Performance evaluation of SIFT against common image deformations on Iban plaited mat motif images

Silvia Joseph, Irwandi Hipiny, Hamimah Ujir, Sarah Flora Samson Juan, Jacey-Lynn Minoi
Faculty of Computer Science and Information Technology, Universiti Malaysia Sarawak, Malaysia

ABSTRACT

Decorative plaited mat is one of the many examples of rich plait work often seen on Borneo handicraft products. The plaited mats are decorated with simple and complex motif designs; each has its own special meaning and taboos. The motif designs are used as a reflection of environment and the traditional beliefs in the Iban community. In line with efforts from UNESCO’s and Sarawak Government’s, digitization, and the use of IR4.0 technologies to preserve and promote this cultural heritage is encouraged. Toward this end goal, we present a novel image dataset containing 10 Iban plaited mat motif classes. The plaited mat motifs are made of diagonal and symmetrical shapes, as well as geometric and non-geometric patterns. Classification’s accuracy using scale-invariant feature transform (SIFT) features was evaluated against 6 common image deformations: zoom+rotation, viewpoint, image blur, JPEG compression, scale and illumination, across multiple threshold values. Varying degrees of each deformation were applied to a digitally cleaned (and cropped) image of each mat motif class. We used RANSAC to remove outliers from the noisy SIFT matching result. The optimal threshold value is 2.0e-2 with a reported 100.0% matching accuracy for the scale change and zoom+rotation set.

Keywords: Iban plaited mat motifs, Pattern matching, Scale-invariant feature transform, RANSAC

This is an open access article under the CC BY-SA license.

Corresponding Author:
Silvia Joseph
Faculty of Computer Science and Information Technology
Universiti Malaysia Sarawak
94300 Kota Samarahan, Sarawak, Malaysia
Email: silvia.joe8@gmail.com

1. INTRODUCTION

Plait work is one of the oldest non-lithic crafts still being practiced by the many indigenous communities of the world, including here in Malaysia. According to [1], there is minimal work on documenting the rich history of plait work in Borneo. In Sarawak, the complex plaited mat motifs are believed by the Iban people to be magical and describe the philosophical meaning of life (and humankind), therefore must be treated with the utmost respect. Efforts towards preserving and promoting Borneo plait work to the global community as well as the younger generation are aplenty. One interesting direction is to develop educational apps utilizing the smartphone’s camera to capture images of these plaited mats. Recognition of the mat motifs would be an amazing feature to have. As reported by [2], the use of image recognition apps among users on their mobile phones or smart devices, such as Pinterest lens and Google lens has emerged as a prevailing trend. These apps work by extracting useful information via an analysis of the visual elements present inside the captured image.

Our work focuses on the Iban plaited mat motifs; Iban being one of the many Sarawak’s indigenous communities. These plaited mats contain unique motifs that are either simple or complex. Patterns incorporated in these mat motifs, each has an embedded meaning representing traditional Iban beliefs with
profound social and ritual significance [3]. The work in [4] detailed out the naming process of Iban plaited mat motifs after natural phenomena. However, the motifs are mostly stylized beyond recognition and are based on loose categories. The motifs include natural elements such as plants, animals, firmaments, and faerie. Besides the classics, new patterns may also be revealed to the craftswoman via her dream. In such cases, only she can name and plait the pattern. The standard rectangular plaited mat, see Figure 1, consists of an outer edge and two or more inner frames of transposed twill. Typically, small decorative motifs are incorporated as fillers. The centre of the mat, otherwise known as the pulau, may display a single main motif as one specific instance or in a repetitive pattern. The pulau may also contains multiple motifs. The main motif plays the dominant role of the specific meaning, whereas the smaller motifs only serve as fillers for filling in the empty area surrounding the main motif. The classification task is further complexified by the variety of positions, scales, and rotations of the major and minor motifs.

