A Local Binary Pattern Image Feature Point Matching Algorithm Postprint

Authors: Wang Qiang, Li Bolin, Luo Jianqiao, Chen Xiaoyan

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

Abstract

This paper investigates the image matching problem and proposes an improved feature point matching algorithm based on BRIEF. The algorithm leverages the relationship between difference magnitude and amplitude of random points and feature points to generate feature point descriptors. To address the noise sensitivity of BRIEF, as small pixel amplitude differences are more susceptible to noise, a small pixel difference threshold is introduced for noise suppression. Differences within the threshold are designated as uncertain bits, whose values are subsequently determined by the mean of their neighboring region. Feature point matching is performed by comparing the Hamming distance between descriptors. Experimental comparisons with BRIEF and ORB algorithms demonstrate that the proposed descriptor exhibits higher discriminability, simple computation, excellent noise suppression performance, fast running speed, and higher matching accuracy.

Full Text

Preamble

Title: Local Binary Pattern Image Feature Point Matching Algorithm

Authors: Wang Qiang[1,2], Li Bailin[1], Luo Jianqiao[1], Chen Xiaoyan[3]

Affiliations: 1. School of Mechanical Engineering, Southwest Jiaotong University, Chengdu 610031, China 2. College of Mechanical Engineering, Chengdu Technological University, Chengdu 611730, China 3. Dept. of Aviation Manufacturing Engineering, Chengdu Aeronautic Polytechnic, Chengdu 610100, China

Abstract: This paper investigates the image matching problem and proposes an improved feature point matching algorithm based on the BRIEF algorithm. The proposed algorithm generates feature point descriptors using both the difference signs and magnitudes between random points and feature points. To address BRIEF' s sensitivity to noise—since small pixel magnitude differences are more susceptible to noise contamination—we introduce a small pixel difference threshold. Differences within this threshold are marked as uncertain bits, whose values are determined by their neighborhood means. Feature point matching is performed by comparing Hamming distances between descriptors. Experimental comparisons with BRIEF and ORB demonstrate that our operator exhibits higher discriminative power, computational simplicity, excellent noise suppression performance, faster runtime, and improved matching accuracy.

Keywords: BRIEF; ORB; feature point matching operator; local binary pattern

0 Introduction

Feature point matching represents a fundamental technique in computer vision with widespread applications in object detection [1], face recognition [2–4], image classification [5], scene recognition [6], moving object tracking [7], and texture classification [8–12]. As smart devices proliferate, image matching algorithms face increasingly stringent requirements for stability, noise robustness, low memory consumption, and high computational speed, necessitating more sophisticated feature point matching operators.

Early work by Moravec [14] introduced feature point-based image matching methods, followed by Harris et al.' s [15] corner detection approach, though matching performance remained limited. The SIFT algorithm [16] employed a complex descriptor formulation, computing gradient orientation histograms across 8 directions for each sub-region within a feature point' s neighborhood to produce a 128-dimensional vector. While robust, this approach suffers from low efficiency. Dimensionality reduction techniques like PCA [17] or LDA [18] can partially alleviate this computational burden. SURF [19–21] improved upon SIFT by using Haar wavelet responses to compute 64-dimensional descriptors, achieving moderate speed improvements. Subsequent algorithms such as BRISK [22] and FREAK [23] offered substantial speed enhancements but at the cost of reduced matching accuracy.

The BRIEF algorithm [13] provided a simpler alternative by selecting random point pairs within a feature point' s neighborhood and comparing their pixel intensities to generate a binary string descriptor. XOR operations enable rapid Hamming distance computation, dramatically improving speed. However, single-pixel comparisons render BRIEF highly noise-sensitive, requiring Gaussian filtering prior to descriptor computation. ORB [24] addressed these limitations through two key modifications: first, using image integrals to compute sums over sub-windows, effectively performing mean filtering for partial noise suppression; second, computing a dominant orientation for rotation invariance. Unfortunately, orientation estimation increases correlation between random point pairs, reducing descriptor discriminative power. Both algorithms exclusively consider difference signs between neighborhood points while ignoring magnitude relationships, resulting in incomplete representation of the joint difference distribution between feature points and their neighborhoods.

