https://doi.org/10.31449/inf.v45i7.3715 Informatica **45** (2021) 97–105 **97**

# Effective Image Representation using Double Colour Histogram for Content-Based Image Retrieval

Ezekiel Mensah Martey, Hang Lei and Xiaoyu Li
School of Information & Software Engineering,
University of Electronic Science and Technology of China, China
E-mail: martey003@std.uestc.edu.cn, hlei@uestc.edu.cn, xiaoyu33521@163.com

Obed Appiah
Department of Computer Science and Informatics,
University of Energy and Natural Resource, Ghana
Email: obed.appiah@uenr.edu.gh

**Keywords:** stacked colour histogram, conventional colour histogram, texton, image representation, content-based image retrieval

**Received:** September 1, 2021

*Image representation is critical to the successful realisation of Content-Based Image Retrieval (CBIR) systems. The choice of features to represent the image affects retrieval performance. Nowadays, image databases are heterogeneous and different feature types can be used for appropriate descriptions. This paper proposes an image representation for CBIR that combines Stacked Colour Histogram (SCH) and Conventional Colour Histogram (CCH) to improve image retrieval precision. This presented technique is designed to capture the colour and texture information of the image. The colour properties of an image are represented by CCH and that of texture by SCH. The weighted similarity measure is used to estimate the proportion of similarity values in the retrieval task. The novel descriptor has been widely tested on four standard image datasets, namely Batik, Coil100, Corel10K and Outext. Batik, Coil100 and Outext are used to assess texture discrimination. Corel10K is used to assess the discrimination of heterogeneous images. Experimental results and comparisons with SCH, CMTH, MTH, TCM, CTM and NRFUCTM demonstrate that the proposed descriptor has superior retrieval performance.*

*Povzetek: V članku je podana učinkovita predstavitev slik z barvnim diagramom z namenom iskanja vsebine na slikah.*

## 1 Introduction

Image representation is crucial in achieving high precision retrieval results. The image is represented based on low-level features that are colour, texture and shape. Among them, colour is the most commonly used to retrieve images. Colour feature has extensive applications in numerous fields such as internet image search, image analysis, remote sensing, medical image detection, and video surveillance. Conventional Colour Histogram (CCH)[1], Colour Coherent Vector (CCV) [2], Colour Moment (CM)[3] and Colour Correlogram (CC)[4] are standard colour feature extraction techniques that are used for indexing images in CBIR applications. CCH is much more flexible to use among the colour feature extraction techniques because of its simplicity and low computational time complexity. However, the CCH lacks information about the spatial layout of colour distributions in an image, making it challenging to represent texture or shape characters of the image [5]. Notwithstanding, CCH and other histogram image representation techniques offer a great advantage of effectively handling scaling, rotation, reflection and translation transformation that an image may go through [6]. This advantage has resulted in the use of CCH and its variants still active in the domain of CBIR.

To extract colours and their neighbourhood information, image texture feature extractions are largely employed. Most CBIR techniques will use this approach because it tends to deal with the weaknesses of the CCH and hence performs better mostly when the application is not so much interested in the distribution of pixels but the location of pixels, especially in the medical image retrieval applications [7]. Texture defines a spatial layout of pixels in an image area, and many researchers have studied it [8]. Edge Histogram Descriptor (EHD)[9], Local Binary Patterns-based descriptors (LBP) [10,11], Gray-Level Co-occurrence Matrix (GLCM) [12], Gabor filter [13,14], Curvelet transform[15] and Hidden Markov random field [16] are some texture descriptors used for CBIR. Texture descriptors do not always yield optimal results because they turn to be invariant to image transformations.

The popular term texton proposed in [17] has played a significant role in texture analysis. The purpose of textons is to use basic primitives to describe complex structures. Textons have been applied in a variety of ways, such as seen in Multi-Texton Histogram (MTH) [18], Texton Co-occurrence Matrix (TCM) [19], Complete Texton Matrix (CTM) [20] and Complete Multi Texton Histogram [21] with generally good results. However, the rigid nature of the textons makes it undesirable at some point [22]. This is generally because when the orientation of an image changes, the texton feature extractions also turn to produce different feature maps or vectors for the same image. The SCH approach proposed in [23] employed recurrent blurring to extract inherent textural information, thus solving the rigid nature of the textons.

A single feature extracted from colour or texture cannot be sufficient enough to provide the best retrieval results. Therefore, hybrid feature extraction techniques have become an adequate alternative in dealing with such problems associated with individual feature extraction issues. As the images dataset increases in size and variety, a single feature extraction technique for indexing images for images search is not enough. One approach to increase the CBIR performance is to utilise fusion methods. Liu and Yang [19] fused colour feature information and pixel orientation into an image representation vector while [24–28] integrated the colour and texture information in a single descriptor using textons.

In this paper, we advanced the scheme Stacked Colour Scheme (SCH) [23] by integrating it with Conventional Colour Histogram (CCH) as these features are robust and compact for heterogeneous image representation and invariant to transformations (rotation and translation) [29, 30].

To this end, we propose a novel feature representation technique that joins two features into a single vector for indexing images. Therefore, our proposed technique provides; 1) a double colour histogram for an image representation vector that effectively deals with colour and heterogeneity of texture images. 2) compact information about colour and texture features with high precision. To test the effectiveness of the proposed technique, we compared with well-known image descriptors for CBIR.

# 2 Proposed methodology

The proposed technique is based on combining two colour histograms, CCH and SCH. Detailed information on the integration of CCH and SCH features are discussed in the subsections.

## 2.1 Conventional Colour Histogram

CCH [1] is the most utilised method to represent the colour distribution of an image. The colour space is required to compute the CCH. The CCH is constructed in RGB colour spaces. “The colour bins are represented in histograms, and the total number of bins in RGB histograms can be determined based on RGB planes” [31- 33].

