Applied Soft Computing 13 (2013) 1832–1852

Contents lists available at SciVerse ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

Novel initialization scheme for Fuzzy C-Means algorithm on color image segmentation

Khang Siang Tan, Wei Hong Lim, Nor Ashidi Mat Isa *

Imaging and Intelligent Systems Research Team (ISRT), School of Electrical and Electronic Engineering, Engineering Campus, Universiti Sains Malaysia, 14300 Nibong Tebal, Penang, Malaysia

ARTICLE INFO

Article history:
Received 14 June 2012
Received in revised form 7 December 2012
Accepted 30 December 2012
Available online 18 January 2013

Keywords:
Fuzzy C-Means (FCM)
Hierarchical Approach (HA)
Initialization scheme
Splitting and merging

ABSTRACT

This paper presents a novel initialization scheme to determine the cluster number and obtain the initial cluster centers for Fuzzy C-Means (FCM) algorithm to segment any kind of color images, captured using different consumer electronic products or machine vision systems. The proposed initialization scheme, called Hierarchical Approach (HA), integrates the splitting and merging techniques to obtain the initialization condition for FCM algorithm. Initially, the splitting technique is applied to split the color image into multiple homogeneous regions. Then, the merging technique is employed to obtain the reasonable cluster number for any kind of input images. In addition, the initial cluster centers for FCM algorithm are also obtained. Experimental results demonstrate the proposed HA initialization scheme substantially outperforms other state-of-the-art initialization schemes by obtaining better initialization condition for FCM algorithm.

1. Introduction

In digital image processing, image segmentation is a process of partitioning an image into non-overlapped, consistent regions which are homogeneous with respect to some characteristics [1]. It serves as a critical and essential component of image analysis and pattern recognition system because it determines the quality of the final result of analysis [2]. Thus, it has been widely used in many image analyses and pattern recognition applications such as object recognition [3–5], optical character recognition [6,7], face recognition [8], fingerprint recognition [9,10], medical image processing [11,12], industrial automation [13] and content based image retrieval [14]. Due to its clustering validity and simplicity of implementation, FCM algorithm has long been a popular image segmentation algorithm [15].

The FCM algorithm was introduced by Ruspini [16] and then improved by Dunn [17] and Bezdek [18]. It is a segmentation algorithm that is based on the idea of finding cluster centers by iteratively adjusting their position and evaluation of an objective function. The iterative optimization of the FCM algorithm is essentially a local searching method, which is used to minimize the distance among the image pixels in corresponding clusters and maximize the distance between cluster centers. Its success mainly attributes to the introduction of partial memberships for the belongingness of each image pixel to all available clusters identified by its centers. The partial membership is proportional to the probability that a pixel belongs to a specific cluster where the probability is only dependent on the distance between the pixel and each cluster center. By iteratively adjusting the cluster centers, the objective function of the FCM algorithm can reach the global minimum when pixels nearly the center of corresponding cluster are assigned to higher membership while those far from the center of corresponding cluster are assigned lower membership.

However, the FCM algorithm is very sensitive to the initialization condition of cluster number and initial cluster centers [19]. The initialization conditions of cluster number and initial cluster centers have their impacts on the segmentation quality. For instance, the cluster number imposes significant impact on segment area whereas the initial cluster centers affect the classification accuracy of the FCM algorithm as different selection of initial cluster centers can potentially lead to different local optimal or different partition. In general, to segment an image into F clusters, there are three most popular initialization schemes as reported by Bezdek et al. [20], which can be described as follow:

1. Using F image pixels randomly selected from the image.
2. Using the first F distinct image pixels in the image.
3. Using F image pixels uniformly distributed across the image.

Among these initialization schemes, using F image pixels randomly selected from the image as the initial cluster centers has

* Corresponding author. Tel.: +60 45996093; fax: +60 45941023.
E-mail addresses: khangsiang85@hotmail.com (K.S. Tan), limweihong87@yahoo.com (W.H. Lim), ashidi@eng.usm.my (N.A.M. Isa).

1568-4946/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.asoc.2012.12.022

been proven to be the best initialization scheme [21]. This initialization scheme is well known as randomly initialization scheme. Wang and Lu [22] proposed a new FCM variant with enhanced algorithm's speed and noise immunity. Accordingly, the algorithms initial cluster centroids could be extracted from the peak values of images' gray histogram and this approach could reduce number of iteration required by their algorithm to achieve the satisfactory clustering results. To increase the algorithm's robustness toward the noise signals (e.g. Gaussian noise, salt and pepper noise and mixed noise), the local spatial information of images are incorporated into their FCM variants. More specifically, a new objective function is proposed to take both the neighbor mean and median values into account during the clustering of center pixels. According to Guo et al. [23], neutrosophy [24] is a new powerful tool to describe the image with uncertainty. Thus, they adopt the neutrosophic set (NS) approach into the FCM technique to perform the image segmentation. In their proposed approach, the input images are first transformed into the NS domain, which is described by three membership sets of white set (T), indeterminate set (I), and non-white set (F). A α-mean operation is performed on the NS images prior the FCM clustering process, in order to increase the uniformity and homogeneity of NS images by reducing the NS's indeterminacy. Motivated by the dependency of color clustering results on the color model and similarity functions used, Patrascu [25] proposed a FCM variant that works based on a new measure of similarity. Accordingly, the new measure is defined in a new perceptual system of *hsl* and the clustering results based on this new measure are promising. Despite achieving different improvements over the original FCM technique, the prior determination of cluster numbers in these techniques is set manually by users. Nevertheless, the process of determining the cluster numbers is laborious, especially for natural color images due to their complexity and diversity. Additionally, it is impractical to expect all users to have sufficient domain knowledge in determining the accurate cluster numbers. As the initialization scheme has substantial impact on the FCM's clustering performance, wrong determination of cluster numbers by users could deliver poor segmentation results.

To alleviate the abovementioned drawback, an initialization scheme called Agglomerated Just Noticeable Difference Histogram (AJNDH) was proposed to automatically determine the cluster number and initial cluster centers with the aim to initialize the FCM algorithm [26]. In this initialization scheme, the Just Noticeable Difference (JND) histogram is first constructed to obtain enough number of histogram bins without compromising the visual image content before applying the agglomeration to obtain the initialization condition for FCM algorithm. With similar cluster number, this initialization scheme was proven to be able to outperform the randomly initialization scheme by obtaining better initial cluster centers and hence producing better segmentation result.

In addition, Ant Colony Optimization (ACO) initialization scheme was also proposed to automatically obtain the initialization condition of cluster number and initial cluster centers to initialize the FCM algorithm [27]. In this initialization scheme, an improved Ant System (AS) includes a cluster merging step to automatically keep a reasonable cluster number for all kinds of input images. In addition, the improved AS is also applied for intelligent initialization of cluster centers. This initialization scheme was also proven to be able to outperform the randomly initialization scheme by obtaining better initial cluster centers and hence producing better segmentation result, both with similar cluster number.

Chen et al. [28] proposed a FCM-based segmentation technique by performing fusion on multi-color space components. Accordingly, the input images are first transformed into the color spaces of grayscale, HSV, YIQ, YCvCr, LAB and LUV, respectively. A peak-finding algorithm is then applied on the components of gray, V, I, Cr, B, and U, from the corresponding color space to determine their respective initial cluster centroids and cluster numbers. The spatial FCM (SFCM) algorithm [29] is applied on these six selected components with different cluster numbers to generate six different initial segmentation results. A fusion process, implemented by the SFCM algorithm, is applied again on these six initial segmentation results to obtain the final cluster number. Bahght et al. [30] proposed a new validity index to access the validity of a cluster. In their approach, a multi-degree entropy algorithm is proposed to perform partition on the input image into different level of intensities using the multi-degree immersion process. Based on the predefined validity function criteria, a merging process is applied on the output of aforementioned process to obtain the image's final cluster numbers. Meanwhile, Sowmya and Sheela Rani [31] investigate the capability of FCM, Possibilistic FCM (PFCM) [32], and competitive neural network [33] in performing the image segmentation. In their studies, a self-estimation algorithm proposed in [34] is adopted to automatically determine the cluster numbers.

In this paper, we propose a novel initialization scheme called Hierarchical Approach (HA) to automatically determine the cluster number and to obtain the initial cluster centers, which are used to initialize the FCM algorithm. In the proposed initialization scheme, the color image is split hierarchically into multiple homogeneous regions before merging is carried out to obtain the initialization condition for the FCM algorithm.

The rest of the paper is organized as follows: Section 2 presents the working of FCM algorithm. Section 3 will describe the proposed initialization scheme. Section 4 will analyze the segmentation results produced by the FCM algorithm using the proposed initialization scheme and at the same time comparing it with those using other initialization schemes. Finally, Section 5 concludes the work of this paper.

2. Fuzzy C-Means algorithm

FCM algorithm is an unsupervised classification technique, thus there is no need for prior knowledge about the pixels set. It is a segmentation algorithm that is based on clustering similar pixels in an iterative way where the cluster centers are adjusted during the iteration. It attempts to partition the pixels into a collection of F fuzzy clusters. Based on the minimization of the objective function, the conventional FCM algorithm yields extremely good segmentation results. Typically, the objective function of the conventional FCM algorithm is defined as:

$$W^{(q)} = \sum_{j=1}^{F}\sum_{i=1}^{N} u_{ij}^{m} ||x_i - c_j||^2 \tag{1}$$

where N is the number of image pixels, μ_{ij} is the membership of pixel x_i to the jth cluster identified by its center c_j, and m is a constant that defines the fuzziness of the resulting partition. $||x_i - c_j||$ denotes the Euclidean distance between x_i and c_j. The parameter m controls the fuzziness of the membership. The value of m is manually determined by the user. In general, most users choose m in the range [1.5, 2.5], with $m = 2.0$ being an overwhelming favorite. The membership of pixel x_i to the jth cluster identified by its center c_j is defined as:

$$u_{ij} = \sum_{k=1}^{F} \left(\frac{||x_i - c_j||}{||x_i - c_k||} \right)^{-2/m-1} \tag{2}$$

where μ_{ij} indicates the strength of the association between x_i and c_j and has the value in the range [0, 1]. In FCM algorithm, the cluster centers are iteratively updated as:

$$c_j = \frac{\sum_{i=1}^{N} u_{ij}^m x_i}{\sum_{i=1}^{N} u_{ij}^m} \quad (3)$$

The procedure of FCM algorithm is illustrated in Fig. 1 as follow:

(*Notes*: ε is a constant and its value is manually determined by the user. In general, most users choose ε in the range of [0.01, 0.0001], with $\varepsilon = 0.001$ being an overwhelming favorite.)

