Jurnal Ilmiah KURSOR
Menuju Solusi Teknologi Informasi

Vol. 7, No. 4, Desember 2014 ISSN 0216 – 0544

IMPRESSION DETERMINATION OF BATIK IMAGE CLOTH BY MULTILABEL ENSEMBLE CLASSIFICATION USING COLOR DIFFERENCE HISTOGRAM FEATURE EXTRACTION

[a]Hani Ramadhan, [b]Isye Arieshanti, [c]Anny Yuniarti, [d] Nanik Suciati
[a,b,c,d] Informatics Engineering, Faculty of Information Technology, Institut Teknologi Sepuluh Nopember (ITS)
E-Mail: hani.its.042@gmail.com

Abstrak

Hampir setiap orang akan memperhatikan impresi busana yang dipakai, termasuk busana dengan motif batik. Namun, perpaduan berbagai motif dan warna batik memberikan impresi yang beragam. Sehingga, penentuan impresi dari satu kain batik menjadi sulit. Untuk membantu seseorang dalam menentukan impresi dari busana batik yang dipilih, dibutuhkan sistem yang mampu mengklasifikasikan impresi citra kain batik secara otomatis. Akan tetapi, pembuatan sistem klasifikasi label jamak merupakan memiliki tantangan tersendiri. Penelitian sebelumnya membuktikan bahwa metode klasifikasi ansambel label jamak dengan pencarian threshold mampu menjawab tantangan tersebut dengan kehandalannya dalam menangani himpunan data label jamak. Studi ini bertujuan untuk mengembangkan sistem yang menerapkan metode klasifikasi ansambel label jamak untuk menentukan impresi citra kain batik. Sistem ini memanfaatkan fitur tekstur dan warna yang dihasilkan dari Histogram Perbedaan Warna. Hasil uji coba metode ini memberikan performa yang baik dalam evaluasi label jamak. Nilai evaluasi tersebut antara lain Hamming Loss sebesar 0,173 dan Average Precision 0,866.

Kata kunci: Histogram Perbedaan Warna, Impresi Citra Kain Batik, Klasifikasi Label Jamak

Abstract

Many people will consider the fashion products' impression that will be worn, including the one with batik motif. Unfortunately, diverse impressions could be produced from combinations of the motif and color from a single batik cloth. Therefore, impression determination becomes a difficult case. To overcome this difficulty, an automatic batik cloth multi-impression classification system should be necessary to aid in choosing certain batik cloth. Nevertheless, this system implementation has its own intriguing challenge. Previous researches implied that multilabel ensemble classification method could deal with the problem against the highly imbalanced dataset. Thus, the aim of this study is to develop the multilabel classification system, which features come from the color and texture feature by Color Difference Histogram. From the test, this method demonstrated good performance by several multilabel evaluations, which are 0.173 by Hamming Loss and 0.866 by Average Precision.

Keywords: Color Difference Histogram, Batik Cloth Image Impression, Multi-Label Classification.

INTRODUCTION

Batik is a motif of a cloth that is produced by a coloring technique using wax. Batik is listed as Representative List of the Intangible Cultural Heritage of Humanity by UNESCO in October 2nd 2009 as an original culture of Indonesia. Batik fashion has been traded everywhere. The types of batik includes modern batiks (current Indonesian batik), Chinese batiks, Dutch batiks, and Hokokai Javanese [1] [2].

Beyond its production, batik has many elements which builds the motif and differs them by the others. Those elements include motif, color, shape, and production techniques variation. The motifs have philosophical meaning that gives certain impression to its wearer.

A system that recommends suitable impression for the customers personality would be helpful. The personality can be determined from the impression that emerged from the batik motif. However, batik motif is not only composed by a single motif. It can be composed by two or more motifs based on the creativity of the creator. Thus, it is needed to have a system which capable to classify a batik into one or more impression labels. The system is called multi-label classifier.

The characteristic of batik motif can be extracted by its composing features, such as colors, motif, and shapes variation. In previous study[3], it is shown that the variation of motif component is highly influenced by texture. Hence, the features used in this study are color and texture features, which extracted by Color Difference Histogram[4]. Color Difference Histogram utilizes the L*a*b* color space and edge orientation which are close to the human visual system. Then, those features are classified by their impression of batik.

