Colour application on mammography image segmentation

R Embong, N M Nik Ab. Aziz , A H Abd Karim & M R Ibrahim
Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia
E-mail: rohana@fskm.uitm.edu.my

Abstract. The segmentation process is one of the most important steps in image processing and computer vision since it is vital in the initial stage of image analysis. Segmentation of medical images involves complex structures and it requires precise segmentation result which is necessary for clinical diagnosis such as the detection of tumour, oedema, and necrotic tissues. Since mammography images are grayscale, researchers are looking at the effect of colour in the segmentation process of medical images. Colour is known to play a significant role in the perception of object boundaries in non-medical colour images. Processing colour images require handling more data, hence providing a richer description of objects in the scene. Colour images contain ten percent (10%) additional edge information as compared to their grayscale counterparts. Nevertheless, edge detection in colour image is more challenging than grayscale image as colour space is considered as a vector space. In this study, we implemented red, green, yellow, and blue colour maps to grayscale mammography images with the purpose of testing the effect of colours on the segmentation of abnormality regions in the mammography images. We applied the segmentation process using the Fuzzy C-means algorithm and evaluated the percentage of average relative error of area for each colour type. The results showed that all segmentation with the colour map can be done successfully even for blurred and noisy images. Also the size of the area of the abnormality region is reduced when compare to the segmentation area without the colour map. The green colour map segmentation produced the smallest percentage of average relative error (10.009%) while yellow colour map segmentation gave the largest percentage of relative error (11.367%).

1. Introduction
Mammography has become the main cancer detection strategy since the cause of cancer is still remain uncertain and it is widely available in most screening centres in many countries. There are various researches done on mammography images which involve the classification of abnormalities or breast density ([1], [2], [3]), detection of abnormalities ([4], [5]), or enhancement of images ([6], [7]). Researches relating to classification and detection mostly require the segmentation process. Segmentation of mammography images is a process of extracting suspicious regions containing breast abnormalities known as regions of interest, from its background. The principle goal of segmentation is to separate an image into homogenous regions of spatially related cluster of pixels called classes, pertaining to some characteristics or features such as texture, motion, etc. The union of any two neighbouring homogenous regions produces a heterogeneous region [8]. Some difficulties in the segmentation of mammography images are due to poor visualization and artefacts present in the images. Hence, the pre-processing is vital in the medical image processing to produce better image quality for segmentation and feature extractions [6].

Mammography images can be considered as grayscale images. Grayscale is a range of shades of grey without apparent colour. White is the lightest shade which allows the total transmission or reflection of light at all visible wavelengths while black is the darkest shade with total absence of any
light. The intermediate shades of grey are represented by equal brightness levels of the three primary colours (red, green and blue) for transmitted light, or equal amounts of the three primary pigments (cyan, magenta and yellow) for reflected light [9]. Colour image segmentation is more functional than grayscale image segmentation because of its ability to enhance the image analysis process, hence improving the segmentation result [10]. The awareness of object boundaries is significantly higher in colour images compared to grayscale images. Also processing colour images provides a richer description of image objects, with almost 10% more edge information [11], even though it is more challenging because it requires managing more data, since the colour space is considered as a vector space [12]. Colour image segmentation is valuable in applications such as in multimedia, text extraction from a colour image, skin tumour feature identification, segmentation of colour topographic maps and many more [10].

Since most of the algorithms involving medical images are in grayscale, little is known about the effect of colour in the segmentation process of grayscale mammography images. In a medical image processing research, blurred grey colour medical images such as mammography images can reduce the effectiveness of the segmentation process, hence reducing the accuracy of the result. The main purpose of this study is to test the effect of colours on the segmentation of grayscale mammography images. The accuracy of the segmentation result is measured based on the percentage relative error of the area of the abnormality in the ROI images. Implementation of this study is done using matrix laboratory (MATLAB) programming.

2. Methodology

2.1. Fuzzy C-Means algorithm

Fuzzy C-means clustering (FCM) algorithm is an extension of the hard C-means algorithm. Its goal is to find the optimum fuzzy clustering to minimize the value of an objective function [13]. Bezdek [14] defines the objective function of the Fuzzy C-means clustering as follows,

\[
J_m(U,V) = \sum_{i=1}^{n} \sum_{j=1}^{c} (u_{ij})^m \left\| x_i - v_j \right\|^2
\]  

(1)

Where \( X = \{x_1, x_2, x_3, \ldots, x_n\} \subseteq R^p \) is the sample set.
\( x_i = \{x_{i1}, x_{i2}, \ldots, x_{ip}\} \) each sample includes \( p \) features.
\( n \) is the number of samples in \( X \).
\( c \) (\( 2 \leq c \leq n \)) is the number of clusters.
\( U = (u_{ij})_{n \times c} \) is a fuzzy \( c \) partition matrix of a sample set \( X \).
\( u_{ij} \) represents the relative degree (membership value) between the \( i \)th sample and the \( j \)th cluster.
\( V = \{v_1, v_2, \ldots, v_c\} \) is the set of clustering center (centroid).
\( v_j \in R^p \) represents the clustering vector of the \( j \)th cluster.
\( \| \cdot \| \) is the standard Euclidean distance between the \( i \)th sample \( x_i \) and the clustering center of the \( j \)th cluster \( v_j \). However, \( \| \cdot \| \) can be defined by different distance formula according to one actual needs.
\( m \) is the fuzzy weight exponent or the smoothing parameter which determines the degree of the fuzzy partition matrix.