3. Proposed Hierarchical Approach

Earlier studies reveal that FCM clustering results are highly dependent on the initializations of the cluster centers and number of clusters. A good initialization conditions can only be obtained by running repetitive experiments based on certain experiences, if there is no initialization scheme that can automatically determine both the cluster numbers and centers. Nevertheless, this approach is laborious and the clustering results obtained are not always promising.

In this study, we proposed the HA module as an initialization scheme for the FCM technique. In other words, the HA module allows users to automatically and adaptively determine both the cluster numbers and centroids, which serves as the initialization conditions for the FCM technique. Compared with the widely used random initialization scheme, the HA module requires a less laborious process and consistently produces good initialization conditions for the FCM technique. The capability of the HA module to determine cluster numbers and centroids automatically and adaptively is due to ability of the HA module to detect them based on the global information in the histogram of the input images. Different types of input images have different types of global information. Therefore, the HA module eventually detects different cluster numbers and centroids, depending on the numbers and intensity values of the significant peaks that are present in the histograms of the input images.

The proposed Hierarchical Approach (HA) initialization scheme consists of two modules namely the splitting module and the merging module. The splitting module is employed to split the color image into multiple small homogeneous regions whereas the merging module is applied to merge those regions that are perceptually close to each other to determine cluster number and to obtain the initial cluster centers for FCM algorithm. Each of the modules is presented in the following sections respectively.

3.1. Splitting module

Previous literatures [35–37] revealed that both of the color space and the similarity/dissimilarity function used to measure the color separation/differences play significant roles in the image analysis's results. Although the RGB color space is widely used to represent the image data in image processing research, it is not a perceptually uniform color space and thus incapable to model the way in which humans perceive color. More specifically, the color differences obtained (i.e. Euclidean distance) from the three dimensional (3-D) RGB color space do not correspond to the human perception of such differences [38,39]. This inhibits the application of RGB color space in various image processing tasks.

Being an approximately uniform color space, HSI achieves better approximation of human perception on the color differences, as compared to the RGB color space [39]. In color images, HSI color space separates the color information (represented by Hue and Saturation) from its intensity information. In HSI, Hue represents basic colors and is determined by the dominant wavelength in the spectral distribution of light wavelengths. Saturation is a measure of the purity of the color and signifies the amount of white light mixed with Hue. Intensity is a measure of the image brightness and is determined by the amount of light. In general, there are two popular variants of HSI color space, namely the HSL and HSV color spaces. The HSI color space has been used extensively for image processing [40–42] due to their intuitive appeal to human's perception and their provision for isolating the luminance component [41]. Thorough experimental studies have been presented in [39] to measure the impact of the use of different color spaces on the performance of color texture analysis methods such as segmentation or classification. It is reported that the approximately uniform space of HSI outperforms the perceptually non-uniform RGB space in both the noise-free and noisy conditions, suggesting that the HSI could be a superior color space compared to RGB for image analysis [39].

Despite having significant advantages over the RGB space, the conventional HSI space encounters some undesirable issue that

FCM Algorithm

Initialize the fuzzy cluster number, F and the cluster centers, $c = \{c_1,...,c_j,...c_F\}$.
Set iteration time $q = 0$;
while $\left\|W^{(q)} - W^{(q-1)}\right\| \geq \varepsilon$ **do**
  for $j = 1$ to F^{th} cluster **do**
    for $i = 1$ to N^{th} image pixel **do**
      Calculate μ_{ij}, i.e. the membership of pixel x_i to the j^{th} cluster (represented by c_j), using Equation (2);
      if $\left\|x_i - c_j\right\| = 0$ **then**
        $\mu_{ij} = 1$;
        reset other membership of pixel i to 0;
      end if
    end for
    Update c_j according to Equation (3);
  end for
  Calculate objective function $W^{(q)}$ using Equation (1);
  $q = q + 1$;
end while

Fig. 1. Implementation of the FCM algorithm.

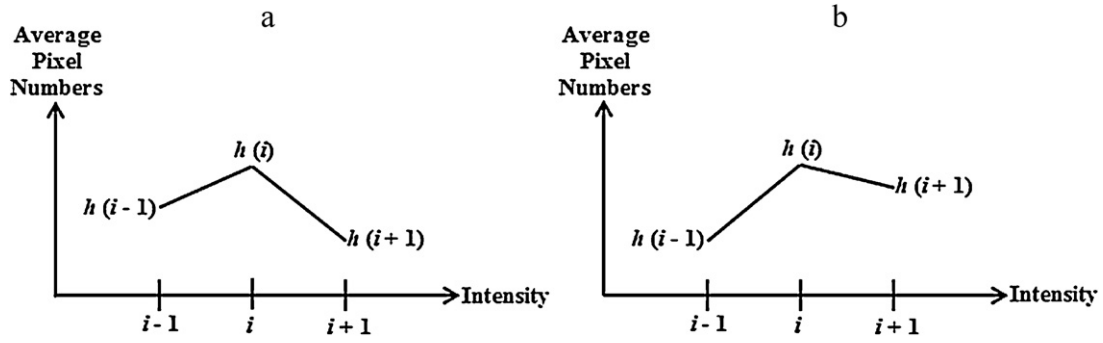

Fig. 2. Possible scenario for intensity level i to be identified as peak by Eq. (8), (a) $h(i-1) > h(i+1)$ and (b) $h(i-1) < h(i+1)$.

limits its application in image processing task. Recall that the Saturation in both the HSL and HSV spaces represents percentages of the maximum saturation obtainable for a given lightness [43]. In other words, for the case where Saturation range is small, large variation of Intensity ranges could be observed and this leads to the occurrence of some undesirable artifacts (e.g. noise) in the Saturation image [43]. To encounter the aforementioned drawback, Hanbury [43] proposes to remove the saturation normalization by lightness in HSI space, thereby derives a unified model of cylindrical coordinate HSL color space, where the Hue, Saturation and Lightness are defined respectively as [43]:

$$H = \arctan\left(\frac{\sqrt{3}(G-B)}{2R-G-B}\right) \tag{4}$$

$$S = \max(R,G,B) - \min(R,G,B) \tag{5}$$

$$L = \frac{R+G+B}{3} \tag{6}$$

In this study, we adopt the new HSL color space proposed by Hanbury [43] to hierarchically split the color image, backed by its good approximation on the human's color perception (compared to RGB space) and its achievement in overcoming the drawback of conventional HSI space.

In the proposed HA initialization scheme, the splitting module consists of 3 phases. In each phase, a peak finding algorithm, namely the Peak Finding Histogram Analysis (PFHA) algorithm, is proposed to obtain the modes in the histogram before the valleys can be successfully identified. For the histogram analysis, the histogram of gray scale image can be separated into several numbers of modes, each corresponds to one region, and there is a threshold value corresponding to the valley between two adjacent modes [44]. The valley between two adjacent modes can be used as the boundary for gray scale image segmentation. In the PFHA algorithm, at the beginning, a moving average filter is applied to remove the unnecessary peaks and valleys in the histogram. It is done by taking the average number of pixels among the $2L+1$ (where L is a positive integer number) adjacent discrete levels in the histogram as:

$$h(i) = \frac{[n(i-L) + n(i-(L-1)) + \ldots + n(i) + \ldots + n(i+(L-1)) + n(i+L)]}{2L+1} \tag{7}$$

where $n(i)$ is the number of pixels associated with ith level in the histogram and $h(i)$ is the value of ith level after applying the moving average.

(*Note*: In this study, the span of the moving average filter can be set from 2 to 8, based on the analysis done using numerous images. However, the span of the moving filter that is smaller than 2 is not capable to eliminate the non-significant peaks reside in the histogram of each color channel respectively. Meanwhile, the span of the moving average filter that is larger than 8 could potentially remove certain significant peaks reside in the histogram of each color channel respectively, which is not desirable. Thus, we set the span of the moving average filter as 5, by taking the average value between these two extreme values.)

Despite the smoothing process performed by the moving average filter, there are still some insignificant peaks and valleys detected in the new histogram and they are identified by Eqs. (8) and (9), respectively. The possible scenarios for a given intensity level i to be identified as the peak and valley are illustrated in Figs. 2 and 3 respectively.

$$Peak = (i, h(i) | h(i) > h(i-1) \& h(i) > h(i+1)) \tag{8}$$

$$Valley = (i, h(i) | h(i) < h(i-1) \& h(i) < h(i+1)) \tag{9}$$

To remove all these insignificant detected peaks and valleys, we propose a set of IF-THEN rule base, represented by Eqs. 10(a)–10(d). Generally speaking, if a particular intensity level of i is identified as peak (i.e. Fig. 2), we first examine the average pixel numbers of its two nearest neighbor, i.e. $h(i-1)$ and $h(i+1)$. To remove the insignificant peak that contributed by i, we replace the average pixel numbers of i, i.e. $h(i)$, with $h(i-1)$ or $h(i+1)$, whichever has the larger value. For intensity level of i that identified as valley (i.e. Fig. 3), we replace $h(i)$ with $h(i-1)$ or $h(i+1)$, whichever has

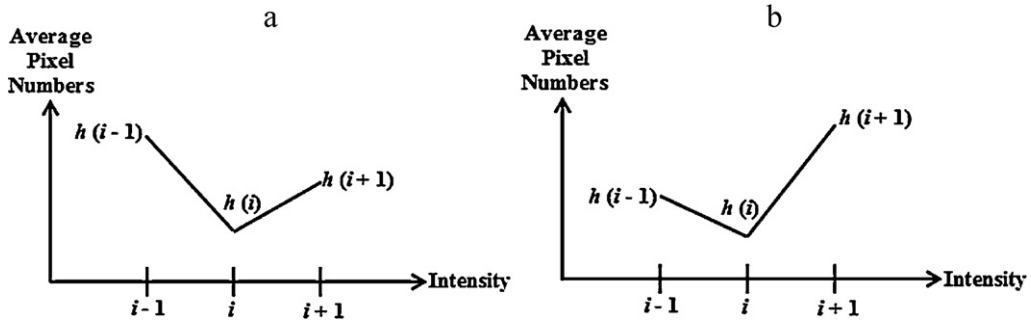

Fig. 3. Possible scenario for intensity level i to be identified as valley by Eq. (9), (a) $h(i-1) > h(i+1)$ and (b) $h(i-1) < h(i+1)$.

