# Color Segmentation of Breast Thermograms: A Comparative Study

**Mahnaz Etehadtavakol and Eddie Y.K. Ng**

**Abstract** Color segmentation of breast thermograms can have a crucial performance in tumor detection. There is a relation between blood vessel activity and the surrounding area temperature. Once a cancer increases blood vessel activity, the cancer cells and their surrounding tissue become hotter than normal tissue. Pre-cancer and cancer cells need plenty of nutrients to multiply and survive consequently; they are highly metabolic tissue and have different thermal patterns compared to the normal one. In this paper, a comparison work is presented for three modeled color segmentation approaches: $K$-means, mean shift (MS), and fuzzy $c$-means (FCM) applied to infrared breast images. There are some drawbacks for $K$-means and MS approaches. Almost empty clusters may be obtained in the segmentation results using $K$-means algorithm. In addition, we frequently confront almost empty clusters with MS algorithm due to its sensitiveness to the window size parameter. Choosing an appropriate window size parameter is not an easy task. On the other hand, the fuzzy inherent breast thermal images aid the FCM technique to obtain more precise outcomes. Malignant tumors show hotter thermal patterns than healthy tissues and even with benign tissues. Segmenting different parts of two breasts in terms of their temperature has potential helping to identify abnormal breast tissues.

M. Etehadtavakol (✉)
Isfahan University of Medical Sciences, Isfahan, Iran
e-mail: mahetehad@gmail.com

E.Y.K. Ng
School of Mechanical and Aerospace Engineering, College of Engineering,
Nanyang Technological University, Singapore, Singapore

# 1 Introduction

Over 2000 years ago Hippocrates layered mud over a patient's body. He noticed the mud dried quicker where the internal disease generated heat. Actually he understood the diagnostic significance of body temperature changes.

Infrared energy wavelength is extended from 700 nm to 1 mm which human eyes sensors are not able to detect the all range, so it is not visible. Infrared energy is a sector of the electromagnetic spectrum that one recognizes as heat. All objects with a temperature above absolute zero radiate infrared electromagnetic energy. Objects with higher temperature radiate larger infrared radiations. One infrared camera can capture the heat pattern emitted from the object.

Breast thermal imaging is a radiation-free updating screening procedure that captures heat pattern of the breast helping early detection of cancer. The infrared radiation from an individual as an outcome of surface temperature is related to the amount of blood circulation. As a tumor grows it develops a blood supply that emits more heat [1].

Breast thermal imaging is a harmless physiological examination that presents the chance of earlier detection of breast disease than has been attainable through breast self-examination, and other breast screening tests alone.

Breast cancer if detected early can be taken out and stopped. Once a woman who has cancer is too late to start early detection, but if identified early enough, it can be removed in initial phases before it metastasizes and spreads to the neighboring area.

Breast thermal imaging is a painless, noninvasive, fast, low cost, passive, updating clinical examination without any risk of radiation. All subjects with any age with any breast size can benefit of this technique. Also women with dense breast as well as nursing or pregnant women can use breast thermal imaging [2–6].

Lawson declared in 1956 that the surface temperature of cancerous breast was higher than the normal one. Therefore, Lawson claimed that the cancerous cells can be recognized as hot regions in thermograms [7]. In infrared pseudo-color images distinct colors show distinct rate of hotness. Hence, breast thermograms' color segmentation can be beneficial in identifying the suspecting zones.

This study is arranged in this way: Pseudo-coloring of gray images is explained in Sect. 2. MS technique is discussed in Sect. 3. Numerical outcomes and discussions are given in Sect. 4 and finally the findings are achieved in Sect. 5.

# 2 Color Conversion of Gray-Scale Images

Pseudo-coloring or color conversion of gray-scale images is one of the possible means of upgrading digital image. The human eye is more sensitive to differences in color than to differences in shades of gray. It is a common task in image processing with the aim of upgrading visual quality, presenting a visual appeal and emphasizing specific features in the image.

