ISSN 2348-375X

Unique Journal of Engineering and Advanced Sciences

Available online: www.ujconline.net

Research Article

DETECTION OF SKIN CANCER BY USING NUCLEAR-TO-CYTOPLASMIC RATIO OF IN VIVO VIRTUAL BIOPSY IMAGES

Sumitradevi KS[1*], Sangeetha KK[2]

[1]M.E-Applied Electronics, Nandha Engineering College, Perundurai Road, Erode-52, Tamil nadu, India
[2]Assistant Professor, Nandha Engineering College, Perundurai Road, Erode-52, Tamil nadu, India

Received: 24-02-2014; Revised: 22-03-2014; Accepted: 21-04-2014

*Corresponding Author: **K.S. Sumitradevi.**
M.E-Applied Electronics, Nandha Engineering College Perundurai Road, Erode-52,Tamil nadu, India.
Email:sumitra808@gmail.com

ABSTRACT

Pathology diagnosis for early detection of skin cancer is achieved primarily through physical biopsy, which involves the removal of suspected tissue. In addition to being painful to the patient, these procedures risk infection or even spreading of cancer cells. So, in recent works, microscopic images has been successfully and safely used for non invasive in vivo virtual optical biopsy on skin to obtain exhaustive tissue images without removing tissues. In existing system watershed transform is used, which involves over segmentation and sensitive to false edges. But, the proposed system overcomes those problems and reveals high accuracy for cell segmentation and has dramatic potential for non invasive analysis of cell nuclear-to-cytoplasmic ratio (NC ratio), which is important in identifying or detecting early symptoms of diseases with abnormal NC ratios, such as skin cancers during clinical diagnosis. In this system k-means clustering technique is used for the segmentation of the cell with less sensitive to false edges and over segmentation is avoided. The main application of this proposed system is for detecting aged skin and skin cancer diagnosis. MATLAB 2011a is used for the implementation of these techniques.

Keywords: Nuclei detection, Cell segmentation, cytoplasm detection, Nuclear-to-cytoplasmic ratio, K-means clustering algorithm.

INTRODUCTION

An early warning of skin cancer is a new, suspicious looking lesion on the skin. Fortunately, there is no reason to panic, since most skin blemishes are not cancerous, but a skin cancer diagnosis should still be performed to be sure. To determine if the lesion is cancerous or not, the doctor or nurse will usually first discuss the medical history to determine the risk factors, including the history of skin cancer in the patient's family and the number of prior sunburns[1]. A skin examination will follow, during which the doctor will note the size, shape, color, and texture of the suspicious area. He or she will then examine the lymph glands to check for swelling, a potential sign of cancer. The only way to definitely diagnose the various types of skin cancer is to biopsy suspicious-looking lesions[2]. Useful information, such as tumor depth, can only be obtained by biopsy. A biopsy is a medical test commonly performed by a surgeon or an interventional radiologist involving sampling of cells or tissue for examination. It is the medical removal of tissue from a living subject to determine the presence or extent of a disease. The tissue is generally examined under a microscope by a pathologist, and can also be analysed chemically. This invasive biopsy procedure includes pain, bleeding, scar formation, infection, and the risk of cancer cell spreading[3]. Optical virtual biopsy techniques for cells and tissues imaging provides capable microscopic details about the benign and malignant lesions without tissue removal. This non-invasive in vivo virtual biopsy avoids or minimizes the above mentioned disadvantages involved in virtual biopsy procedure.

It also reduces the cost and time consumption in traditional biopsy procedures. So, we are disabled to study living bio-activities since the sampled tissue is no longer alive. In order to make a definitive diagnosis in the early stage, the least invasive in vivo microscopic technique that keeps the advantages of physical biopsy is highly required[4]. In microbiology in vivo is often used to refer to experimentation done in live isolated cells rather than in a whole organism, for example, cultured cells derived from biopsies. The use of an imaging modality(e.g., ultrasonography, functional MRI, or virtual biopsy colonoscopy-a direct visualisation technique coupled with a probe-based Confocal Laser Endomicroscope-

pCLE) to localise a tumour and assess its physical features is called virtual biopsy[5].

CELL SEGMENTATION TECHNIQUES

Nuclear-to-cytoplasm ratio analysis is a method to identify early symptoms of diseases like skin cancer, whose ratio is generally larger in skin cells than in normal cells[6]. Therefore, there are several cell segmentation techniques, which help us to measure the NC ratio. Some of the techniques have been discussed below[7].