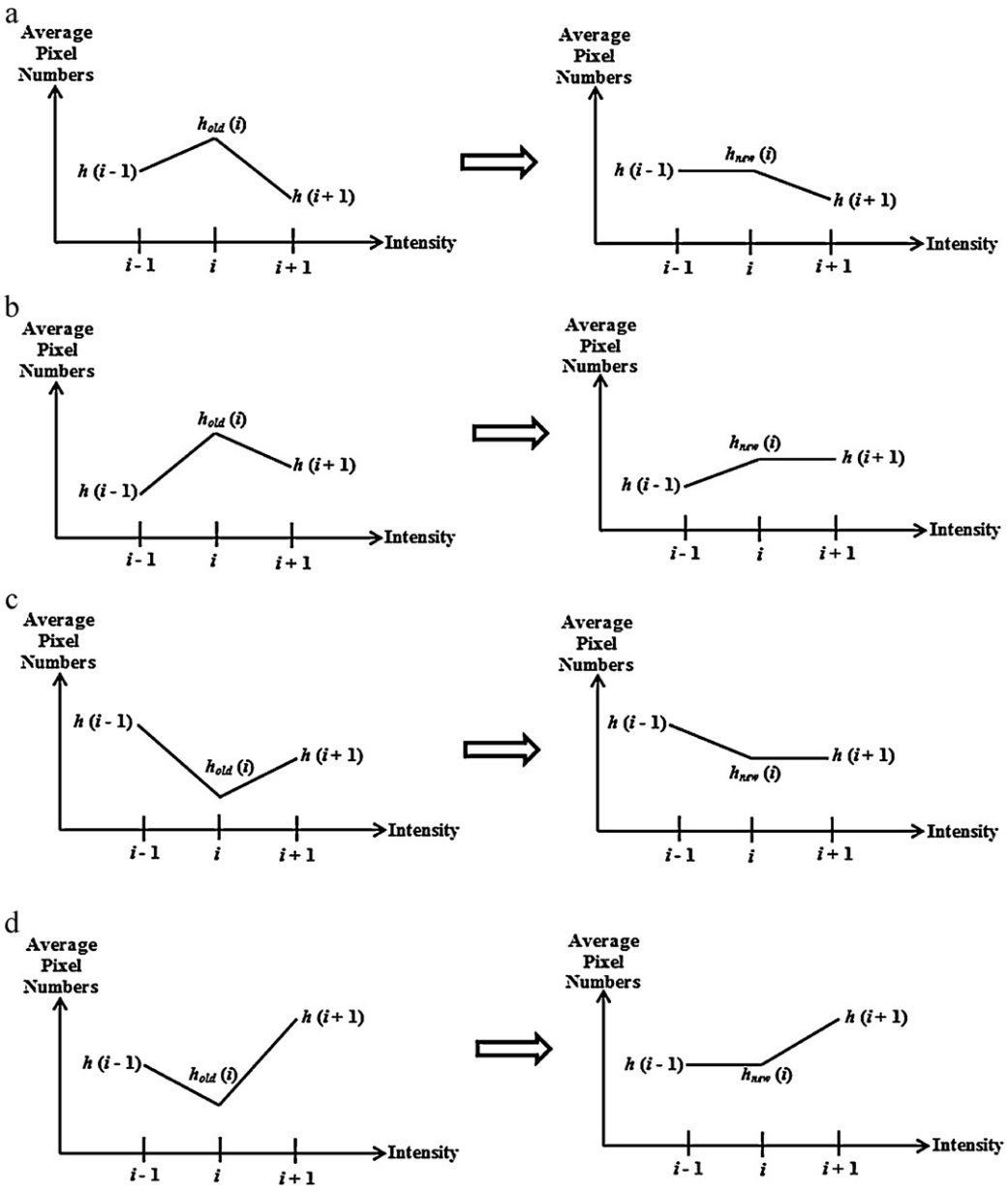

Fig. 4. Illustration of the removal of insignificant peak and valley by (a) Eq. (10a), (b) Eq. (10b), (c) Eq. (10c), and (d) Eq. (10d), respectively.

the smaller value, in order to remove the insignificant valley that contributed by i. Fig. 4 illustrates the mechanisms of Eq. (10) in removing the insignificant peak and valley that contributed by the intensity level i.

$$IF(i \in Peak) AND (h(i+1) > h(i-1)) \quad THEN(h(i) = h(i+1)) \tag{10a}$$

$$IF(i \in Peak) AND (h(i+1) < h(i-1)) \quad THEN(h(i) = h(i-1)) \tag{10b}$$

$$IF(i \in Valley) AND (h(i+1) < h(i-1)) \quad THEN(h(i) = h(i+1)) \tag{10c}$$

$$IF(i \in Valley) AND (h(i+1) > h(i-1)) \quad THEN(h(i) = h(i-1)) \tag{10d}$$

As a result, a new smoothed histogram, without any insignificant peaks and valleys, could be obtained. In order to obtain the modes in the new histogram, we first examine the gradient change, *grad* (i) exhibited by each intensity level i in the new histogram. Mathematically, *grad* (i) represents the variation of h (i) to h ($i+1$) with respect to the intensity level i and it is represented as follows:

$$grad(i) = h(i) - h(i-1) \tag{11}$$

By successively examining the *grad* (i) values of each intensity level i, we could identify the intensity level i as mode in the new histogram, if the conditions of h (i) > h ($i-1$) and h ($i+1$) < h (i) are fulfilled. In other words, the intensity level i in the new histogram is assigned as the modes, if the *grad* (i) shows positive value (i.e. h (i) > h ($i-1$)), whilst the *grad* ($i+1$) shows negative value (i.e. h ($i+1$) < h (i)). Mathematically, the modes identification mechanism is represented as follows:

$$Mode = (i, grad(i) \,|\, grad(i) \geq 0 \;\&\; grad(i+1) < 0) \tag{12}$$

After the modes in the new histogram are detected, the valleys can be obtained by taking the minimum value between any two adjacent modes in the new histogram. They are used as the

PFHA Algorithm

1: Apply a moving average filter on the input histogram to remove the unnecessary peaks and valleys using Equation (7) ;
2: Identify insignificant peaks and valley using Equations (8) and (9) respectively ;
3: Remove insignificant peaks and valley in new histogram using Equations (10) ;
4: Identify modes in new smoothed histogram using Equations (11) and (12) ;
5: Identify valleys in new smoothed histogram by taking minimum value between any adjacent modes in the histogram ;

Fig. 5. Implementation of the PFHA algorithm.

boundaries for segmentation in this paper. Generally, the implementation of the PFHA algorithm is illustrated in Fig. 5 as follows:

In the first phase of the proposed HA initialization scheme (i.e. the splitting module), 3 steps are involved. First, Hue histogram for all the image pixels is constructed as Hue can be easily distinguished, while the perception of different Saturation and Lightness does not imply the recognition of different colors in the human vision system. Then, the proposed PFHA algorithm is used to locate the modes in the histogram. Finally, the valleys can be obtained after the modes are located and they are used as the boundaries for segmentation. Although each region obtained is homogeneous with respect to Hue, the pixels of each region may consist of different Saturation as Hue is invariant to certain types of highlights, shading and shadows. Thus, for all the regions obtained in the first phase of splitting module, 3 steps involved in the first phase are again used in Saturation histogram in the second phase of splitting module. Thus, the regions are further split into multiple homogenous regions with respect to both Hue and Saturation. However, these resultant regions play little role in distinguish colors if the Lightness of the color lies close to black or white. Thus, for all the regions obtained in the second phase of splitting module, 3 steps involved in the first phase of splitting module are again used in Lightness histogram in the third phase of splitting module. Thus, the regions are further split into multiple homogenous regions with respect to Hue, Saturation and Lightness. The resultant regions obtained after the third phase of splitting module are represented by its respective cluster center. The cluster centers are obtained as:

$$c_j = \frac{1}{|R_j|} \sum_{x_i \in R_j} x_i \tag{13}$$

where R_j is the pixel set that are assigned to jth cluster center, $|R_j|$ is the number of pixels assigned to jth cluster.

It is worth mentioning that, our main purpose to employ the PFHA algorithm on Hue, Saturation, and Lightness's histograms is to identify the significant peaks and valleys on these histograms, thereby obtain the boundaries for segmentation. It is evident that output images produced by the PFHA algorithm consist of smoothed histograms of Hue, Saturation, and Lightness, due to the averaging process (i.e. Eq. (7)), as well as the elimination of insignificant peaks and valley (i.e. Eqs. (8)–(10)). However, reader should be aware that such smoothed images will not be employed for the further processing. To prevent the blurring effect on the input images, we apply both of the significant peaks and valleys, obtained from the PFHA algorithm, on the original images for the subsequent steps.

3.2. Merging module

The number of cluster formed by the homogeneous regions is usually large and is not suitable for post-processing, especially in image segmentation by the FCM algorithm. To obtain reasonable cluster number to initialize the FCM algorithm, a merging technique is applied to merge the clusters that are perceptually close to each other.

In the merging technique, there are 2 steps involved. First, the Manhattan distance between any two cluster centers, $D_{shortest}$ is computed. The employment of the Manhattan distance to measure the color similarity between any two cluster centers is motivated by the experimental findings reported by Loo and Tan [45]. Accordingly, Manhattan distance is a better distance measurement than the Euclidean distance as the former exhibit a more stable visual color similarity, whilst the latter tends to produce wider variation of color perception with the same color distance [45].