**Algorithm 1: Conventional Colour Histogram Algorithm**

1. *I ← Read an image*
2. *q(I) //quantized colour*
3. *cH(I) ← Extract histogram of the image I*
4. *cH(I) //Normalize colour histogram*
5. *Store cH as an Index for the read image.*

## 2.2 Stacked Colour Histogram (SCH)

SCH is the bin-by-bin summation of several histograms generated from the recurrent transformation of an image [23]. Mean filter is used to transform images iteratively. The filtering, which employs a window size of $n \times n$, is performed N times [34]. In each instance of the filtering process, the output of the previous filtering is used as input to that of the current filtering processes. Histograms are generated from filtered images and later added together to estimate the SCH. The filtering process uses neighbourhood information and therefore estimated histogram contains information about the nature of pixels' spatial distribution in images.

**Algorithm 2: Stacked Colour Histogram Algorithm**

1. *I ← Read an image*
2. *sH (J) // Histogram with J bins. Initialize with Zeros.*
3. *count ← N // N is the number of times the filtering of image is to be done.*
4. *while (count > 1)*
5. *H(J) ← Extract histogram of image I*
6. *sH(binIndex) = sH(binIndex) + H(binIndex)*
7. *while (binIndex > 1)*
8. *sH(binIndex) = sH(binIndex) + H(binIndex)*
9. *binIndex = binIndex − 1*
10. *I ← Mean Filter (I) and assign to I. //Perform mean filtering of the image in variable I*
11. *count = count – 1*
12. *Store sH as Index for the read image.*

## 2.3 Proposed joint histogram based on SCH and CH

The diagrammatic representation of the proposed approach based on CCH and SCH features is shown in Figure 1. Concise descriptions of the key steps are as follows:

**Image indexing**

1. Read the image (I)
2. Quantised image into 64 colours
3. Compute CCH of (I)
4. Compute SCH of (I)
5. Combine CCH and SCH into a single vector and store as an image's Index. The index vector should have two parts, LowerIndex which contains the CCH, and UpperIndex, which stores SCH

**Image retrieval**

1. Read the image (I) Query Image
2. Quantised image into 64 colours
3. Compute CCH of (I)
4. Compute SCH of (I)
5. Combine CCH and SCH into a single vector

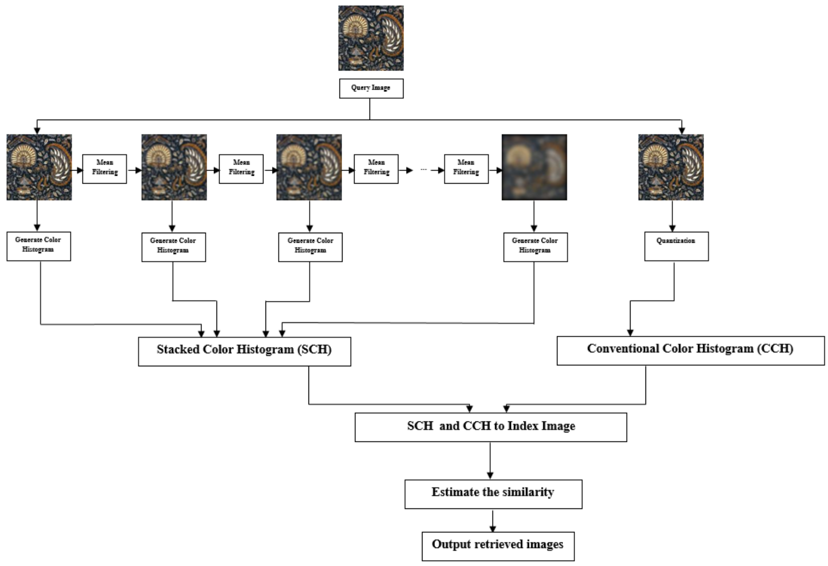


Figure 1: Block diagram of the proposed technique.

6. Estimate the similarity of LowerIndex of Query image with the LowerIndex of Target image using the Euclidean distance and store value as LI
7. Estimate the similarity of UpperIndex of Query image with the UpperIndex of Target image using the Euclidean distance and store value as U.I.

Using weighted values $(\alpha, \beta)$, estimate the final similarity value by equation 1

$$\text{Similarity} = \alpha * LI + \beta * \text{UI} \quad (1)$$

The final similarity is used to retrieve images that are closer in the distance to the query image. $\alpha$ *and* $\beta$ *are selected using equation* 2 :

$$\alpha + \beta = 1 \quad (2)$$

such that $0 \leq \alpha \leq 1$

$$0 \leq \beta \leq 1$$

The Euclidean distance is presented in equation 3:

$$d_E = \sqrt{\sum_{i=1}^{n} (p_i - q_i)^2} \quad (3)$$

where:

$d_E$ = distance metric
$P_i$ = target image
$q_i$ = query image

# 3 Experimental evaluation

For evaluation of our experiment, we used four popular datasets used in [23]—Batik, Corel10K, Outext and Coil100. The Batik, Coil100 and Outext datasets are used to assess the proposed descriptor in differentiating texture images. The Corel10K dataset is used to assess the heterogeneous image recognition capabilities of the proposed technique.

Below is a concise description of datasets.

- ❖ Batik [35] dataset contains 300 images in total. The image database is categorized into 50 classes, and each class contains six images. The images are of 128x128 pixels sizes in JPEG format.
- ❖ Coil100 [36] dataset contains 7200 colour images categorised into 100 classes. There are 72 images of each object in different poses in each class.
- ❖ Outext [37] dataset contains 11484 total images. The image database is categorised into 29 classes.
- ❖ Corel10K [38] comprises 10000 images in total. There are 100 classes, and each class contains 100 images. The images are sized 128 × 192 pixels in JPEG format.

Samples images from each dataset are illustrated in Figure 2.

The initial experiment evaluated window sizes and recurrent transformations on accuracy using different classifiers and different dataset features. The proposed method implemented here used a window size of 7x7 and 11x11 with recurrent transformations of 5, 10, 15, 20 and 25 to extract features for this experiment.