The objective function of FCM, Equation (1) yields the equations for cluster centres \( v_j \) and elements of the fuzzy \( c \) partition matrix \( u_{ij} \).

\[ v_j = \frac{\sum_{i=1}^{n} u_{ij}^m x_i}{\sum_{i=1}^{n} u_{ij}^m}, \text{ for } j = 1, \ldots, c \]  

(2)

\[ u_{ij} = \frac{1}{\sum_{k=1}^{c} \left( \frac{||x_i - v_j||}{||x_i - v_k||} \right)^{m-1}}, \text{ for } i = 1, \ldots, n \]  

(3)

2.2. Implementation

The implementation of this study consists of four main parts, namely input image, colour map implementation, image segmentation, and evaluation. The process is summarized by the flow chart in Figure 1. Input images are mammography images obtained from the National Cancer Society of Malaysia. Four colour maps (red, blue, green and yellow) are applied on the mammography images. The colour map is an \( m \)-by-3 matrix of real numbers between 0 and 1. A colour vector defines each colour on each row. The segmentation process applies Fuzzy C-means (FCM) clustering algorithm, and the segmentation areas of the abnormalities are calculated. Evaluation is based on the percentage relative error compared to the ground truth area.

![Flow chart of the methodology.](image)

The first and second part of the implementation of the methodology is shown in Figure 2. Four blurred mammography images and four different colour maps, blue, red, green and yellow are applied to the original mammography images. A total of nine images are tested.
When applying the colour maps on the blurred mammography images, each colour map creates boundary layer(s), thus producing clearer region of abnormality without going through any de-noising process. Therefore, all nine blurred images tested using colour maps are successfully segmented using FCM. The Fuzzy C-means (FCM) clustering algorithm is applied in the segmentation process of each image with four different colour maps. The area of the abnormality is calculated. The results are shown in Figure 3.

**Figure 2.** Implementation of four colour maps: blue, red, green and yellow.
3. Result and discussion
Table 1 lists the percentage of relative error of the area of the abnormalities in the segmentation of images with colour map using the Fuzzy C-means (FCM) clustering algorithm on nine images. Formula to calculate the percentage of relative error of the area of the abnormality is given by:

\[
\text{Percentage of relative error of area} = \left( \frac{\text{Area colour FCM} - \text{Original area}}{\text{Original area}} \right) \times 100
\]

Most of the areas of the abnormalities in the colour map images produced smaller areas, this is indicated by the negative values listed in Table 1. The average percentage of relative error of blue, red, green and yellow colour maps resulted in 3.2495%, 3.2150%, 4.7264% and 4.3469%, respectively. The biggest relative error 4.7264% is from the green colour map image segmentation by FCM, while the smallest relative error 3.2150% is from the red colour map images. The pie chart of the average relative error of each colour map is shown in Figure 4.
Table 1. Percentage relative errors of the area of the abnormalities in colour map images.

| Image | Blue FCM  | Red FCM  | Green FCM | Yellow FCM |
|-------|-----------|----------|-----------|------------|
| 1     | −2.0634   | −1.4475  | 0.5236    | 2.0840     |
| 2     | −5.2890   | −4.4048  | −3.7825   | −3.2913    |
| 3     | 0.1140    | 0        | 9.8041    | −8.5708    |
| 4     | 0.1570    | −0.7737  | −2.0295   | −2.5172    |
| 5     | −1.1839   | −0.9226  | −0.3429   | −0.3429    |
| 6     | −0.9544   | −1.3256  | −2.4390   | −2.9162    |
| 7     | 3.4611    | 2.6983   | −4.8341   | 0.0667     |
| 8     | −5.8676   | −6.3589  | −7.1098   | −7.4898    |
| 9     | −10.1552  | −11.0038 | −11.6718  | −11.8433   |
|       | Average   |          |           |            |
|       | relative  |          |           |            |
|       | error     |          |           |            |