Table 1
Categories of color similarity in terms of the Manhattan distance [45].

dc	Visual inspection result
10–30	Same color
31–70	Same color, low intensity variance
71–90	Same color series
91–120	Same color series, low intensity variance
121–150	Different colors, small color range
151–190	Different colors, wider color range
Above 190	Very randomly occurring color

The two nearest clusters are then merged and the new cluster centers are updated by Eq. (13) if their Manhattan distance, is less than predefined threshold, *dc*. As shown in Table 1, Loo and Tan [45] revealed that the Manhattan distance is below 70, where the same color is observed with a very low intensity variance. Thus, we set *dc* to be 70 to merge the perceptually close cluster centers.

In the merging technique, these steps are repeated until no minimum Manhattan distance between two nearest cluster centers is less than the predefined threshold *dc*. As a result, the cluster number and the initial cluster centers could be obtained for the FCM algorithm. The implementation of the merging module is presented in Fig. 6.

3.3. Illustration of the implementation procedure

In order to provide the clear understanding of the fundamental concept of the proposed HA initialization scheme, we present the implementation illustration of the proposed HA initialization scheme in detail in this section. The test image *House* as shown in Fig. 7(a) is adopted for the implementation illustration of the proposed HA initialization scheme.

As explained in the previous section, the HA initialization scheme consists of three phases to hierarchically split the test image *House* into multiple homogenous regions in the HSL color space. In the first phase of the HA initialization scheme, the Hue histogram of the test image *House* will be computed, and then followed by the application of the PFHA algorithm to locate the modes in the histogram. The valleys between the two adjacent modes in the Hue histogram will be used as the segmentation boundaries in the first phase of the HA initialization scheme. The segmented result after the first phase of the HA initialization scheme, which consists of 15 types of clusters is shown in Fig. 7(b). In this figure, it is clearly seen that the edge of the house roof has been mistakenly assigned to be part of the sky due to the reason that Hue is invariant to highlights, shadings, and shadows.

In order to form the regions that are variant to highlights, shadings, and shadows, the Saturation histogram for every regions obtained in the first phase of the HA initialization scheme is computed. In this second phase of the HA initialization scheme, the PFHA algorithm is again applied to obtain the modes and valleys of the Saturation histogram. The segmented result after the second phase of the HA initialization scheme, which consists of 44 types of clusters is shown in Fig. 7(c). It can clearly be observed that, unlike the segmented results produced during the first phase of the HA initialization scheme, the regions obtained in the second phase are distinguishable to highlights, shadings, and shadows.

Despite the fact that the segmented results produced in the second phase are distinguishable to highlights, shadings, and shadows,

Merging Module

1: Calculate all the M cluster centers for all the regions obtained after the third phase of Splitting Module using Equation (13) ;
2: **while** $D_{shortest} < dc$ **do**
3: Calculate the Manhattan distance, D between any two out of the M cluster centers ;
4: Identify the shortest distance between two nearest cluster centers, $D_{shortest}$;
5: Merge the two nearest clusters ;
6: Refresh the overall cluster centers using Equation (13) ;
7: Reduce the cluster number M by one ;
8: **end while**

Fig. 6. Implementation of the merging module.

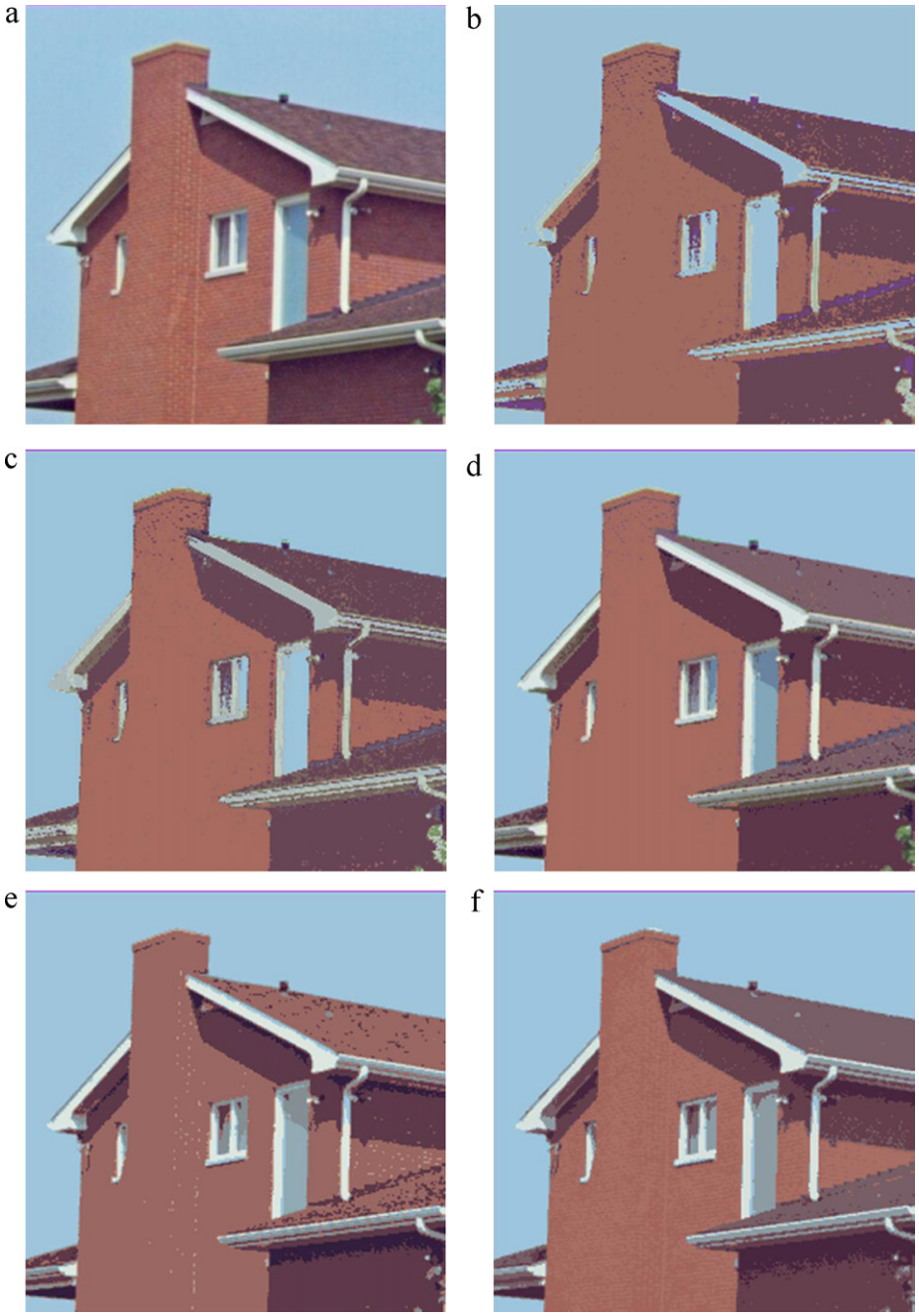

Fig. 7. Illustration of segmentation result using the proposed HA initialization scheme: (a) test image *House*, (b) segmented result after first phase, (c) segmented result after second phase, (d) segmented result after third phase, (e) segmented result after merging process and (f) final segmentation result.

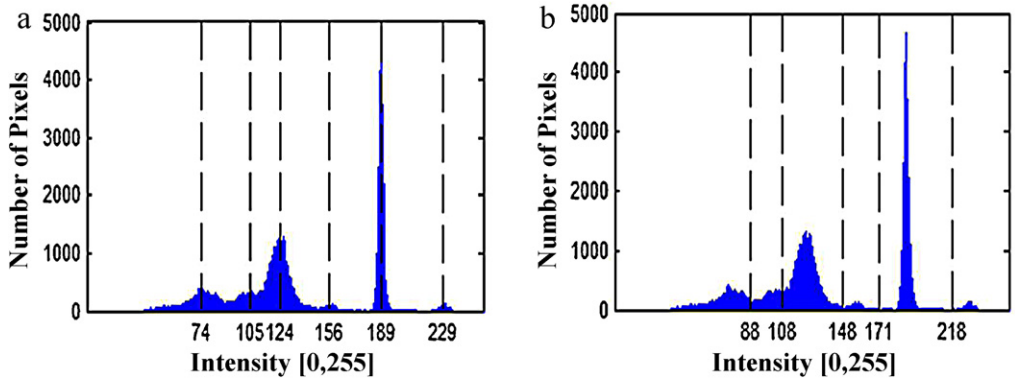

Fig. 8. The peaks and valleys located on the histogram for image *House*, (a) peaks located on its histogram and (b) valleys located on its histogram.

yet they still play insignificant role in distinguishing regions if the intensity of the regions lie close to white or black. Thus, in the third phase of the HA initialization scheme, the Lightness histogram for every regions obtained in the second phase of the HA initialization scheme are constructed. Similar to the first and second phases, the PFHA algorithm is applied again to locate the modes in the Lightness histograms, followed by the identification of the valleys in the corresponding histograms. The segmented result after the third phase of the HA initialization scheme consists of 101 types of clusters as shown in Fig. 7(d).

In order to obtain the reasonable number of cluster used to initialize the classical FCM algorithms, a merging process is employed to merge the similar regions. The segmented result after the merging process consists of 9 types of clusters as shown in Fig. 7(e). The number of cluster and their corresponding centers obtained after the merging process are used as the initialization condition for the conventional FCM algorithm. Finally, the conventional FCM algorithm is applied to perform color image segmentation and the final segmented result is illustrated in Fig. 7(f).

4. Results and discussion

This section begins with the investigation of the proposed PFHA algorithm's capability in detecting the histograms' modes. As explained in the previous section, it can be observed that proposed PFHA algorithm has played a vital role in locating the modes (representing the homogenous regions) of the Hue, Saturation, and Lightness histograms. The performance of the proposed PFHA algorithm has huge impact on the clustering quality as the final cluster number produced is greatly dependent on the number of significant modes identified in the early stage. Thus, in the first part of this section, we are particularly interested to evaluate the capability of the proposed PFHA algorithm in identifying and locating the modes of the histogram. A total of 100 grayscale images, taken from the public segmentation database are adopted to evaluate the capability of the proposed PFHA algorithm.