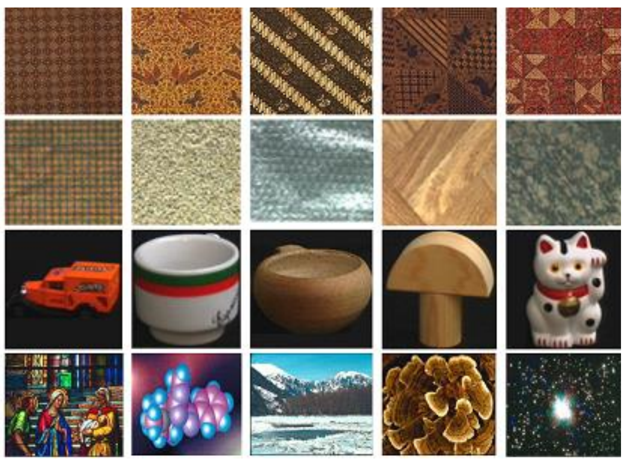

Figure 2: Samples images from each dataset (a) Batik (b) Outext (c) Coil100 and (d) Corel10K.

**Machine learning approach for evaluation**
K-fold cross-validation was performed to avert the over-fitting of the data during the process of training. The different classifiers with K-fold cross-validation (k = 5) were used to evaluate the performance on multiple and different features of datasets. Using Principal Component Analysis (PCA), we reduced the feature sizes to speed up computation, setting the threshold to 95%.

Five different classifiers, namely Discriminant Analysis Classifier (DAC), Support Vector Machine (SVM), Naive Bayes (NB), Decision Tree (DT) and KNN (k-Nearest Neighbours), were used to classify images from each dataset. Classification accuracy metric is used to measure the performance of classifiers. Equation (4) indicates mathematical expression for classification accuracy.

$$Accuracy = \frac{correctly\ classified\ images}{total\ images} \quad (4)$$

## 3.1 Image retrieval

The performance of the proposed descriptor is evaluated in the classic CBIR task. Precision and Recall metrics are used to evaluate the performance of the proposed descriptor in the CBIR task. The precision determines the number of correctly retrieved images. Equations (5) and (6) present the Precision and Recall metrics formulae, respectively.

$$Precision = \frac{number\ of\ retrieved\ relevant\ images}{Number\ of\ retrieved\ images} \quad (5)$$

$$Recall = \frac{number\ of\ retrieved\ relevant\ images}{Number\ of\ relevant\ images\ in\ the\ dataset} \quad 6)$$

**We first present the results of the machine learning approach used for the experiment.**

Table 1 presents the mean accuracy values of the KNN classifier with different values of k on three datasets. The mean accuracies when k = 1, 98%, 38.28%, 79.48% were recorded for Coil100, Corel10K, and Outext datasets, respectively. When k = 3, mean accuracy percentages of 96.24%, 37.26%, 78.08% were recorded for Coil100, Corel10K, and Outext datasets, respectively. The value k = 5 produced accuracy values of 95.35%, 39.82%, 76.64% for Coil100, Corel10K, and Outext datasets, respectively. Finally, using k = 9 recorded 95.15%, 40.34%, 77.94% for Coil100, Corel10K, and Outext datasets, respectively. The mean accuracies give a good indication of the classification efficiency of the proposed scheme.

The results obtained for the KNN classifier were compared with features in [29]. The accuracy percentages of our proposed descriptor show improved performance in representing heterogeneous texture images. Also, our method performs very well for texture images. Table 2 compares the performance of our method with other approaches for different values of k (KNN's parameter)

To experimentally validate if our proposed method keep up with other classifiers as implemented in our earlier work [29], we reconduct evaluation using Decision Tree (DT), Discriminant Analysis Classifier (DAC), Support Vector Machine (SVM), and Naive Bayes (NB). The best accuracy values achieved by the DT classifier on all datasets are shown in table 3. The Coil100, Corel10K, Outext and batik datasets recorded best accuracy values of 87.5%, 20.5 %, 66.3% and 51.6% respectively. Mean accuracies of 84.78%, 23.3%, 65.38% and 39.18% were

| k-Value | Coil100 | Corel10K | Outext |
|---|---|---|---|
| k=1 | 0.98 | 0.3828 | 0.7948 |
| k=3 | 0.9624 | 0.3726 | 0.7808 |
| k=5 | 0.9535 | 0.3982 | 0.7664 |
| k=9 | 0.9515 | 0.4034 | 0.7794 |

Table 1: accuracy values for KNN classifier with different values of k on three datasets.

| | | **Outext** | **Corel10K** |
|---|---|---|---|
| k = 1 | TCM | 72.49% ± 2.8% | 16.71% ± 0.1% |
| | MTH | 51.92% ± 2.3% | 23.68% ± 0.1% |
| | CTM | 72.31% ± 2.8% | 21.22% ± 0.2% |
| | NRFU | 36.72% ± 1.7% | 5.38% ± 0.0% |
| | CMTH | 78.40% ± 3.0% | 31.76%± 0.1% |
| | SCH | 77.63%± 3.0% | 31.80%± 0.2% |
| | Proposed Technique | 79.48%± 2.0% | 38.28%± 0.1% |
| k = 3 | TCM | 76.42% ± 3.0% | 16.29 % ± 0.1% |
| | MTH | 52.30% ± 2.4% | 22.68 % ± 0.2% |
| | CTM | 70.18% ± 2.8% | 20.02 % ± 0.2% |
| | NRFU | 36.89% ± 1.8% | 5.26 % ± 0.0% |
| | CMTH | 76.42% ± 3.0% | 30.89 %± 0.2% |
| | SCH | 77.33%± 3.0% | 31.35%± 0.2% |
| | Proposed Technique | 78.08%± 2.0% | 37.26%± 0.1% |
| k = 5 | TCM | 68.41 % ± 2.7% | 17.51 % ± 0.1% |
| | MTH | 52.68 % ± 2.5% | 25.16 % ± 0.2% |
| | CTM | 68.93 % ± 2.9% | 20.97 % ± 0.2% |
| | NRFU | 39.49 % ± 2.1% | 6.00 % ± 0.08% |
| | CMTH | 74.83 % ± 3.0% | 33.16 % ± 0.2% |
| | SCH | 76.04%± 3.0% | 34.03% ± 0.2% |
| | Proposed Technique | 78.08% ± 2.0% | 39.82% ± 0.1% |
| k = 9 | TCM | 72.49% ± 2.8% | 16.71% ± 0.1% |
| | MTH | 51.92% ± 2.3% | 23.68% ± 0.1% |
| | CTM | 72.31% ± 2.8% | 21.22% ± 0.2% |
| | NRFU | 36.72% ± 1.7% | 5.38% ± 0.0% |
| | CMTH | 78.40% ± 3.0% | 31.76% ± 0.1% |
| | SCH | 77.19%± 3.0% | 31.93% ± 0.2% |
| | Proposed Technique | 77.94% ± 2.0% | 40.34% ± 0.1% |

Table 2: Results of proposed technique compared with other approaches for different values of k (KNN’s parameter).

recorded for Coil100, Corel10K, and Outext datasets respectively.

