UDC 004.62

RESEARCH ON TEXTURE IMAGE FEATURE EXTRACTION METHOD

J.K. CHEN, J.X. FU

Belarusian State University of Informatics and Radioelectronics, Republic of Belarus

Received February 22, 2023

Abstract. In this paper, we give several classical feature extraction methods, including grayscale co-generation matrix, Gabor and wavelet transform features, and local binary pattern series features. We introduce the basic principles of these feature extraction algorithms and some derivative methods respectively. Finally, we analyze the advantages and disadvantages of the existing feature extraction methods: grayscale covariance matrix can analyze the arrangement rules of image texture and extract local spatial features of the image, filtering methods and local feature extraction methods are widely used, but the extracted features do not provide a good description of the image structure; and the multi-feature fusion operation brings huge computational effort. Therefore, the future developable directions are proposed based on the existing problems and difficulties in processing texture images.

Keywords: texture image segmentation, feature extraction, image processing.

Introduction

Image, as a visual description of things, has many attributes. Such as chromaticity, brightness, saturation, etc., where texture is an important property of an image. Texture mainly represents the structural features of physical surfaces, which are complex and have many properties. Texture has a basic unit called texture primitive, and texture is a structure composed of a large number of texture primitives arranged according to a given law. Because each texture is arranged differently, the texture of each image is different and has variability. Textures are mainly classified into natural and artificial textures, and almost all images have different texture information.

And the image texture feature extraction is a very important part in image texture classification [1], texture segmentation, texture synthesis, etc. A good texture feature should have the advantages of small computational effort, small feature dimension and strong differentiation ability. In this paper, we introduce several existing feature extraction methods for texture images and analyze the problems and improvement aspects of each method.

Gray Level Co-occurrence Matrix

The texture of an image is formed by the recurrence of gray level distribution in spatial locations, so there is spatial correlation between different pixel gray levels. The Gray Level Co-occurrence Matrix (GLCM) is a classical texture feature matrix, which is computed in the spatial domain based on the following assumptions: The texture information in image I is contained in the overall or average spatial relationship between the gray levels in the image and each other.

The grayscale co-generation matrix is a calculation of four closely related joint and conditional probability density functions, and these four calculated values represent the texture features of the image. Among them, the second-order joint conditional probability density function $P(i, j, d, \theta)$ represents the number of occurrences of the combination of gray levels i and j in an image for two pixels with gray levels i and j, under the condition that the distance is d and the directional angles differ by θ. For an image with one gray level, GLCM is a $G \times G$ matrix. For an image of size 4×4 with four gray levels (0 to 3), as shown in Figure 1, *a*. The general form of the image grayscale covariance matrix is given in figure 1. According to Figure 1, *b*, it can be seen that for an image with four gray levels, its grayscale covariance matrix is a 4×4 matrix, and the grayscale covariance matrix varies at different angles. In the calculation of the grayscale covariance matrix, the distance d and the angle θ are two important parameters, and in general, the value of d is taken as 1. The four plots in figure 2 represent the values of probability density functions under the conditions of distance d as 1 and θ as 0°, 45°, 90°, 135°, respectively.

0	0	1	1
0	0	1	1
0	2	2	2
2	2	3	3

a

	0	1	2	3
0	#(0,0)	#(0,1)	#(0,2)	#(0,3)
1	#(1,0)	#(1,1)	#(1,2)	#(1,3)
2	#(2,0)	#(2,1)	#(2,2)	#(2,3)
3	#(3,0)	#(3,1)	#(3,2)	#(3,3)

b

Figure 1. 4×4 image with four gray-level values and general form of any GLCM for image with value 0-3: *a* – 4 gray level images; *b* – General form of GLCM

4	2	1	0
2	4	0	0
1	0	6	1
0	0	1	2

6	0	2	0
0	4	2	0
2	2	2	2
0	0	2	0

2	1	3	0
1	2	1	0
3	1	0	2
0	0	2	0

4	1	0	0
1	2	2	0
0	2	4	1
0	0	1	0

Figure 2. Four coeval matrices with distance 1

When the gray level G is relatively large, the size of the gray co-occurrence matrix will be very large, which makes the subsequent computation increase dramatically. Therefore, we can analyze the histogram of the texture image, perform appropriate grayscale transformation to achieve the purpose of compressing the gray level without affecting the texture quality as much as possible, and then calculate the grayscale co-occurrence matrix.