On the other hand, for the second part of this section, the performance of the proposed HA initialization scheme will be compared with the randomly initialization scheme, and several latest adaptive initialization schemes, namely the Ant Colony Optimization (ACO) and the Agglomerated Just Noticeable Difference Histogram (AJNDH) initialization schemes. The randomly initialization scheme is chosen for the performance comparison as it has successfully proven its outstanding performance over the other two initialization techniques as reported by Bezdek et al. [20], among the C-means family [21]. As the randomly initialization scheme could not decide the number of cluster for each image adaptively, we assign the number of cluster of the randomly initialization scheme with the final cluster number obtained from the HA initialization scheme for the purpose of fair comparison.

For the ACO initialization scheme, it applies the intelligent searching capability of the Ant Colony Algorithm (ACA) to automatically determine the number of cluster and their corresponding cluster centers, which are required in the subsequence stage of the FCM algorithm. The ACO initialization scheme is chosen for the comparison as the previous literature [27] has demonstrated its excellent performances over the other recently proposed techniques such as the X-means, Mean Shift, Normalized Cut, and Han and Shi's methods. Meanwhile, the AJNDH initialization scheme, which is one of the latest introduced initialization schemes, is also chosen for comparisons as we are interested to see whether our proposed HA initialization scheme could generally outperform the AJNDH initialization scheme. In the second part of this section, a total of 140 color images obtained from the public segmentation

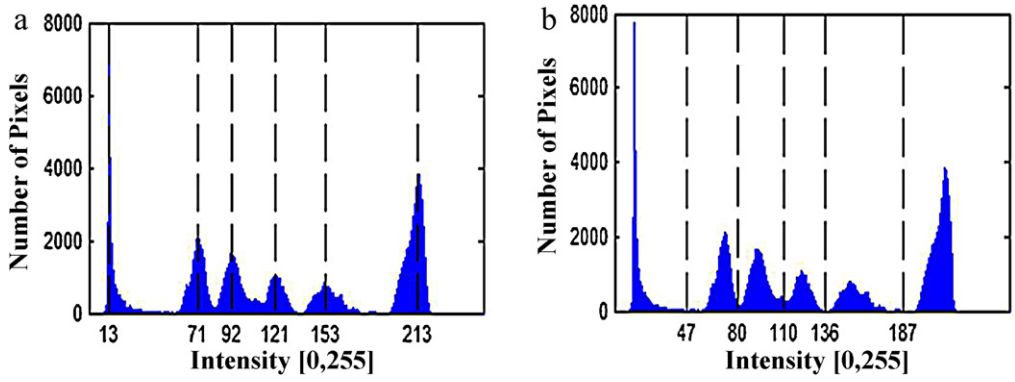

Fig. 9. The peaks and valleys located on the histogram for image *Mountain*, (a) peaks located on its histogram and (b) valleys located on its histogram.

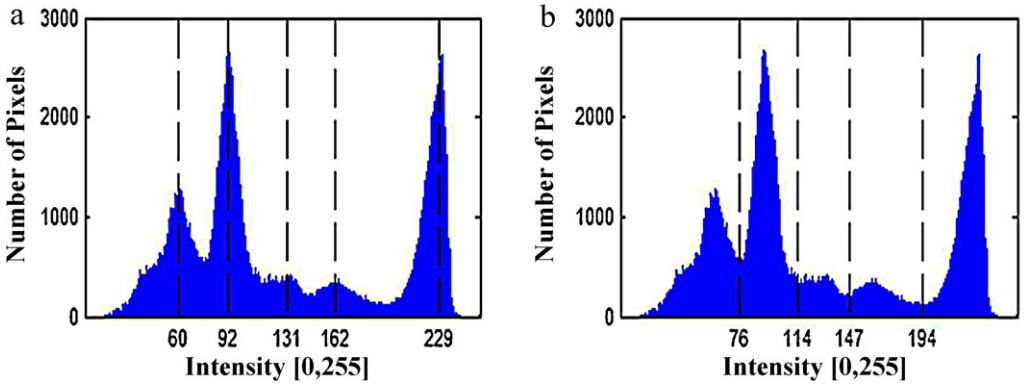

Fig. 10. The peaks and valleys located on the histogram for image *Beach*, (a) peaks located on its histogram and (b) valleys located on its histogram.

databases are used as tested dataset for the performance comparison between the proposed HA and other state-of-the-art initialization schemes.

4.1. Evaluation on the proposed PFHA algorithm

In this subsection, three grayscale test images namely *House*, *Mountain* and *Beach* are evaluated in details to highlight the capability of the proposed PFHA algorithm in identifying and locating the modes in the histogram. Generally, as shown in Figs. 8–10, the proposed PFHA algorithm can successfully locate the modes in the histograms and then the valleys (i.e. minimum value between two adjacent modes) on the histograms can be found. These values of modes and valleys for the grayscale images of *House*, *Mountain* and *Beach* are shown in Figs. 8(a), (b)–10(a), (b), respectively.

In this study, the valleys between two adjacent modes on the histograms have been proven to play a significant role, i.e. used as boundaries for the segmentation as shown in Figs. 11–13. For example, based on this concept, notice for the grayscale image *House*, the sky as well as the wall of the house are successfully segmented as homogeneous regions respectively as shown in Fig. 11. For the grayscale image *Mountain* as illustrated in Fig. 12, the mountains are successfully separated as different clusters and we could observe the different layers of mountains as shown in the segmented image. Similarly, for the grayscale image of *Beach*, both of the beach and sky are segmented as two different homogenous regions as shown in Fig. 13. These promising results obtained show that the proposed PFHA algorithm could successfully locate the modes in the histogram and the valleys identified in the next stage could play a significant role in segmenting the input images into the homogeneous regions of interest. As a result, we conclude that the proposed PFHA algorithm could be applied in the splitting module of the proposed HA initialization scheme to split any input color images into the multiple homogenous regions.

4.2. Performance comparison with unsupervised initialization schemes

As mentioned previously, the proposed HA initialization scheme will be compared with the randomly initialization scheme and several latest adaptive unsupervised initialization scheme, namely the ACO and AJNDH initialization schemes. In this section, the segmentation results produced by the randomly initialized, ACO, AJNDH, and HA initialization schemes are first evaluated qualitatively, i.e. through the naked eye, in term of the homogeneity of the segmented areas. In addition, the accuracy of the classification will be used to visually evaluate the performance of these algorithms as it is affected by the initial cluster centers. Meanwhile, as the conventional FCM algorithm is very sensitive to the cluster number produced in the initialization stage, it is worth to evaluate and discuss the capability of the abovementioned initialization schemes in producing the cluster number as each of these initialization schemes are able to initialize the cluster centers distribution adaptively.

In addition, we are also interested to investigate the cluster quality produced by the randomly initialized, ACO, AJNDH, and HA initialization schemes as the final segmentation results in our work is significantly dependant on the cluster quality. Several important cluster validity criteria have been featured by the previous work on the fuzzy clustering for the evaluation of the cluster quality. In this

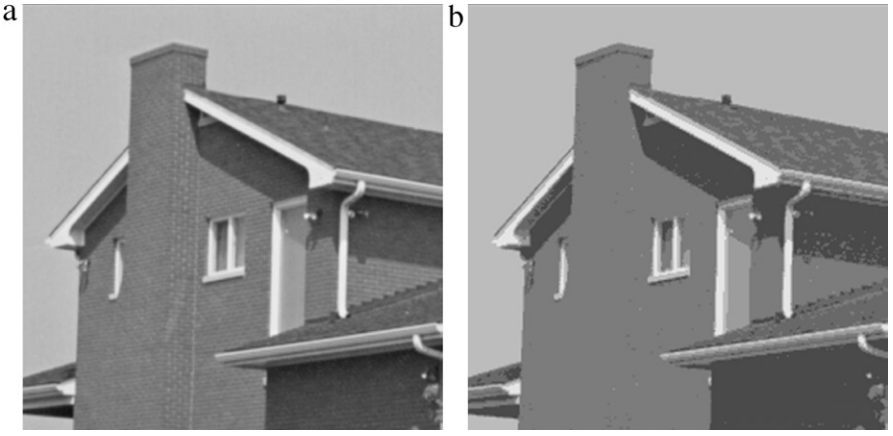

Fig. 11. Result on segmentation for image *House*, (a) original image and (b) resultant segmented image.

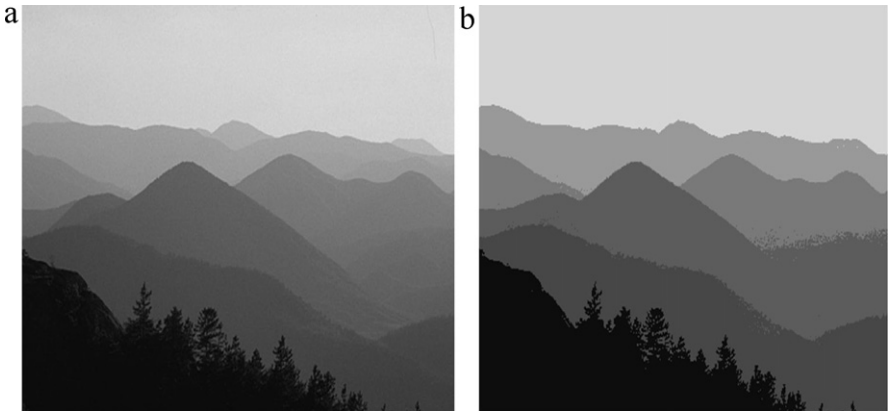

Fig. 12. Result on segmentation for image *Mountain*, (a) original image and (b) resultant segmented image.

paper, we adopt the mean square error (*MSE*) analysis to serve as the benchmark function that could be used to evaluate the cluster quality, which is described as follow:

$$MSE = \frac{1}{N}\sum_{j=1}^{M}\sum_{i\in S_i}||x_i - c_j||^2 \tag{14}$$

where N represents the number of output image pixels and M is the number of cluster produced during the clustering process. On the other hand, S_j represents the set of pixels in jth cluster, c_j denotes the pixel's intensity levels of the jth cluster center, and x_i is the intensity levels of ith pixel in jth cluster center. The concept of the *MSE* analysis is quite simple: for a fixed number of cluster, the cluster centers should be placed in such a way that they reduce the distance to data pieces as much as possible, in order to generate a good clustering results, which has small distortion.

