https://doi.org/10.2352/ISSN.2470-1173.2017.13.IPAS-199

Texture representations in different basis functions for image synthesis using system criteria analysis

V. Voronin[a], V. Ryzhov[b], V. Marchuk[a], S. Makov[a]; [a]Don State Technical university, Dept. of Radio-Electronics Systems, Gagarina 1, Rostov on Don, Russian Federation; [b]Southern Federal University, Nekrasovski 44, Taganrog, Russian Federation;

Abstract

This paper introduces a texture representation suitable for image synthesis of textured surfaces. An efficient representation for natural images is of fundamental importance in image processing and analysis. The automated analysis of texture is widely applied in a number of real-world applications, e.g., image and video retrieval, object recognition and classification. For texture representation we consider the orthogonal decomposition of two-dimensional signals (images) using spectral transform in the different basis functions. This paper focuses on the analysis of the following basis functions Fourier, Walsh, Haar, Hartley and cosine transform using system criteria analysis. This criterion includes error signal representation and computational cost. For correct calculation of the components of the system criterion we use statistical averaging. It is shown that the Haar transform can represent textural patches more efficiently with smaller average risk than other basis functions. The texture representations results compare favourably against other state-of-the-art directional representations.

Introduction

Texture plays an important role in numerous computer vision applications. Image texture represents an image area contributing repetition of patterns of pixel intensities arranged in some structural way. Textures are prominent in natural images (as in grasslands, brick walls, fabrics, etc.). Many useful properties for image description and interpretation are gained through texture observation and analysis such as granularity, smoothness, coarseness, periodicity, geometric structure, orientation etc.

This paper introduces a texture representation suitable for image synthesis of textured surfaces. An efficient representation for natural images is of fundamental importance in image processing and analysis. The automated analysis of texture is widely applied in a number of real-world applications, e.g., image and video retrieval, object recognition and classification. For texture representation we consider the orthogonal decomposition of two-dimensional signals (images) using spectral transform in the different basis functions.

The visual perception of textures has been an area of interest spanning a wide variety of disciplines from art to computer science. The fields of computer vision, perception, and graphics have each made significant contributions to our overall understanding of texture perception and representation, albeit in quite different ways. An efficient representation for natural images is of fundamental importance in image processing and analysis. The commonly used separable transforms such as wavelets axe not best suited for images due to their inability to exploit directional regularities such as edges and oriented textural patterns; while most of the recently proposed directional schemes cannot represent these two types of features in a unified transform.

The quality of the final reconstruction produced by known algorithms informs us as to the utility of both the representation used for the original texture and the process by which that representation is used to generate novel images. However, for us to truly feel confident in relating the computational procedure used for texture synthesis to human perceptual processes it is helpful if the algorithm uses representations employed by the human visual system. Synthesis requires a time-consuming search process through the sample provided for analysis.

Related Work

The automated analysis of image textures has been the topic of extensive research in the past years. Any methods for analyzing of texture have been proposed in literature. There are well known approaches for texture feature extraction operating in the spatial domain (for e.g. gray level co-occurrence matrices), in the frequency domain (for e.g. Fourier spectrum measurements), or in the spatial-frequency domain (for e.g. energy of wavelet coefficients or contourlet coefficients).

Existing techniques for modeling texture include co-occurrence statistics [1, 2], filter banks [3], and random fields [4, 5].

Proposed work relates to texture analysis ([6-9] and references therein), perception ([10]), and synthesis ([11]). Some work has been done in summarizing images (epitome [12]) and video [13, 14]. These methods do not handle textures explicitly and as a result, their reconstructed textures suffer. Other schemes aim to compact the spectral energy into few coefficients [15-17].

Motivated by successes in spatial texture research, approaches have been proposed that uniformly treat a diverse set of dynamic patterns based on aggregate statistics of local descriptors. A seminal example of this approach was based on extracting first- and second-order statistics of motion flow field-based features, assumed to be captured by estimated normal low.