Two limitations with gray-scale images which are related to specific characteristics of the individual discernable system lead to suggest the concept of pseudo-coloring. First, as the size of the feature composed of the alternate gray shade is reduced, a human's ability to discern gray-shade differences in an image decreases. As the size of the feature decreases, it takes larger differences in gray-scale intensity for a person to be able to see the alternate shade of gray. Second, humans are only capable of discerning between 60 and 90 separate shades of gray, suggesting much of the information in a typical 256 gray-shade image is lost on the viewer. Furthermore, individuals are capable of discerning more than 500 shades of color [8, 9].

Although it is controversial that the viewer allocates greater knowledge by pseudo-coloring of gray images, thoroughly produced pseudo-color image indeed allows tumor identification that is as good as gray scale and enhances fulfillment of other jobs like recognition and interpretation of a tumor. In the interest of rehearsing a thoroughly produced pseudo-coloring which can keep the whole knowledge of grey-scale images and derives no deformity in images, uniform color space is suggested [10].

In a uniform color space, identical alterations in the component of color space correlate to identical noticeable alterations in color chroma and color hue angle. However, many elements cause some difficulties. Some of these elements are as follows: (1) the color of an area relies on neighborhood colors; (2) chromatic accommodation impacts are influenced by eye motions, because colors are a head or come after other colors in time; and (3) color discriminability evaluates on the angular stretch of the investigation area. The 1976 CIELAB color space is suggested by the International Commission on Illumination (CIE).

The space of CIELAB is shown in Fig. 1.

It is an attempt to provide an intuitively uniform color space. In this color space, the amount of dissimilarities of the colors in terms of luminance, chroma, and hue can be nearly described by the two-point separation length. In addition, tristimulus values XYZ can obtain CIELAB coordinates ($L*$, $a*$, $b*$) as expressed by the

**Fig. 1** CIELAB color space [11]
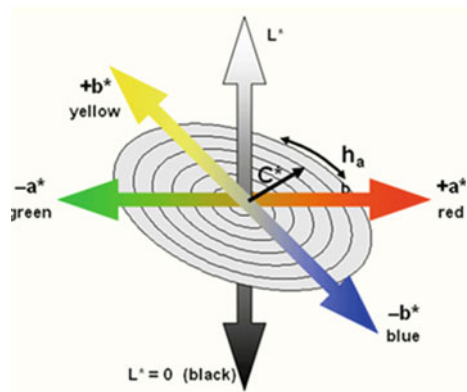
formulas in Eq. (1) where the white point values are denoted by the subscript $n$ [12, 13]:

$$L^* = 116g\left(\frac{Y}{Y_n}\right) - 16$$
$$a^* = 500\left[g\left(\frac{X}{X_n}\right) - g\left(\frac{Y}{Y_n}\right)\right] \tag{1}$$
$$b^* = 200\left[g\left(\frac{Y}{Y_n}\right) - g\left(\frac{Z}{Z_n}\right)\right],$$

where

$$g(m) = m^{1/3} \quad m > \left(\frac{6}{29}\right)^3 \tag{2}$$

$$g(m) = \frac{1}{3}\left(\frac{29}{6}\right)^2 m + \frac{4}{29} \quad m \le \left(\frac{6}{29}\right)^3. \tag{3}$$

## 3 mean shift clustering (ms)

MS is a nonparametric technique that was introduced in 1975, [14]. This technique does not need knowing about the number of clustering and does not force any format of the clusters in advance [15–18]. Supposing a set of feature vectors, $y_i$ ($i = 1,…, n$), be in a feature space of dimension $d$. The relation between the density of any point $y$ in terms of the kernel density evaluator $K(y)$ and the size of window $h$ is obtained by Eq. (4):

$$f_{h,K}(y) = \frac{c_{k,d}}{nh^d} \sum_{i=1}^{n} K\left(\left\|\frac{y - y_i}{h}\right\|^2\right). \tag{4}$$