Image filtering method

Tuceryan and Jain summarized five major categories of texture features [2], which are: statistical-based, geometric-based, structure-based, model-based, and signal processing-based features. Image filtering methods mainly extract signal processing-based features, mainly including Laws texture template, ring, wedge filter, binary Gabor filter, wavelet transform, discrete cosine transform, optimized Gabor filter, optimized finite impulse response filter, etc. These methods extract the local energy of filter response as features. It has been proved by previous experiments that these methods are more effective than statistical and model-based methods for segmentation. The basic assumption of most filtering methods is that the distribution of energy in the frequency domain identifies textures, so that the spectral energy features of different textures are different if the frequency domain of a texture image is decomposed into a sufficient number of subbands. We present two common methods in the following:

1. Gabor Filters.

Jain and Farrokhnia proposed a set of Gabor filters (also known as Gaussian-shaped bandpass filters) for binary coverage of the radial spatial frequency range and multiple directions, and the designed binary filter bank captures image texture information well due to the joint optimal resolution of the filters in time and frequency. The basic even-symmetric Gabor filter is a band-pass filter with unit impulse response in the 0° direction:

$$h(k,l)=e^{\frac{1}{2}\left(\frac{k^2}{\sigma_x^2}+\frac{l^2}{\sigma_y^2}\right)}\cos(2\pi f_0 k) \tag{1}$$

In equation (1), f_0 is the radial center frequency, (k, l) is the reference coordinate system, the filter response in other directions can be obtained by rotating the reference coordinate system (k, l). The filter has an infinite unit impulse response, but in practical experiments it is approximated as a finite length filter. Jain et al. used five radial frequencies and four directions for an image of size 256 × 256, where the discrete radial center frequencies are $\frac{\sqrt{6}}{2^6},\frac{\sqrt{6}}{2^5},\frac{\sqrt{6}}{2^4},\frac{\sqrt{6}}{2^3},\frac{\sqrt{6}}{2^2}$, angles of 0°, 45°, 90°, 135°.

2. Wavelet transforms.

Wavelet transforms include discrete wavelet transforms, orthogonal wavelet transforms, two-dimensional wavelet transforms, and wavelet packet transforms. Transforms like discrete wavelet correspond to critical sampling filter banks with specific filter parameters and sub-band decomposition; therefore, wavelet transform methods are filter bank methods. The application of wavelet transforms and its derivatives to texture image recognition has received extensive attention in the related literature, where Mallat applies the standard discrete wavelet transform to feature extraction, i.e., critical extraction with a binary subband structure. All wavelet transforms are obtained by stretching and translating the wavelet basis function. The difference between wavelet and Fourier transform is that: the wavelet transform replaces the infinitely long trigonometric basis with a finite length wavelet basis that decays. The wavelet formula is:

$$WT(a,\tau)=\frac{1}{\sqrt{a}}\int_{-\infty}^{+\infty} f(t)\varphi\left[\frac{t-\tau}{a}\right]dt \tag{2}$$

There are two variables in equation (2): scale an and translation t. Scale a controls the scaling of the wavelet function and translation t controls the translation of the wavelet function. Each wavelet transform has a mother wavelet and a father wavelet, and the basis function of the wavelet transform is formed by the scaling and translation of the mother wavelet and the father wavelet. In general, the scaling multiplier is of the order of 2, and the size of the translation is related to the degree of scaling. The basic functions derived from different mother and father wavelets are also different. In digital image processing, the discrete wavelet transform and discrete wavelet packet transform are strictly sampled multirate filter banks that decompose the signal at different scales, for which the choice of scale is determined according to different situations, to obtain low-frequency information and high-frequency information. Among them, the low-frequency information is important and contains the signal characteristics, while the high-frequency information contains the details and differences of the signal. However, strictly sampled filter banks usually imply inaccurate texture edge localization. Later, related researchers have used a complete wavelet representation, i.e., wavelet frames, to alleviate this problem and improve the final results.

