The Third Information Systems International Conference

Batik Image Classification Using SIFT Feature Extraction, Bag of Features and Support Vector Machine

Ryfial Azhar[a], Desmin Tuwohingide[a,b], Dasrit Kamudi[a,b], Sarimuddin[a,c], Nanik Suciati[a]

[a]Department of Informatics, Institut Teknologi Sepuluh Nopember Surabaya, Indonesia
[b]Department of Information Systems, Politeknik Negeri Nusa Utara Sulawesi Utara, Indonesia
[c] Department of Informatics, Universitas Sembilanbelas November Kolaka, Indonesia

Abstract

Batik is a traditional fabric of Indonesian cultural heritage. Automatic batik image classification is required to preserve the wealth of traditional art of Indonesia. In such classification, a method to extract unique characteristics of batik image is important. Combination of Bag of Features (BOF) extracted using Scale-Invariant Feature Transform (SIFT) and Support Vector Machine (SVM) classifier which had been successfully implemented in various classification tasks such as hand gesture, natural images, vehicle images, is applied to batik image classification in this study. The experimental results show that average accuracy of this method reaches 97.67%, 95.47% and 79% in normal image, rotated image and scaled image, respectively.

Keywords: Batik Image Classification, Bag of features, Scale-Invariant Feature Transform,Support vector machine ;

1. Introduction

Batik is a traditional fabric which has been admitted by UNESCO as one of the cultural heritage of Indonesia since 2 October 2009. Batik has a variety of patterns (motifs) appearing repeatedly on a piece of cloth [1]. At this time, there are hundreds number of batik motif. In addition to classifying batik based on its motif, batik fabric can also be grouped based on its origin [2].

Previously, there were some researches in batik image recognition and classification that have been reported. A research by Rangkuti et al, that developed batik image classification system using Fuzzy Neural Network (FNN) classifier and wavelet-based feature extraction, showed good result in normal testing dataset, without rotation and scale variation [3]. Batik image classification by Kurniawardhani et

al, which implemented combination of Rotation - Invariant Local Binary Pattern (LBPROT) and Completed Robust Local Binary Pattern (CRLBP), reported good result in testing dataset with rotation and scaling [4]. In some researches [5]–[7], batik image recognition using Scale-Invariant Feature Transform (SIFT) feature extraction method developed by David G. Lowe [8], which was scale-invariant showed satisfactory results. The results proved that SIFT method was robust against scale, rotation, and illumination condition. In those researches, the distance calculation between closer key points was employed for finding appropriate batik image in learning dataset.

Bag of features is one of the algorithm models which is often used in computer vision. There are three main steps in the using of bag of features in image classification, i.e. (i) extracting features by using a certain feature extraction method, such as SIFT, (ii) clustering the features, and (iii) constructing a bag of features, which calculated the number of features that are entered on each cluster [9]. A previous research by Lou et al reported that image classification using bag of features model was more efficient compared with other algorithm models [10]. Some researches [9]–[13] also reported that image classification using bag of features and SVM as classifier can significantly improve the accuracy of classification. The good accuracy result is mainly because of the method can describe features properly with reducing the dimensionality of the feature space.

Digitizing batik fabrics which have variety of patterns (motifs) into batik images, can be performed by using camera. The process which might be done under variation of illumination condition, rotation, and scale, results some different images for one batik fabric. Classification of such kind of image dataset requires a robust method both in features extraction and classification method. In this study, combination of Bag of Features (BOF) extracted using Scale-Invariant Feature Transform (SIFT) and Support Vector Machine (SVM) classifier which had been successfully implemented in various classification tasks such as hand gesture, natural images, vehicle images, is applied to batik image classification.

2. Methodology

2.1. Dataset

The data used in this study are batik images which consist of 50 classes. There are 6 sub images with size 128 x 128 in each class. Thus, the total number of sub images are 300, which is divided into 200 sub images for learning and other 100 images for testing.

2.2. Overall System

In this study, there are 2 main steps which are learning and testing. In learning step, batik image dataset will be learned to result classification model which is used in testing. There are some substantial steps i.e. preprocessing (image transformation from RGB image to grayscale image), constructing BOF (feature extraction using SIFT, vocabulary construction using K-means and histogram construction) and classification using SVM. In testing step, it is similar with the steps before. The result of this step is class label for any existing test data.