Extracting and quantifying the texture features of image is central to texture-oriented image retrieval [18]. The texture features mainly include coarseness, directionality, contrast, line likeness, regularity, and roughness [19]. To analyze image texture, several approaches such as the statistical method and the structural method [20] have been proposed. The statistical method extracts the texture's characteristics and the relations between them and parameters according to the statistical information of peels gray degree. This method is usually used to analyze unregulated objects such as wood and lawn. The structural method describes the texture's structure and the relations between them and parameters according to the texture cell and their arranging orderliness. It could be used to analyze some regularly formed patterns such as cloth.

A recent research trend is the use of statistical generative models to jointly capture the spatial appearance and dynamics of a pattern. Recognition is realized by comparing the similarity between the estimated model parameters. Several variants of this approach have appeared, including: autoregressive (AR) models [21-23] and multi-resolution schemes [24-25]. By far the most popular of these approaches for recognition is the join photometric-dynamic, AR-based Linear Dynamic System (LDS) model, proposed in [26].

For some images such as building and car, which need to integrate the shape and texture features, the traditional methods such as gray co-occurrence matrix based algorithm cannot extract effectively the shape information in the images.

Some algorithms model textures as a set of features, and generate new images by matching the features in an example texture [27- 29]. These algorithms are usually more efficient than Markov Random Field algorithms. Heeger and Bergen [27] model textures by matching marginal histograms of image pyramids. Their technique succeeds on highly stochastic textures but fails on more structured ones. Simoncelli and Portilla [29] generate textures by matching the joint statistics of the image pyramids. Their method can successfully capture global textural structures but fails to preserve local patterns.

For image processing, there are two methods [30]: one is to process the image in the space domain and the other is to change the image from the space domain to the frequency domain and, after processing, change it back from the frequency domain to the space domain.

It follows that images with high space frequency characterize tiny changes or detailed contents, whereas images with low space frequency characterize the outline of a big object or trend of change. This is the basis of image processing in the frequency domain.

Many techniques apply processing in the frequency domain. One difficulty with the Fourier transform is that it has relatively poor spatial resolution, as Fourier coefficients depend on the entire image. Methods based on the Fourier transform do not perform well in practice, because it lacks spatial localization. The classical way of introducing spatial dependency into Fourier analysis is through the windowed Fourier transform [31].

Gabor filters gives better spatial localization; but, their usefulness is limited in practice as there does not exist a single filter resolution at which one can localize a spatial structure in natural textures [32]. Carrying similar properties to the Gabor transform, wavelet transform representations have also been widely used for texture analysis. Gabor filters and wavelet-based techniques on the other hand compute the textural characteristic by first transforming the image into the frequency domain and then dividing the domain into several frequency subbands.

The distribution of energy in each of these subbands is used as the basis for distinguishing different textures. The difference between the two techniques lies on the way the frequency domain is divided, as well as on the types of the filter used. The wavelet transform has emerged to provide a more formal, solid and unified framework for multiscale signal analysis, with implementations that are generally more efficient than existing equivalent methods [33].

In digital signal processing systems is widely used spectral representation. It is good agreement with the algorithms of digital processing and allows a compact description of signal. To minimize signal descriptions and hence computational complexity needed based on some criteria to choose a system of basis functions with regard to the properties of the signals. It should also take into account the possibility of a simple hardware or software implementation. The basis functions selection largely determines the efficiency of image processing algorithms. Despite the permanent progress of computer technology, the task of reducing the description of signals and saving computing resources is relevant. The most commonly used functions are trigonometric, Walsh, Haar and other orthogonal systems of functions for digital spectral analysis. Of course, the problem of the choice of basis expansions of signals depends of properties of the set of analyzed signals. But the choice of the optimal basis for minimum error criterion submission hampered by the lack of sufficient a priori information. So, Karhunen-Loeve basis is optimal for the stationary random processes for criterion of minimum mean squared error. It is used only when known the correlation function, which rarely corresponds to the real situation.

