$$\beta_4 = \ln\left(|V_{zz(t)}|\, |V_{yy(t)}|^{-\eta}\right), \qquad \beta_{11} = \left(|V_{zz(t)}|\, |V_{yy(t)}|^{-\eta}\right)^{\frac{1}{2C}},$$
$$\beta_5 = \mathrm{tr}\left\{V_{yy(t)}\ \lambda_t^{-1}\right\}, \qquad \beta_{12} = \sqrt{|V_{yy(t)}|\, |\lambda_t|^{-1}},$$

where μ is the mean value of Y_r. The dissimilarity between two feature vectors of two textures is computed using fuzzy contrast [8] in its symmetrical form FC_3.

Table 1. Classification accuracy for 2DCAR features and variable scale inconsistency between test and learning texture scales computed separately on all RGB spectral channels.

Scale	Test											
Train	50	55	60	65	70	75	80	85	90	95	100	∅
50	100	100	100	100	99.24	96.97	95.45	88.64	79.55	68.94	64.39	90.29
55	100	100	100	100	100	99.24	99.24	99.24	92.42	87.12	78.79	96.01
60	100	100	100	100	100	100	100	99.24	98.48	94.70	89.39	98.35
65	100	100	100	100	100	100	100	100	100	99.24	96.97	99.66
70	100	100	100	100	100	100	100	100	100	100	100	100
75	99.24	100	100	100	100	100	100	100	100	100	100	99.93
80	94.70	99.24	100	100	100	100	100	100	100	100	100	99.45
85	84.09	95.45	100	100	100	100	100	100	100	100	100	98.14
90	75.00	90.15	96.97	100	100	100	100	100	100	100	100	96.56
95	66.67	82.58	93.18	97.73	99.24	100	100	100	100	100	100	94.49
100	61.36	72.73	88.64	94.70	98.48	99.24	100	100	100	100	100	92.29

2.2 Local Binary Patterns

Local Binary Patterns [6] are histograms of texture micro patterns. For each pixel, a circular neighbourhood around the pixel is sampled, P is the number of samples and R is the radius of circle. The sampled point values are thresholded by the central pixel value and the pattern number is formed:

$$LBP_{P,R} = \sum_{s=0}^{P-1} \operatorname{sgn}(Y_s - Y_c)\, 2^s, \tag{1}$$

where *sgn* is the sign function, Y_s is the grey value of the sampled pixel, and Y_c is the grey value of the central pixel. Subsequently, the histogram of patterns is computed. Because of the thresholding, the features are invariant to any monotonic grey-scale change. The multiresolution analysis is done by growing of the circular neighbourhood size. All LBP histograms were normalised to have unit L_1 norm. The similarity between LBP feature vectors is measured by means of Kullback-Leibler divergence as the authors suggested. We have tested features $\mathrm{LBP}_{8,1+8,3}$, which are combination of features with radii 1 and 3. They were computed either on gray images or on each spectral plane of color image and concatenated. We also tested uniform version $\mathrm{LBP}^{u}_{16,2}$, but their results were inferior.

2.3 Gabor Features

The Gabor filters [1,7] can be considered as orientation and scale tunable edge and line (bar) detectors and statistics of Gabor filter responses in a given region

Table 2. Classification accuracy for LBP features on gray-scale textures.

Scale	Test											
Train	50	55	60	65	70	75	80	85	90	95	100	∅
50	100	100	99.24	93.94	75.76	63.64	46.21	35.61	26.52	20.45	19.70	61.91
55	100	100	100	99.24	93.94	79.55	65.15	45.45	36.36	28.79	26.52	70.45
60	99.24	100	100	100	99.24	95.45	81.82	65.91	46.21	38.64	38.64	78.65
65	86.36	99.24	100	100	100	100	96.21	83.33	66.67	50.00	46.97	84.44
70	71.97	91.67	100	100	100	100	100	98.48	85.61	71.97	63.64	89.39
75	48.48	71.21	93.94	100	100	100	100	100	98.48	87.12	83.33	89.33
80	39.39	51.52	72.73	96.97	100	100	100	100	100	99.24	93.94	86.71
85	27.27	38.64	53.03	74.24	96.97	100	100	100	100	100	95.45	80.51
90	22.73	28.79	40.15	57.58	77.27	98.48	100	100	100	100	96.97	74.72
95	18.18	21.97	28.79	40.91	59.85	84.09	98.48	100	100	100	96.21	68.04
100	16.67	20.45	28.03	37.12	52.27	76.52	89.39	93.94	96.21	96.21	100	64.26