A multi-label classifier is used to handle the various impression in a batik motif. Multi-label classifier determine an object labels by a subset of labels. The ensemble technique is used to approach the imbalance dataset problem [5]. Multi-label ensemble classification consists of some base classifiers. This study uses two methods as base classifiers, Ensemble of Classifier Chains (ECC) and Multi-label K-Nearest Neighbor.

The flow of the proposed method is depicted in Figure .

Color Difference Histogram

Color Difference Histogram [4] utilizes the CIEL*a*b* (in short L*a*b*) color space which is close to human visual perception [6]. Composing process of Color Difference Histogram is started from the conversion of standard RGB (Red-Green-Blue) color space to L*a*b* color space. This conversion takes the XYZ coordinate mapping by the matrix multiplication in (1). The result of mapping is used to conversion process to L*a*b* color space considering the illumination point (white reference point) of $[\boldsymbol{X_n}, \boldsymbol{Y_n}, \boldsymbol{Z_n}]$ = [0,950450, 1,000000, 1,088754], of D65 illumination. D65 is a illumination standard defined by International Commission on Illumination (CIE) [10]. An L*a*b* image is produced after (2), (3), (4), and (5) operation of each color dimension.

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0{,}412453 & 0{,}357580 & 0{,}180423 \\ 0{,}212671 & 0{,}715160 & 0{,}072169 \\ 0{,}019334 & 0{,}119193 & 0{,}950227 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{1}$$

$$\begin{cases} L* = 116\left(\frac{Y}{Y_n}\right)^{1/3} - 16 & \text{for } \frac{Y}{Y_n} > 0{,}008856 \\ L* = 903{,}3\left(\frac{Y}{Y_n}\right) & \text{for } \frac{Y}{Y_n} \le 0{,}008856 \end{cases} \tag{2}$$

$$\begin{cases} f(u) = u^{1/3} & \text{for } u > 0{,}008856 \\ f(u) = 7{,}787\,u + \frac{16}{116} & \text{for } u \le 0{,}008856 \end{cases} \tag{3}$$

$$a* = 500\left(f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right)\right) \tag{4}$$

$$b* = 500\left(f\left(\frac{X}{X_n}\right) - f\left(\frac{Z}{Z_n}\right)\right) \tag{5}$$

$$u = \frac{\partial L*}{\partial x} + \frac{\partial a*}{\partial x} + \frac{\partial b*}{\partial x} \tag{6}$$

After the L*a*b* conversion, an edge detection which produces edge quantization map of the L*a*b* image is performed. The process begins with edge orientation detection which yields horizontal and vertical vectors in (6) and (7), respectively. This process could be performed by a corresponding Sobel filter. The next process is the gradient vector finding of horizontal-horizontal vectors, vertical-vertical vectors, and horizontal-vertical vectors which respectively defined by (8), (9), and (10).

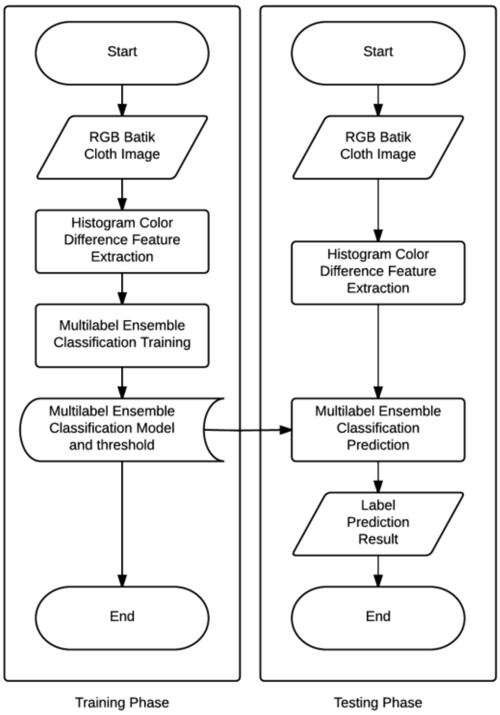


Figure 1. Proposed Method Flowchart

$$v = \frac{\partial L*}{\partial y} + \frac{\partial a*}{\partial y} + \frac{\partial b*}{\partial y} \quad (7)$$

$$g_{xx} = u^T u = \left|\frac{\partial L*}{\partial x}\right|^2 + \left|\frac{\partial a*}{\partial x}\right|^2 + \left|\frac{\partial b*}{\partial x}\right|^2 \quad (8)$$

$$g_{yy} = v^T v = \left|\frac{\partial L*}{\partial y}\right|^2 + \left|\frac{\partial a*}{\partial y}\right|^2 + \left|\frac{\partial b*}{\partial y}\right|^2 \quad (9)$$

$$g_{xy} = u^T v = \frac{\partial L*}{\partial x}\frac{\partial L*}{\partial y} + \frac{\partial a*}{\partial x}\frac{\partial a*}{\partial y} + \frac{\partial b*}{\partial x}\frac{\partial b*}{\partial y} \quad (10)$$