A. Image Thresholding

Image thresholding is one of the cell segmentation methods used to segment the objects out of background and it does not separate the touching nuclei. In cell images global threshold minimizes the variations between different images which are important for separating the touching or overlapping nuclei but local threshold provides the most suitable value for segmentation with better adaptability[8]. Bin Fang et al.(2003) proposed a two-stage tumour cells identification strategy, in its first stage it uses local adaptive threshold for automatic potential tumour cell segmentation. Here, the global threshold is very expensive and the clumped cells cannot be considered, because it uses single threshold level for entire image[9]. In global threshold, 'brighter' backgrounds are misclassified as cells and 'darker' cell regions are misclassified as background. By minimizing the negative effect of background noise, local adaptive threshold segments the region of interest from background. Local threshold allows different values, which have been applied to each pixel[10].

B. Convergence Index Filter

For cytoplasm segmentation convergence index filter is used[11]. Convergence index is a measure of how strongly the gradient vector point towards the pixel of interest. Convergence index filter degree of convergence is based on the distribution of the gradient vector not on their magnitude[12]. This filter evaluates the degree of convergence of the gradient vectors in the neighbourhood of each pixel of interest, i.e., the distribution of the directions of the gradient vectors with respect to the pixel of interest is evaluated. Hidefumi Kobatake et al. (1999) proposed a unique filter called an iris filter, which evaluates the degree of convergence of the gradient vectors within its region of support toward a pixel of interest[13]. This filter changes its shape and size based on the gradient vectors distribution in the pixel of interest. Regardless of the contrast of its background, it is possible to enhance indistinct boundaries and detect rounded convex regions in an image. It is based on the maximization of the convergence index at each point of the spatial co-ordinates[14].

Some of the convergence index filters are Coin filter (CF), Adaptive ring filter (ARF), iris filter, sliding band filter (SBF). Support region, is the major difference among these filters [15]. The CF uses a circle with variable radius as support region, the IF maximizes the convergence index by adapting the circle's radius value on each direction and the ARF uses a ring shaped region with fixed width and varying radius[15]. Finally, the SBF combines the ideas of IF and ARF by defining a support region formed by a band of fixed width, whose position is changed in each direction to allow the maximization of the convergence index at each point. The SBF filter is normally well suited for cell detection because of its overall convex shape[16].

C. K-Means Clustering

Clustering is the process of partitioning or grouping a given set of patterns in the same cluster are alike and patterns belonging to two different clusters are different. Clustering has been a widely studied problem in a variety of application domains including neural network, AI, and statistics. Several algorithms have been proposed in the literature for clustering: ISODATA, CLARA, CLARANS, DBSCAN, BIRCH and GRIDCLUS. The k-means method has been effective in producing good clustering results for many practical applications. However, a direct algorithm of k-means method requires time proportional to the product of number of patterns and number of cluster per iteration. This is computationally very expensive especially for large datasets.

PROPOSED SYSTEM

Among the several cell segmentation methods, k-means clustering is the efficient method, so the proposed system uses this method. The proposed system has two parts, such as, cell segmentation and Nuclear-to-Cytoplasmic ratio (NC ratio) analysis. This system uses some sets of in vivo virtual biopsy images for automatic cell segmentation, which is based on the watershed-based approach and the concept of convergence index filter[17]. And the analysis of NC ratio is a good indicator for identifying early symptoms of skin cancer in addition to revealing the type and stages of the developing diseases[18]. In microbiology in vivo is often used to refer to experimentation done in live isolated cells rather than in a whole organism, for example, cultured cells derived from biopsies. The use of an imaging modality (e.g.,) ultrasonography, functional MRI, or virtual biopsy colonoscopy-a direct visualization technique coupled with a probe-based Confocal Laser Endomicroscope-pCLE)) to localise a tumour and assess its physical features is called virtual biopsy[19]. The block diagram is divided into two parts: cell segmentation and nuclei-to-cytoplasmic ratio analysis. Among the several algorithms for cell segmentation, image thresholding is one of the method for high-speed cell segmentation, but it produces good results only for images with high contrast between object and background[20]. Then watershed transform is the common technique, uses the concept of morphological image processing and considers contextual information to produce stable segmentation results. Much of cell biology experimental research is based on microscopy image analysis of cell culture. In many cases the use of different fluorescence dies or proteins is used to enable the collection of multivariate images which contain information on different aspects of each cell[21]. Although analysis of such images can be performed manually, it is time consuming, exhausting and prone to human error, requiring frequent repetitions to validate results. These factors motivate the development of automatic cell analysis tools, to identify each individual cell and extract relevant cell characteristics[22]. This process involves the separation of each individual cell from all others cells and from the background. Multivariate imaging is widely used in cell microscopy to obtain separated cell nucleus and cytoplasm information by using different fluorescence markers, originating two image channels.