The zero gradient of Eq. (4) produces Eq. (5)

$$= \frac{2c_{k,d}}{nh^{(d+2)}} \sum_{i=1}^{n} q\left(\left\|\frac{y - y_i}{h}\right\|^2\right) \left(\frac{\sum_{i=1}^{n} y_i q\left(\left\|\frac{y-y_i}{h}\right\|^2\right)}{\sum_{i=1}^{n} q\left(\left\|\frac{y-y_i}{h}\right\|^2\right)} - y\right), \tag{5}$$

where $q(s) = -K'(s)$, supposing function $q$ be the derivate of function $K$. Equation (5) is a product of two expressions. The first expression is proportional to the density estimate at $y$, and the MS vector which is the second expression is introduced by Eq. (6):

**Fig. 2** Demonstration of mean shift algorithm; the *red arrow* shows the MS vector; the *green circle* shows the updated center of the cluster
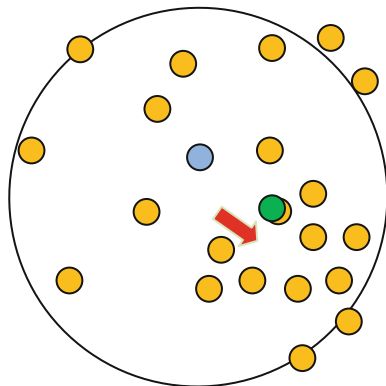
$$m_h(y) = \frac{\sum_{i=1}^{n} y_i q\left(\left\|\frac{y-y_i}{h}\right\|^2\right)}{\sum_{i=1}^{n} q\left(\left\|\frac{y-y_i}{h}\right\|^2\right)} - y. \tag{6}$$

It is noticed that the direction of the greatest amount of growth in the density is certainly pointed by the MS vector. The MS algorithm is demonstrated in Fig. 2. The red arrow displays the MS vector while the updated center of the cluster is indicated by the green circle. It is noteworthy to mention that the window size variable $h$ is the only variable of the MS technique. It has remarkable impact on the acquired results.

## 4 Experimental Results

In this study $K$-means [19] and FCM [20] were applied for color segmentation of breast thermograms [21, 22]. However, the three techniques $K$-means, MS, and FCM are applied and compared. Sixty breast thermal images accessible from five different breast thermography centers [23–27] were investigated in this work.

These three techniques put all cases into practice. With $K$-means algorithm, almost empty clusters are appeared in some experiments. Besides, if $K$-means is redone some times, the outcomes for distinct experiments are non-identical in accordance with the drawbacks of the $K$-means technique [28–31]. Figure 3b shows the experiment in which two almost empty clusters are obtained for a cancerous individual as shown in Fig. 3a. Nevertheless by performing the FCM technique (Fig. 3c), no empty cluster is obtained because fuzzy $c$-means technique measures the colors in a comparative way and categorize them in groups that are not with rigid borders. Consequently, data point is able to associate to more than one group although in hard clustering ($K$-means) data is collected to rigid groups suggesting each data point associates to particularly one group [32–34]. Also performance of
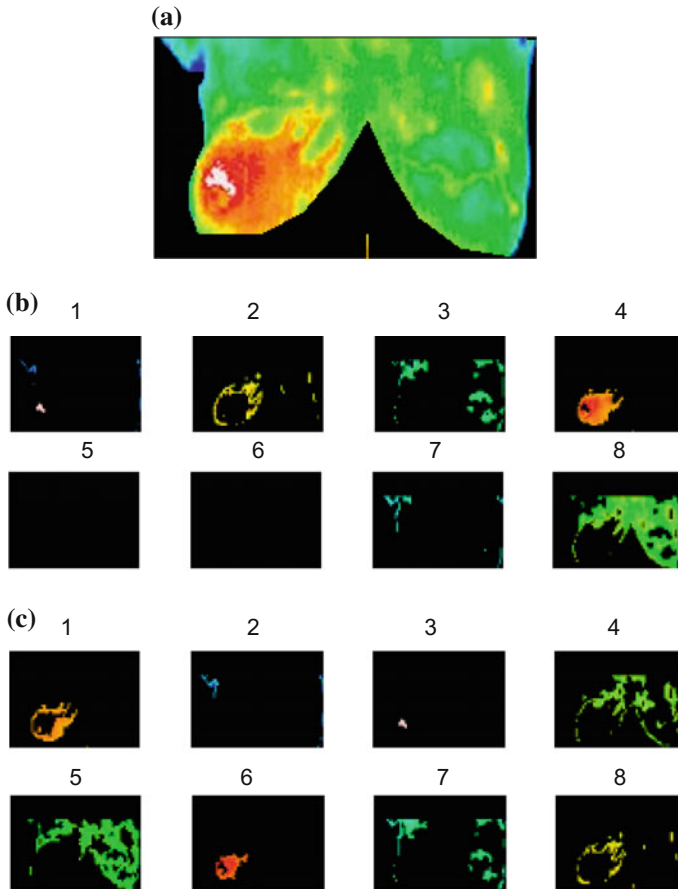
**Fig. 3 a** An inflammatory cancer breast thermogram. **b** Two almost empty clusters out of eight clusters obtained by *K*-means technique. **c** Eight non-empty clusters obtained by FCM clustering

