Postprint of Texture Similarity Computation Using Multi-channel Multi-modal Fused LBP Features

Authors: Liu Tao, Zhou Xianchun, Yan Xijun

Date: 2018-05-20T00:00:00+00:00

Abstract

Texture similarity calculation constitutes one of the fundamental approaches in big data analysis and data mining. To address the issue of insufficient discriminative capability of existing texture features for color images, this paper proposes a texture similarity calculation method based on improved LBP features. The method introduces three feature fusion modes—extreme value mode, summation mode, and encoding mode—to fuse LBP features extracted from the H, S, and V channels of color images, thereby obtaining texture description features for color images. Fusion operations are performed at three stages: neighborhood pixel LBP calculation, center pixel LBP calculation, and histogram feature extraction, which enhances the discriminative capability of the features. Texture similarity calculation experiments conducted on the VisTex texture database demonstrate that the false acceptance rate, false rejection rate, and equal error rate of the proposed method are significantly lower than those of the methods described in references [7, 8, 9].

Full Text

Texture Similarity Calculation with Multi-Channel Multi-Mode Fused LBP Features

Liu Tao[1], Zhou Xianchun[2], Yan Xijun[3]
[1]School of Information & Mechanical and Electrical Engineering, Jiangsu Open University, Nanjing 210017, China
[2]School of Electronic & Information Engineering, Nanjing University of Information Science & Technology, Nanjing 210044, China
[3]College of Computer & Information, Hohai University, Nanjing 210098, China

Abstract

Texture similarity calculation is one of the fundamental techniques in big data analysis and data mining. To address the limited discriminative power of existing texture features for color images, this paper proposes an improved texture similarity calculation method based on fused Local Binary Pattern (LBP) features. The method introduces three feature fusion modes—extreme mode, summation mode, and encoding mode—to fuse LBP features extracted from the H, S, and V channels of color images, yielding comprehensive texture description features for color images. Fusion operations are performed at three stages: LBP calculation of neighborhood pixels, LBP calculation of central pixels, and histogram feature extraction, thereby enhancing feature discriminability. Experimental results on the VisTex texture database demonstrate that the proposed method achieves significantly lower false acceptance rate (FAR), false rejection rate (FRR), and equal error rate (EER) compared to the methods described in references [7, 8, 9].

Keywords: texture similarity; local binary pattern; multi-channel; similarity metric; feature fusion

---

0 Introduction

Texture features are commonly used visual features for image representation and description, characterizing the surface structure and organization of objects with periodic or gradual variations. As quantized visual features, the similarity between texture features can directly reflect the similarity between corresponding images. Consequently, texture similarity calculation serves as a fundamental tool in big data analysis and data mining, enabling classification, retrieval, matching, and correlation of different texture images [1,2]. Texture features form the basis for such calculations, with commonly employed features including gray-level co-occurrence matrices, Gabor filters, and Local Binary Patterns (LBP). However, these features primarily describe grayscale image textures and prove insufficient for color images, exhibiting weak discriminative power [3-6].

Numerous effective methods have emerged for color image texture similarity calculation. Reference [7] proposed a statistical model-based approach for color image texture feature extraction, which applies a pyramidal dual-tree directional filter bank for image filtering, estimates Gamma distribution parameters using moment estimation methods, and employs these parameters as color image texture features with KL divergence for similarity measurement. Reference [8] introduced an illumination-robust color texture descriptor called the Intensity-Color Contrast Descriptor, constructed by fusing LBP features with color contrast features, and utilized cosine similarity for texture comparison. Reference [9] presented a quaternion rotation-invariant texture feature that computes the Fourier spectral norm of color quaternions based on how object colors vary with illuminant spectral power distribution, lighting, and viewing angles. While these methods can describe color image textures, their discriminative capabilities remain limited.

To address this limitation, this paper proposes a multi-channel multi-mode fused LBP feature for texture similarity calculation in color images. The primary innovation lies in introducing three fusion modes—extreme mode, summation mode, and encoding mode—specifically designed for multi-channel LBP features, enabling comprehensive calculation and fusion of LBP features across the H, S, and V color channels.

---

1 Overview of Texture Similarity Calculation

Texture similarity calculation characterizes the similarity between different textures based on their feature similarities, encompassing two key components: feature selection and similarity measurement.