Direction calculation from the maximum rate of change $\varphi(x, y)$ from arbitrary vector $I(x, y)$ in $\boldsymbol{L^* a^* b^*}$ color space is shown by (11). The result value of the rate of change at (x, y) is defined by $G(x, y)$ as seen in (12).

$$\varphi(x, y) = \frac{1}{2}\arctan\left(\frac{2g_{xy}}{g_{xx} - g_{yy}}\right) \quad (11)$$

$$G(x, y) = \left\{\frac{1}{2}\left[\left(g_{xx} + g_{yy}\right) + \left(g_{xx} - g_{yy}\right)\cos 2\varphi + 2g_{xy}\sin 2\varphi\right]\right\}^{\frac{1}{2}} \quad (12)$$

If φ_0 is a solution of (11), thus another solution of $\varphi_0 \pm \pi/2$ is also possible because of the result of arctan operation. Then, the double values of rate of change at (x, y) are defined by $G_1(x, y)$ and $G_2(x, y)$ which computed by (13) and (14). Thus, the value of $\varphi(x, y)$ which represents the maximum gradient is computed by (15).

$$G_1 = \left\{\frac{1}{2}\left[\left(g_{xx} + g_{yy}\right) + \left(g_{xx} - g_{yy}\right)\cos 2\varphi_0 + 2g_{xy}\sin 2\varphi_0\right]\right\}^{\frac{1}{2}} \quad (13)$$

$$G_2 = \left\{\frac{1}{2}\left[\left(g_{xx} + g_{yy}\right) + \left(g_{xx} - g_{yy}\right)\cos 2\left(\varphi_0 + \frac{\pi}{2}\right) + 2g_{xy}\sin 2\left(\varphi_0 + \frac{\pi}{2}\right)\right]\right\}^{\frac{1}{2}} \quad (14)$$

$$\varphi(x, y) = \begin{cases} \varphi_0 & \text{if } \max\left(G_1(x, y), G_2(x, y)\right) = G_1(x, y) \\ \varphi_0 + \frac{\pi}{2} & \text{if } \max\left(G_1(x, y), G_2(x, y)\right) = G_2(x, y) \end{cases} \quad (15)$$

The value of $\varphi(x, y)$ will be quantized to m bins. An edge orientation map $\theta(x, y)$, which maximum value is V, and $\theta(x, y) = \phi, \phi \in 0,1, \ldots, V - 1$. If the value of V is 6, then all edge orientation will be uniformly quantized into range of 0°, 60°, 120°, 180°, 240°, and 300°. In this case, the number of bins is 18.

In the next step, the L*a*b* image will be quantized by color. The color quantization has benefits of limiting the color variation, as the human capacity of only 12 level of grayscale differentiation [7]. Generally, the steps of color quantization is chosen and applying a limited set of color for a color image in maximum fidelity [6]. In this case of quantization, the bin number of each color dimension of L*, a*, and b* are 10, 3 and 3 respectively.

The pixel value of color quantization is named$\boldsymbol{C(x, y)}$, where $\boldsymbol{0} < x < M$, and $\boldsymbol{0} < y < N$, M is the length of the image while N is the width. For example, if an image has 3 attributes in its color space, and each attributes are quantized using 10, 3, and 3 bins, then $\boldsymbol{C(x, y)}$ has the value ranged between 0 to 89, as 10×3×3=90, which then represented by $\boldsymbol{W}$ minus 1.