Local Binary Patterns

Local Binary Patterns (LBP) was first proposed as an efficient texture description operator, and it has been widely used due to its excellent ability to depict local texture features in images. They are invariant to monotonic grayscale variations. The image is scanned line by line, and for each pixel point in the image, the grayscale of the point is used as a threshold to binarize its surrounding 3×3 8-neighborhoods, and the result of the binarization is formed into an 8-bit binary number in a certain order, and the value of this binary number (0 to 255) is used as the response of the point. Based on the traditional LBP, many variants of the method have been proposed, among which the more famous ones are the grayscale and rotation invariant binary patterns proposed by Ojala et al. in 2002 [3, 4]. Subsequently, the local ternary mode LTP was also used to describe the texture structure by encoding the differences between pixels into three classes. Liao et al. then suggested to use the most frequently occurring mode [5], called the principal local binary mode DLBP, as texture features, however, only the principal mode frequencies were considered in DLBP, while the types of information modes were discarded. The features obtained by the LBP family of methods are a local descriptor, which cannot capture the larger scale texture structure features. Therefore, patch sampling and geometric sampling structures have been proposed again to encode the structural information. To address the noise sensitivity of LBP, many approaches have been tried, including the use of local averaging, frequency or transform domain components, or error correction mechanisms to mitigate this problem. In addition, alternative coding rules have been designed for specific applications, such as the complete local binary counting method CLBC for texture classification and the local directional digit LDN for facial recognition. Combining information from different aspects of an image to generate a histogram can lead to a feature representation with strong discriminative power, for example, Guo et al. proposed the complete local binary pattern CLBP by encoding three complementary components, i.e., the central pixel as well as the sign and size of the local differences [6, 7]. This method adds the central pixel and the size value of the local differences compared with the traditional LBP method, and the information from the three aspects form a more complete feature histogram that contains more local information of the texture. Compared with LBP, the ability to recognize images is much better. In addition, there is also the use of locally enhanced binary code LEBC to enhance the binary code LBP for texture representation and good results were obtained. Other works include applying the idea of LBP to encode the neighborhood information of Gabor features and single gene signal features to enhance face recognition. These algorithms also enhance the discriminative power of the algorithm by expanding the LBP codes.

Conclusion and Outlook

From the introduction of the methods above, it is clear that there are many methods for feature extraction. Grayscale co-occurrence matrix can analyze the arrangement rules of image texture and extract local spatial features of the image. In addition, there are filtering methods as well as local feature extraction methods. In fact, most of the texture segmentation methods extract features from local image patches and then provide them to general clustering or segmentation model algorithms. Various descriptors have been devised by various researchers to characterize texture appearance, and widely used filtering operations are based on filters and statistical models, where filters are used to decompose an image into a set of sub-bands using filter banks, and the model attributes texture to some underlying probability distribution. Although these features can describe the image well, a single texture descriptor alone does not describe the structure of the image well enough and tends to ignore the important information of the image. The simple multi-feature fusion operation, in turn, leads to high feature dimensionality, which causes high computational complexity and expensive costs.

Recent work in texture analysis has shown that texture descriptors constructed by convolving an image with a bunch of filters, based on the local distribution of filter responses, show promising texture recognition performance. Such descriptors can then be combined with well-established segmentation methods to segment texture images. However, there are two main problems with this processing: The first problem stems from the high feature dimensionality of multiple filter responses and their distributional representations. Many widely used segmentation methods, such as graph segmentation methods, curve evolution, and mean shift, rely heavily on measuring the distance between local features, and thus applying them to the distance calculation operation of texture descriptors requires high computational cost. Moreover, choosing the appropriate distance metric for a high-dimensional space is always a tricky problem. Although dimensionality reduction techniques can be used, the suitability of the technique for features usually lacks theoretical justification support. Contending with the above problems, possible improvement aspects are: optimizing and proposing some image segmentation algorithms based on feature extraction for alleviating the current dilemma faced; proposing new segmentation models and algorithms that filter out redundant information in the ensemble of signs and extract a CMI set of features that better describes the image.

References

1. Wang L. [et al.] // Signal Processing, 2018, P. 27–35.
2. Jain A., Farrokhnia F. // Pattern Recognition, 1991, Vol. 24. P. 1167–1186.
3. Ojala T., Pietikainen M. // Pattern Recognition, 1996, Vol. 29. P. 51–59.
4. Ojala T. [et al.] // IEEE Transaction on Pattern Analysis and Machine Intelligence, 2002, Vol. 24. P. 971–984.
5. Liao S., Law M., Chung A. // IEEE Transs Image Proces, 2009, Vol. 18. P. 1107–1118.
6. Sengur A., Guo Y. // Computer Vision and Image Understanding, 2011, Vol. 115. P. 1134–1144.
7. Guo Z., Zhang L., Zhang D. // IEEE Transactions on Image Processing, 2010, Vol. 19. P. 1657–1663.