1.1 Feature Selection

Current texture features fall into four categories: statistical, model-based, structural, and signal-processing features. Statistical features describe pixel gray-level properties within neighborhoods, with gray-level co-occurrence matrices being representative. These matrices estimate second-order joint conditional probability densities, describing gray-level co-occurrence probabilities for pixel pairs at specific directions and distances. While offering good discriminative ability, they suffer from low computational efficiency. Model-based features characterize texture distributions by estimating model parameters, commonly employing random field models and fractal models. However, parameter estimation requires large sample sizes, which is often impractical. Structural features treat texture elements as primitives, describing textures through primitive types, orientations, and quantities. Signal-processing features derive from time-domain, frequency-domain, and multi-scale analyses, typically applying transformations (e.g., LBP, DCT, Gabor) to images or sub-blocks and extracting features from the transformed domains. These transformed domain features generally exhibit stronger discriminative power and are widely used for texture description.

Given that our primary contribution involves multi-channel multi-mode fusion of LBP, we focus on LBP feature selection. For a pixel point (x, y) in a grayscale image, traditional LBP calculation employs a circular neighborhood centered at (x, y) with radius R, comparing N neighboring pixels with the central pixel. The comparison results (binary values, 0 or 1) are concatenated in a specific order to obtain the LBP value:

$$LBP(x, y) = \sum_{n=0}^{N-1} 2^n \cdot f(I_n - I_c)$$

where $f(x) = \begin{cases} 1, & \text{if } x \geq 0 \\ 0, & \text{otherwise} \end{cases}$, I_c represents the gray value of the central pixel (x, y), and I_n denotes the gray value of the n-th neighboring pixel. The number of neighboring pixels increases with radius R, as illustrated in [Figure 1: see original paper]. In this paper, we set $R = 2$.

After computing LBP values for all pixels, the original grayscale image transforms into an LBP image, from which global or local histogram features are extracted to generate the LBP feature vector.

1.2 Similarity Measurement

Texture similarity between images can be represented by the similarity between their corresponding feature vectors. Let $\mathbf{X} = (x_1, x_2, \ldots, x_n)$ and $\mathbf{Y} = (y_1, y_2, \ldots, y_n)$ be two n-dimensional feature vectors. Common similarity measures include:

a) Euclidean Distance: The most widely used distance metric, defined as:

$$d(\mathbf{X}, \mathbf{Y}) = \sqrt{\sum_{k=1}^{n} (x_k - y_k)^2}$$

b) Manhattan Distance: Also known as city block distance:

$$d(\mathbf{X}, \mathbf{Y}) = \sum_{k=1}^{n} |x_k - y_k|$$

c) Chebyshev Distance: Inspired by king’ s moves in chess:

$$d(\mathbf{X}, \mathbf{Y}) = \max_k |x_k - y_k| = \lim_{p \to \infty} \left(\sum_{k=1}^{n} |x_k - y_k|^p \right)^{1/p}$$

d) Minkowski Distance: A generalization of the above distances:

$$d(\mathbf{X}, \mathbf{Y}) = \left(\sum_{k=1}^{n} |x_k - y_k|^p \right)^{1/p}$$

When $p = 1$, it becomes Manhattan distance; when $p = 2$, Euclidean distance; and as $p \to \infty$, Chebyshev distance.

e) Cosine Similarity: Measures directional difference between feature vectors. Larger values indicate smaller angles and higher similarity:

$$\cos\theta = \frac{\sum_{k=1}^{n} x_k y_k}{\sqrt{\sum_{k=1}^{n} x_k^2} \sqrt{\sum_{k=1}^{n} y_k^2}}$$

2 Multi-Channel Multi-Mode Fusion of Local Binary Patterns

LBP is a widely used texture feature with advantages including intensity invariance, rotation invariance, and illumination insensitivity, making it popular for texture similarity calculation in face recognition and expression analysis. However, LBP is typically extracted from grayscale images, inevitably losing discriminative information for color images. Some researchers have proposed computing LBP independently on each color channel (e.g., R, G, B) and concatenating features at either the LBP image level or histogram level. Yet this approach neglects inter-channel relationships at each pixel, resulting in high redundancy and limited discriminative improvement, thereby reducing reliability in color image texture similarity calculation.

To overcome this limitation, we propose a multi-channel multi-mode fused LBP feature that performs fusion at three stages: neighborhood pixel LBP calculation, central pixel LBP calculation, and histogram feature extraction. This yields more comprehensive local binary patterns, enhancing inter-texture discrimination and similarity calculation reliability.