2.3. Scale-Invariant Feature Transform

The novel image feature extraction was described at 1999 by David G. Lowe who is researches from University of British Colombia. This method is named as Scale-Invariant feature Transform (SIFT)[8]. SIFT is a method that is invariant to scale, rotation, and illumination condition. For our research object (batik image), this method is fit to extract local features that can describe the different of the batik types which have almost same characteristic. For the detailed

1. This step aims to find extreme values in scale space to locate the position of keypoints on image. The pixel will be checked and compared with 26 pixels, which are 8 neighboring pixels in the same scale and 9×2 neighboring pixels in scale before and after.

processes of the method refer to [8]. The SIFT algorithm is divided into several parts: Determine and delete bad keypoints, which are edges and low contrast regions via fitting three dimensional quadratic functions, meanwhile this step will make algorithm robust and efficient.

2. Each keypoint direction and magnitude will be measured via the direction of gradient of its neighboring pixels. After that choose and find the most prominent orientation in the region as the region of the keypoint. This effectively cancels out the effect of orientation, making it rotation invariant.

3. At this step, SIFT feature vector is generated. To increase the stability of matching, Lowe suggests to describe each keypoint using 4×4 seed points in actual calculation. Thus, 128 data points, i.e. a 128-dimensional SIFT vector, are extracted for each keypoint. Now SIFT vector is free from the influence of geometric transformations such as scale changes and rotation.

It is true that in some researches [5]–[7] SIFT features can describe well the unique characteristic in the image with a 128-dimensional for each keypoint. Therefore, those features can easily be distinguished one another.

2.4. Bag of Features

SIFT features has high dimensionality as describe in the previous section. For image classification, SIFT features need to be reduce the dimensionality of the feature space. BOF framework illustrated in Fig 1. Results vocabulary that can be used to explain each unique image as histogram with implementing clustering method in feature extraction. This method is used to compute histogram which is locate feature relative which is known as the Spatial Pyramid method[14].

Basically there are four stages of BOF stages:
- Detection and description of image features using SIFT
- Grouping descriptor to the set of clusters (vocabulary) with vector quantization algorithm using K-means
- Construction of a bag of features, which calculates the number of features that are entered on each cluster
- Classification, training bag of features as feature vectors, and determine category of the image

Ideally this step is designed to maximize classification accuracy and minimize computation. Therefore, not only rich enough to carry enough information to be discriminative at the category level but also the descriptors extracted in the first step should be invariant to variations that are irrelevant to the categorization task (image transformations, lighting variations and occlusions). In the second step, the vocabulary should be fit (not really big so it can distinguish relevant changes in image parts and not so small then it can distinguish irrelevant variations such as noise). In text categorization, the quantized feature vectors (cluster centers) as "keypoints" are assumed as "keywords". The main goal is to employee vocabulary which is possible to rise categorization perform in learning process[15]. Repeatable motifs and pattern on batik image base will impact on repeated feature extracted by SIFT so that this method can reduce the feature for better computation time.
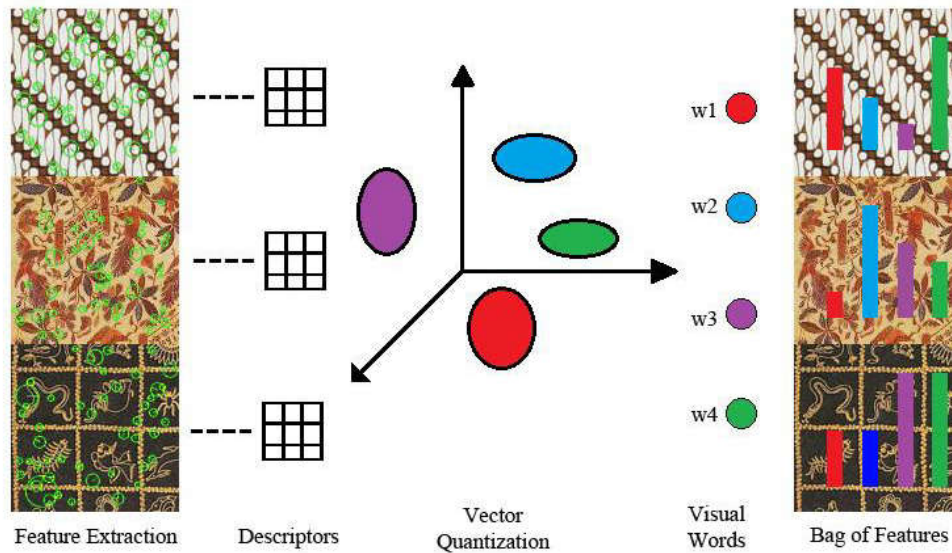
Fig. 1 Framework Bag of Features[16]

2.5. Support Vector Machine

Support Vector Machine (SVM) is popular classification method nowadays. It works well with high dimension data and SVM can use kernel function that can map original data to be higher dimensionality. In contrast to another classification method, SVM not use all of data to be learned in the learning process, but just several chosen data is contributed to build model in learning proses. This research use SVM because features used have big dimensionality depending on the number of vocabulary. For detail classification process using SVM, the process can be seen at [17].