Additional fluorescence markers (and channels) can be used to express other phenomena or objects such as parasites or different nuclei acid concentrations. Analysis of cell nuclear-to-cytoplasmic ratio(NC ratio), which is important in identifying or detecting early symptoms of diseases with abnormal NC ratios, such as skin cancers during clinical diagnosis. The NC ratio in the epidermis is a good indicator for early symptoms of skin diseases. The nucleus-to-cytoplasm ratio is the ratio of the nucleus size to the cytoplasm size[23]. During cell division, the nucleus makes up the larger portion of the cell.

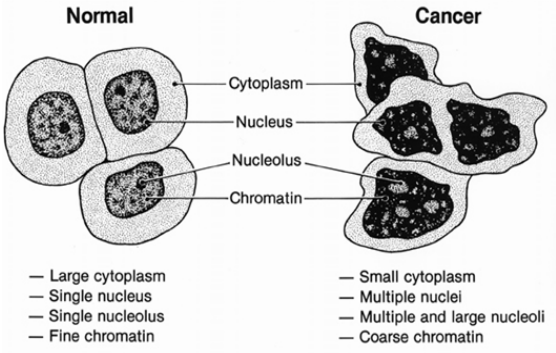


Figure 1: Normal cells And Cancer Cells

The cell nuclear-to-cytoplasmic ratios (NC ratios) of epidermis are generally larger in skin cancer than in normal cells in Fig. 1. Nuclei segmentation is performed using gradient watershed transform with marker-controlled strategy, blob detection, and consideration of shape descriptors to obtain accurate segmented nuclei. The nucleus-to-cytoplasm ratio is the ratio of the nucleus size to the cytoplasm size. During cell division, the nucleus makes up the larger portion of the cell. The ratio gets decreases when the cell start matures and the cancerous cell does not get mature and healthy. So, the nucleus-to-cytoplasm ratio is always larger in mature normal cells[24]. This NC ratio can be used to diagnose the cancer cells in some tissues, by using this ratio we can also determine the skin age and it is used for many medical applications. Input image is the microscopic image of the skin cell[25]. The algorithms such as image thresholding, watershed transform, another common technique, uses the concept of morphological image processing and considers contextual information to produce stable segmentation results. It uses region minimum as starting points.

For nuclei segmentation, watershed algorithm is the most widely used one. Watershed algorithm usually leads to over-segmentation, because it is difficult to have one to one correspondence between regional minima and nuclei. When nuclei are clustered, it becomes worse. Watershed algorithm is used to segment the touching objects, sometimes leads to over segmentation. So, the proposed methodology incorporates k-means algorithm for cell segmentation. The use of the conventional watershed algorithm for cell segmentation analysis is widespread because of its advantages, such as always being able to produce a complete division of the image[26]. However, its drawbacks include over-segmentation and sensitivity to false edges. We address the drawbacks of the conventional watershed algorithm when it is applied to medical images by using k-means clustering to produce a primary segmentation of the image before we apply our improved watershed segmentation algorithm to it[27]. The k-means clustering is an unsupervised learning algorithm, while the improved watershed segmentation algorithm makes use of automated thresholding on the gradient magnitude map and post-segmentation merging on the initial partitions to reduce the number of false edges and over-segmentation.