For color images, we compute LBP features in the HSV color space due to its independent intensity, hue, and saturation channels, which produce weakly correlated features with minimal mutual interference during fusion. We calculate LBP on the H, S, and V channels using the aforementioned formula, denoting the LBP value for pixel (x, y) in channel t as $LBP_t(x, y)$, where $t \in \{H, S, V\}$.

We design three multi-channel fusion modes operating on each neighborhood pixel' s LBP values across the three channels:

a) Extreme Mode: Takes the maximum value among the three channel LBP values as the fused result:

$$M_m(x, y) = \max_{t \in \{H,S,V\}} LBP_t(x, y)$$

b) Summation Mode: Sums the three channel LBP values:

$$M_a(x, y) = \sum_{t \in \{H,S,V\}} LBP_t(x, y)$$

c) Encoding Mode: Encodes the three binary LBP values (0 or 1) in H-S-V order as a binary number:

$$M_e(x, y) = LBP_H(x, y) \times 4 + LBP_S(x, y) \times 2 + LBP_V(x, y)$$

provides examples of these three fusion modes applied to H, S, and V channel LBP values. The extreme mode yields binary values in $[0, 1]$, summation mode produces values in $[0, 3]$, and encoding mode generates values in $[0, 7]$.

The mapping functions for these modes can be expressed as:

$$M_m(x,y) = \begin{cases} 1, & \text{if } \max_t LBP_t(x,y) = 1 \\ 0, & \text{otherwise} \end{cases}$$

$$M_a(x,y) = \begin{cases} 1, & \text{if } \sum_t LBP_t(x,y) \geq 1 \\ 0, & \text{otherwise} \end{cases}$$

$$M_e(x,y) = \begin{cases} 1, & \text{if } LBP_H \times 4 + LBP_S \times 2 + LBP_V = \varepsilon \\ 0, & \text{otherwise} \end{cases}$$

where ε represents the encoded value. For each pixel' s three channel LBP values, we generate 2 extreme mode outputs, 4 summation mode outputs, and 8 encoding mode outputs.

Similarly, the fused LBP values for central pixels are computed as:

$$LBP_m(x,y) = \sum_{n=0}^{N-1} 2^n \cdot M_m(x,y)$$

$$LBP_a(x,y) = \sum_{n=0}^{N-1} 2^n \cdot M_a(x,y)$$

$$LBP_e(x,y) = \sum_{n=0}^{N-1} 2^n \cdot M_e(x,y)$$

Thus, each central pixel yields 14 fused LBP values, transforming a color image into 14 fused LBP images.

To obtain the multi-channel fused feature vector, we first compute histograms for the 2 extreme-mode fused LBP images:

$$H_m(\varepsilon) = \frac{1}{WH} \sum_{x=0}^{W-1} \sum_{y=0}^{H-1} \varphi(LBP_m(x,y), \varepsilon), \quad \varepsilon \in [0,1]$$

where W and H are image width and height, and $\varphi(a,b) = 1$ if $a = b$, otherwise 0.

Similarly, histograms for the 4 summation-mode and 8 encoding-mode fused LBP images are:

$$H_a(\varepsilon) = \frac{1}{WH} \sum_{x=0}^{W-1} \sum_{y=0}^{H-1} \varphi(LBP_a(x,y), \varepsilon), \quad \varepsilon \in [0,3]$$

$$H_e(\varepsilon) = \frac{1}{WH} \sum_{x=0}^{W-1} \sum_{y=0}^{H-1} \varphi(LBP_e(x,y), \varepsilon), \quad \varepsilon \in [0,7]$$

After normalization, we concatenate the three histogram feature vectors:

$$\mathbf{H} = [\mathbf{H}_m, \mathbf{H}_a, \mathbf{H}_e]$$

In summary, our multi-channel multi-mode fused LBP feature begins fusion at the neighborhood pixel LBP calculation stage, proceeding through central pixel fusion, individual mode histogram fusion, and final histogram concatenation. Compared to traditional channel-independent LBP calculation with simple concatenation, our approach produces features with lower redundancy and stronger discriminative power.

---

3 Simulation Experiments

Our innovation is a multi-channel multi-mode fused LBP feature extraction method for color image texture similarity calculation. We evaluate its performance through comparative experiments against existing methods.