are used to characterise the underlying texture information. A two dimensional Gabor function $g(r) : \mathbb{R}^2 \to \mathbb{C}$ can be specified as

$$g(r) = \frac{1}{2\pi\sigma_{r_1}\sigma_{r_2}} \exp\left[-\frac{1}{2}\left(\frac{r_1^2}{\sigma_{r_1}^2} + \frac{r_2^2}{\sigma_{r_2}^2}\right) + 2\pi i V r_1\right],$$

where $\sigma_{r_1}, \sigma_{r_2}, V$ are filter parameters. The convolution of a texture image and Gabor filter extracts edges of given frequency and orientation range. The whole filter set was obtained by four dilatations and six rotations of the function $g(r)$. The filter set is designed so that Fourier transformations of filters cover most of image spectrum, see [5] for details. Finally, given a single spectral image with values $Y_{r,j},\ r \in I,\ j = 1$, its Gabor wavelet transform is defined as

$$W_{k\phi,j}(r_1, r_2) = \int\limits_{u_1,u_2\in\mathbb{R}} Y_{r,j}\ g^*_{k\phi}(r_1 - u_1, r_2 - u_2)\ du_1\, du_2,$$

where $(\cdot)^*$ indicates the complex conjugate, ϕ and k are orientation and scale of the filter. The Gabor features [5] are defined as the mean μ_j and the standard deviation σ_j of the magnitude of filter responses W. The Gabor features of colour images have been computed either on grey images or on each spectral plane separately and concatenated to form a feature vector. The distance between two textures T, S is measured as the sum:

$$L1_{STD}(T, S) = \sum_{i=0}^{p} \left| \frac{f_i^{(T)} - f_i^{(S)}}{\sigma(f_i)} \right|, \tag{2}$$

where $\sigma(f_i)$ is standard deviations of a feature f_i computed over all database, and p is the size of the feature vector. Another extension of the Gabor filters to colour textures [4] is based on adding a chromatic antagonism, while the spatial

antagonism is modelled by Gabor filters themselves. Opponent Gabor features consists of the monochrome part of features: $\eta_{i,m,n} = \sqrt{\sum_r W^2_{i,m,n}(r)}$, where $W_{i,m,n}$ is the response to Gabor filter of orientation m and scale n, i is $i-$th spectral band of the colour texture T. The opponent part of features is:

$$\psi_{i,j,m,m',n} = \sqrt{\sum_r \left(\frac{W_{i,m,n}(r)}{\eta_{i,m,n}} - \frac{W_{j,m',n}(r)}{\eta_{j,m',n}} \right)^2},$$

for all i, j with $i \neq j$ and $|m - m'| \leq 1$. (Opponent features could be also expressed as correlation between spectral planes responses.) The distance between textures T, S using the Opponent Gabor features is measured as sum

$$L2_{STD}(T,S) = \sqrt{\sum_{i=0}^{p} \left(\frac{f_i^{(T)} - f_i^{(S)}}{\sigma(f_i)} \right)^2}, \tag{3}$$

where $\sigma(f_i)$ is standard deviations of feature f_i computed over all database, and p is the size of feature vector.

2.4 Histogram Based Features

The simplest features used in this study are based on histograms of colours or intensity values. Although, these features cannot be considered as proper textural features, because they are not able to describe spatial relations which are the key texture properties, their advantage is robustness to various geometrical transformations, fast, and easy implementation. The cumulative histogram proposed in [9] is defined as the distribution function of an image histogram. The i-th bin H_i is computed as $H_i = \sum_{\ell \leq i} h_\ell$, where h_ℓ is the ℓ-th bin of ordinary histogram. The distance between two cumulative histograms is computed in L_1 metric. The cumulative histogram is more robust than the ordinary histogram, because a small intensity change characterised by a one-bin shift in the ordinary histogram, have only negligible effect on the cumulative histogram.

3 Experiments

The scale sensitivity of the selected textural features was tested on the wood database, which contains veneers from varied European and exotic wood species, each with two sample images only. The original images of 66 wood species were acquired by a colour scanner device. These images were scaled down to $95\%, 90\%, 85\%, \ldots, 50\%$ of their original size. Finally, regions with the same resolution were cropped out, see examples in Fig. 1. As a consequence, image of scale 50% covers doubles size of the original texture image, but with half of details than scale 100%. In the experiment, the training set was composed of images of a single selected scale and the classification accuracy was tested for all scales, separately. To avoid training and testing on the same images the scaled

Table 3. Classification accuracy for colour histogram and opponent Gabor features.