Color Difference Histogram is composed by color and edge detection quantization which has been discussed before. They are $\boldsymbol{C(x, y)}$, which value is in the range $w \in 0,1, \ldots, W - 1$, and $\theta(x, y)$, which value is in the range $v \in 0,1, \ldots, V - 1$. Color Difference Histogram considers the pixel relationship in position (x, y) and its neighbor position (x', y') as far as D. Thus, the quantization of each color and edge of neighboring pixel are $C(x', y')$ and

$\theta(x', y')$. The computation of Color Difference Histogram based on color quantization $H_{color}(C(x,y))$ is shown in (16), while the one which based on edge quantization $H_{ori}(\theta(x,y))$ is shown in $H_{ori}(\theta(x,y))$(17)(17). The results of both operation is combined using (18). Hence, the feature vector produced is a $W+V$ dimension vector, which is taken account in the impression classification.

$$H_{color}(C(x,y)) = \sum\sum\sqrt{(\Delta L)^2 + (\Delta a)^2 + (\Delta b)^2} \tag{16}$$

which $\theta(x,y) = \theta(x',y')$ and $\max(|x-x'|, |y-y'|) = D$

$$H_{ori}(\theta(x,y)) = \sum\sum\sqrt{(\Delta L)^2 + (\Delta a)^2 + (\Delta b)^2} \tag{17}$$

which $C(x,y) = C(x',y')$ and $\max(|x-x'|, |y-y'|) = D$

$$H = [H_{color}(0),.., H_{color}(W-1), H_{ori}(0),.., H_{ori}(V-1)] \tag{18}$$

Ensemble of Classifier Chain

Ensemble of Classifier Chain (ECC) combines some models of Classifier Chain to predict multi-label in a dataset. ECC is able to predict multi-label in a dataset.ECC trains m Classifier Chains $\mathrm{h}_1, \dots, \mathrm{h}_\mathrm{m}$. Each classifier has a random order and trained by a random ordering of datasets with N data. The binary classification results in classifier chain will be combined into a multi-label [8].

Multi-label K-Nearest Neighbour

Multi-label K-nearest Neighbour (MLKNN) predicts the label membership of an instance using statistical information such as membership counting [9]. This method is the elaboration of the popular k-Nearest Neighbour algorithm. It consists of two main methods. For each test instance, k neighbours from the train dataset will be identified. Then, maximum a posteriori probabilistic will be identified for a test instance using statistical information.

Ensemble of Multi-label Classifier

Combination of q multi-label classifiers set which is defined by $H = \{H_1, H_2, \dots, H_q\}$ is performed using MEAN, MAX, or MIN. Those combiners are the most simple and popular to combine continuous probabilistic output score that comes from each classifier [10].

In order to improve the performance, several studies use threshold selection for multi-label classifier [11] [12]. A threshold t will be chosen using (19) by the X_{train} training set and X_{test} test set. The threshold is used to give a final prediction of multi-label. $LCard$ (Label Cardinality) is a standard measurement of multi-labeledness [13]. It is calculated by the average number of relevant labels of each instace. $LCard$ of a dataset X is defined by $LCard(X) = \frac{\sum_{i=1}^{|X|}|E_i|}{|X|}$ whee E_i is the actual labels from the training set and set of predicted labels by the thresholding of t towards the testing set.

$$t = \underset{\{t\in 0{,}00,\ 0{,}01,\ \dots,\ 1{,}00\}}{\arg\min} \left| LCard(X_{train}) - LCard(H_t(X_{test})) \right| \tag{19}$$

RESULT AND DISCUSSION

Dataset

The dataset used in is study is the dataset of batik cloth image from the study of Harfiani [3]. The dataset consists of 102 batik cloth image that has repetitive squares, kawung, parang, lereng, and buketan motifs. Each motifs has its own philosophical meaning and impression as defined in Table 1.

Table 1. Batik motifs and its philosophy and impression [14]

Motif	Philosophy	Impression
Repetitive square	Varies, depends on its ornament, and shows wisdom and prosperity	mature, calm
Kawung	Shows hope, wisdom and guidance	warm, calm, mature
Parang	Shows changes, dynamics, and advantages.	dynamic, masculin
Lereng	Shows changes, dynamics, and advantages.	dynamic, masculine
Buketan	Express the beauty	feminine

This multi-label dataset has six labels, they are mature, calm, dynamic, masculine, and feminine. The examples of the batik cloth images which used in this study are shown in Figure 2.

Evaluation Measurements

In this study, 5 types of measurements are used. They are Hamming Loss, One Error, Coverage, Ranking Loss, and Average Precision [9][11].