In solving computation complexity applications as texture synthesis it seems appropriate to make a choice of basis spectral decomposition using system criteria, including the error signal representation and computational cost of spectral analysis procedures.

The objective of our work is to analysis texture representations in different basis functions for image synthesis using system criteria analysis

Proposed analysis

Decomposition of the image $f(x, y)$ in a Fourier series in basis $u_{u,v}^{(m)}(x, \mathrm{y})$ is written in the form:

$$f(x, y)^{(k)} = \sum_{u=1}^{\infty} \sum_{v=1}^{\infty} c_{u,v}^{(k)} u_{u,v}^{(m)}(x, \mathrm{y}),$$

where $c_{u,v}^{(k)}$ are the Fourier coefficients.

If number of series is limited then approximation error presented as:

$$\varepsilon_m^{(k)}(M, N) = \rho\left[f^{(k)}(x, y); \hat{f}^{(k)}(x, y; M, N)\right],$$

where $\rho[\cdot]$ is a distance of some metric, $\hat{f}^{(k)}(x, y; M, N)$ is a partial sum M and N terms of the series, k is a number of a process.

Calculating the number of Fourier coefficients in the basis $u_{u,v}^{(m)}(x, \mathrm{y})$ associated with computational cost $Q_m(M, N)$.

If number of series is limited then cost function presented as:

$$W_m^{(k)} = \Psi\left[\varepsilon_m^{(k)}(M, N); Q_m^{(k)}(M, N)\right]$$

The total cost function includes booth approximation error and computation complexity.

Averaging this function, we obtain value of conditional risk which depends from the basis:

$$R_m = < W_m^{(k)} > = \int\int \Psi\left[\varepsilon_m^{(k)}(M, N); Q_m^{(k)}(M, N)\right] w(x, y) dx dy$$

where $w(x, y)$ is the probability for every subclass signal, $< \ldots >$ is statistical averaging.

The value of the conditional risk depends from the subclass of signals and basis. It is calculated by averaging the cost function $\overline{R}_m =< W_{\ell h}^{(k)} >_k$.

The average risk is determined by averaging the conditional risk for every subclass signals:

$$\overline{R}_m = \sum_{l=1}^{L} P_{lm} R_m ,$$

where P_{lm} is the probability for every subclass signal, L is the number of subclasses of signals.

Experiments

Proposed approach has been used in the spectral representation of the image in the bases of different types of functions (Fourier, Hartley, cosine transform), and the step functions (Walsh and Haar). Modelling was performed using MatLab simulation.

A simplified mathematical model of the original image is a two-dimensional discrete sequence of eight bits $f_{i,j}, i=\overline{1,N}, j=\overline{1,M}$, where N is a number of rows, and M is a number of columns. In our experiments we generated 128 x 128 pixel images of two texture models.

a) Textures «Gaussian» (fig. 1).

For generate «Gaussian» textures we use the expression:

$$S(u,v,\sigma) = \frac{1}{2\pi\sigma^2} \exp(-\frac{u^2+v^2}{2\sigma^2}) ,$$

where σ is a standard deviation.

б) Textures «Clouds» (fig. 2).

We choose cost function in the different forms:

$$W^{(k)}{}_m = \varepsilon_m^2 Q_m , \quad (1)$$

$$W^{(k)}{}_m = \varepsilon_m Q_m , \quad (2)$$

where ε_m is the approximation error, Q_m is the computation cost, k is the number of a process.

Analytical expressions of algorithmic complexity of 2-D transformation algorithms present in table 1.