FCM for a normal individual (Fig. 4a) is demonstrated in Fig. 4b. MS technique is easily affected by parameter *h*, the window size. In this work, MS algorithm was performed for a normal case. Figure 4b shows the results for $h = 12$ with eleven empty clusters while Fig. 4c for $h = 15$ with four empty clusters.

In this work, FCM technique has potential to identify the two first hottest areas by measuring their colors with respect to the provided palette spectrum colors. By color segmenting of breast thermograms, we are able to identify the two hottest clusters. Consequently providing some helpful data from the doubtful areas by measuring the color of the two hottest areas with the color of the provided palette spectrum is feasible.

**(a)**



**(b)**



**(c)**
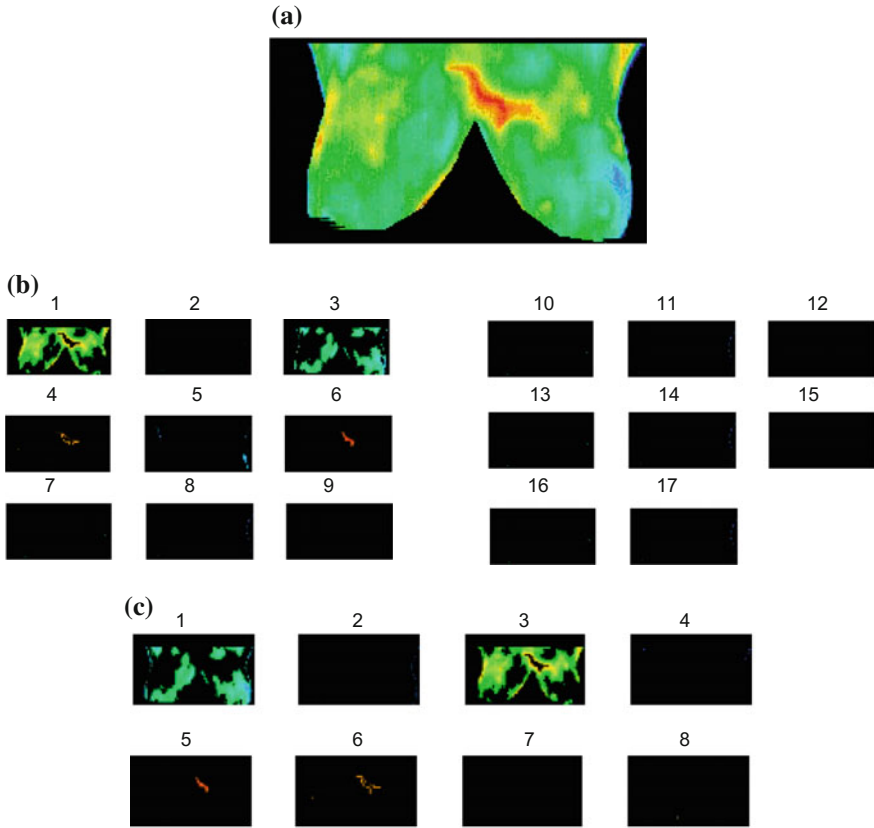


**Fig. 4  a** A normal breast thermogram. **b** Eleven almost empty clusters out of seventeen clusters obtained by MS with h = 12. **c** Four almost empty clusters out of eight clusters obtained by MS with h = 15