Hamming Loss measurements, which is defined by **Error! Reference source not found.**, measures how many the misclassified label-instance. The performance of the multi-label classifier is perfect if the Hamming Loss value is 0, the performance is considered better when the value is smaller. Hamming Loss refers to binary result of the classification.

$$hloss(h) = \frac{1}{X}\sum_{i=1}^{|X|}\frac{1}{|L|}\left|\{h(x_i)_j \neq y_j, j = 1,...,|L| \,|\, y \in L\}\right| \quad (20)$$

The One-error evaluation, which is defined by (20), measures the number of labels in top ranks which do not exist in the actual labels of each instance.

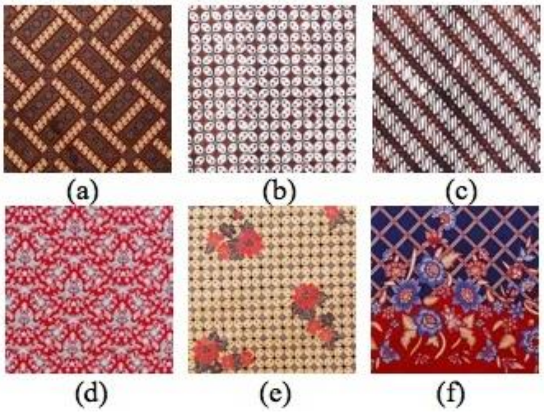


Figure 2. Batik images within its impressions (a) mature, calm (b) warm, calm, mature (c) dynamic, masculine (d) feminine (e) warm, calm, mature, feminine (f) deminine, mature, calm

The performance of the multi-label classifier is considered perfect if the One-error value is 0, the performance is considered better when the value is smaller. One-error refers to real results of the classification. Note that $l_{top}(x_i) \notin Y_i$ gives the top-most rank of i-th instance incorrect label. Using π notationπ, $[\![\pi]\!]$ gives value of 1 is it has member and 0 if it is empty.

$$\text{One} - \text{error}(h) = \frac{1}{|X|}\sum_{i=1}^{|X|}[\![l_{top}(x_i) \notin Y_i]\!] \quad (20)$$

The Coverage evaluation, which is defined by (21), measures the average of the searching effort of all label rank list which fulfills the correct labels of an instance. Coverage has weak relation with precision if it is compared by the perfect recall. The performance of the multi-label classifier is considered better when the value of Coverage is smaller.

$$Coverage(h) = \frac{1}{|X|}\sum_{i=1}^{|X|}\left(\max_{y \in Y_i} rank(x_i, y) - 1\right) \quad (21)$$

$$RankingLoss(h) = \frac{1}{|X|}\sum_{i=1}^{|X|}\frac{1}{|Y_i||\bar{Y_i}|}\left|\{(y_1, y_2) \,|\, rank(x_i, y_1) > rank(x_i, y_2), (y_1, y_2) \in Y_i \times \bar{Y_i}\}\right| \quad (22)$$

The Ranking Loss evaluation, which is defined by (22), measures the average fraction of label pairs that ranked reversely of an instance. The performance of the multi-label classifier is considered perfect if the Ranking Loss value is 0, the performance is considered better when the value is smaller.

The Average Precision metrics, *AP(h)*, evaluates the average fraction of ranked label which is placed above correct labels $y \in Y$ (which means y should be the correct labels). The performance of the multi-label classifier is considered perfect if the value of Average Precision is 1, the performance is considered better when the value is bigger. The formula of Average Precision is shown by (23), which written as.

$$AP(h) = \frac{1}{|X|}\sum_{i=1}^{|X|}\frac{1}{|Y_i|}\times\sum_{y \in Y_i}\frac{\left|\{y' \,|\, rank(x_i, y') \leq rank(x_i, y), y' \in Y_i\}\right|}{rank(x_i, y)} \quad (23)$$

Experiment Setting

Experiments is performed in certain computational environment and tools. The environment of the experiment is a 64-bit system with 4 GB RAM and Intel® Core i5 2.3 GHz processor. The tool used in this experiment is MATLAB R2008a (7.6.0.324).

In this experiment, ECC uses some parameters of base classifier. The base classifier of the ECC is the Sequential Minimal Optimization (SMO) [15]. This study uses SMO which has default parameters of WEKA. Those parameters are adapted to MATLAB using the polynomial kernel function and 1 order.