2 Proposed Method

2.1 Theoretical Analysis

Ojala et al. [25,26] demonstrated that image texture can be characterized through relationships between center pixels and their neighborhoods, forming the basis of Local Binary Pattern (LBP) encoding. Treating the center pixel as a feature point, LBP can be viewed as a feature point descriptor leveraging difference distributions between neighborhood and center pixels. A descriptor expressing sufficiently complete difference relationships achieves higher discriminative power.

The relationship between a neighborhood pixel and center pixel can be fully described by both difference sign and magnitude. While difference signs capture most information—utilized by both LBP and BRIEF—difference magnitudes also convey valuable information [8]. Maitre [27] proved that pixel differences follow Gaussian or Laplacian distributions, with the sign relationship following a Bernoulli distribution:

$$p_b = \begin{cases} 1 & \text{if } z \geq 0 \\ -1 & \text{if } z < 0 \end{cases}$$

where $z = p_c - p_g$ represents the pixel difference between center pixel p_c and neighborhood pixel p_g. The magnitude component $p_m = |z|$ follows a Laplacian distribution. If we reconstruct the difference signal $\hat{z}$ using only sign information, the reconstruction error E_b is:

$$E_b = E[(\hat{z} - z)^2] = 2\sigma^2$$

Reconstruction using only magnitude information yields error $E_m = 4\sigma^2$. This analysis reveals that difference magnitude carries approximately 20% of the original image information [8]. Integrating magnitude information significantly enhances BRIEF' s discriminative capability.

2.2 Method Description

To more completely express difference relationships and improve discriminative power, we jointly encode both difference signs and magnitudes. Leveraging the Gaussian distribution of pixel differences [27], we generate random points within a feature point' s $N \times N$ neighborhood using a Gaussian distribution and construct binary descriptors from relationships between these random points and the feature point.

Figure 1 [Figure 1: see original paper] illustrates the descriptor generation process. Figure 1(a) shows the feature point neighborhood, (b) depicts pixel differences between neighborhood and center, (c) represents difference sign encoding (where "-1" is replaced by "0" in actual coding), and (d) shows difference magnitude encoding.

Let f_c denote the feature point pixel value and f_i represent neighborhood pixels. The joint distribution is:

$$T = (f_c, f_1, f_2, ..., f_k)$$

For grayscale invariance, we subtract the center pixel value:

$$T' = (f_c, f_1 - f_c, f_2 - f_c, ..., f_k - f_c)$$

The sign function is defined as:

$$s(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

The sign-based descriptor component is:

$$\tau_s = (s(f_1 - f_c), s(f_2 - f_c), ..., s(f_k - f_c))$$

To preserve magnitude information, we compute the mean absolute difference:

$$m_T = \frac{1}{k} \sum_{i=1}^{k} |f_i - f_c|$$

and encode magnitudes by comparing each absolute difference against this mean:

$$\tau_m = (s(|f_1 - f_c| - m_T), s(|f_2 - f_c| - m_T), ..., s(|f_k - f_c| - m_T))$$

The final descriptor concatenates both components: $\tau = (\tau_s, \tau_m)$.

2.3 Noise Suppression

Single-pixel comparisons are vulnerable to noise. We suppress noise using neighborhood mean thresholding and handling small differences as uncertain bits. Let m_f be the neighborhood mean:

$$m_f = \frac{1}{9}\sum_{i=1}^{9} f_i \quad \text{for a } 3 \times 3 \text{ region}$$

Small pixel differences are particularly noise-sensitive [2]. We introduce threshold t to identify uncertain bits where $|f_i - m_f| \leq t$. For these uncertain points p, we compute the mean of their 3×3 neighborhood $f_{p,m}$ and replace f_p with $f_{p,m}$ before comparison, effectively performing local mean filtering.

The modified sign function becomes:

$$s'(x) = \begin{cases} 1 & \text{if } x > t \\ 0 & \text{if } x < -t \\ \text{uncertain} & \text{if } |x| \leq t \end{cases}$$

Uncertain bits are resolved by comparing the neighborhood-averaged value against the threshold. This approach eliminates noise from small pixel differences while preserving descriptor discriminative power.