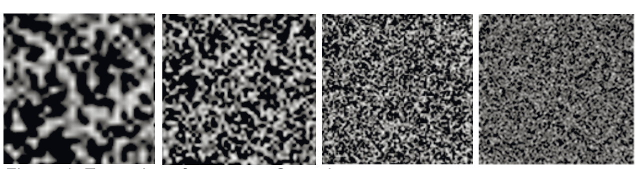

Figure 1. Examples of textures «Gaussian»

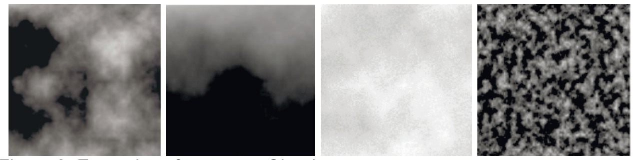

Figure 2. Examples of textures «Clouds»

We calculate the root mean square error as the approximation error:

$$\varepsilon_m = \sqrt{\sum_{x=0}^{N-1}\sum_{y=0}^{M-1}\left(f(x,y) - \hat{f}(x,y)\right)^2 \Big/ N \cdot M} ,$$

where $f(x,y)$ is the original image, $\hat{f}(x,y)$ is the reconstructed image by the truncated series, $i=\overline{1,N}, j=\overline{1,M}$.

Table 1. Analytical expressions of algorithmic complexity of 2-D transformation algorithms

2-D transformation	The number of equivalent additions for image $N \times N$
Fourier	$24N^2 \log_2 N$
Cosine	$10N^2 \log_2 N$
Hartley	$12N^2 \log_2 N + 5N^2$
Walsh	$2N^2 \log_2 N$
Haar	$4N^2 - 2$

The dependence of the root mean square error of approximation of the textures from the number of terms of the series present in table 2.

Using choosing cost functions (1) and (2) we calculate the average risk of approximation of the texture "Gaussian" and "Clouds" for different basis (tables 3 and 4).

Of course, it is impossible when making decisions guided only by the amount of the average risk; it should be set a limit on the allowable error signal representation.

The experimental results show that the lowest amount of average risk for the chosen cost function derived for the Haar function. But this conclusion should not be absolute, because we take into account only the operation of the spectral decomposition. It will be appreciated that there are filtering algorithms and other signal processing functions nonharmonic bases. However, the use of a complex criterion allows most effectively solving many problems in image processing.

Table 2. The dependence of the root mean square error of approximation of the textures from the number of terms of the series

2-D transformation	The number of terms of the series					
	2	4	8	16	32	64
Texture "Gaussian"						
Fourier	0,671	0,265	0,093	0,067	0,043	0,022
Cosine	0,884	0,408	0,098	0,070	0,044	0,022
Hartley	0,704	0,307	0,094	0,068	0,043	0,022
Walsh	0,883	0,596	0,303	0,165	0,091	0,044
Haar	0,883	0,596	0,303	0,165	0,091	0,044
Texture "Clouds						
Fourier	0,230	0,214	0,181	0,133	0,068	0,023
Cosine	0,230	0,224	0,198	0,143	0,074	0,025
Hartley	0,230	0,213	0,183	0,139	0,072	0,024
Walsh	0,230	0,226	0,204	0,158	0,112	0,060
Haar	0,230	0,226	0,204	0,158	0,112	0,060

Table 3. The average risk of approximation of the texture "Gaussian" for different basis and cost functions

2-D transformation	The number of terms of the series					
	2	4	8	16	32	64
Cost function $W^{(k)}{}_m = \varepsilon_m^2 Q_m$						
Fourier	43,2	54,2	40,2	110,4	225,6	277,0
Cosine	31,2	53,2	18,6	50,0	98,9	119,1
Hartley	33,7	43,6	23,4	62,0	123,6	149,6
Walsh	6,2	22,8	35,2	55,9	84,6	93,8
Haar	10,9	22,0	23,2	27,9	33,8	31,3
Cost function $W^{(k)}{}_m = \varepsilon_m Q_m$						
Fourier	64,4	203,8	430,1	1647,4	5265	12782,2
Cosine	35,3	130,4	188,8	715,5	2250	5410,1
Hartley	47,9	142,2	247,5	917,5	2868	6870,2
Walsh	7,1	38,2	116,2	338,5	930,9	2147,9
Haar	12,4	36,9	76,8	168,9	372,2	715,9