The MLKNN needs some parameters to measure the distance of an instance to its neighbor. Based on the study of Liu dan Yang [4], the suitable distance measurement of the Color Difference Histogram feature extraction is the modified Canberra distance which is shown by (24) and (25).

$$d(x_i, x_j) = \sum_{k=1}^{m} \frac{\left|x_{i_k} - x_{j_k}\right|}{\left|x_{i_k} - u_i\right| + \left|x_{j_k} - u_j\right|}, i \neq j \quad (24)$$

$$u_i = \frac{\sum_{k=1}^{m} x_{ik}}{m}, u_j = \frac{\sum_{k=1}^{m} x_{jk}}{m} \quad (25)$$

Based on some internal experiments, this experiment applies other parameters in the multi-label classification. Those parameters are the number of neighbours (k), which is 4, and the number of Classifier Chain within the ECC, which is 45. In addition, a configuration of ensemble composition is also chosen by the internal experiment. The ensemble configuration used is pair of 1 MLKNN model-4 ECC models, which combined by MEAN combiner.

The proportion of the training and testing set is 65:35. The experiment is performed 100 times randomly considering the proportion of impression labels, each training set should have at least one true label of each impression label. This action is performed to avoid the error caused in SMO training, which prohibits all training data to have the same class.

Experimental Results

The experiment compares the performance of the Color Difference Histogram (CDH) with other feature extractions taht is used the color and texture features. The other considered feature extractions are Microstructure Descriptor (MSD) by Harfiani [3], Multi-texton Histogram (MTH) by Liu dan Zhang [16], and the combination of Color Co-occurrence Matrix (CCM), Difference between Pixels of Scan Pattern (DBPSP), and Color Histogram for K-Means (CHKM) by Pratomo [17]. Performance evaluation used in this experiment is based on the classification result from the ensemble of multi-label classification using threshold selection, which has explained in previous section. The ensemble of multi-label classification method using threshold selection is applied to Color Difference Histogram feature extraction and the other tested feature extractions.

Table 2. Experiments evaluation of multi-label classification for various feature extractions

Feature Extraction	HL	OE	Cov	RL	AP
MTH	0.252	0.375	1.720	0.406	0.761
MSD	0.289	0.402	1.783	0.355	0.744
CCM, DBPSP, CHKM	0.190	0.263	1.270	0.350	0.855
CDH	0.173	0.228	1.260	0.361	0.866

The performance of the random 100 experiments, which is represented by Hamming Loss (HL), One Error (OE), Coverage (Cov), Ranking Loss (RL) and Average Precision (AP), is shown in Table . The best result of each evaluation metric is bold formatted.

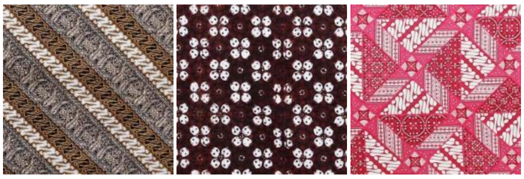
Then, an experiment to compare the impression classification result by the combination of CCM, DBPSP, and CHKM, then, MSD, and after that, CDH feature extraction method is also performed. The comparison of sampled images' impression classification result is shown in Error! Reference source not found., while the sampled images are depicted in Figure .

Discussion

Table shows the Color Difference Histogram feature extraction gives the best performance in multi-label classification. The better Ranking Loss of combination of CCM, DBPSP, and CHKM feature extraction shows that the Color Difference Histogram is not powerful enough to calculate the number of errors from the two label events, which should be correctly and incorrectly classified.

The other characteristics which can be seen by the small difference performance of Color Difference Histogram feature extraction and one that proposed by Pratomo [17] (combination of CCM, DBPSP and CHKM). It showed the nearly draw performance but in really different aspects. Those aspects are the texture features which represented by scan pattern in CCM and DBPSP, the RGB color space, the color quantization which represented by CHKM, and the number of features extracted, which is 71 features by combination of CCM, DBPSP, and CHKM, while Color Difference Histogram extracts 108 features.

Meanwhile, the MTH [16] and MSD [3] feature extractions also has slightly difference performance compare to that of Color Difference Histogram. Those three feature extractions used the same edge detection based



(A) (B) (C)

Figure 3. The sample images to be classified

on chromatic changes as defined in (6) until (12). However, MTH used the “texton” concept to build its texture features which also developed from the same edge detection. MTH and MSD also implemented same color quantization but uses different color space. MSD used HSV color space while MTH used RGB color space.