Scale	Test											
Train	50	55	60	65	70	75	80	85	90	95	100	∅
Colour histogram features												
50	100	100	98.48	96.21	93.94	91.67	87.12	82.58	79.55	76.52	65.91	88.36
55	100	100	100	98.48	96.97	95.45	92.42	88.64	85.61	81.82	70.45	91.80
60	98.48	100	100	100	99.24	99.24	96.21	93.18	87.12	86.36	74.24	94.01
65	96.21	99.24	100	100	100	100	99.24	97.73	92.42	90.91	78.79	95.87
70	93.18	98.48	100	100	100	100	100	97.73	96.21	94.70	82.58	96.63
75	90.15	93.94	98.48	100	100	100	100	100	97.73	95.45	84.85	96.42
80	84.09	92.42	93.18	99.24	100	100	100	100	100	96.97	85.61	95.59
85	82.58	87.12	90.15	93.94	99.24	100	100	100	100	100	87.12	94.56
90	78.79	82.58	85.61	90.91	95.45	98.48	99.24	100	100	100	86.36	92.49
95	74.24	78.03	82.58	85.61	92.42	95.45	98.48	99.24	100	100	87.12	90.29
100	65.91	73.48	76.52	78.79	78.79	83.33	85.61	87.12	86.36	86.36	100	82.02
Opponent gabor features												
50	100	100	100	98.48	96.21	92.42	87.12	82.58	75.00	67.42	62.12	87.40
55	100	100	100	100	100	99.24	96.97	93.18	87.12	80.30	70.45	93.39
60	100	100	100	100	100	100	98.48	98.48	96.97	93.18	85.61	97.52
65	100	100	100	100	100	100	100	98.48	98.48	96.21	90.15	98.48
70	100	100	100	100	100	100	100	100	100	96.97	95.45	99.31
75	99.24	100	100	100	100	100	100	100	100	99.24	96.97	99.59
80	95.45	99.24	100	100	100	100	100	100	100	100	97.73	99.31
85	87.88	96.21	99.24	100	100	100	100	100	100	100	98.48	98.35
90	81.06	90.91	98.48	100	100	100	100	100	100	99.24	97.73	97.04
95	72.73	85.61	93.18	97.73	99.24	100	100	100	100	100	98.48	95.18
100	62.12	75.00	87.88	93.94	98.48	98.48	99.24	99.24	99.24	99.24	100	92.08

images were split to upper half and lower half with 812×1034 pixel resolution. The final result were computed as an average of classification accuracy tested on lower and upper halfs (training was performed on the other half). The computed feature vectors were compared using the suggested distances and classified with the Nearest Neighbour (1-NN) classifier. Scale sensitivity of textural features is visualised in the graph on Fig. 2, which is created from two experiment results. The left part is based on training on scale 100% and testing on downscaled images $100\%, 95\%, \ldots, 50\%$, while the right part (scaling factor 1–2) uses training on scale 50% and testing on $55\%, 60\%, \ldots, 100\%$, which simulates upscaled images which avoids heavy interpolation. Results on Tables 1, 2, 3 and Fig. 2 illustrate the best and the most robust scale invariant performance of the 2DCAR features, which benefits from multiscale approach, followed by opponent

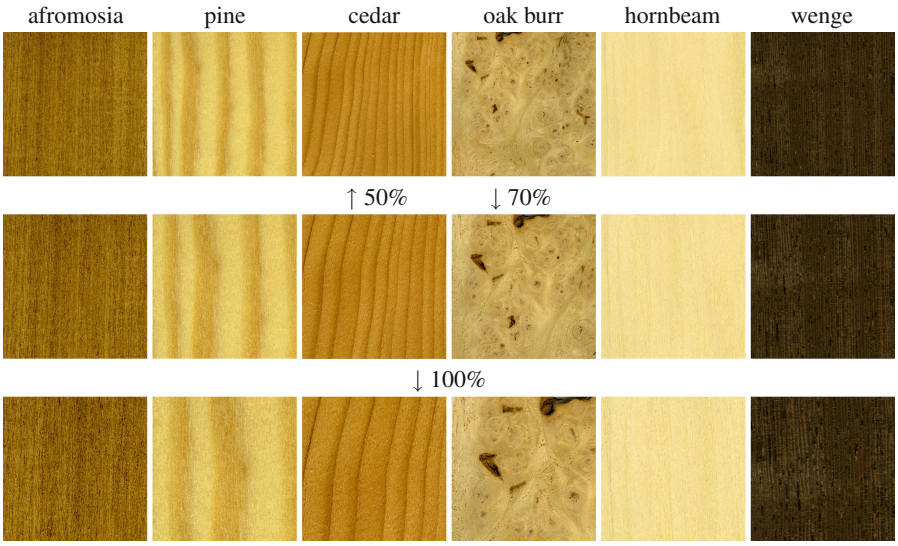

Fig. 1. The afromosia, pine, cedar, oak burr, hornbeam, and wenge veneer textures, respectively. The rows correspond to different resolution (50% top, 70% middle, 100% bottom) setups.