Experimental Setup

We select the publicly available VisTex natural texture image database, which contains 640 images across 40 texture categories (16 images per category) at 128×128 resolution with 24-bit color depth. Sample images are shown in [Figure 2: see original paper].

Evaluation metrics include False Acceptance Rate (FAR), False Rejection Rate (FRR), and Equal Error Rate (EER). FAR measures the proportion of inter-class textures incorrectly accepted as matching, while FRR measures the proportion of intra-class textures incorrectly rejected. EER is the rate at which FAR equals FRR. These metrics are threshold-dependent: higher similarity thresholds decrease FAR but increase FRR.

We compare our method against those in references [7-9] using three similarity measures: Euclidean distance, Manhattan distance, and cosine similarity.

Results and Analysis

[Figure 3: see original paper], [Figure 4: see original paper], and [Figure 5: see original paper] present performance comparisons under Euclidean, Manhattan, and cosine similarity measures, respectively. Three key conclusions emerge:

a) FAR and FRR Comparison: Across all similarity measures, our multi-channel multi-mode fused LBP feature consistently outperforms the three comparison methods. For equivalent FRR values, our method yields lower FAR; conversely, for equivalent FAR values, our method produces lower FRR.

b) EER Comparison: Regardless of similarity measure, our method achieves the lowest EER values, followed by reference [9], with reference [7] performing worst.

c) Measure Sensitivity: For all four methods, cosine similarity produces the lowest EER, Manhattan distance the highest, with Euclidean distance close to cosine similarity. Reference [9]' s method shows the greatest sensitivity to similarity measure choice, while our method exhibits the least.

These results confirm that combining our multi-channel multi-mode fused LBP features with cosine similarity yields optimal texture similarity calculation performance.

---

4 Conclusion

Addressing the limited discriminative power of features in color image texture similarity calculation, this paper proposes a multi-channel multi-mode fused LBP feature. Unlike traditional channel-independent LBP calculation with simple concatenation, we introduce three fusion modes—extreme, summation, and encoding—applied across three stages: neighborhood pixel LBP calculation, central pixel LBP calculation, and histogram extraction. The resulting color image texture descriptors exhibit low redundancy and strong discriminative capability, significantly improving texture similarity calculation performance.

References

[1] Pappas T N, Neuhoff D L, Ridder H D, et al. Image analysis: focus on texture similarity [J]. Proceedings of the IEEE, 2013, 101(9): 2044-2057.

[2] Zhai Y, Neuhoff D L, Pappas T N. Local radius index: a new texture similarity feature [C]// Proc of IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE, 2013: 1434-1438.

[3] Zujovic J, Pappas T N, Neuhoff D L. Structural texture similarity metrics for image analysis and retrieval [J]. IEEE Trans on Image Processing, 2013, 22(7): 2545-2558.

[4] Naser K, Ricordel V, Callet P L. Experimenting texture similarity metric STSIM for intra prediction mode selection and block partitioning in HEVC [C]// Proc of International Conference on Digital Signal Processing. 2014.

[5] Porebski A, Vandenbroucke N, Hamad D. LBP histogram selection for supervised color texture classification [C]// Proc of IEEE International Conference on Image Processing. 2014: 3239-3243.

[6] 冀中, 聂林红. 基于抗噪声局部二值模式的纹理图像分类 [J]. 计算机研究与发展, 2016, 53(5): 1128-1135.

[7] 杨君, 尚赵伟. 基于 PDTDFB 变换的彩色纹理图像检索 [J]. 计算机工程, 2012, 38(20): 204-208.

[8] Cusano C, Napoletano P, Schettini R. Illuminant Invariant Descriptors for Color Texture Classification [C]// Proc of International Conference on Computational Color Imaging. Springer-Verlag, 2013: 239-249.

[9] Joshua A. Tucker. Rotation invariant features for color texture classification and retrieval under varying illumination [J]. Pattern Analysis & Applications, 2013, 16(1): 69-81.

[10] Chandrakala D, Sumathi S. Application of Artificial Bee Colony Optimization Algorithm for Image Classification Using Color and Texture Feature Similarity Fusion [J]. ISRN Artificial Intelligence, 2012, 2012(3).

[11] Dash J K, Mukhopadhyay S. Similarity learning for texture image retrieval using multiple classifier system [J]. Multimedia Tools & Applications, 2016, 8(7): 1-25.