In this work, the accuracy of Lowe’s SIFT [5] descriptor was evaluated across multiple threshold values: 0.02, 0.04, 0.06 and 0.08. Using SIFT matches, we classified a dataset of mat motif images produced by varying the degree of image deformations on the original mat motif image. A robust geometric estimation algorithm, i.e. namely random sample consensus algorithm (RANSAC) [6] was implemented to remove outliers. RANSAC works by identifying the homography matrix relating two given images. For the evaluation criterion, we use a similar method to the one proposed in [7]. A visualized explanation of this study is shown in Figure 2.

Figure 1. An example of (a) Buah Talam Kumang (Malaysian language) plaited mat with two “pulau” and (b) Buah Nabau Ngelantar (Malaysian language) plaited mat with a single large “pulau”. The (2); decorative edge is positioned between two strengthening rows of transposed twill, i.e., “sapa” (1, 3).

Figure 2. A visualized explanation of our proposed method

We observed that each motif has its own distinctive features. Thus, to recognize the plaited mat motif, we need to build the descriptor using these discriminative features. To date, a significant amount of research had been done to perform robust detection, matching and recognition of discriminative features points inside an image. Feature detectors and matching algorithms had been developed for various purposes,

Performance evaluation of SIFT against common image deformations on iban plaited... (Silvia Joseph)
such as pattern recognition [8]-[12], features recognition [13]-[16], and gesture recognition [17], among others. This work aims to evaluate the performance of SIFT [18] against common image deformations applied to a clean synthetic mat motif image. The deformations are blur, scale, viewpoint, JPEG compression and zoom+rotation. SIFT algorithm is chosen due to its robustness against changes in scale, viewpoint, and illumination [19]. According to [20], SIFT algorithm is a good fit for image matching and object recognition. However, a classification task might return a false match due to noisy SIFT feature matching [21]. Thus, the RANSAC algorithm proposed by [6] is implemented to eliminate false matching features. RANSAC was used in several past studies and had been proven to improve the matching accuracy [22]-[24]. The experimental results in this study are presented in the form of quantitative comparison, i.e., number of detected key points inside the query and reference image, number of matching key point, and number of correct matches. To the best of our knowledge, no existing work had evaluated the use of SIFT and RANSAC in the context of Iban plaited mat motifs. The rest of this paper is organized as follows: Section 2 detailed out the data collection and pre-processing process, Section 3 outlined the proposed method. The results and analysis are presented in Section 4 whilst the conclusion is provided in Section 5.

2. DATA COLLECTION AND PRE-PROCESSING

This work was implemented on a laptop with the following specifications: Intel Core i7-8550U @ 1.80GHz processor and 8 GB RAM. We used the SIFT and RANSAC implementation in VLFeat Version 0.9.21. The open-source library was implemented in Matlab 2017a. VLFeat provides ready implementations of popular computer vision algorithms for image understanding, as well as for local features extraction and matching [25].

Images of the decorative plaited mat motifs were captured indoor using a setup consisting of a downward-facing camera. In total, 50 motif classes were collected. Nevertheless, for this initial work of ours, only 10 motifs were used in the experiments. The full dataset will be released to the public in our future publications. We labelled the class by the most prominent motif on the plaited mat. The images were then resized to a resolution of 800 X 355 pixels (original resolution is 5184 x 3456 pixels) for performance reason.

The synthetic plaited mat motif images, as seen in Figure 3 (Set 1 to Set 6), are the test sequence for each class in our dataset. The Level 1 images (top-most) were used as the query images during the matching task for evaluating SIFT performance against image blur, illumination changes, JPEG compression, viewpoint changes, zoom+rotation and scale changes. For each motif class, we prepared a test sequence for every image deformation. A test sequence contains five images exhibiting gradual increments of geometric and/or photometric transformation. In total, the dataset contains 10 motif classes x 6 image deformation types x 5 deformation levels, yielding a total of 300 images, see Figure 3. The gradual incremental values for each image deformation are tabulated in Table 1.