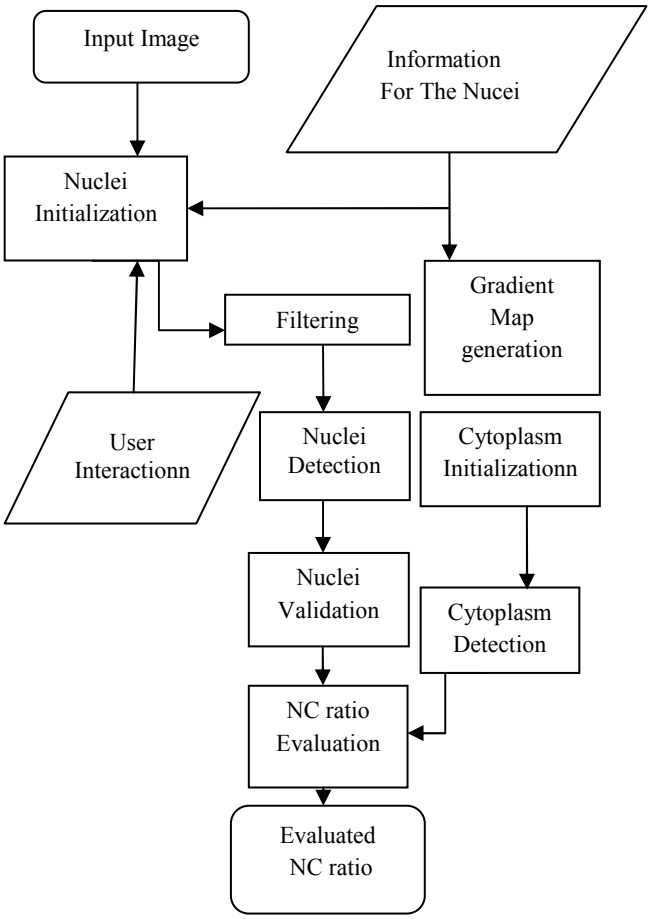


Figure 2: Block Diagram of Nuclear-to-cytoplasmic ratio

A. Nuclei initialization

Nuclei segmentation is performed by using K-means clustering, blob detection, and consideration of shape descriptors to obtain accurate segmented nuclei in Fig. 2. Blob detection refers to mathematical methods that are aimed at detecting regions in a digital image that differ in properties, such as brightness or color, compared to areas surrounding those regions[28]. Informally, a blob is a region of a digital image in which some properties are constant or vary within a prescribed range of values; all the points in a blob can be considered in some sense to be similar to each other[29]. A distance transform, also known as distance map or distance field, is a derived representation of a digital image. The choice of the term depends on the point of view on the object in question: whether the initial image is transformed into another representation, or it is simply endowed with an additional map

or field. Images are often corrupted by random variations in intensity, illumination, or have poor contrast and can't be used directly. Filtering transforms the pixel intensity values to reveal certain image characteristics. An outlier may be due to variability in the measurement or it may indicate experimental error; the latter are sometimes excluded from the data set[30]. Compactness indicates irregularity associated with cancer cells, is utitlized in this stage of nuclei validation and is defined as, Compactness=A/P^2. Where A represents the area of the object and P represents the perimeter of the object.

B. K-means clustering

Clustering is a way to separate groups of objects. K-means clustering treats each objects as having a location in space. It finds partitions such that objects within each cluster are as close to each other as possible, and as far from objects in other clusters as possible. K-means clustering requires that you specify the number of clusters to be partitioned and a distance metric to quantify how close two objects are to each other. Color-based segmentation using K-means clustering. The goal is to segment colors in an automated fashion using the L*a*b color space and K-means clustering.

Step 1: Read image

Step 2: Convert image from RGB color space to L*a*b color space

Step 3:Classify the colors in 'a*b*' space using K-means clustering

Step 4: Label every pixel in the image using the results from K-means

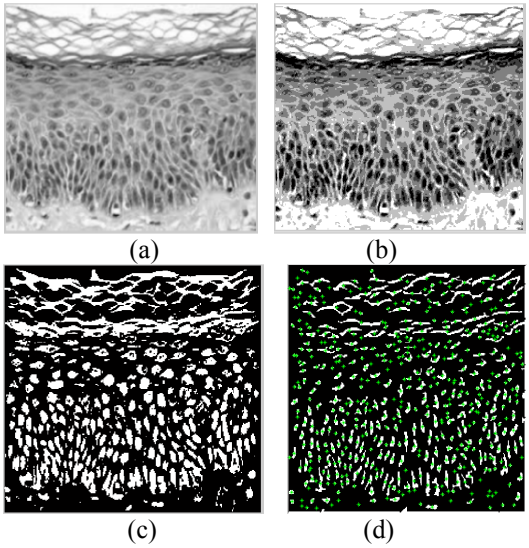
Step 5: Create images that segment the H&E image by color