Table 4. The average risk of approximation of the texture "Clouds" for different basis and cost functions

2-D transformation	The number of terms of the series					
	2	4	8	16	32	64
Cost function $W^{(k)}{}_m = \varepsilon_m^2 Q_m$						
Fourier	5,0	35,0	151,5	436,7	567,4	315,9
Cosine	2,1	15,9	75,3	209,7	277,1	146,9
Hartley	3,6	21,1	87,5	261,4	348,1	176,7
Walsh	0,4	3,3	15,9	51,3	129,0	173,9
Haar	0,7	3,2	10,5	25,6	51,6	57,9
Cost function $W^{(k)}{}_m = \varepsilon_m Q_m$						
Fourier	22,0	164,1	835,5	3275,9	8349,8	13651,5
Cosine	9,1	71,5	380,1	1465,5	3766,4	6010,2
Hartley	15,6	98,9	479,2	1883,2	4813,8	7465,0
Walsh	1,8	14,4	78,2	324,2	1149,5	2924,4
Haar	3,2	13,9	51,7	161,8	459,6	974,7

Conclusions

We present a novel texture representation approach suitable for image synthesis of textured surfaces. For texture representation we consider the orthogonal decomposition of two-dimensional signals (images) using spectral transform in the different basis functions. This paper focuses on the analysis of the following basis functions Fourier, Walsh, Haar, Hartley and cosine transform using system criteria analysis. For system criteria analysis we use error signal representation and computational cost. It is shown that the Haar transform can represent textural patches more efficiently with smaller average risk than the other basis functions.

Acknowledgment

The reported study was supported by the Russian Foundation for Basic research (RFBR), research projects №15-01-09092 and №17-57-53192.

References

[1] R. Haralick. Statistical and structural approaches to texture. Proceedings of the IEEE, 67: pp. 786–804, 1979.

[2] B. Julesz. Visual pattern discrimination. IRE Transactions on Information Theory, IT-8: pp. 84–92, 1962.

[3] J. Malik and P. Perona. Preattentive texture discrimination with early vision mechanisms. J. Opt. Soc. Am. A, 7(5): pp. 923–932, 1990.

[4] J. Mao and A. Jain. Texture classification and segmentation using multi resolution simultaneous autoregressive models. Pattern Recognition, 25: pp. 173–188, 1992.

[5] J. Zhang, P. Fieguth, and D. Wang. Random field models. In A. Bovik, editor, Handbook of Image and Video Processing, pp. 301–312. Academic Press, San Diego, CA, 2000.

[6] B.S. Manjunath and W. Y. Ma. Texture features for browsing and retrieval of image data. PAMI, 2002.

[7] S. Lazebnik, C. Schmid, and J. Ponce. A sparse texture repesentation using local affine regions. PAMI, 2005.

[8] Kokkinos, G. Evangelopoulos, and P. Maragos. Texture analysis and segmentation using modulation features, generative models, and weighted curve evolution. PAMI, 31(1): pp. 142–157, 2009.

[9] Y. Liu, W.-C. Lin, and J. H. Hays. Near-regular texture analysis and manipulation. SIGGRAPH, 2004.

[10] Malik and P. Perona. Pre attentive texture discrimination with early vision mechanisms. JOSAA, 1990.

[11] M. Cimpoi, S. Maji, I. Kokkinos, S. Mohamed, and A. Vedaldi. Describing textures in the wild. CVPR, 2014.

[12] N. Jojic, B. J. Frey, and A. Kannan. Epitomic analysis of appearance and shape. ICCV, 2003.

[13] V. Cheung, B.J. Frey, and N. Jojic. Video epitomes. IJCV, 2008.

[14] Wexler, E. Shechtman, and M. Irani. Space-time completion of video. PAMI, 2007.

[15] C. Beers, M. Agrawala, and N. Chaddha. Rendering comcompressed textures. In ACM SIGGRAPH, 1996.

[16] S. Fenney. Texture compression using low-frequency signal modulation. In ACM SIGGRAPH, 2003.