## 5  Conclusion

Although it is controversial that with pseudo-colored gray images more data can be allowed to the observer, but precisely produced pseudo-color image indeed is capable of tumor identification that is equivalent to the gray scale and upgrades accomplishment of other jobs like perception and judgment of a tumor. In this research, three techniques for breast thermograms color segmentation were used. MS technique is easily affected by parameter $h$, the window size, and accordingly, in this study, we usually challenged empty clusters and selecting a suitable $h$ was not straight forward. $K$-means technique reduces the sum of within cluster outcomes numerically and clusters are tight and fully divided and produce precise outcomes. Due to the $K$-means drawbacks, nevertheless, for many cases, $K$-means technique is not practicable where clusters are not separated and an element in an image may

associate to various clusters. Choosing appropriate initial cluster centers is crucial for *K*-means clustering. Inadequate initial centers may produce empty clusters. In FCM technique, elements of the image are included in different clusters with disparate levels of association. By reason of the fuzzy nature of the breast thermograms, more faultless outcomes were resulted by FCM method to segment breast thermograms.

In this study, 60 breast thermograms were studied. By implementing FCM technique, we provided the two first hottest areas for each subject where convenient features are obtainable. There are some conveniences of breast thermograms color segmentation by FCM such as identification of irregular subjects by contrasting the analogous clusters from the contralateral breasts (over the symmetrical line), determination of level of malignity with identifying the two first hottest areas, and extraction of some useful features from them.

# References

1. Etehadtavakol, M., Chandran, V., Ng, E.Y.K., Kafieh, R.: Breast cancer detection from thermal images using bispectral invariant features. Int. J. Thermal Sci. **69**, 21–36 (2013)
2. Jones, B.F.: A reappraisal of the use of infrared thermal image analysis in medicine. IEEE Trans. Med. Imaging **17**(6), 1019–1027 (1998). doi:10.1109/42.746635
3. Ng, E.Y.K.: A review of thermography as promising non- invasive detection modality for breast tumour. Int. J. Therm. Sci. **48**(5), 849–859 (2009)
4. Diakides, N., Bronzino, J.D.: Medical Infrared Imaging, 3rd edn. CRC, Taylor & Francis, New York (2008)
5. Ng, E.Y.K., Kee, E.C.: Advanced integrated technique in breast cancer thermography. J. Med. Eng. Technol. **32**(2), 103–114 (2008). doi:10.1080/03091900600562040
6. Etehadtavakol, M., Ng E.Y.K., Breast thermography as a potential non-contact method in early detection of cancer: a review. J. Mech. Med. Biol. **13**(2), 1330001-1–1330001-20 (2013)
7. Lawson, R.N.: Implications of surface temperature in the diagnosis of breast cancer. Can. Med. Assoc. J. **75**, 4309–4310 (1956)
8. Vaughn G., "Image Processing Class Notes", Texas Tech University Interdisciplinary Engineering Masters Program, 2007
9. Weeks, A.R., Fundamentals of Electronic Image Processing, 3rd edn. SPIE Press, Bellingham, Washington (2004)
10. Zhou, X., Zhang, C.: A perceptive uniform pseudo-color coding method of SAR images, Radar, CIE. International Conference, Oct. 2006 IEEE, pp. 1–4 (2006)
11. http://www.flexoglobal.com/flexomag/08-September/flexomag-ploumidis.htm
12. Li, H., Burgess, A.E., Evaluation of signal detection performance with pseudo-color display and lumpy backgrounds, JrnlID 10916_ArtID 9213_Proof# 1 - 08/09/2008. In: Kundel, H.L. (ed.) SPIE, Medical Imaging: Image Perception, vol. 3036, pp. 143–149. Newport Beach CA, USA (1997)
13. Connolly, C., Fliess, T.: A study of efficiency and accuracy in the transformation from RGB to CIELAB color space. IEEE Trans. Image Process. **6**, 1046–1048 (1997). doi:10.1109/83.597279
14. Fukunaga, K., Hostetler, L.D.: The estimation of the gradient of a density function, with applications in pattern recognition. IEEE Trans. Inf. Theory 2, 32–40 (1975)