Finally, a total of three benchmark analyses will be used to evalute the segmentation results of the randomly initialized, ACO, AJNDH, and HA initialization schemes quantitatively. These benchmark analyses are chosen as the standard quantiative tests for color segmentation to evaluate the goodness of the segmentation results based on some human characterization about the properties of ideal segmentation [27,46]. All of these three benchmark functions are used to penalize the segmentation that form too many regions and having non-homogenous regions by giving a larger values. These three benchmark function are described as follows:

$F(I)$ proposed by Liu and Yang [47],

$$F(I) = \frac{\sqrt{M}\sum_{j=1}^{M} e_j^2}{\sqrt{N_j}} \tag{15}$$

$F'(I)$ proposed by Borsotti et al. [48],

$$F'(I) = \frac{\sum_{j=1}^{M} e_j^2 \sqrt{\sum_{a=1}^{MaxArea}[S(a)]^{1+1/a}}}{(1000\times N)\sqrt{N_j}} \tag{16}$$

$Q(I)$ further refined from $F(I)$ by Borsotti et al. [48],

$$Q(I) = \frac{\sqrt{M}}{1000\times N}\sum_{j=1}^{M}\left[\frac{e_j^2}{1+\log N_j} + \left(\frac{S(N_j)}{N_j}\right)^2\right] \tag{17}$$

where I is an image and N is the total pixels in I. The segmentation can be described as an assignment of pixels in the image I into M regions. Let C_j denotes the set of pixels in region j, $N_j = |C_j|$ denotes the number of pixels in C_j. The value of e_j, which represents the homogeneity within a region, is defined as the Euclidean distances between the RGB color vectors of the pixels of region j and the color

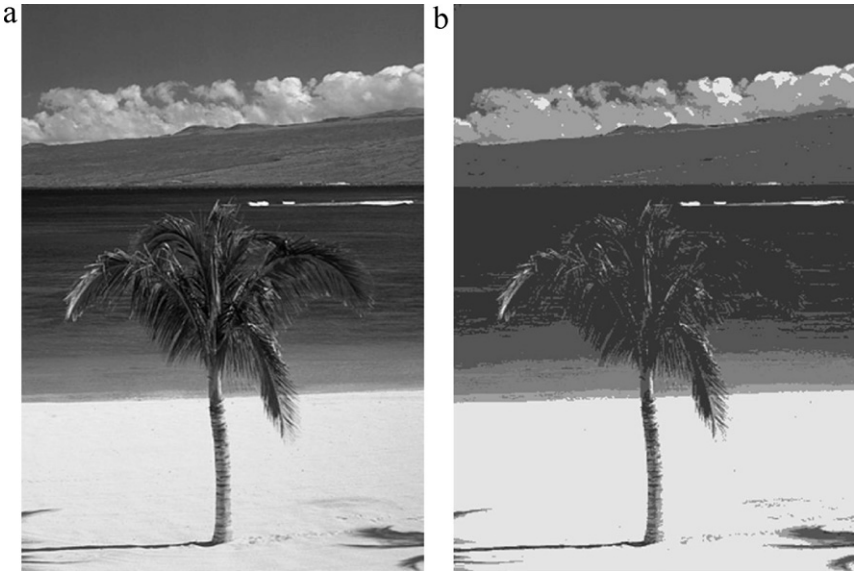

Fig. 13. Result on segmentation for image *Beach*, (a) original image and (b) resultant segmented image.

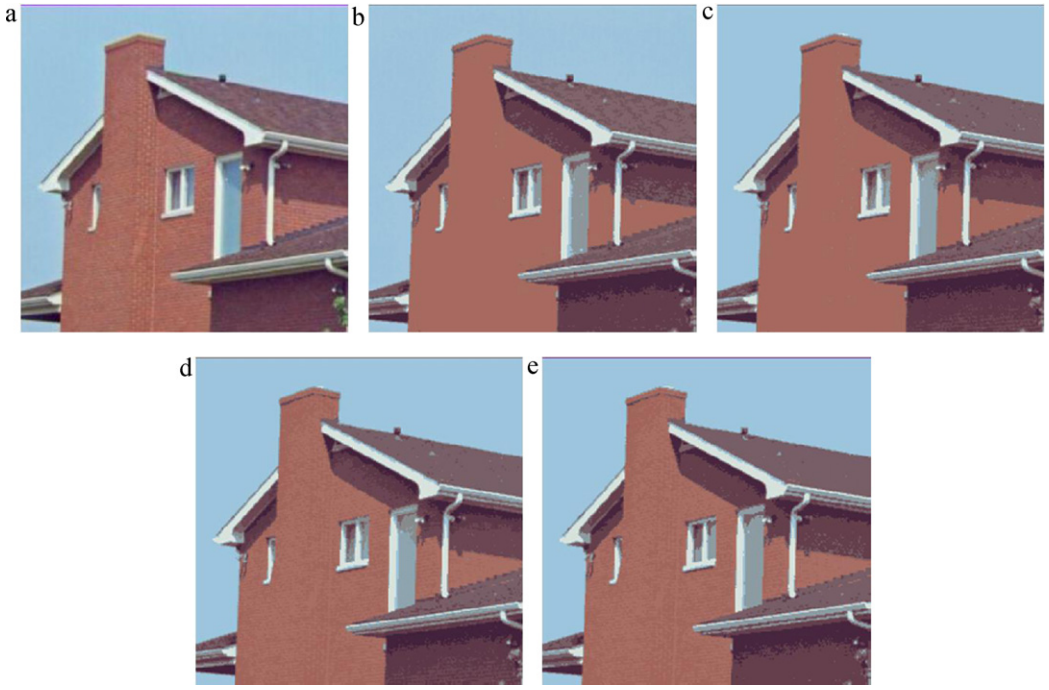

Fig. 14. The image *House*; (a) original image, and the rest are segmentation results of the test image by various initialization schemes (b) randomly initialized (c) ACO, (d) AJNDH, and (e) HA.

vector attributed to region j in the segmented image. Finally, $S(a)$ denotes the number of regions in the image I that has an area of exactly a and *MaxArea* denotes the largest region in the segmented image. It is worth to point out that these evaluation functions do not require any prior knowledge of correct segmentation. Hence, the users do not need to set any parameter or threshold values for the quantitative evaluation of the segmentation performance on color images. In these evaluation functions, the smaller value indicates the better segmentation result.

4.2.1. Qualitative evaluation on segmentation results

In this section, the results of the proposed HA initialization scheme and the randomly initialized, ACO, and AJNDH initialization schemes are evaluated visually using 7 out of the 140 tested images. The segmentation results for those images, namely *House*, *Capsicums*, *Nemo*, *Fire Fighter*, *Birds*, *Swimmer*, and *Moon* are shown in Figs. 14–20, respectively. Generally, the proposed HA initialization scheme produces better segmentation results as compared to the other compared initialization schemes for all images as illustrated in Figs. 14–20. The proposed HA initialization scheme is able to produce more homogeneous segmented regions as compared to the randomly initialized, ACO, and AJNDH initialization schemes. In addition, the proposed HA initialization scheme also achieves superior performance in preserving the salient features of the input color images as it possess an excellent accuracy of classification.

As for image *House* illustrated in Fig. 14, there is an obvious classification error in the segmentation results produced by the randomly initialized, ACO, and AJNDH initialization schemes as shown in Fig. 14(b)–(d), respectively by mistakenly assigned the uppermost line in purple color as part of the house. However, the proposed HA initialization scheme successfully prevents this classification error from occuring and manages to segment the purple line to be a single region as shown in Fig. 14(e). In addition, the HA initilization scheme also outperforms the randomly initialized, ACO, and AJNDH initialization schemes by producing more homogenous house roof and wall as shown in Fig. 14(b)–(e), respectively.

Similarly, for the image *Capsicums* as illustrated in Fig. 15, the proposed HA initialization scheme also produces better segmentation results than the randomly initialized, ACO, and AJNDH initialization schemes. This is because the proposed HA initialization scheme is capable to produce a significantly more homogenous segmented regions on the red capsicum's and the green capsicum's surfaces as compared to the other initialization schemes. As for image *Nemo* shown in Fig. 16, it is interesting to see that there is an classification error where the body of Nemo fish is mistakenly assigned to the background by the randomly initialized and ACO initialization schemes as shown in Fig. 16(b) and (c), respectively. Both of the AJNDH and HA initialization schemes, however, have successfully avoided this classification error by clustering the Nemo fish's body and the background into two distintive clusters. Furthermore, the proposed HA initialization scheme successfully proves its superior performance over the AJNDH initialization scheme by producing a more homogenoues region on the Nemo fish's body as illustrated in Fig. 16(e) and (d), respectively.

Notice for the image *Fire Fighter* as shown in Fig. 17, the randomly initialized, AJNDH, and HA initialization schemes produce better segmentation results than the ACO initialization scheme by assigning the helmet of the fire fighter into the desired blue color. Although the randomly initialized and AJNDH initialization schemes do not suffer the severe classification error as demonstrated by the ACO initialization scheme in Fig. 17(c), there are considerable number of pixels in the blue helmet are falsely assigned as the background as can be seen in the resultant images produced by the randomly initialized and AJNDH initialization schemes in Fig. 17(b) and (d), respectively. In addition, for the resultant segmented image produced by the AJNDH initialization scheme as shown in Fig. 17(d), some pixels in the yellow color strip are falsely clustered as the part of the orange color uniform. While all the compared initialization schemes have demonstated different level of classification errors, the proposed HA initialization scheme, on the other hand, has successfully avoided these classification errors by assigning the blue helmet, background, orange uniform, and yellow strip as separate clusters as shown in Fig. 17(e). This

Fig. 15. The image *Capsicums*; (a) original image, and the rest are segmentation results of the test image by various initialization schemes (b) randomly initialized (c) ACO, (d) AJNDH, and (e) HA.

observation again proves the superior accuracy of the proposed HA initialization scheme in classifying the objects into different clusters of interest.