![Figure 3. Selected test images; from left to right showing the clean synthetic query images of Buah Balan Sebayan (Malaysian language) sequence for blur changes, Buah Mata Punai (Malaysian language) sequence for illumination change, Buah Engkatak Tingkap Pama Merap (Malaysian language) sequence for viewpoint change, Buah Panchar Matalari (Malaysian language) sequence for JPEG compression change, Buah Burung Kenyalang (Malaysian language) sequence for zoom+rotation change, and Buah Nabau Besundang (Malaysian language) sequence for scale change.](image_url)

We decided to use clean synthetic images instead of actual photographed images because we wish to prevent the background noise from affecting the resulting image descriptors. Furthermore, synthetic datasets improve the validation and training of machine learning models, according to [26], [27].
ing. The Iban’s plaited mat motif recognition is expected since SIFT is invariant. This observation is attributable to a higher degree of similarity across image and refining the matrix based on pairs of detected SIFT key points inside the query vs. reference image, the number of detected SIFT key points decreases, the number of SIFT keypoints was higher on geometric counterparts (i.e., Buah Bulan Sebayan, Buah Engkatak Tingkap Pama Merapat, Buah Panchar Matahari, Buah Talam Kuma) compared to their non-geometric counterparts (i.e., Buah Burung Kenyalang, Buah Nabau Besundang, Buah Nabau Ngelantar and Buah Bung Midang).

Table 1. Incremental values for each image deformation

<table>
<thead>
<tr>
<th>Image Type</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image Blur</td>
<td>5%, 35%, 50%, 65%, 80%</td>
</tr>
<tr>
<td>Illumination</td>
<td>5%, 20%, 40%, 60%, 90%</td>
</tr>
<tr>
<td>JPEG compression</td>
<td>5%, 30%, 60%, 90%, 100%</td>
</tr>
<tr>
<td>Viewpoint</td>
<td>Fronto-parallel surface 10º, Slanted surface 10º, 30º, 60º, 90º</td>
</tr>
<tr>
<td>Scale</td>
<td>95%, 65%, 35%, 15%, 5%</td>
</tr>
</tbody>
</table>

3. OUR METHOD

The two main steps in a recognition task are feature extraction and feature matching [28]. The first step is for finding the distinctive key points and constructing the descriptor around each discovered key point. The resulting descriptors are then used for image matching. The Iban’s plaited mat motif recognition is performed using pairwise matching based on the nearest neighbor distance ratio (NNDR), as in [5]. NNDR has been widely used, for example in [29], [30], due to its proven effectiveness in local feature matching task. NNDR is implemented using (1), as in [5]. $D_2$ denotes a descriptor of the reference image, while $D_1$ and $D_2$ are the nearest, and second nearest neighbor to $D_A$ respectively. The ratio value in (1) is duly calculated and we reject the match if the ratio value is less than a specified threshold. The threshold values are 0.02, 0.04, 0.06 and 0.08. Similarity scores are normalized to a range of 0.0 to 1.0.

\[
\frac{\|D_A - D_1\|}{\|D_A - D_2\|} \leq t
\]

(1)

Once we have the SIFT matches (above the threshold value), we use RANSAC to find the homography matrix to eliminate any mismatches or possible false key points. RANSAC works by repeatedly selecting a random and small sets of data points in query and reference image and refining the matrix based on their correspondence. RANSAC algorithm allowed the identification of a projection matrix and the key points fitting the pair image relationship as much as possible [31].