Table 4 demonstrates the outcomes of the SVM classifier on the datasets. The best accuracy result of 97.4% is obtained for Coil100. The transformations of 20 produced the best classification result of 50.3 % for Corel10K. Accuracy percentages of 78.9% and 91.7% are produced as best for the Outext and batik datasets, respectively. Mean accuracy percentages of 96.4%, 49.57%, 77.58% and 86.98% were obtained for Coil100, Corel10K, Outext and batik datasets, respectively.

Table 5 illustrates the outcome of the NB classifier analysed on the four datasets. The classifier produced the highest accuracy of 93.7% for Coil100. 41 % is generated as the highest for Corel10K. 63.9% is produced as the best accuracy percentage for the Outext dataset. Mean accuracies of 92.7%, 40.20%, and 63.12% were recorded for Coil100, Corel10K and Outext datasets, respectively.

The results obtained by the DAC classifier are shown in table 6. The accuracy values of 93.2%, 44.2 %, and 70.2% for the respective Coil100, Corel10K, and Outext datasets demonstrate the classification efficiency of the proposed scheme. Mean accuracies of 92.2%, 43.91%, and 69.22% were recorded for Coil100, Corel10K, and Outext datasets respectively.

Table 7 presents the results compared for four different classifiers for well-established SCH and texton based feature extraction techniques. The texton-based features and SCH lost performances in representing heterogeneous images. Our proposed method outperforms the SCH and textons. (i.e., SCH yields 32.72% with Corel10K. However, our proposed technique outperforms with 43.91% for DAC).

| Window Size − Recurrent Number | Coil100 | Corel10K | Outext |
|---|---|---|---|
| 7x7 - 5 | 0.834 | 0.23 | 0.781 |
| 7x7 - 10 | 0.845 | 0.232 | 0.771 |
| 7x7 - 15 | 0.84 | 0.231 | 0.762 |
| 7x7 - 20 | 0.845 | 0.231 | 0.722 |
| 7x7 - 25 | 0.875 | 0.241 | 0.783 |

Table 3: accuracy results for DT classifier on four datasets.

| Window Size − Recurrent Number | Coil100 | Corel10K | Outext |
|---|---|---|---|
| 7x7 - 5 | 0.947 | 0.486 | 0.758 |
| 7x7 - 10 | 0.96 | 0.4975 | 0.758 |
| 7x7 - 15 | 0.969 | 0.497 | 0.788 |
| 7x7 - 20 | 0.974 | 0.503 | 0.789 |
| 7x7 - 25 | 0.972 | 0.495 | 0.786 |

Table 4: accuracy results for SVM classifier on all datasets.

| Window Size − Recurrent Number | Coil100 | Corel10K | Outext |
|---|---|---|---|
| 7x7 - 5 | 0.913 | 0.4 | 0.614 |
| 7x7 - 10 | 0.925 | 0.402 | 0.628 |
| 7x7 - 15 | 0.928 | 0.396 | 0.639 |
| 7x7 - 20 | 0.932 | 0.402 | 0.64 |
| 7x7 - 25 | 0.937 | 0.41 | 0.635 |

Table 5: accuracy results for NB classifier on all datasets.

**We present the results of the image retrieval performance**

The proposed technique is primarily proposed for indexing images for CBIR. We evaluated the proposed descriptor in a classic CBIR task. The experiment initially evaluated the effect of several window sizes and the number of recurrent transformations on Precision and Recall on the identified dataset. The 7x7 window size with

| Window Size − Recurrent Number | Coil100 | Corel10K | Outext |
|---|---|---|---|
| 7x7 - 5 | 0.911 | 0.441 | 0.669 |
| 7x7 - 10 | 0.911 | 0.436 | 0.692 |
| 7x7 - 15 | 0.932 | 0.438 | 0.699 |
| 7x7 - 20 | 0.929 | 0.438 | 0.702 |
| 7x7 - 25 | 0.928 | 0.442 | 0.699 |

Table 6: accuracy results for DAC classifier on all datasets.

| | | Outext | Corel10K |
|---|---|---|---|
| DAC | TCM | 41.96 % ± 2.5% | 16.71% ± 0.1% |
| | MTH | 33.85 % ± 2.4% | 23.68% ± 0.1% |
| | CTM | 33.65 % ± 1.8% | 21.22% ± 0.2% |
| | NRFU | 39.83 % ± 2.5% | 5.38% ± 0.0% |
| | CMTH | 42.34 % ± 2.6% | 31.76% ± 0.1% |
| | SCH | 64.24%± 2.5% | 32.72%± 0.1% |
| | Proposed Technique | 69.22%± 0.1% | 43.91%± 0.1% |
| SVM | TCM | 42.95 % ± 3.0% | 20.58 % ± 0.1% |
| | MTH | 45.31 % ± 2.6% | 33.02 % ± 0.1% |
| | CTM | 27.43 % ± 1.3% | 21.15 % ± 0.2% |
| | NRFU | 31.51 % ± 1.9% | 11.69 % ± 0.1% |
| | CMTH | 51.01 % ± 1.6% | 40.61 % ± 0.2% |
| | SCH | 77.08% ± 0.1% | 40.96% ± 0.1% |
| | Proposed Technique | 77.58% ± 0.1% | 49.57% ± 0.1% |
| NB | TCM | 37.24 % ± 1.7% | 23.62 % ± 0.2% |
| | MTH | 25.54 % ± 1.1% | 24.68 % ± 0.2% |
| | CTM | 34.97 % ± 1.5% | 23.06 % ± 0.2% |
| | NRFU | 41.01 % ± 1.5% | 22.71 % ± 0.2% |
| | CMTH | 38.71 % ± 1.5% | 34.22 % ± 0.2% |
| | SCH | 61.08% ± 0.1% | 34.28% ± 0.2% |
| | Proposed Technique | 63.12% ± 0.1% | 40.20% ± 0.1% |
| DT | TCM | 64.63 % ± 2.5 | 18.33 % ± 0.1% |
| | MTH | 47.70 % ± 2.3% | 17.23 % ± 0.1% |
| | CTM | 67.38 % ± 2.7% | 19.69 % ± 0.1% |
| | NRFU | 60.81 % ± 2.6% | 15.96 % ± 0.1% |
| | CMTH | 64.45 % ± 2.8% | 22.94 % ± 0.1% |
| | SCH | 76.32 % ± 0.1% | 22.96% ± 0.1% |
| | Proposed Technique | 76.38% ± 0.1% | 23.3% ± 0.1% |