Step 6: Segment the nuclei into a separate image

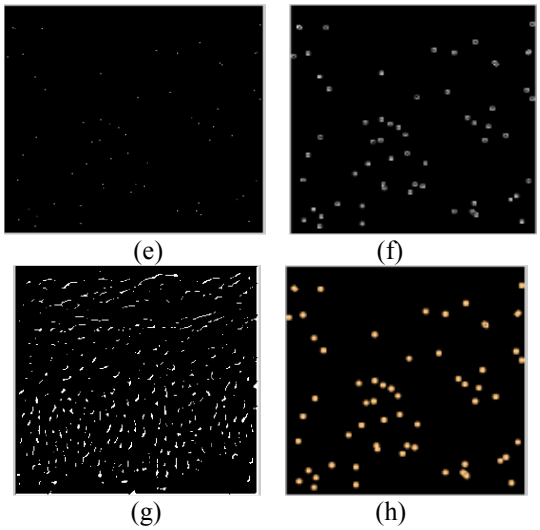
Finally, this algorithm aims at minimizing an objective function, in this case a squared error function. The objective function

$$J=\sum_{J=1}^{k}\sum_{I=1}^{n}\left\|x_i^{(j)}-c_j\right\|^2,$$

Where $\left\|x_i^{(j)}-c_j\right\|^2$ is a chosen distance measure between a data point $x_i^{(J)}$ and the cluster center c_j, is an indicator of the distance of the n data points from their respective cluster centers.



(a) (b) (c) (d)



(e) (f) (g) (h)

Figure 4: (a) Gray image. (b) K-means clustering. (c) Threshold image. (d) Blob detection. (e) Internal markers. (f) Segmented nuclei. (g) Euclidean distance map. (h) Segmented cells.

Figure 4. (a) shows the gray scale image of the skin cell. The techniques of cell segmentation is desired in the field of medical research to aid the analysis of biomedical images by delineating round objects and obtaining information about objects' size, area, or shape to locate their positions or measure useful properties. Several algorithms for cell segmentation have been published. The algorithms such as image thresholding, watershed transform and many other techniques can be used. Fig. 4. (b) K-means clustering is the another common technique to produce stable segmentation results. In Fig. 4. (c) the intensity image is converted into a binary image. It creates a black and white version of a grayscale image by specifying a single threshold value; pixels below this value become black and above this value are white. Fig. 4.(d) refers to mathematical methods that are aimed at detecting regions in a digital image that differ in properties, such as brightness or color, compared to areas surrounding those regions. In Fig. 4.(e) the groups of connected pixels inside each region where the potential nucleus are segmented. In Fig. 4.(f) the nuclei is segmented. Fig.4.(g) shows the Euclidean distance maps of the internal-marker. The Euclidean distance map represents the original binary image by labelling each pixel with the Euclidean distance between that pixel and the nearest non-zero pixel in a binary image, which generates a useful representation in the area of cell segmentation. Fig. 4. (h) shows the segmented cells, obtained after several segmentation process, which is used for calculating the NC ratio, inorder to assess the cancer cell and normal cell.

C. Cell Size and NC Ratio Evaluation

Cellular size and nuclear size are indicators not only for the developing status of some diseases but of skin aging and other quantifiable physical factors[14].

For example, cellular and nuclear size in the layers of basal cells in forearm skin has been found to increase with age. The NC ratio, which is defined as the volume ratio of nucleus to cytoplasm, is commonly used in diagnosis.

A protocol has been developed to obtain accurate NC ratios. Although the NC ratio is defined as a volume ratio, it can be approximated by an area ratio of nucleus to cytoplasm[15].

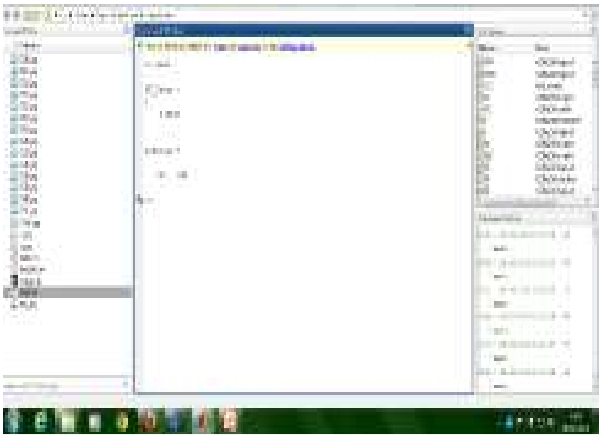


Figure 5: Nuclear-to-Cytoplasmic Ratio

Fig. 5 shows the Nuclear-to-Cytoplasmic Ratio obtained by segmenting nucleus and cytoplasm with the help of K-means clustering algorithm.