15. Comaniciu, D., Meer, P.: Mean shift: a robust approach toward feature space analysis. IEEE Trans. Pattern Anal. Mach. Intell. 24(5), 603–619 (2002)
16. Cheng, Y.: Mean shift, mode seeking, and clustering. IEEE Trans. Pattern Anal. Mach. Intell. 17(8), 790–799 (1995)
17. Jin, H., Tao, W., Zhang, Y.: Color image segmentation based on mean shift and normalized cuts. IEEE Trans. Syst. Man Cyberetics 37(5), 1382–1389 (2007)
18. Mayer, A., Greenspan, H., Segmentation of brain MRI by adaptive mean shift, biomedical imaging: nano to micro 3rd IEEE International Symposium, pp. 319–322 (2006)
19. MacQueen, J.B., Some methods for classification and analysis of multivariate observations. In: Proceedings of Fifth Berkeley Symposium on Mathematical Statistical and Probability, vol. 1, pp. 281–297. University of California Press (1967)
20. Bezdek, J.C., Keller, J., Krisnapuram, R., Pal, N.R.: Fuzzy models and algorithms for pattern recognition and image processing. Kluwer, Norwell, MA (1999)
21. Forgy, E.W.: Cluster analysis of multivariate data: efficiency versus interpretability of classifications. Biometrics 21, 768–769 (1965)
22. Etehadtavakol, M., Sadri, S, Ng, E.Y.K., Application of K- and fuzzy c-means for color segmentation of thermal infrared breast images. J. Med. Syst. **34**(1), 35–42 (2010). doi:10. 1007/s10916-008-9213-1
23. AAT: http://aathermography.com. Access Aug 2015
24. MII: http://www.breastthermography.com/case_studies. Access Aug 2015
25. ACCT: www.thermologyonline.org/Breast/breast_thermography_what.htm. Access Aug 2015
26. http://www.thermographyofiowa.com/casestudies.htm. Access Aug 2015
27. ST Imaging: http://www.stimaging.com.au/page2. Access Aug 2015
28. Deelers, S., Auwatanamongkol, S.: Enhancing K-means algorithm with initial cluster centers derived from data partitioning along the data axis with the highest variance. Int. J. Comput. Sci. **2**, 4247–4252 (2007)
29. Bradley, P.S., Fayyad, U.M., Refining initial points for K-Means clustering. In: Proceedings of 15th International Conference on Machine Learning, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, pp. 91–99 (1998)
30. Krishna, K., Narasimha, M., Genetic K-Means algorithm. IEEE Trans. Syst. Man Cybern. **29** (3), 433–439 (1999). doi:10.1109/3477.764879
31. Golestani, N., Etehadtavakol, M., Ng, E.Y.K.: Level set method for segmentation of infrared breast thermograms. Exp. Clin. Sci. **13**, 241–251 (2014)
32. Sowmya, B., Bhattacharya, S., Color image segmentation using fuzzy clustering techniques. IEEE Indicon, Conference, Chennai, India, pp: 41–45, Dec (2005)
33. Solvenia, K. R., Fuzzy C-means clustering and facility location problems. In: Proceeding of Artificial Intelligence and Soft Computing, Palma de Mallorca, Spain, p. 544 (2006)
34. Acharya, U.R., Ng, E.Y.K., Tan, J.H., et al.: An integrated index for the identification of diabetic retinopathy stages using texture parameters. J. Med. Syst. **36**(3), 2011–2020 (2012)