For the image *Birds* as shown in Fig. 18, it can be observed that all the segmented images produced by the randomly initialized, ACO, and AJNDH initialization schemes are experiencing a severe classification error as all of these techniques mistakenly assign the white feather of the bird into the blue color sky as illustrated in Fig. 18(b)–(d), respectively. Despite the fact that the HA initialization scheme produces more clusters at the background, it has successfully avoided the abovementioned classification error by assigning the white feather of bird and blue color of sky into two different clusters as shown in Fig. 18(e). Meanwhile, for image *Swimmer* as illustrated in Fig. 19, we could observe that both of the randomly initialized and ACO initialization schemes fail to classify the pixels in the swimming trunk into the desired red color as can be seen in the resultant images in Fig. 19(b) and (c), respectively. Both of the AJNDH and HA initialization schemes, on the other hand, manage to prevent themselves from suffering in such classification errors as shown in Fig. 19(d) and (e), respectively.

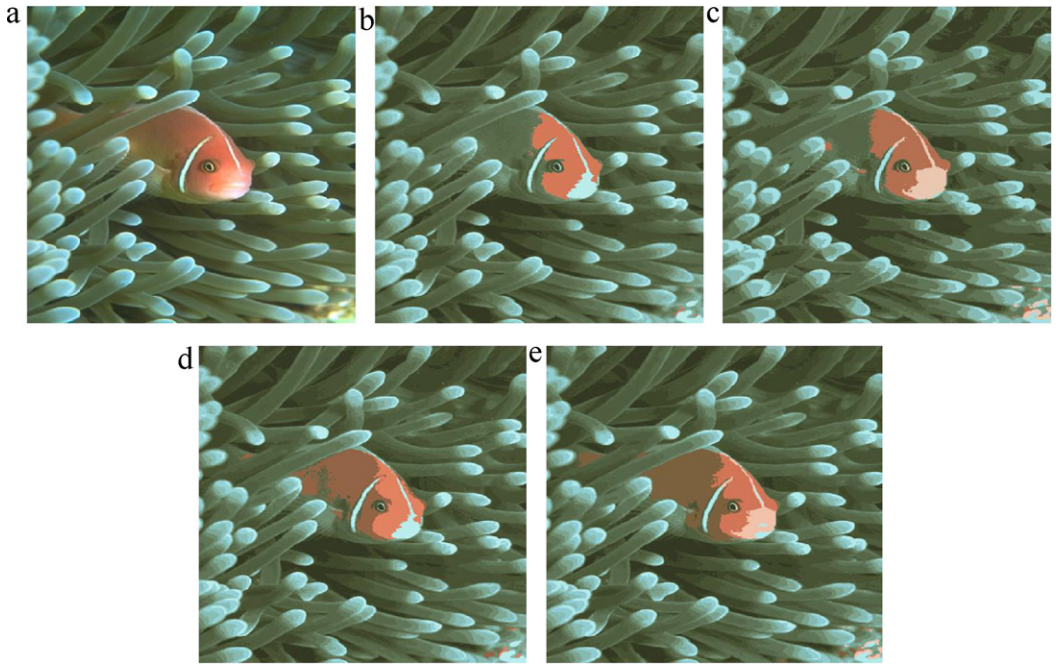

Fig. 16. The image *Nemo*; (a) original image, and the rest are segmentation results of the test image by various initialization schemes (b) randomly initialized (c) ACO, (d) AJNDH, and (e) HA.

Fig. 17. The image *Fire Fighter*; (a) original image, and the rest are segmentation results of the test image by various initialization schemes (b) randomly initialized (c) ACO, (d) AJNDH, and (e) HA.

However, the proposed HA initialization scheme performs slightly better than the AJNDH initialization scheme as the proposed HA initialization scheme is able to produce a more homogeneous region on the swimmer's leg, with less clusters as compared to the AJNDH initialization scheme.

Finally, for the image *Moon* as illustrated in Fig. 20, it is interesting to see that there is an obvious classification error where the moon in the image has been falsely clustered to the sky by the randomly initialized, ACO, and AJNDH initialization schemes in Fig. 20(b)–(d) respectively, although some of them are able to produce a more homogeneous background with fewer clusters. The proposed HA initialization scheme, on the other hand, has successfully avoid this classification error by assigning the moon and sky as separate clusters as shown in Fig. 20(e).

Apart from the visual inspection results on those 7 images, an additional of 20 supplementary images are analyzed as well to support the abovementioned findings. The findings obtained from these 20 supplementary images are illustrated in Fig. A.1 in Appendix A. Based on the visual inspection performed in Appendix A, it can be proven that the proposed HA initialization scheme outperforms the randomly initialized, ACO, and AJNDH initialization schemes as it is able to produce significantly better segmentation results, with more homogenous segmented regions and less classification errors.

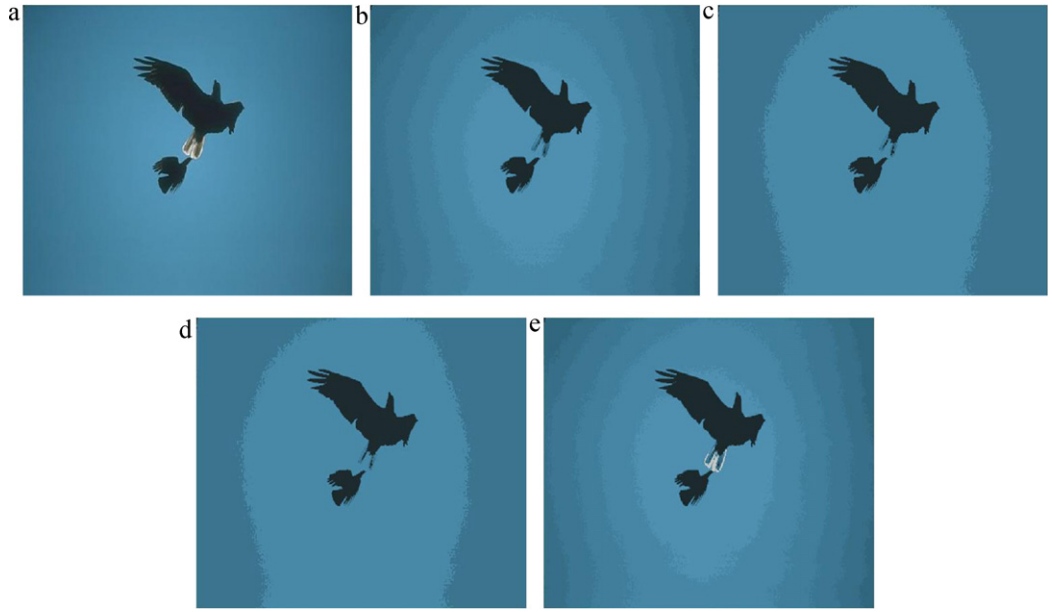

Fig. 18. The image *Birds*; (a) original image, and the rest are segmentation results of the test image by various initialization schemes (b) randomly initialized (c) ACO, (d) AJNDH, and (e) HA.

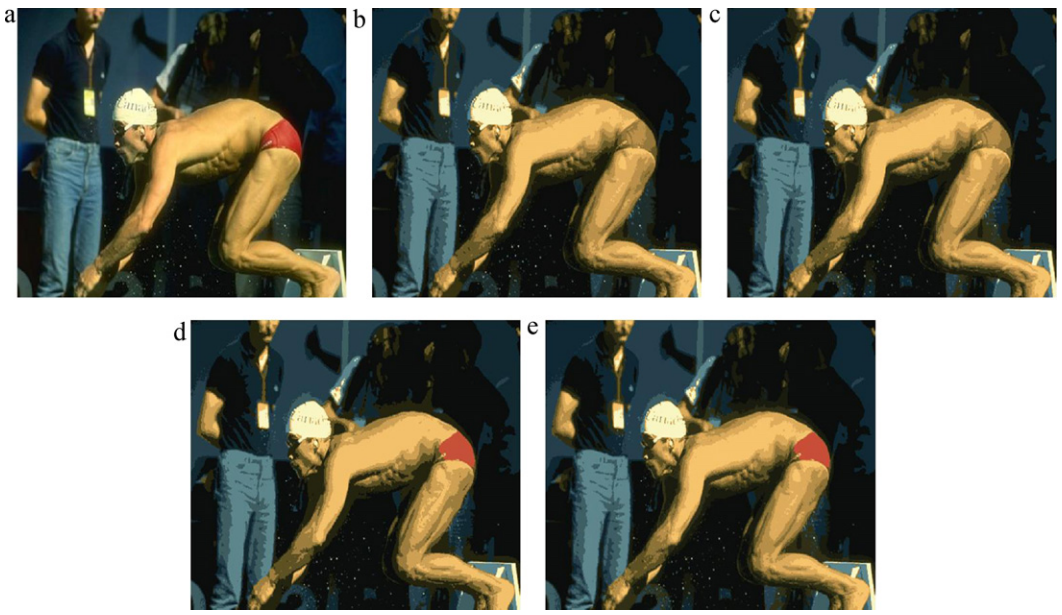

Fig. 19. The image *Swimmer*; (a) original image, and the rest are segmentation results of the test image by various initialization schemes (b) randomly initialized (c) ACO, (d) AJNDH, and (e) HA.

4.2.2. Evaluation on cluster number

In the qualitative results shown in the previous section, we observe that all of the randomly initialized, ACO, AJNDH, and HA initialization schemes have their own unique mechanism to adaptively initialize the cluster center distribution and the centroid values. A good clustering result is highly dependent on the cluster center initialization mechanism as a good initialization scheme can guarantee a high accuracy of classification and the advantage of less distortion during the segmentation process. Thus, in this section, we are interested to know the relationship between the quality of segmentation results and the number of cluster produced by each initialization scheme.

The number of cluster produced by each initialization scheme is tabulated in Table 2. It is interesting to observe that the cluster

Table 2
Number of cluster produced by different initialization schemes.

Images	Initialization schemes			
	Randomly initialized	ACO	AJNDH	HA
House	**9**	**9**	10	**9**
Capsicums	15	15	**14**	15
Nemo	14	**9**	13	14
Fire Fighter	19	**9**	19	19
Birds	8	**3**	3	8
Swimmer	14	13	**12**	14
Moon	8	**3**	**3**	8

Note: The bold values represent the best results obtained for the comparison.