RESULTS

The Table 1 shows the profile of each segmented cell containing the cell index, NC ratio and position of the cell.
Moreover, user interaction with medical doctors or medical staff can be adopted to exclude mistakenly segmented cells further with the profiles that record the information of each segmented cell or select specific cells of interest to enhance the performance of cell segmentation and NC ratio evaluation in clinical diagnosis. So, the Table 1 gives the clear view about the NC ratio, position of the cell and type of the cell, whether it is a normal cell or carcinoma cell.

Table 1: Profile Of Each Segmented Cell

Cell Index	NC Ratio	Position (Rows, Column)	Type of Cell
1	0.25	(151 , 152)	Normal
2	0.54	(86 , 174)	Carcinoma
3	0.55	(113 , 39)	Carcinoma
4	0.27	(101 , 68)	Normal
5	0.52	(223 , 158)	Carcinoma
6	0.28	(19 , 190)	Normal
7	0.55	(229 , 69)	Carcinoma
8	0.57	(237 , 107)	Carcinoma
9	0.27	(219 , 243)	Normal
10	0.23	(141 , 3)	Normal

The Table 2 shows the comparison of mean, standard deviation, range, cell segmentation and processing time of existing and proposed method.

Table 2: Comparison with a Previous Study

Measure	Obtained by Proposed Method	Obtained by Existing Method
Range	0.25-0.30	0.32-0.38
Segmentation	Over Segmentation Avoided	Over Segmentation
Sensitivity	Not Sensitive To False Edges	Sensitivity To False Edges

The statistical tendency of estimated NC ratios in both automatic and manual approaches is robust and objective to provide enough medical information for medical doctor's clinical diagnosis. The automatic cell segmentation and NC ratio analysis on healthy volunteers and the estimated results not only locate in the acceptable range of NC ratios for healthy subjects.
But also meet the statistical tendency of NC ratios in both automatic and manual approaches with medical doctor's interpretation.

CONCLUSION

The proposed method provides objective cell segmentation results with high efficiency and consistent accuracy. The determination of NC ratios of skin cells using the proposed automatic algorithm is more objective and robust than that using manual approach, and hence medical doctors can diagnose potential diseases without the influence of any subjective factor. In addition, the evaluated NC ratio values can help medical doctors to noninvasively and immediately

identify early symptoms of diseases, especially fatal diseases like cancer, that involve abnormal NC ratios. It can be useful international cosmetics companies that asses and examine how a given product may affect the cells of potential users. This method can also be implemented for the detection of breast cancer and many other diseases related to skin.

