Batik Classification using Transformation-Invariant Features

Review on recent related researches

Yohanes Gultom - 1506706345

October 5, 2016

IKO61181 Advance Image Processing

Table of contents

Introduction

Introduction

- **Batik** is one of the most profound cultural heritage in Indonesia so continuous **research is necessary**

Introduction

- **Batik** is one of the most profound cultural heritage in Indonesia so continuous **research is necessary**
- Current classification methods are **robust** enough to noise addition, compression and retouching **but not to variance in transformations** (translation, rotation, scaling) [Nurhaida et al., 2015]

Introduction

- **Batik** is one of the most profound cultural heritage in Indonesia so continuous **research is necessary**
- Current classification methods are **robust** enough to noise addition, compression and retouching **but not to variance in transformations** (translation, rotation, scaling) [Nurhaida et al., 2015]
- Recent improvements motivated by transformation invariance feature extractors such as SIFT & SURF and deep learning

Literature Reviews

Nurhaida et al. [2015]

Automatic Indonesian's Batik Pattern Recognition Using SIFT Approach (2015)

Nurhaida, Ida and Noviyanto, Ary and Manurung, Ruli and Arymurthy, Aniati M

- Using **Scale-Invariant Feature Transform (SIFT) descriptors** to calculate similarity between Batik images

Nurhaida et al. [2015]

Automatic Indonesian's Batik Pattern Recognition Using SIFT Approach (2015)

Nurhaida, Ida and Noviyanto, Ary and Manurung, Ruli and Arymurthy, Aniati M

- Using **Scale-Invariant Feature Transform (SIFT) descriptors** to calculate similarity between Batik images
- **Voting Hough Transform** was applied to the descriptors to eliminate mismatched keypoint candidates

Nurhaida et al. [2015]

Automatic Indonesian's Batik Pattern Recognition Using SIFT Approach (2015)

Nurhaida, Ida and Noviyanto, Ary and Manurung, Ruli and Arymurthy, Aniati M

- Using **Scale-Invariant Feature Transform (SIFT) descriptors** to calculate similarity between Batik images
- **Voting Hough Transform** was applied to the descriptors to eliminate mismatched keypoint candidates
- Achieve **91.53%** accuracy on 20 batik patterns each with 6 transformation variations

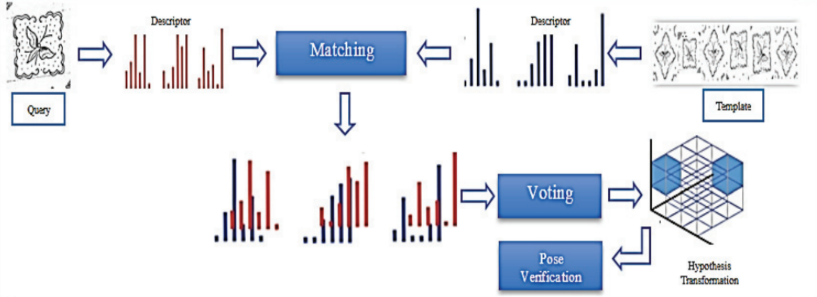

Figure 1: SIFT with hough voting method for Batik classification

Lowe [2004]

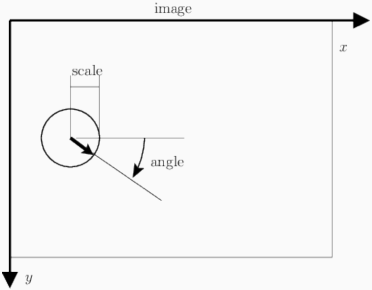

Figure 2: SIFT Keypoint

Figure 3: SIFT Keypoints in Batik Parang

Lowe [2004]

Distinctive image features from scale-invariant keypoints (2004)

Lowe, David G

- SIFT keypoint is an **point of interest** in image obtained by detecting extrema of Difference of Gaussian (DoG) pyramid

Lowe [2004]

Distinctive image features from scale-invariant keypoints (2004)