4. RESULTS AND ANALYSIS

We recorded the number of detected SIFT key points inside the query vs. reference image, the number of correspondences, and number of correct matches after RANSAC. For image blurring, when the level of distortion increases, the number of detected SIFT key points decrease dramatically. However, the decrease in number does not correlate with loss of accuracy. False matches only started at the 5th level of deformation for image blur, see Table 2. As for illumination changes, the number of SIFT key points decrease dramatically starting at the 5th level of deformation. SIFT managed to detect a considerable amount of key points at increasing degree of Zoom+Rotation. We found that SIFT does not perform well in frontoparallel surface and slanted surface images, at a value ranging from 30º to 90º. We observed that the number of SIFT keypoints was higher on geometric-shaped mat motifs (i.e., Buah Bulan Sebayan, Buah Engkatak Tingkap Pama Merapat, Buah Panchar Matahari, Buah Talam Kumang and Buah Tungku) compared to their non-geometric counterparts (i.e., Buah Burung Kenyalang, Buah Nabau Besundang, Buah Nabau Ngelantar and Buah Bung Midang). This observation is attributable to a higher degree of similarity and the more repetitive patterns inside geometric-shaped motifs. SIFT managed to detect a sufficient number of SIFT key points, even at the worst deformation level for zoom+rotation. The average drop in number of SIFT key points, from Level 1 to Level 5 was less than 70% across classes. This is expected since SIFT is rotational-invariant. SIFT also performed well in JPEG compression, up to 90%. As for scale changes, the average drop in number of SIFT key points from Level 1 to Level 5 was less than 70% across classes, before and after RANSAC, see Table 3. Visualised examples of SIFT keypoints and the resulting descriptors are shown in Figure 4. Figure 5 shows the dramatic improvement of SIFT matching with RANSAC.

Table 2. Failure point of SIFT, at optimal threshold of 0.02, for each image deformation category

<table>
<thead>
<tr>
<th>Level of image deformation</th>
<th>Blur</th>
<th>Illumination</th>
<th>JPEG_Compression</th>
<th>Viewpoint</th>
<th>Zoom_rotation</th>
<th>Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image1</td>
<td>TP</td>
<td>TP</td>
<td>TP</td>
<td>FP</td>
<td>TP</td>
<td>TP</td>
</tr>
<tr>
<td>Image2</td>
<td>TP</td>
<td>TP</td>
<td>TP</td>
<td>TP</td>
<td>TP</td>
<td>TP</td>
</tr>
<tr>
<td>Image3</td>
<td>TP</td>
<td>TP</td>
<td>TP</td>
<td>FP</td>
<td>TP</td>
<td>TP</td>
</tr>
<tr>
<td>Image4</td>
<td>TP</td>
<td>TP</td>
<td>TP</td>
<td>FP</td>
<td>TP</td>
<td>TP</td>
</tr>
<tr>
<td>Image5</td>
<td>FP</td>
<td>FP</td>
<td>FP</td>
<td>FP</td>
<td>TP</td>
<td>TP</td>
</tr>
</tbody>
</table>
Table 3. Comparison in number of SIFT matches between deformation level 1 vs. 5 (optimal threshold of 0.02)

<table>
<thead>
<tr>
<th>Image Deformation</th>
<th>SIFT matches</th>
<th>Average drop in number (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>w/o RANSAC</td>
<td>w/ RANSAC</td>
</tr>
<tr>
<td>Illumination1</td>
<td>607</td>
<td>555</td>
</tr>
<tr>
<td>Illumination5</td>
<td>25</td>
<td>18</td>
</tr>
<tr>
<td>Image Blur1</td>
<td>625</td>
<td>582</td>
</tr>
<tr>
<td>Image Blur5</td>
<td>41</td>
<td>24</td>
</tr>
<tr>
<td>Viewpoint1</td>
<td>507</td>
<td>471</td>
</tr>
<tr>
<td>Viewpoint5</td>
<td>93</td>
<td>22</td>
</tr>
<tr>
<td>JPEG_Compression1</td>
<td>784</td>
<td>774</td>
</tr>
<tr>
<td>JPEG_Compression5</td>
<td>185</td>
<td>100</td>
</tr>
<tr>
<td>zoom_rotation1</td>
<td>524</td>
<td>452</td>
</tr>
<tr>
<td>zoom_rotation5</td>
<td>180</td>
<td>138</td>
</tr>
<tr>
<td>Scale1</td>
<td>540</td>
<td>471</td>
</tr>
<tr>
<td>Scale5</td>
<td>178</td>
<td>144</td>
</tr>
</tbody>
</table>