The complete algorithm proceeds as follows: 1. Generate random point positions using a Gaussian distribution within the $N \times N$ neighborhood 2. For each feature point: - Compute neighborhood mean m_f - Set threshold t and compare random points against m_f to obtain sign bits - For points with differences within $[-t, t]$, mark as uncertain and determine values using 3×3 neighborhood means - Compute magnitude bits by comparing absolute differences against the mean magnitude 3. Concatenate sign and magnitude bits to form the final noise-resistant binary descriptor

---

3 Experimental Results

We evaluated our algorithm using standard image matching benchmark datasets, comparing against BRIEF and ORB. Figure 2 [Figure 2: see original paper] shows five test sequences: wall (viewpoint change), trees (blur), light (illumination change), JPG (compression), and boat (rotation and scale). Each sequence contains six images. We matched the first image against the remaining five, then repeated with Gaussian and salt-and-pepper noise added to the second image.

Implementation used C++ with OpenCV 2.4.9. CenSurE feature points were employed with neighborhood size $N = 17$. Small difference threshold t was set to 5, which is appropriate for moderate noise levels (typically 2-5 for low noise, increasing for stronger noise). BRIEF and ORB used 512-bit descriptors.

Tables 1–5 [TABLE:1–5] show matching accuracy for original images. Our algorithm achieves the highest matching rates for illumination, blur, and compression changes. For rotation and scale changes, it outperforms ORB when transformations are moderate, though ORB' s dominant orientation calculation yields better performance at large rotations/scales. However, ORB's orientation estimation increases random point correlation, reducing discriminative power as evidenced in Tables 1–4.

Figure 3 [Figure 3: see original paper] demonstrates matching performance across different neighborhood sizes using the trees sequence. Accuracy improves with larger neighborhoods up to $N = 17$, beyond which gains diminish while computational cost increases. $N = 17$ provides a good balance between accuracy and efficiency.

Figure 4 [Figure 4: see original paper] shows Gaussian noise examples ($\sigma = 0.05, 0.10, 0.15, 0.20$), and Figure 5 [Figure 5: see original paper] shows salt-and-pepper noise examples ($p = 0.05, 0.10, 0.15, 0.20$). Matching accuracy decreases with increasing noise, but our method consistently outperforms both BRIEF and ORB across all noise levels and image types (Tables 6–15 [TABLE:6–15]).

BRIEF's Gaussian filtering provides strong noise suppression, making it superior to ORB under noise. ORB' s integral image approach accelerates computation but its orientation estimation reduces discriminative power. Our method' s integration of magnitude information and uncertain bit handling yields stronger descriptors without requiring explicit filtering.

Computational complexity is $O(n)$ for n feature points, with one loop computing k comparisons per point. Descriptor generation time is slightly higher than ORB due to thresholding operations and doubled descriptor length (sign + magnitude). However, both sign and magnitude components can be computed efficiently using integral images. Table 16 compares descriptor computation times on an Intel i5 2.5 GHz processor for 875 feature points, showing our method is comparable to ORB and faster than BRIEF (which requires expensive Gaussian filtering).

---

4 Conclusion

BRIEF achieves remarkable efficiency while maintaining reasonable matching accuracy. ORB introduces rotation invariance but increases descriptor correlation, reducing discriminative power. Both algorithms exhibit noise sensitivity –BRIEF requires Gaussian filtering, while ORB uses integral images for mean filtering and faster computation. However, neither fully exploits relationships between feature points and neighborhoods, nor do they utilize difference magnitude information.

Our proposed binary descriptor jointly encodes difference signs and magnitudes, completely representing the joint difference distribution and achieving higher discriminative power. Noise suppression through neighborhood mean thresholding and uncertain bit handling eliminates the need for explicit filtering. Hamming distance computation via XOR operations maintains computational simplicity and speed. Experimental results demonstrate superior matching accuracy across various transformations and noise conditions, with computational efficiency comparable to ORB. The method is particularly suitable for applications involving moderate rotation and scale variations where high discriminative power and noise robustness are paramount.

References

[1] Satpathy A, Jiang X, Eng H L. LBP-based edge-texture features for object recognition [J]. IEEE Trans on image Processing, 2014, 23(5): 1953-1964.

[2] Ren J, Jiang X, Yuan J, et al. LBP encoding schemes jointly utilizing the information of current bit and other LBP bits [J]. IEEE Signal Processing Letters, 2015, 22(12): 2373-2377.

[3] Chen D, Cao X, Wen F, et al. Blessing of dimensionality: high-dimensional feature and its efficient compression for face verification [C]// Proc of IEEE Conference on Computer Vision and Pattern Recognition. 2013: 3025-3032.