Lowe, David G

- SIFT keypoint is an **point of interest** in image obtained by detecting extrema of Difference of Gaussian (DoG) pyramid
- A keypoint has 4 parameters: center **coordinates** x and y, **scale** and its **orientation** (an angle expressed in radians) (Figure 2)

Lowe [2004]

Distinctive image features from scale-invariant keypoints (2004)

Lowe, David G

- SIFT keypoint is an **point of interest** in image obtained by detecting extrema of Difference of Gaussian (DoG) pyramid
- A keypoint has 4 parameters: center **coordinates** x and y, **scale** and its **orientation** (an angle expressed in radians) (Figure 2)
- An image has **multiple keypoints** (Figure 3)

Lowe [2004]

Distinctive image features from scale-invariant keypoints (2004)

Lowe, David G

- SIFT keypoint is an **point of interest** in image obtained by detecting extrema of Difference of Gaussian (DoG) pyramid
- A keypoint has 4 parameters: center **coordinates** x and y, **scale** and its **orientation** (an angle expressed in radians) (Figure 2)
- An image has **multiple keypoints** (Figure 3)
- A keypoint descriptor is is a 3-dimensional (**128-elements sparse array**) spatial **histogram of the image gradients** characterizing a SIFT keypoint

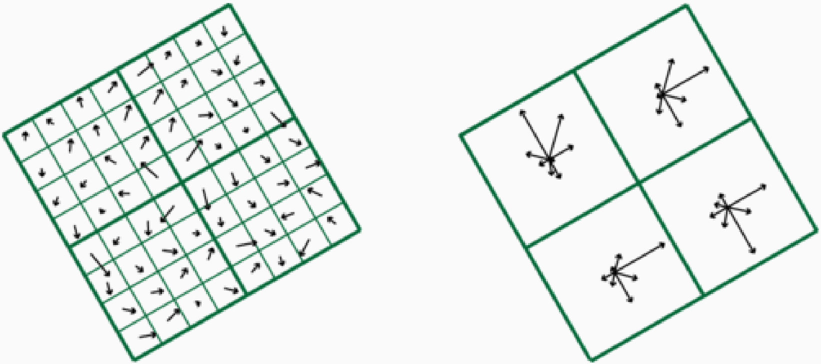

Figure 4: SIFT descriptor

Azhar et al. [2015]

Batik Image Classification Using SIFT Feature Extraction, Bag of Features and Support Vector Machine (2015)

Azhar, Ryfial and Tuwohingide, Desmin and Kamudi, Dasrit and Suciati, Nanik

- Classification using **support vector machine (SVM)** fed by **bag of words (BOF)** features extracted using **SIFT** descriptors

Azhar et al. [2015]

Batik Image Classification Using SIFT Feature Extraction, Bag of Features and Support Vector Machine (2015)

Azhar, Ryfial and Tuwohingide, Desmin and Kamudi, Dasrit and Suciati, Nanik

- Classification using **support vector machine (SVM)** fed by **bag of words (BOF)** features extracted using **SIFT** descriptors
- SIFT descriptors were clustered using **k-means vector quantization** algorithm to build vocabularies for BOF

Azhar et al. [2015]

Batik Image Classification Using SIFT Feature Extraction, Bag of Features and Support Vector Machine (2015)

Azhar, Ryfial and Tuwohingide, Desmin and Kamudi, Dasrit and Suciati, Nanik

- Classification using **support vector machine (SVM)** fed by **bag of words (BOF)** features extracted using **SIFT** descriptors
- SIFT descriptors were clustered using **k-means vector quantization** algorithm to build vocabularies for BOF
- Similar to [Nurhaida et al., 2015], SIFT descriptors are **not used directly** in matching

Azhar et al. [2015]

Batik Image Classification Using SIFT Feature Extraction, Bag of Features and Support Vector Machine (2015)