REFERENCES

1. Gwo GL, Huan HL, Ming RT, Sin YC, Wen JL, Yi HL, Chi KS, Chun FC. Automatic Cell Segmentation and Nuclear-to-Cytoplasmic Ratio Analysis for Third Harmonic Generated Microscopy Medical Images. 2013; 7.
2. Chen SY, Chen SU, Wu HY, Lee WJ, Liao YH, Sun CK. In Vivo virtual biopsy of human skin by using noninvasive higher harmonic generation microscopy, IEEE J. Sel. Topics Quantum Electron. 2010; 16(3): 478.
3. Chen SY, Wu HY, Sun CK. Invivo harmoni generation biopsy of Human skin,J. Biomed. Opt. 2009; 14(6).
4. Cheng YH, Lin CF, Sun CK. A novel intra vital multi-Harmonic generation microscope for early diagnosis of oral cancer at the Optical BiopsyConf. Photonics West,San Francisco, CA, USA, paper.2013.
5. Gonzalez RC, Woods RE. Digital Image Processing. Englewood Cliffs, NJ, USA: Prentice-Hall. 2002.
6. Liao YH, Chen SY, Chou SY, Wang PH, Tsai MR, Sun CK. Invivo harmonic generation biopsy for quantitative evaluation in chronological aged skin keratinocytes, Biomed. Opt. Expr. 2013; 4(1): 77.
7. Mac EAM. Compactness of geographic shape: Comparison and evaluation of measures, Geografiska Annaler. Human Geography. 1985.
8. MaurerC, QR, Raghavan V. A linear time algorithm for computing exact Euclidean distance transforms of binary images in arbitrary dimensions,IEEE Trans. Pattern Anal. Mach. Intell. 2003; 25.
9. Meyer F. Topographic distance and watershed lines, Signal Process. 1994; 38: 113.
10. Oram Y, Turhan O, Din EA. Diagnostic value of cytology in basal cell And Squamous Cell Carcinomas,Int. J. Dermatol. 1997; 36(2): 156.
11. Paglieroni D. Distance transforms: Properties and machine vision applications, Comput. Vis. Graph. Image Process.: Graph. Models Image Process. 1992; 54.
12. Rosenfeld A, PfaltzJ. Sequential operations in digital picture processing, J. Assoc. Comput. Mach. 1996; 13.
13. SoilleP. Morphological Image Analysis: Principles and Applications. 2003.
14. Sun CK. Higher harmonic generation microscopy,Adv. BiochemEng/Biotechnol. 2006; 95L: 17.
15. Tsai MR, Chen SY, ShiehDB, Lou PJ, Sun CK. Invivooptical Virtual biopsy of human oral mucosa with harmonic generation microscopy,Biomed Opt. Expr. 2011; 2(8): 2317.
16. Tsai MR, Liao YH, Sun CK. Differential diagnosis of Pigmented skin lesions based on harmonic generation microscopy, presented at the Photonics in Dermatology and Plastic Surgery Conf. Photonics West, San Francisco, CA, USA. 2013.
17. Tsai MR, Shieh DB, Lou PJ, Lin CF, Sun C. Characterization of oral Squamous Cell Carcinoma based on higher-harmonic generation microscopy, J. Biophoton. 415.
18. TsangTYF. Optical third-harmonic generation at interfaces, Phys. Rev. 1995;52(5): 4116.
19. Vega ME, Martinez D, Larios N, Waxtein LM, Dominguez SL. Cyto diagnosis of Cutaneous Basal and Squamous Cell Carcinoma," Int. Dermatol. 2000; 39(2): 116.
20. Vincent L, Soille P. Watersheds in digital spaces: An efficient Algorithm Based on immersion simulations, IEEE Trans. Pattern Anal.Machine Intell. 1991; 13: 583.
21. Mat IN, Mashor MY, Othman NH, Sulaiman SN. Application of Moving K-Means Clustering for Pap Smear Image Processing, Proceeding of International Conference on Robotics, Vision, Information and Signal Processing, Penang Malaysia. 2002.
22. Yang X, Weidong Z, Yufei C, Xin F. Image Segmentation With A Fuzzy Clustering Algorithm Based On Ant-Tree,Signal Processing. 2008; 88(10): 2453.
23. MatIsa NA, Samy AS, Ngah UK. Adaptive fuzzy moving Kmeans algorithm for image segmentation, IEEE Transactions on Consumer Electronics. 2009; 55(4): 2145.
24. Boudraa AEO, Mallet JJ, Besson JE, Bouyoucef SE, Champier J. Left Ventricle Automated Detection Method In Gated Isotopic Ventriculography Using Fuzzy Clustering, IEEE Transactions on Medical Imaging. 1993; 12(3): 451.
25. Chen CW, Luo J, Parker KJ. Image Segmentation via Adaptive K-Mean Clustering And Knowledge-Based Morphological Operations With Biomedical Applications, IEEE Transactions on Image Processing. 1998; 7(12): 1673.
26. Mat Isa NA, Mashor MY, Othman M. Comparison of Segmentation Performance of Clustering Algorithms for Pap Smear Images, Proceeding of International Conference on Robotics, Vision, Information and Signal Processing. Penang. 2003.
27. Rezaee MR, Vander Z, Lelieveldt BPE, Vander GRJ, Reiber JHC. A Multiresolution Image Segmentation Technique Based On Pyramidal Segmentation And Fuzzy Clustering, IEEE Transactions on Image Processing. 2000; 9(7): 1238.
28. Mitianoudis Q, Li N, Stathaki T. Spatial kernel K-harmonic means clustering for multi-spectral image segmentation, Image Processing, IET. 2007; 1(2): 156.

29. Ordonez C. Clustering Binary Data Streams with K-means, Proceedings of the 8th ACM SIGMOD workshop on Research issues in data mining and knowledge discovery, ACM: San Diego, California. 2003.
30. Agarwal PK, Mustafa NH. K-means Projective Clustering,Proceedings of the twenty-third ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems. ACM: Paris, France. 2004.

Source of support: Nil, Conflict of interest: None Declared