Table 7: Results of the proposed descriptor with SCH and texton based features using ML classifiers.

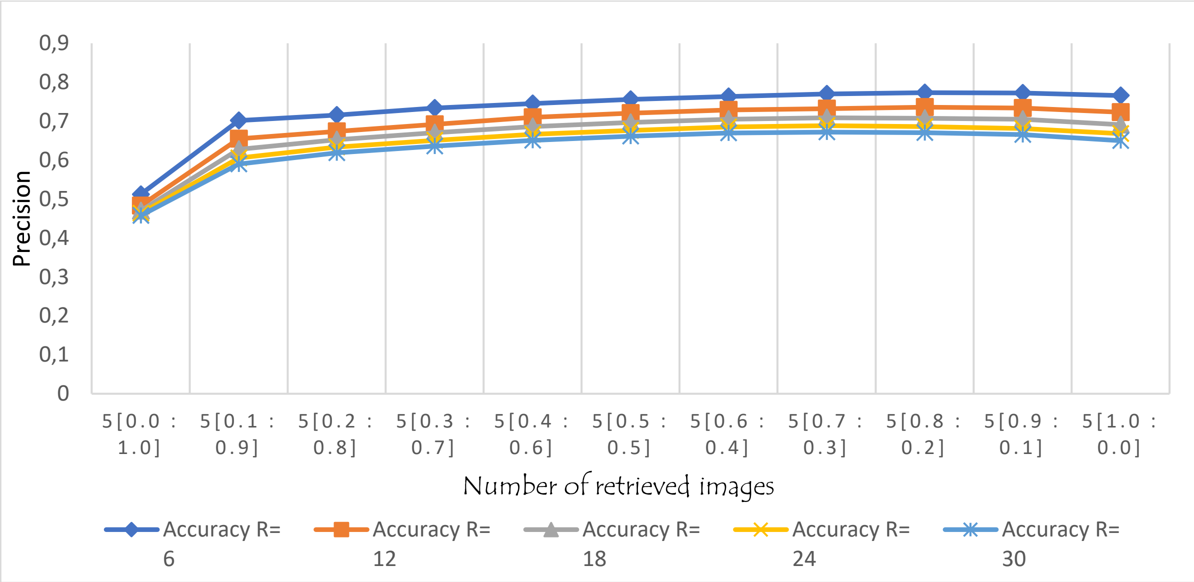


Figure 3: Precision for various retrieved values for the Outext dataset.

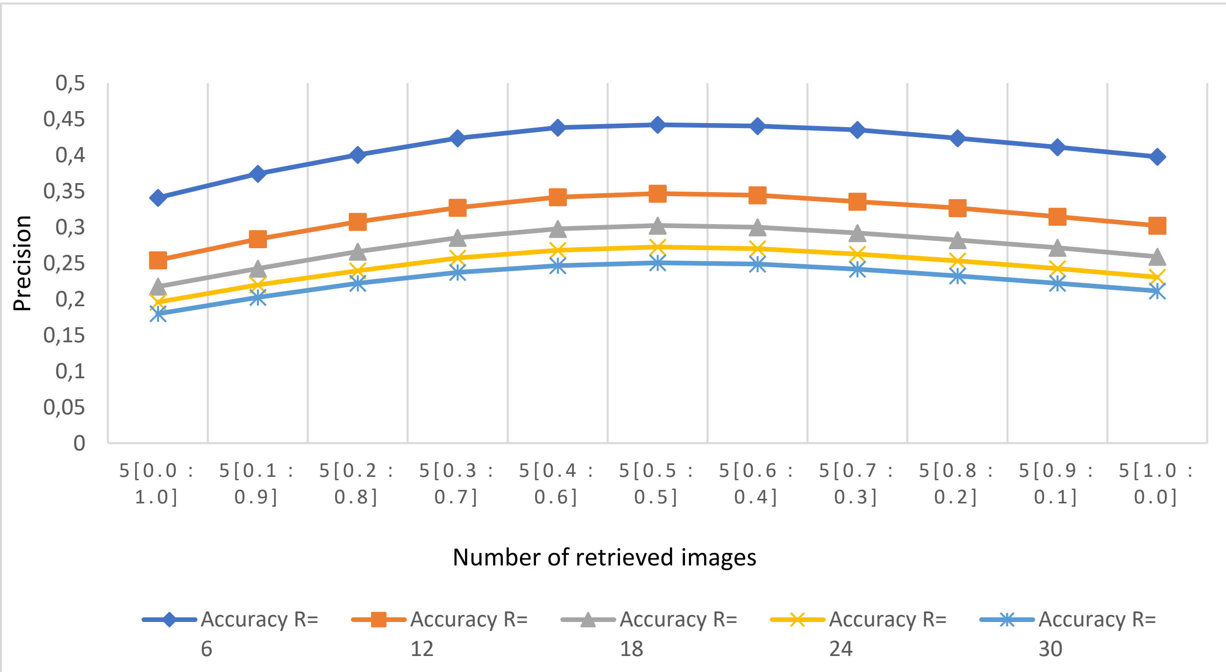


Figure 4: Precision for various retrieved values for Corel dataset.

the recurrent transformation of 5 provided good results hence adopted the same for extracting features for this experiment. Again, the number of retrieved images from each dataset was 6, 12, 18, 24 and 30.