Table 3. Sample Image Impression Classification

Image	Feature Extraction	Impression
A	Actual Impression	warm, calm, mature
	MSD	mature, warm, calm, feminine
	CCM, DBPSP, CHKM	warm, calm, mature
	CDH	mature, warm, calm, feminine
B	Actual Impression	dynamic, masculine
	MSD	dynamic, masculine
	CCM, DBPSP, CHKM	feminine
	CDH	dynamic, masculine
C	Actual Impression	mature, calm
	MSD	mature, calm, feminine, warm
	CCM, DBPSP, CHKM	feminine
	CDH	mature, calm

CONCLUSION

In this study, the ensemble multi-label classification (Ensemble of Classifier Chain and Multi-label K-Nearest Neighbour) and Color Difference Histogram has been developed for the system of batik impression determination. The experiment shows that the system achieve better performance compare to other system that use other feature extraction method.

For further development, the Color Difference Histogram can be extended by considering the other neighbourhood relation in order to obtain the color difference. Another improvement could be performed by using more images and richer motifs.

REFERENCES

[1] S. Doellah, Batik: Pengaruh Jaman dan Lingkungan, Solo: DanarHadi, 2002.

[2] F. Kerlogue, The Book of Batik, Singapore: Archipelago Press, 2004.

[3] A. Harfiani, N. Suciati, and I. Arieshanti, Implementasi Metode Image-to-Class Distance untuk Klasifikasi Impresi pada Citra Batik, Surabaya: Teknik Informatika ITS, 2014.

[4] G.-H. Liu and L. Zhang, "Image Retrieval Based on Multi-Texton Histogram," *Pattern Recognition*, vol. 43, no. 7, pp. 2380-2389, 2010.

[5] N. Chawla and J. Sylvester, "Exploiting Diversity in Ensembles: Improving the Performance on Unbalanced Datasets," *in Proceedings of Multiple Classifier Systems*, Heidelberg, Springer, 2007, pp. 397-406.

[6] W. Burger and M. J. Burge, Principles of Digital Image Processing, London: Springer-Verlag, 2009.

[7] R. C. Gonzalez and R. E. Woods, Digital Image Processing, 3rd ed., Prentice Hall, 2008.

[8] J. Read, B. Pfahringer, G. Holmes and E. Frank, “Classifier chains for multi-label classification,” Heidelberg: Springer, vol. 5782, pp. 254-269, 2009.

[9] M. L. Zhang and Z. Zhou, "ML-KNN: A lazy learning approach to multi-label learning," *Pattern Recognition*, vol. 40, no. 7, pp. 2038-2048, 2007.

[10] L. Kuncheva, Combining Pattern Classifiers, Chichester: Wiley, 2004.

[11] M. A. Tahir, J. Kittler, K. Mikolajczyk and F. Yan, "Improving Multilabel Classification Performance by Using Ensemble of Multi-label Classifiers," *in Proceedings of Multiple Classifier Systems*, Berlin, Heidelberg: Springer, 2010, pp. 11-21.

[12] Fan and C. Lin, "A study on threshold selection for multi-label classification," National Taiwan University, Taiwan, 2007.

[13] G. Tsoumakas, I. Katakis and I. Vlahavas, “Mining Multi-label Data,” In: Data Mining and Knowledge Discovery Handbook, vol. 2, Heidelberg: Springer, 2009.

[14] A. Nilogiri, Klasifikasi Kansei Multi Label dengan Probabilistic Neural Network pada CItra Batik Meggunakan Kombinasi Fitur Warna, Tekstur, dan Bentuk, Surabaya: ITS, 2012.

[15] J. C. Platt, "Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machine," *Microsoft Research*, 1998.

[16] G.-H. Liu and J.-Y. Yang, "Content-based image retrieval using color difference histogram," *Pattern Recognition*, vol. 46, no. 1, pp. 188-198, 2013.

[17] W. A. Pratomo, N. Suciati and D. Purwitasari, Klasifikasi Kansei Citra Batik Menggunakan Backpropagation Neural Network dan Kombinasi Warna dan Tekstur, Surabaya: Jurusan Teknik Informatika ITS, 2014.