[17] P. Mavridis and G. Papaioannou. Texture compression using wavelet decomposition. Proceedings of Pacific Graphics, 2012.

[18] J.R. Smith, "Image retrieval evaluation", IEEE Workshop on Content-based Access of Image and Video Libraries, pp. 112-113, 1998.

[19] R.M. Haralick, K. Shanmugam, I. Dinstein, "Texture features for image classification", IEEE Transactions on Systems Man and Cybernetics, vol. 3, no. 6, pp. 768-780, 1973.

[20] Yang Yu-Bin, Research and Application on the key techniques of Content-based Image Retrieval, 2003.

[21] M. Szummer and R. Picard, "Temporal texture modeling," in ICIP, pp. 823–826, 1996

[22] G. Doretto, A. Chiuso, Y. Wu, and S. Soatto, "Dynamic textures," IJCV, vol. 51, no. 2, pp. 91–109, 2003.

[23] A. Fitzgibbon, "Stochastic rigidity: Image registration for nowhere-static scenes," in ICCV, 2001, pp. I: 662–669.

[24] D. Heeger and A. Pentland, "Seeing structure through chaos," in Workshop on Motion, 1986, pp. 131–136.

[25] Z. Bar-Joseph, R. El-Yaniv, D. Lischinski, and M. Werman, "Texture mixing and texture movie synthesis using statistical learning," T-VCG, vol. 7, no. 2, pp. 120–135, 2001.

[26] G. Doretto, A. Chiuso, Y. Wu, and S. Soatto, "Dynamic textures," IJCV, vol. 51, no. 2, pp. 91–109, 2003.

[27] D.J. Heegeran, J.R. Bergen. Pyramid-Based texture analysis/synthesis. InR. Cook, editor, SIGGRAPH95 Conference Proceedings, Annual Conference Series, pp. 229–238. ACMSIGGRAPH, Addison Wesley, 1995.

[28] J. De Bonet. Multire solution sampling procedure for analysis and synthesis of texture images. InT. Whitted, editor, SIGGRAPH97 Conference Proceedings, Annual Conference Series ,pp. 361–368, 1997

[29] E. Simoncelli and J.Portilla. Texture characterization via joint statistics of wavelet coefficient magnitudes. In Fifth International Conference on Image Processing, volume1, pp. 62–66,1998.

[30] W.Y. Ma, H.J. Zhang, "Benchmarking of image features for content-based retrieval", Proc. 32nd Conference on Signals Systems and Computers, pp. 253-257, 1998.

[31] A.K. Jain and F. Farrokhnia, "Unsupervised texture segmentation using Gabor filters," in Conference Proceedings., IEEE International Conference on Systems, Man and Cybernetics, pp. 14 19, 1990.

[32] Smita S Patil1, A. A. Junnarkar and. D. V. Gore. Study of Texture Representation Techniques. International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), Volume 3, Issue 3, 2014.

Authors Biography

Viacheslav Voronin was born in Rostov (Russian Federation) in 1985. He received his BS in radio engineering from the South-Russian State University of Economics and Service (2006), his MS in radio engineering from the South-Russian State University of Economics and Service (2008) and his PhD in technics from Southern Federal University (2009). Voronin V. is member of Program Committee of conference SPIE. His research interests include image processing, inpainting and computer vision.

Vladimir Ryzhov is a Doctor of Technical Science. His research interests are include spatial-time signal processing, synthesis of signals with desired spectral properties.

Vladimir Marchuk was born in 1951. He received the D.Tech. degree in technics from Southern Federal University (Russian Federation) in 2006. Since 2006, he has been a Professor. His research interests are in the areas of applied statistical mathematics, signal and image processing.

Sergey Makov was born in 1978. In 2001 he has graduated Don State University of Service as radio engineer. Since then he is working in the engineering company as designer of telecommunication hardware. He has got PhD in 2011. Since then he has worked as a professor assistant in the department of electronic systems of Don State Technical University