To compute the weight values of CCH and SCH features, the α value starts with zero (0) and increases to one (1) with a step of 0.1. The notation [x: y] on the results chart has x as α weight while y represents the β weight used for the retrieval task.

Figures 3–6 present the Precision performances of the proposed technique on Outext, Corel-10K, and COIL100 and Batik datasets.

Figures 3, 4, 5 and 6 present the precision values of the five retrieved values (6, 12, 18, 24 and 30). Generally, precision values for the top five retrieved values describe a gentle distribution curve for all datasets. When the two weight values (α, β) are the same, the retrieval result shows that the visual effect is the best. That is when α is 0.5 and β is also 0.5, the proposed technique generally performs best.

Figure 7 presents the proposed technique results together with SCH and textons implemented by Martey et al. [23] on Corel-10K. Both experiments used retrieval values of 10 to 100 with a step of 10.

From Figure 7, the result of our proposed scheme outperforms TCM, MTH, CTM, NRFUCTM, CMTH and

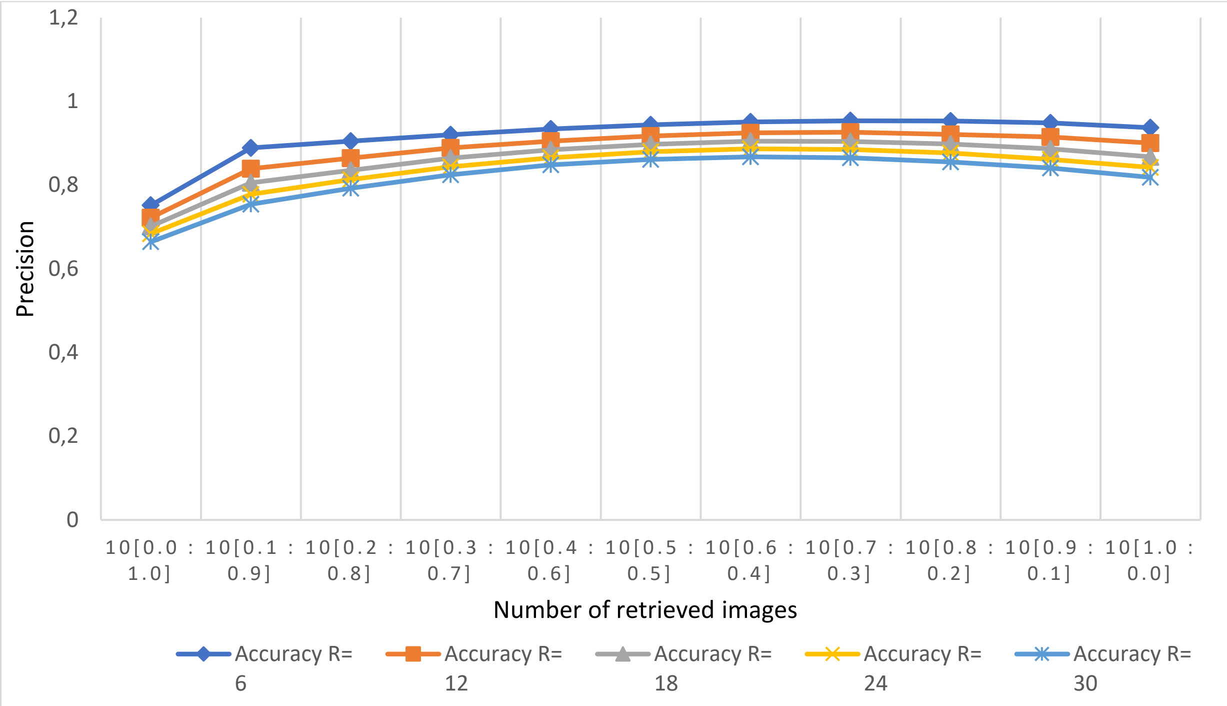


Figure 5: Precision for various retrieved values for Coil dataset.

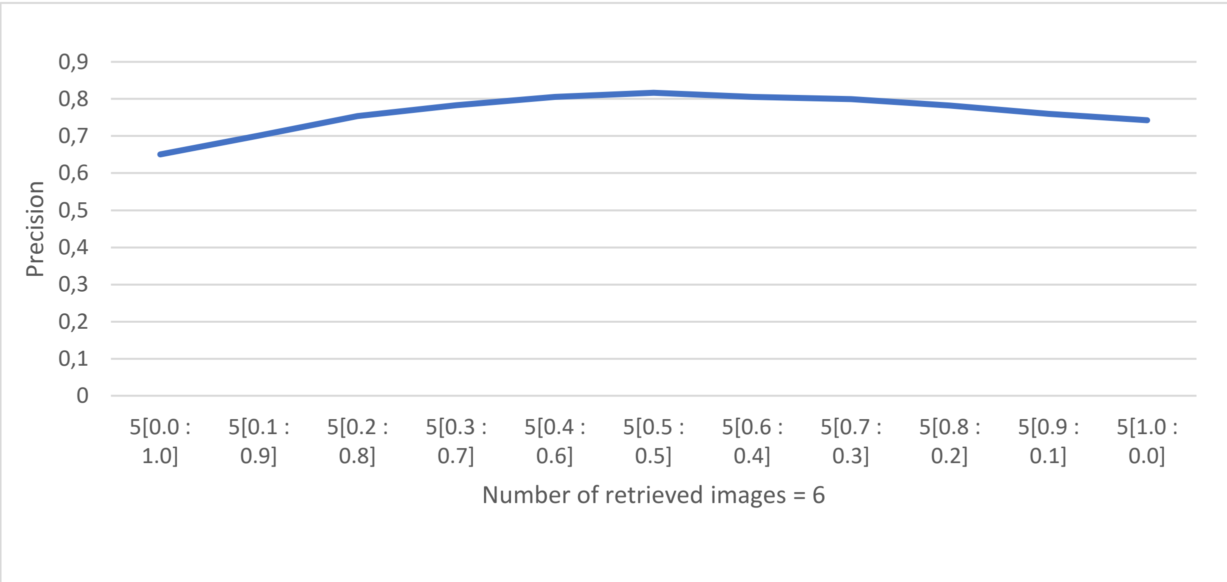


Figure 6: Precision for various retrieved values for Batik dataset.