Figure 4. Visualised example of (a) detection result of SIFT key points and (b) built feature descriptor

Figure 5. Example of a SIFT matching result, before (Left) and after (Right) RANSAC

The average matching accuracy of SIFT at various threshold values is shown in Figure 6. The optimal threshold value is observed to be 0.02 with a comparably higher average matching accuracy compared to the rest. The confusion matrix for each image deformation is shown in Figure 7. The classification results were obtained using the optimal threshold value of 0.02. As can be seen, SIFT when paired with the optimal threshold value, performed quite well across all image deformations, except for the viewpoint sequence.
Figure 6. Average matching accuracy of SIFT, at threshold values of 0.02, 0.04, 0.06, and 0.08; following six image deformations

Figure 7. Confusion matrix from the classification results, at an optimal distance threshold value of 0.02, for (a) blur, (b) viewpoint, (c) illumination, (d) JPEG Compression, (e) Scale and (f) zoom + rotation

5. CONCLUSION
SIFT performed well, at optimal threshold value of 0.02, for all image deformation sequences at varying levels (i.e., degrees). The average accuracy obtained was 88.0%. Nevertheless, SIFT performed poorly on the viewpoint sequence. For future work, we wish to explore the use of global features, using color and edge-based descriptors and extend our work to include a comparative study for evaluating additional feature detectors and descriptors, such as speeded up robust features (SURF), features from accelerated segment test (FAST), maximally stable extremal region detector (MSER), and binary robust invariant scalable key points (BRISK).

ACKNOWLEDGEMENTS
We would like to acknowledge Universiti Malaysia Sarawak for providing the necessary fund towards the publication of this research work.
REFERENCES


Performance evaluation of SIFT against common image deformations on iban plaited… (Silvia Joseph)

Silvia Joseph is currently a postgraduate student at the Faculty of Computer Science and Information Technology (FCSIT), Universiti Malaysia Sarawak (UNIMAS). Working as a System Analyst at Malaysian Administrative Modernisation and Management Planning Unit (MAMPU). She obtained her Master in Advanced IT (MAIT) in 2016. Silvia’s research interests are on features recognition in fields of motifs analysis, image processing applications and image retrieval system.

Dr Irwandi Hipiny is a Senior Lecturer at the Faculty of Computer Science and Information Technology (FCSIT), Universiti Malaysia Sarawak (UNIMAS). He obtained his Ph.D in Computer Vision from University of Bristol in 2014. Irwandi’s research interests are in Computer Vision and Visual Animal Biometrics.

Ts. Dr. Hamimah Ujir is a Senior Lecturer at the Faculty of Computer Science and Information Technology (FCSIT), Universiti Malaysia Sarawak (UNIMAS). She obtained her Ph.D. in the School of Electronic, Electrical and Computer Engineering, University of Birmingham, United Kingdom in 2013. Hamimah’s research interests lay in the interdisciplinary field of computer vision, with related interests being in computer graphics, image processing, and mathematical methods. Her previous and current works include 3D physical simulation and 3D static and dynamic facial expression analysis.

Dr. Sarah F. S. Juan is a Senior Lecturer at the Faculty of Computer Science and Information Technology (FCSIT), Universiti Malaysia Sarawak (UNIMAS). She obtained her PhD in Computer Science from University of Grenoble-Alpes, France in 2015. Her research interests are on natural language processing and automatic speech recognition for under-resourced languages. She is also interested in mathematical modeling and computational thinking.

Dr. Jacey-Lynn Minoi is a Senior Lecturer of Computing and Data Analysis at Faculty of Computer Science and Information Technology (FCSIT), Universiti Malaysia Sarawak (UNIMAS). She obtained her Ph.D. in Visual Information Processing, Imperial College London 2009. Her research focusses on the development of multivariate tensor-based and multivariate statistical shape analysis algorithms for shape modelling across multiple dimensions and for pattern recognition in fields of face analysis, face recognition and image processing applications.