Azhar, Ryfial and Tuwohingide, Desmin and Kamudi, Dasrit and Suciati, Nanik

- Classification using **support vector machine (SVM)** fed by **bag of words (BOF)** features extracted using **SIFT** descriptors
- SIFT descriptors were clustered using **k-means vector quantization** algorithm to build vocabularies for BOF
- Similar to [Nurhaida et al., 2015], SIFT descriptors are **not used directly** in matching
- **Very good** average accuracy of **97.67%** for normal images, **95.47%** for rotated images and **79%** for scaled images

Azhar et al. [2015]

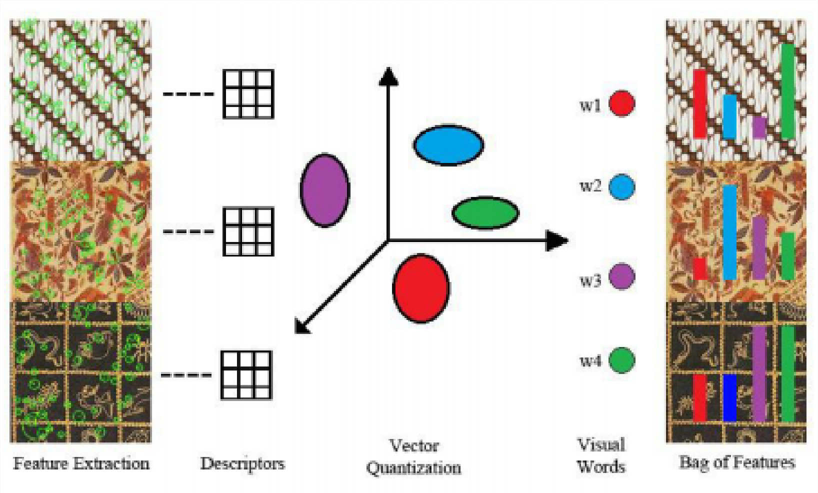

Figure 5: SIFT for building bag of words visual vocabularies

Willy et al. [2013]

Evaluation of SIFT and SURF features in the songket recognition (2013)

Willy, Dominikus and Noviyanto, Ary and Arymurthy, Aniati Murni

- Proved that **SURF were faster than SIFT** for Songket classification but had less accuracy

Willy et al. [2013]

Evaluation of SIFT and SURF features in the songket recognition (2013)

Willy, Dominikus and Noviyanto, Ary and Arymurthy, Aniati Murni

- Proved that **SURF were faster than SIFT** for Songket classification but had less accuracy
- Unlike Moertini and Sitohang [2005] and Azhar et al. [2015] **SIFT and SURF features were directly** used to compute the matching scores

Willy et al. [2013]

Evaluation of SIFT and SURF features in the songket recognition (2013)

Willy, Dominikus and Noviyanto, Ary and Arymurthy, Aniati Murni

- Proved that **SURF were faster than SIFT** for Songket classification but had less accuracy
- Unlike Moertini and Sitohang [2005] and Azhar et al. [2015] **SIFT and SURF features were directly** used to compute the matching scores
- Matching scores are calculated by (1) the number of **matched keypoints** and (2) the **average total distance of the n-nearest keypoints**

Willy et al. [2013]

Evaluation of SIFT and SURF features in the songket recognition (2013)

Willy, Dominikus and Noviyanto, Ary and Arymurthy, Aniati Murni

- Proved that **SURF were faster than SIFT** for Songket classification but had less accuracy
- Unlike Moertini and Sitohang [2005] and Azhar et al. [2015] **SIFT and SURF features were directly** used to compute the matching scores
- Matching scores are calculated by (1) the number of **matched keypoints** and (2) the **average total distance of the n-nearest keypoints**
- Accuracy with SIFT features was **92-100%** and **65-97%** with SURF.

Willy et al. [2013] (continued)

Evaluation of SIFT and SURF features in the songket recognition (2013)

Willy, Dominikus and Noviyanto, Ary and Arymurthy, Aniati Murni

- With SURF features, the **accuracy dropped** quite significant if salt and pepper noises were added while SIFT was more stable.