SCH used to index the Corel 10K. The experimental result shows improved retrieval performance of the proposed technique compared with the SCH and texton approaches for extracting image texture features. SCH and Textons were chosen because literature has demonstrated that they are very effective for extracting image textural information [27, 35, 39] and hence can be used as a benchmark for evaluating similar feature extraction schemes. The result demonstrates the proposed technique leverage the transformational (rotational, scaling, translation and deformation) invariant nature of the SCH and CCH.

# 4 Conclusion

This paper presents a double colour histogram that effectively improves image feature representation for colour and texture images in a given database. The present descriptor combines SCH and CCH into a single vector for indexing images. The indexing scheme used a vector dimension of 128 and was very effective for image retrieval. The descriptor demonstrates high robust feature set used to represent the colour and texture of the image. The final experiments were carried out over four public database sets. Three image databases, mainly from Batik, Coil100 and Outext, were used to assess the colour and texture representation capabilities of our proposed

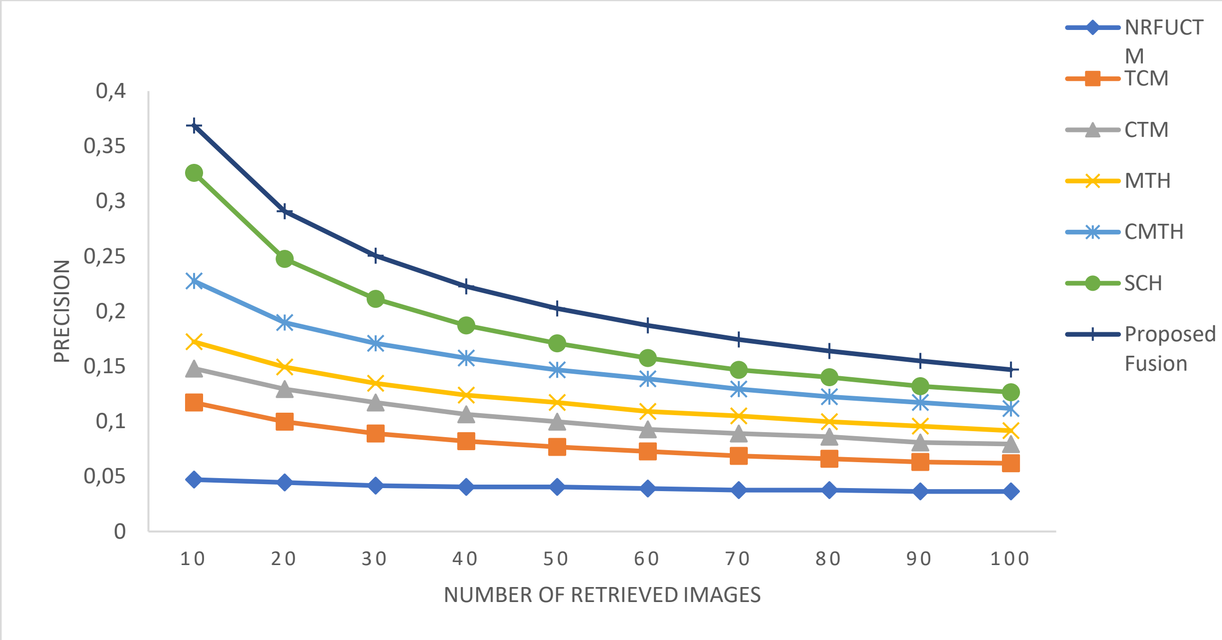


Figure 7: Performance of proposed technique compared with SCH and textons on Corel-10K.

descriptor. Corel10K was used to assess its ability to represent heterogeneous images. Experimental results have shown that our proposed method has strong texture and colour discrimination power and outperforms SCH and well-established texton-based features.

**Funding statement:** This study is supported by the Provincial Science and Technology Key Research & Development Sector (2019). “Establishment and practice of cancer patient assessment assistant intelligent decision system” (Grant No. 19ZDYF0794).

**Conflicts of interest:** The authors declare no conflict of interest concerning this work.

## References


[1] T. Deselaers, D. Keysers and H. Ney. Features for image retrieval: an experimental comparison. Inf. Retrieval 11(2):77–107, 2008.

[2] G. Pass and R. Zabith . Histogram refinement for content-based image retrieval, in Proceedings of IEEE Workshop on Applications of Computer Vision, 96–102, 1996.

[3] P. Srivastava, N. T. Binh, and A. Khare. Content-based image retrieval using moments. In International Conference on Context-Aware Systems and Applications, 228-237, 2013.

[4] J. Huang, S. Kumar, M. Mitra, W. J. Zhu and R. Zabih. Image indexing using colour correlograms. In Proceedings of IEEE computer society conference on Computer Vision and Pattern Recognition, 762-768, 1997.

[5] V. Tyagi . Content-based image retrieval: ideas, influences, and current trends. Springer Nature, 133-159, 2018.

[6] A. Rao, R. K. Srihari and Z. Zhang. Spatial colour histograms for content-based image retrieval. In Proceedings 11th International Conference on Tools with Artificial Intelligence, 183-186, 1999.

[7] C. Kurtz, A. Depeursinge, S. Napel, C. F. Beaulieu and D. L. Rubin. On combining image-based and ontological semantic dissimilarities for medical image retrieval applications. Medical image analysis, 18(7):1082-1100, 2014.

[8] M. Tuceryan and A. K. Jain. Texture analysis. Handbook of pattern recognition and computer vision, 235-276, 1993.

[9] D. K. Park, Y. S. Jeon and C. S. Won. Efficient use of local edge histogram descriptor. In Proceedings of the 2000 ACM workshops on Multimedia, 51-54, 2000.