[4] Deshmukh A, Sirpotdar P, Sheikh J, et al. High accuracy face recognition system based on SIFT [J]. International Journal of Advanced Research in Computer Engineering & Technology, 2015, 4(6): 2626-2628.

[5] Li W, Chen C, Su H, et al. Local binary patterns and extreme learning machine for hyperspectral imagery classification [J]. IEEE Trans on Geoscience and Remote Sensing, 2015, 53(7): 3681-3693.

[6] Xiao Y, Wu J, Yuan J, et al. mCENTRIST: a multi-channel feature generation mechanism for scene categorization [J]. IEEE Trans on Image Processing, 2014, 23(2): 823-836.

[7] Zhao G, Ahonen T, Matas J, et al. Rotation-invariant image and video description with local binary pattern features [J]. IEEE Trans on Image Processing, 2012, 21(4): 1465-1477.

[8] Guo Z, Zhang L, Zhang D, et al. A completed modeling of local binary pattern operator for texture classification [J]. IEEE Trans on Image Processing, 2010, 19(6): 1657-1663.

[9] Song T, Li H, Meng F, et al. Noise-robust texture description using local contrast patterns via global measures [J]. IEEE Signal Process. Letters, 2014, 21(1): 93-96.

[10] Wang K, Bichot C E, Zhu C, et al. Pixel to patch sampling structure and local neighboring intensity relationship patterns for texture classification [J]. IEEE Signal Processing Letters, 2013, 20(9): 853-856.

[11] Liao S, Law M W K, Chung A C S, et al. Dominant local binary patterns for texture classification [J]. IEEE Trans on Image Processing, 2009, 18(5): 1107-1118.

[12] Kwak J T, Xu S, Wood B J, et al. Efficient data mining for local binary pattern in texture image analysis [J]. Expert Systems with Applications, 2015, 42(9): 4529–4539.

[13] Calonder M, Lepetit V, Ozuysal M, et al. BRIEF: computing a local binary descriptor very fast [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2012, 34(7): 1281-1298.

[14] Moravec H P. Rover visual obstacle avoidance [C]// Proc of the 7th International Joint Conference on Artificial Intelligence. 1981: 785-790.

[15] Harris C, Stephens M. A combined corner and edge detector [C]// Proc of the 4th Alvey Vision Conference. 1988: 147-151.

[16] Lowe D G. Distinctive image features from scale-invariant keypoints [J]. International Journal of Computer Vision, 2004, 60(2): 91-110.

[17] Mikolajczyk K, Schmid C. A performance evaluation of local descriptors [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2005, 27(10): 1615-1630.

[18] Hua G, Brown M, Winder S. Discriminant embedding for local image descriptors [C]// Proc of the 11th IEEE International Conference on Computer Vision. 2007: 1-8.

[19] Bay H, Tuytelaars T, Van Gool L. Surf: speeded up robust features [C]// Proc of International Conference on Computer Vision. Berlin: Springer, 2006: 404-417.

[20] Chu S Y, Xi Z H. Digital image stabilization technology combined with SURF [J]. Journal of Computer-Aided Design & Computer Graphics, 2014, 26(2): 241-247.

[21] Chen X D, Du Y R, Gao X B, et al. A fast image feature point matching algorithm based on SURF [J]. Journal of Yangzhou University: Natural Science Edition, 2012, 15(4): 64-67.

[22] Leutenegger S, Chli M, Siegwart R Y. BRISK: binary robust invariant scalable keypoints [C]// Proc of International Conference on Computer Vision. 2011: 2548-2555.

[23] Alahi A, Ortiz R, Vandergheynst P. Freak: fast retina keypoint [C]// Proc of IEEE Conference on Computer Vision and Pattern Recognition. 2012: 510-517.

[24] Rublee E, Rabaud V, Konolige K, et al. ORB: an efficient alternative to SIFT or SURF [C]// Proc of International Conference on Computer Vision. 2011: 2564-2571.

[25] Ojala T, Pietikainen M, Harwood D. A comparative study of texture measures with classification based on featured distributions [J]. Pattern Recognition, 1996, 29(1): 51-59.

[26] Ojala T, Pietikainen M, Maenpaa T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2002, 24(7): 971-987.

[27] Weinberger M J, Rissanen J J, Arps R B, et al. Applications of universal context modeling to lossless compression of gray-scale images [J]. IEEE Trans on Image Processing, 1996, 5(4): 575-586.

Note: Figure translations are in progress. See original paper for figures.