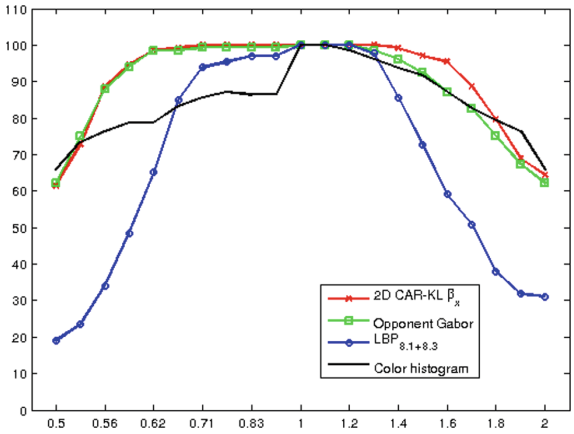

Fig. 2. Classification accuracy (y: [%]) for test x: scaling factor(0.5–2), training at scale factor 1, and average over 66 classes.

Gabor features. Histogram feature are the least sensitive to extreme scale variation, but they simultaneously they lack sufficient discriminability for recognition at similar scales. LBP are very sensitive to scale changes, because they have small support. Multiscale approach could probably improve their performance, but it will increase the already large number of LBP features.

4 Conclusion

The presented results indicate that Markovian illumination invariant texture features (2DCAR), based on Markovian descriptive model, are the most robust textural features for texture classification when learning and classified textures differ in scale. The 2DCAR features outperformed tested textural features, i.e., the LBP, Gabor or histogram texture features, respectively. Their additional advantage is their fast and numerically robust estimation. Additionally, our colour Markovian textural features were successfully applied in recognition of wood veneers using a smart phone camera. The method's correct recognition accuracy improvements are between 20% and 40%, compared to the Local Binary Patterns (LBP) features and up to 8% compared to the opponent Gabor features which is the second best alternative from all tested textural features.

It is worth to note that the presented results apply for recognition with bounded scale variation. If the expected scale variation would go to extremes, the fully scale invariant textural features should be considered. On the other hand, fully invariant features usually losses some discriminability, therefore each application need to carefully balance invariance to expected variability and discriminability. In the future, the presented results on wood recognition will be extended to a larger and more general texture database.

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

References

1. Bovik, A.: Analysis of multichannel narrow-band filters for image texture segmentation. IEEE Trans. Sig. Process. **39**(9), 2025–2043 (1991)
2. 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). chap. 14
3. Haindl, M., Filip, J.: Visual Texture. Advances in Computer Vision and Pattern Recognition. Springer-Verlag, London (2013)
4. Jain, A.K., Healey, G.: A multiscale representation including opponent color features for texture recognition. IEEE Trans. Image Process. **7**(1), 124–128 (1998)
5. Manjunath, B.S., Ma, W.Y.: Texture features for browsing and retrieval of image data. IEEE Trans. Pattern Anal. Mach. Intell. **18**(8), 837–842 (1996)
6. Ojala, T., Pietikäinen, M., Mäenpää, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans. Pattern Anal. Mach. Intell. **24**(7), 971–987 (2002)
7. Randen, T., Husøy, J.H.: Filtering for texture classification: a comparative study. IEEE Trans. Pattern Anal. Mach. Intell. **21**(4), 291–310 (1999)
8. Santini, S., Jain, R.: Similarity measures. IEEE Trans. Pattern Anal. Mach. Intell. **21**(9), 871–883 (1999)
9. Stricker, M.A., Orengo, M.: Similarity of color images. In: SPIE, vol. 2420, pp. 381–392 (1995)

10. Vácha, P., Haindl, M.: Natural material recognition with illumination invariant textural features. In: Proceedings of the 20th International Conference on Pattern Recognition, ICPR 2010, pp. 858–861. IEEE Computer Society CPS, Los Alamitos (August 2010)
11. Vacha, P., Haindl, M.: Image retrieval measures based on illumination invariant textural MRF features. In: Proceedings of the 6th ACM International Conference on Image and Video Retrieval, CIVR 2007, pp. 448–454. NY, USA. ACM, New York (2007)