Willy et al. [2013] (continued)

Evaluation of SIFT and SURF features in the songket recognition (2013)

Willy, Dominikus and Noviyanto, Ary and Arymurthy, Aniati Murni

- With SURF features, the **accuracy dropped** quite significant if salt and pepper noises were added while SIFT was more stable.
- Not paying much attention to **transformation variance** unlike Azhar et al. [2015] and Nurhaida et al. [2015]

Willy et al. [2013]

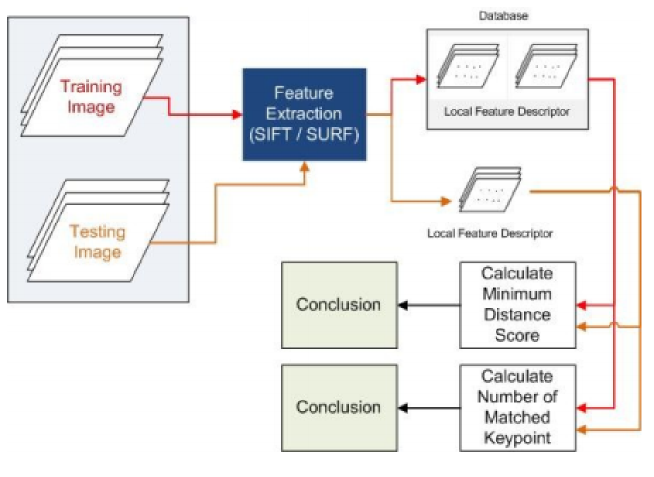

Figure 6: SIFT vs SURF method for Songket classification

LeCun et al. [2015]

Deep learning (2015)

LeCun, Yann and Bengio, Yoshua and Hinton, Geoffrey

- Deep learning is a **multilayer representation learning** in artificial neural network

LeCun et al. [2015]

Deep learning (2015)

LeCun, Yann and Bengio, Yoshua and Hinton, Geoffrey

- Deep learning is a **multilayer representation learning** in artificial neural network
- Representation learning is a method in machine learning to **automatically extract/learn representation** (features) from raw data

LeCun et al. [2015]

Deep learning (2015)

LeCun, Yann and Bengio, Yoshua and Hinton, Geoffrey

- Deep learning is a **multilayer representation learning** in artificial neural network
- Representation learning is a method in machine learning to **automatically extract/learn representation** (features) from raw data
- Basic deep learning architectures: convolutional neural network (ConvNet), deep belief network (DBN), autoencoder (AE) and recurrent neural network (RNN)

LeCun et al. [2015]

Deep learning (2015)

LeCun, Yann and Bengio, Yoshua and Hinton, Geoffrey

- Deep learning is a **multilayer representation learning** in artificial neural network
- Representation learning is a method in machine learning to **automatically extract/learn representation** (features) from raw data
- Basic deep learning architectures: convolutional neural network (ConvNet), deep belief network (DBN), autoencoder (AE) and recurrent neural network (RNN)
- Deep architectures can **extract transformation-variant features** (eg ConvNet for object detection)

Menzata [2014]

Sistem perolehan citra berbasis konten dan klasifikasi citra batik dengan convolutional stacked autoencoder (2014)

Menzata, Remmy Augusta

- Batik classification using **convolutional stacked autoencoder**

Menzata [2014]

Sistem perolehan citra berbasis konten dan klasifikasi citra batik dengan convolutional stacked autoencoder (2014)

Menzata, Remmy Augusta

- Batik classification using **convolutional stacked autoencoder**
- Using convolutional transformations to **reduce the input nodes** of stacked autoencoder

Menzata [2014]

Sistem perolehan citra berbasis konten dan klasifikasi citra batik dengan convolutional stacked autoencoder (2014)