[10] T. M. Ojala and T. M. Pietikainen. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns”. IEEE Trans. Pattern Anal. Mach. Intell, 24: 971–987, 2002.

[11] S. Marula, R.P. Maheshwari and R. Balasubramanian. Local tetra pattern: A new feature descriptor for content-based image retrieval. IEEE transactions on image processing, 21(5): 2874–2886, 2012.

[12] R. M. Haralick, K. Shanmugam and I. Dinstein, Textural Features for Image Classification, In IEEE Transactions on Systems, Man, and Cybernetics, 3(6):610-621, 1973. https://doi.org/10.1109/TSMC.1973.4309314

[13] D. Zhang, A. Wong, M. Indrawan and G. Lu. Content-based image retrieval using Gabor texture features. IEEE Transactions Pami, 3656: 13âĂŞ15. 2000.

[14] B. S. Manjunath and W. Y. Ma. Texture features for browsing and retrieval of image data. IEEE

Transactions on Pattern Analysis & Machine Intelligence, 18(8): 837–842, 1996.

[15] I. J. Sumana, M. M. Islam, D. Zhang and G. Lu. Content-based image retrieval using curvelet transform. In 2008 IEEE 10th workshop on multimedia signal processing, 11-16, 2008.

[16] F.S. Cohen, Z. G. Fan, and M. A. Patel. Classification of rotated and scaled textured images using Gaussian Markov random field models. IEEE Transactions on Pattern Analysis & Machine Intelligence, 13(02):192–202, 1991.

[17] B. Julesz. Texton gradients: The texton theory revisited. Biological Cybernetics, 54(4): 245-251, 1986.

[18] G. H. Liu, L. Zhang, Y. K. Hou, Z. Y. Li and J. Y. Yang. Image retrieval based on multi-texton histogram. Pattern Recognition, 43(7): 2380-2389, 2010.

[19] G. H. Liu and J. Y. Yang. Image retrieval based on the texton co-occurrence matrix. Pattern Recognition, 41(12): 3521-3527, 2008.

[20] Y. S. Kumari, V. V. Kumar and C. Satyanarayana. Texture classification using complete texton matrix. International Journal of Image, Graphics and Signal Processing, 9(10): 60, 2017.

[21] B. Khaldi, O. Aiadi and K. M. Lamine. Image representation using complete multi-texton histogram. Multimedia Tools and Applications, 79(11): 8267-8285, 2020.

[22] V. Tyagi. Content-Based Image Retrieval Using a Short Run Length Descriptor. In Content-Based Image Retrieval; Springer: Singapore, 241–256, 2017.

[23] E. M Martey., H. Lei, X. Li and O. Appiah. Image Representation Using Stacked Colour Histogram. Algorithms, 14(8): 228, 2021.

[24] G. H. Liu, Z. Y. Li, L. Zhang and Y. Xu. Image retrieval based on the micro-structure descriptor. Pattern Recognition. 44(9): 2123–2133, 2011.

[25] G. H. Liu, J. Y. Yang and Z. Li. Content-based image retrieval using computational visual attention model. Pattern Recognition, 48(8):2554–2566, 2015.

[26] G.H. Liu and J.Y. Yang. Image retrieval based on the texton co-occurrence matrix. Pattern Recogn. 41(12):3521–3527, 2008.

[27] G.H .Liu, and J.Y. Yang. Content-based image retrieval using colour difference histogram. Pattern Recognition, 46(1): 188–198, 2013.

[28] W. Xingyuan and W. Zongyu. A novel method for image retrieval based on structure elements descriptor. Journal of Visual Communication and Image Representation, 24(1): 63–74, 2013.

[29] X. Tian L. .Jiao, X. Liu and X. Zhang. Feature integration of eodh and colour-sift: Application to image retrieval based on codebook, Signal Processing: Image Communication, 29( 4): 530–545, 2014.

[30] S. R. Dubey, S. K Singh and R. K Singh. Rotation and scale-invariant hybrid image descriptor and retrieval, Computers & Electrical Engineering, 46: 288–302, 2015.

[31] A. Nazir, and N. Kashif. An efficient image retrieval based on fusion of low-level visual features, arXiv preprint, 2018. https://arxiv.org/abs/1811.12695

[32] B. Zafar, R. Ashraf., N. Ali, M. Ahmed, S. Jabbar and S. A. Chatzichristofis. Image classification by addition of spatial information based on histograms of orthogonal vectors, PloS one, 13(6): e0198175, 2018.

[33] E. M. Martey, Lei H., X Li., O. Appiah and N. S. Awarayi. Evaluation of RGB Quantization Schemes on Histogram-Based Content-Based Image Retrieval. In International Conference on Artificial Intelligence and Security, 736-747, 2021.

[34] O.Appiah, E. M. Martey and E. Quayson. Effect of Window's Shape on Median Filtering. In Proceedings of the 2019 IEEE AFRICAN, Accra, Ghana, 1–8, 2019.

[35] A. E. Minarno, Y. Munarko, F. Bimantoro, A. Kurniawardhani and N. Suciati. Batik image retrieval based on enhanced micro-structure descriptor. In Proceedings of the 2014 Asia-Pacific Conference on Computer-Aided System Engineering (APCASE), South Kuta, Indonesia, 65–70, 2014.

[36] G. Mori, S. Belongie and J. Malik. Shape contexts enable efficient retrieval of similar shapes. In Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Kauai, HI, USA, 1, 2001.

[37] B. Khaldi, O. Aiadi and M. L. Kherfi. Combining colour and grey-level co-occurrence matrix feature: A comparative study. IET Image Process, 13:1401–1410, 2019.

[38] Bala A. and Kaur, T. Local texton XOR patterns: A new feature descriptor for content-based image retrieval. Eng. Sci. Technol. Int. J., 19:101–112, 2016.

[39] M. E. Mensah, X. Li, H. Lei, A. Obed and N. C. Bombie. Improving Performance of Colour-Histogram-Based CBIR Using Bin Matching for Similarity Measure. In International Conference on Artificial Intelligence and Security, 586-596, 2020.