3. Experimental Studies

We implemented the system on the environment, hardware which is used is a laptop with specification Intel Core i5-3317U @1.70 GHz processor and 4 Gb RAM, Windows 7 64 bit is the system operation, and Matlab R2013a application is software used in this study. SIFT feature extraction method in this research is library from VLFeat which is library to implement popular computer vision algorithms specializing in image understanding and local features extraction and matching[18].

In this study, there are 3 scenarios which are implemented to test proposed batik image classification method. The first scenario is to test effect choosing the number of vocabulary used. The next is robustness of system to rotated datasets examined. The last case study is with scaling dataset to test the method. Measurement to notice classification accuracy.

$$Accuracy = \frac{n_b}{N} \times 100\% \quad (1)$$

Where n_b is the sum of test data classified correct, and N is the sum of testing data.

3.1. The number of vocabularies

The number of vocabularies to build clustering model is the substantial factor which has impact to the classification accuracy. The main goal this part is to analyze how many cluster is fit enough to improve the accuracy using SIFT on grayscale image. The numbers of clusters which is implemented are 400, 800, 1200 2000, 2500, 2800 and 3000.

3.2. Rotated dataset

We rotate the testing data in some angles 15^0, 45^0, 90^0, 135^0 and 180^0 in order to prove that the method is invariant to rotation.

3.3. Scaled dataset

In order to prove that the method is invariant to scale, we zoom in the testing data in range 10%, 20%, 30%, 50% and 100% and we zoom out the testing data in range 10%, 20%, 30% and 50%.

4. Result and Discussion

Testing result of the experiment describe on Table 1, Table 2, and Table 3. Every scenario was implemented three times trial.

Table 1. The testing result of the number of clusters

NO.	The Number of Clusters						
	400	800	1200	2000	2500	2800	3000
1	95%	96%	96%	97%	98%	100%	99%
2	96%	97%	99%	99%	99%	100%	100%
3	93%	98%	96%	98%	96%	100%	99%
Maximum	96%	98%	99%	99%	99%	100%	100%
Average	94.67%	97%	97%	98%	97.67%	100%	99.33%

Table 2. The testing result of rotated images

NO.	Image Rotation				
	15^0	45^0	90^0	135^0	180^0
1	94	93	99	94	99
2	97	92	98	90	99
3	99	92	97	91	98
Maximum	99	93	99	94	99
Average	96.67	92.33	98.00	91.67	98.67

Table 3. The testing result of scaled images

NO.	Image Scale	
	Zoom Out	Zoom In

	10%	20%	30%	50%	10%	20%	30%	50%	100%
1	97%	84%	72%	32%	95%	92%	90%	85%	67%
2	98%	84%	67%	38%	96%	95%	91%	85%	60%
3	99%	83%	68%	28%	97%	92%	90%	86%	71%
Maximum	99%	84%	72%	38%	97%	95%	91%	86%	71%
Average	98%	83.67%	69%	32.67%	96%	93%	90.33%	85.33%	66%

In three times trial with normal data (without rotation and scale), the best number of clusters are 2800 clusters. The result shows that the average accuracy is 97.67%. Based on result reached on the first scenario, the number of clusters used are 2800 clusters. This scenario was built to find out the best number of clusters that can impact to the better result. As we said in the previous section the vocabulary should be fit (not really big so it can distinguish relevant changes in image parts and not so small then it can distinguish irrelevant variations such as noise).

The next scenario is to do testing on rotated images with 15°, 45°, 90°, 135° and 180° angle. At table 2 shows that from three times trial, system can give satisfying results. The lowest average result given is 92.33% for rotated image with 45^0 angle while the highest average result is 98.67% for rotated image with 180^0 angle.

The last scenario is to do testing at scaled image (zoom in and out) in the some scale. The highest average research accuracy result given when images zoomed out is 98% for 10% scale, while the lowest average result is 32.67% for images which is zoomed out at 50%. The highest average accuracy result for images zoomed in can reach 96% for 10% scale and the lowest average result is 66% for 100% scale.

It is true with the result above, that extracted SIFT feature effects classification method robustness against rotation and scale.