Menzata, Remmy Augusta

- Batik classification using **convolutional stacked autoencoder**
- Using convolutional transformations to **reduce the input nodes** of stacked autoencoder
- Achieved **81,73%** accuracy (trained using small patches). When noises were added its accuracy dropped to **49%** for gaussian noises, **61%** for rotations, **70%** for scalings and **75%** for illumination noises. Less accurate than Azhar et al. [2015] and Nurhaida et al. [2015]

Menzata [2014]

Sistem perolehan citra berbasis konten dan klasifikasi citra batik dengan convolutional stacked autoencoder (2014)

Menzata, Remmy Augusta

- Batik classification using **convolutional stacked autoencoder**
- Using convolutional transformations to **reduce the input nodes** of stacked autoencoder
- Achieved **81,73%** accuracy (trained using small patches). When noises were added its accuracy dropped to **49%** for gaussian noises, **61%** for rotations, **70%** for scalings and **75%** for illumination noises. Less accurate than Azhar et al. [2015] and Nurhaida et al. [2015]
- While Fischer et al. [2014] showed that **ConvNet outperformed SIFT**

Future Works

Future Works

- **Deep learning** architecture on Batik classification may give better result that current approaches as suggested by Fischer et al. [2014]

Future Works

- **Deep learning** architecture on Batik classification may give better result that current approaches as suggested by Fischer et al. [2014]
- As mentioned in Menzata [2014] suggestion, usage of **other deep architectures** than Convolutional Stacked Autoencoder such as Convulutional Neural Network (ConvNet) may improve result

Future Works

- **Deep learning** architecture on Batik classification may give better result that current approaches as suggested by Fischer et al. [2014]
- As mentioned in Menzata [2014] suggestion, usage of **other deep architectures** than Convolutional Stacked Autoencoder such as Convulutional Neural Network (ConvNet) may improve result
- Consider **replacing SoftMax** layer with other algorithm such as Naive Bayes Classifier (NBC) or Support Vector Machine (SVM).

Future Works

- **Deep learning** architecture on Batik classification may give better result that current approaches as suggested by Fischer et al. [2014]
- As mentioned in Menzata [2014] suggestion, usage of **other deep architectures** than Convolutional Stacked Autoencoder such as Convulutional Neural Network (ConvNet) may improve result
- Consider **replacing SoftMax** layer with other algorithm such as Naive Bayes Classifier (NBC) or Support Vector Machine (SVM).
- **Transfer learning** model from another training session such as from popular VGG-16 by Simonyan and Zisserman [2014]

Research Proposal

Batik classification using convolutional neural network (convnet) and transfer learning

- Using dataset from Menzata [2014]
- Experimenting with ImageNet model from Simonyan and Zisserman [2014]
- Compare method with direct SIFT descriptor matching from Willy et al. [2013]

References

Ryfial Azhar, Desmin Tuwohingide, Dasrit Kamudi, Nanik Suciati, et al. Batik image classification using sift feature extraction, bag of features and support vector machine. *Procedia Computer Science*, 72:24–30, 2015.

Philipp Fischer, Alexey Dosovitskiy, and Thomas Brox. Descriptor matching with convolutional neural networks: a comparison to sift. *arXiv preprint arXiv:1405.5769*, 2014.

Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521(7553):436–444, 2015.

David G Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, 2004.

Remmy Augusta Menzata. Sistem perolehan citra berbasis konten dan klasifikasi citra batik dengan convolutional stacked autoencoder, 2014.

Veronica S Moertini and Benhard Sitohang. Algorithms of clustering and classifying batik images based on color, contrast and motif. *Journal of Engineering and Technological Sciences*, 37(2):141–160, 2005.

Ida Nurhaida, Ary Noviyanto, Ruli Manurung, and Aniati M Arymurthy. Automatic indonesian's batik pattern recognition using sift approach. *Procedia Computer Science*, 59:567–576, 2015.

Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

Dominikus Willy, Ary Noviyanto, and Aniati Murni Arymurthy. Evaluation of sift and surf features in the songket recognition. In *Advanced Computer Science and Information Systems (ICACSIS), 2013 International Conference on*, pages 393–396. IEEE, 2013.

Thank you