| Image | Blue FCM  | Red FCM  | Green FCM | Yellow FCM |
|-------|-----------|----------|-----------|------------|
| 1     | −2.0634   | −1.4475  | 0.5236    | 2.0840     |
| 2     | −5.2890   | −4.4048  | −3.7825   | −3.2913    |
| 3     | 0.1140    | 0        | 9.8041    | −8.5708    |
| 4     | 0.1570    | −0.7737  | −2.0295   | −2.5172    |
| 5     | −1.1839   | −0.9226  | −0.3429   | −0.3429    |
| 6     | −0.9544   | −1.3256  | −2.4390   | −2.9162    |
| 7     | 3.4611    | 2.6983   | −4.8341   | 0.0667     |
| 8     | −5.8676   | −6.3589  | −7.1098   | −7.4898    |
| 9     | −10.1552  | −11.0038 | −11.6718  | −11.8433   |
|       | Average   |          |           |            |
|       | relative  |          |           |            |
|       | error     |          |           |            |

4. Conclusion

In conclusion, segmentation of mammography images with the application of colour maps can be successfully done even though the images are very blurred. Four colour maps, namely blue, red, green and yellow are selected and the segmentation is done using the Fuzzy C-means algorithm. The smallest average percentage relative error of the abnormality areas is from the red colour map, followed by blue, yellow and green. In most images, applying colour maps, reduced the area of the abnormality region. Further research can be done to test whether the reduction of the area may signify the reduction of the execution time. If this is true, then it may benefit many researches in medical image segmentation since processing medical images is very time consuming. Also, by expanding the colour selection may provide the ideal colour for certain purposes. For example, an application of a certain colour map can be a simple and a suitable method of de-noising.
Acknowledgments
This paper is sponsored by grant no. 600-IRMI/DANA 5/3/LESTARI (0128/2016) by Universiti Teknologi MARA (UiTM), Shah Alam, Malaysia.

References
[1] Benmazou S, Merouani H F, Layachi S and Nedjmeddine B 2014 Classification of mammography images based on cellular automata and Haralick parameters Evolving Systems 5(3) 209-216
[2] Arevalo J, González F A, Ramos-Pollán R, Oliveira J L and Lopez M A G 2016 Representation learning for mammography mass lesion classification with convolutional neural networks Computer methods and programs in biomedicine, 127, 248-257
[3] Virmani J, Dey N and Kumar V 2016 PCA-PNN and PCA-SVM based CAD systems for breast density classification In Applications of intelligent optimization in biology and medicine 159-180 Springer International Publishing
[4] Zhang Y D, Wang S H, Liu G and Yang J 2016 Computer-aided diagnosis of abnormal breasts in mammogram images by weighted-type fractional Fourier transform Advances in Mechanical Engineering, 8(2), 1687814016634243
[5] Wigati K T, Cockmartin L, Marshall N W, Soejoko D S and Bosmans H 2016 Comparison of two psychometric functions in analyzing 4-AFC detection results using a task-based structured phantom for digital mammography Physica Medica: European Journal of Medical Physics, 32, 215-216
[6] Makandar A and Halalli B 2017 Mammography image enhancement using linear, nonlinear and wavelet filters with histogram equalization International Journal of Computer (IJC), 25(1), 8-17
[7] Abdallah Y M Y, Hayder A and Wagiallah E 2014 Automatic enhancement of mammography images using contrast algorithm Methods, 7, 8
[8] Embong R, Wan Abd Rahman W E Z, Ahmad T, Mahmud R, Ibrahim A, Abu Bakar Z, Salikin M S and Jumaat A K 2009 Boundary segmentation of tumor regions using snake algorithms in digital mammography images In Proceedings of The 7th International Conference on Robotics, Vision, Signal Processing, & Power Applications (RoViSP 2009), Penang, Malaysia, 19-20 December 2009, ISBN 978-983-3986-52-1
[9] Rouse M 2010, 23 May Grayscale Retrieved June 17, 2017 WhatIs.com: http://whatis.techtarget.com/definition/grayscale
[10] Bora D J, Gupta A K and Khan F A 2015 Comparing the performance of L* A* B* and HSV color spaces with respect to color image segmentation arXiv preprint arXiv:1506.01472
[11] Dutta S and Chaudhuri B B 2009 A colour edge detection algorithm in RGB colour space In Proceedings of Advances in Recent Technologies in Communication and Computing, 337-340
[12] Mittal A, Sofat S and Hancock P 2011 Detection of edges in colour images: A review and evaluative comparison of state-of-the-art techniques Department of Computer Science, University of York, YORK, UK
[13] Jia Z, Zhang C and Sun R 2007 Fuzzy c-means algorithm based on pre-treatment of similarity relationship Dynamics of Continuous, Discrete and Impulsive Systems Series B: Appl. & Algo. 14, 103-111
[14] Bezdek J C 1980 A convergence theorem for the fuzzy ISODATA clustering algorithms IEEE Trans on Pattern Anal. Machine Intell. PAMI-2, 1-8