5. Conclusion

Based on observations during the process of design, implementation, until the testing process and system analysis, we concluded that the system classification batik image is built through several stages of process which are converting RGB image into grayscale image, feature extraction using BOF and SIFT and classification method using SVM. The system built has been able to do identification process of batik image type very well. Based on the analytical results obtained, the amount of accuracy is determined by the number of clusters or the appropriate amount of vocabulary, because the appropriate amount of cluster can describe image properly. Whereas for the overall classification process using the BOF and SIFT shows that this system can perfectly label batik image types. Future experiment using more number of batik data from various regions of Indonesia can be carried out. the selection of the appropriate number of clusters should be done because in this study we still need to do repeatable experiment to fine the optimal number of clusters and the authors hope that the existing method can be developed in terms of speeding-up the computing time.

References

[1]   A. haake, "The Role Of Symmetry In Javanese Batik Patterns," *Comput. Math. Applic.*, vol. 17, no. 4, pp. 815–826, 1989.
[2]   A. E. Minarno, Y. Munarko, A. Kurniawardhani, F. Bimantoro, and N. Suciati, "Texture Feature Extraction Using Co-Occurrence Matrices of Sub-Band Image For Batik Image Classification," in *International Conference on Information and Communication Technology (ICoICT)*, 2014, pp. 249–254.
[3]   A. H. Rangkuti, "Content Based Batik Image Classification Using Wavelet Transform And Fuzzy Neural Network," *J. Comput. Sci.*, vol. 10, no. 4, pp. 604–613, 2014.
[4]   A. Kurniawardhani, N. Suciati, and I. Arieshanti, "Klasifikasi Citra Batik Menggunakan Metode Ekstraksi Ciri Yang Invariant Terhadap Rotasi," *JUTI*, vol. 12, pp. 48–60, 2014.

[5]   D. Willy, A. Noviyanto, and A. M. Arymurthy, "Evaluation Of SIFT And SURF Features In The Songket Recognition," *ICACSIS*, pp. 978–979, 2013.

[6]   D. H. Fals, "Pengenalan Pola Batik Menggunakan Metode SIFT," Universitas Mercu Buana, 2014.

[7]   R. Akta, "Klasifikasi Motif Batik Menggunakan Metode Scale Invariant Feature Transform," Universitas Indonesia, 2012.

[8]   D. G. Lowe, "Distinctive Image Features From Scale-Invariant Keypoints," pp. 1–29, 2004.

[9]   K. Roy, G. Subhramanya, and V. R. K. Rao, "ART Based Clustering Of Bag-Of-Features For Image Classification," *2012 5th Int. Congr. Image Signal Process. CISP 2012*, no. Cisp, pp. 841–846, 2012.

[10]   X. Lou, D. Huang, L. Fan, and A. Xu, "An Image Classification Algorithm Based on Bag of Visual Words and Multi-kernel Learning," *J. Multimed.*, vol. 9, no. 2, pp. 269–277, 2014.

[11]   S. Lazebnik and C. Schmid, "Beyond Bags Of Features : Spatial Pyramid Matching For Recognizing Natural Scene Categories," *CVPR'06*, 2006.

[12]   N. H. Dardas and N. D. Georganas, "Real-Time Hand Gesture Detection And Recognition Using Bag-of-Features and Support Vector Machine Techniques," *3592 IEEE Trans. Instrum. Meas.*, vol. 60, no. 11, pp. 3592–3607, 2011.

[13]   W. Lin, Y. Wu, W. Hung, and C. Tang, "A Study of Real-Time Hand Gesture Recognition Using SIFT on Binary Images," pp. 235–246, 2013.

[14]   J. Kidd, "Experimenting With Frameworks for Visual Object Recognition," University of South Florida.

[15]   G. Csurka, C. R. Dance, L. Fan, J. Willamowski, C. Bray, and D. Maupertuis, "Visual Categorization with Bags of Keypoints," *Xerox Res. Cent. Eur.*, 2AD.

[16]   R. J. López-Sastre, J. Renes-Olalla, P. Gil-Jiménez, and S. Maldonado-Bascón, "Visual Word Aggregation," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 6669 LNCS, pp. 676–683, 2011.

[17]   Christopher J.C. Burges, " A tutorial On Support Vector Machine For Pattern Recognition," Journal Data Mining and Knowledge Discovery, Vol.2 Issue 2, June 1998, pp 121-167.

[18]   "VLFeat Open Source Computer Vision Library." [Online]. Available: http://www.vlfeat.org/.