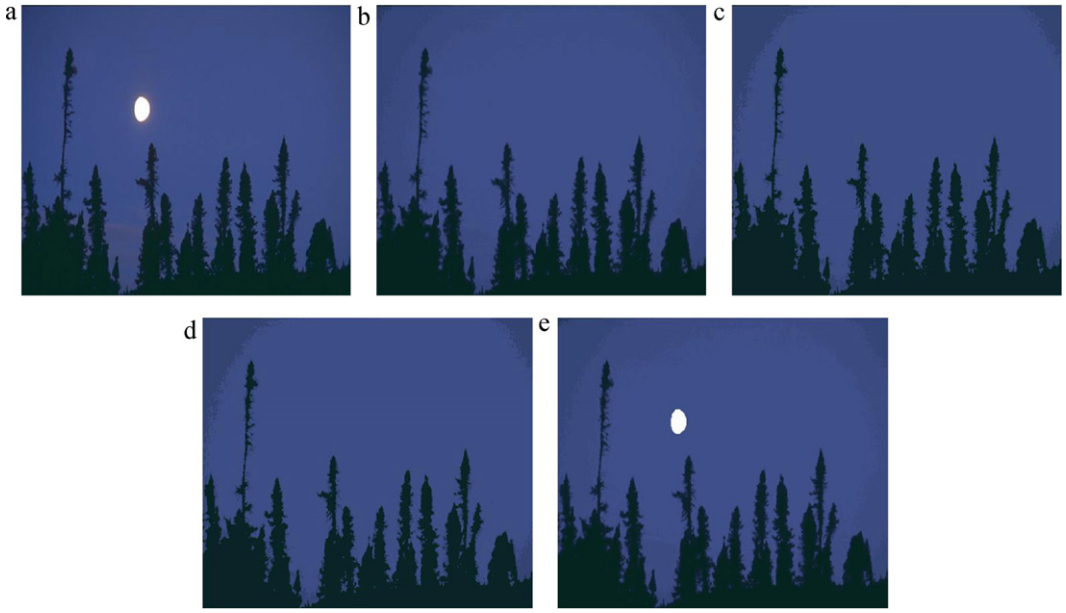

Fig. 20. The image *Moon*; (a) original image, and the rest are segmentation results of the test image by various initialization schemes (b) randomly initialized (c) ACO, (d) AJNDH, and (e) HA.

number produced by each initialization scheme for every image is similar, but not exactly the same. This observation is reasonable as each initialization scheme has their unique mechanism in deciding the cluster number. Meanwhile, we also can observed that the cluster number produced by the randomly initialization scheme is same as the proposed HA initialization scheme. This is because we have made minor modification on the randomly initialization scheme, where the final cluster number is set to that obtained by the proposed HA initialization scheme. This minor modification could guarantee the fair comparison between the randomly initialized and HA initialization schemes.

Another observation that should be highlighted is the possibility of significant qualitative and quantitative differences in the segmentation results produced by different algorithms despite having the same cluster number. While the number of clusters produced is necessary, it is an insufficient performance metric to reveal the overall clustering performances of the aforementioned algorithms. The randomly initialized FCM and HA techniques have the same cluster number. However, Sections 4.2.3 and 4.2.4 show that the other performance metrics used to access the clustering quality (i.e. the *MSE* values) and homogeneity [i.e. the $F(I)$, $F'(I)$, and $Q(I)$ values] of the segmented images are significantly different among the randomly initialized FCM and HA technique.

As shown in Table 2, the proposed HA initialization scheme produces less or comparable number of cluster than the randomly initialized, ACO, and AJNDH initialization schemes for image *House* and *Capsicums*. Thus, by having smaller or comparable number of cluster, the proposed HA initialization scheme could produce larger and better homogenous regions in the segmented image.

Meanwhile, for images *Birds* and *Moon*, although the ACO and AJNDH initialization schemes could produce larger homogenous regions by obtaining fewer number of cluster than the randomly initialized and HA initialization schemes, there are considerable pixels in the segmented images produced by the ACO and AJNDH initialization schemes which are falsely assigned to the background (i.e. sky), leading to the classification error problem. However, this problem has successfully avoided by the proposed HA initialization scheme, as the appropriate number of cluster has been obtained. Although the randomly initialization scheme possess the same cluster number as the proposed HA initialization scheme, it suffers the same classification error as produced by the ACO and AJNDH initialization schemes. This is due to the fact that the randomized initialization mechanism possessed by the randomly initialization scheme, which has the unstable nature, does not properly initialize the initial cluster centers during the initialization stage.

As for image *Nemo* and *Fire Fighter*, although the cluster number produced by the proposed HA initialization scheme is more or comparable than those produced by the randomly initialized, ACO, and AJNDH initialization schemes, the segmented results produced by the proposed HA initialization scheme successfully avoid the classification errors. Unlike the HA initialization scheme, the other compared initialization schemes have experienced the different level of classification errors, where there are certain amount of pixels in the segmented results are mistakenly clustered, as explained in the previous section. Finally, for image *Swimmer*, the cluster number produced by all the initialization schemes is similar. However, the AJNDH and HA initialization schemes perform better as the severe classification error on the swimmer's trunk could be observed in the segmented images produced by the randomly initialized and ACO initialization schemes. Furthermore, the proposed HA initialization scheme performs slightly better than the AJNDH initialization scheme as it produces visually better quality of the segmentation results than the AJNDH initialization scheme.

Based on the results tabulated in Table 2, we conclude that fewer number of cluster does not always guarantee a good segmentation results as there is a tradeoff between the number of cluster produced and the quality of segmentation results. The insufficient number of cluster produced during the initialization stage tends to lead to the classification errors problems as shown in images *Nemo*, *Fire Fighter*, *Birds*, and *Moon*. Thus, instead of emphasizing the small cluster number, we should keep a reasonable number of cluster and achieve more homogeneity within regions in order to obtain good segmentation results.

Table 3
Comparison of clustering quality among the proposed HA and other initialization schemes based on the *MSE* evaluation function.

Images	Initialization schemes			
	Randomly initialized (*1.0e+2)	ACO (*1.0e+2)	AJNDH (*1.0e+2)	HA (*1.0e+2)
House	3.1688	2.9670	***2.7373***	**2.1728**
Capsicum	***4.4297***	4.4379	4.6439	**4.0973**
Nemo	3.3309	3.8545	***2.8995***	**2.4640**
Fire Fighter	2.7603	6.6132	***2.7187***	**2.6126**
Birds	***0.7084***	1.8408	1.8408	**0.5244**
Swimmer	***2.8028***	2.9533	3.1334	**2.6291**
Moon	***2.7987***	3.5934	3.5934	**0.4167**

4.2.3. Evaluation on clustering quality

The *MSE* values of the randomly initialized, ACO, AJNDH, and HA initialization schemes are tabulated in Table 3. In Table 3, the best results obtained are made bold while the second best are made bold and italic. This notation will be employed for the other results presented in this paper.

Based on Table 3, we can observe that the proposed HA initialization scheme has exhibited its superior performance in term of the clustering quality over the other initialization schemes. The proposed HA initialization scheme always produces the smallest *MSE* values (i.e. ranked as the best), showing its capability in producing the clustering results with less distortion as compared to the other initialization schemes. This observation is quite consistent with the visual inspection as explained in Section 4.2.1.

Apart from the images as shown in Figs. 15–20, the *MSE* values of the randomly initialized, ACO, AJNDH, and HA initialization schemes for another 20 supplementary images (as depicted in Fig. A.1 in Appendix A) are tabulated in Table B.1 in Appendix B. Generally, the proposed HA initialization scheme outperforms the randomly initialized, ACO, and AJNDH initialization schemes by consistently producing the relatively smaller *MSE* values (i.e. ranked as the best or second best) in those 20 images. The capability of the proposed HA initialization scheme to consistently produce the smaller *MSE* values successfully proves its advantage in producing the clustering results with better cluster distribution as compared to the other initialization schemes.

4.2.4. Quantitative evaluation on segmentation results

The quantitative results obtained from the $F(I)$, $F'(I)$, and $Q(I)$ evaluation functions are tabulated in Tables 4–6 respectively. From these tables, it can clearly be observed that the proposed HA

Table 4
Comparison of segmentation results based the $F(I)$ evaluation function.

Images	Initialization schemes			
	Randomly initialized (*1.0e+2)	ACO (*1.0e+2)	AJNDH (*1.0e+2)	HA (*1.0e+2)
House	4.3023	3.7738	***2.6004***	**2.0706**
Capsicum	3.6407	***3.6401***	3.6861	**3.3379**
Nemo	6.6320	5.4932	***4.1116***	**3.3812**
Fire Fighter	5.0150	13.7404	***4.2908***	**4.2477**
Birds	***1.3315***	4.7516	4.7516	**1.3098**
Swimmer	***3.7584***	4.0451	4.2976	**3.6715**
Moon	***2.5424***	2.9085	2.9085	**1.5595**

Table 5
Comparison of segmentation results based on the $F'(I)$ evaluation function.

Images	Initialization schemes			
	Randomly initialized (*1.0e+1)	ACO (*1.0e+1)	AJNDH (*1.0e+1)	HA (*1.0e+1)
House	4.3922	3.8584	***2.6542***	**2.1257**
Capsicum	3.7269	***3.7253***	3.7752	**3.4248**
Nemo	6.7685	5.6448	***4.2101***	**3.4610**
Fire Fighter	5.0397	13.8679	***4.3126***	**4.2737**
Birds	***1.3499***	4.9124	4.9124	**1.3299**
Swimmer	***3.8453***	4.1389	4.4134	**3.7577**
Moon	***2.5758***	3.0569	3.0569	**1.5864**

initialization scheme produces the best (smallest) $F(I)$, $F'(I)$, and $Q(I)$ values for the image *House*, *Capsicums*, *Birds*, *Swimmer*, and *Moon*. While for images *Nemo* and *Fire Fighter*, the proposed HA initialization scheme produces the best $F(I)$ and $F'(I)$ values and second best $Q(I)$ value. These results support the promising qualitative findings obtained by the proposed HA initialization scheme in the previous section as the smaller values of $F(I)$, $F'(I)$, and $Q(I)$ evaluation functions tend to produce good segmentation results, which have more homogenous and less distortion segmented regions. Thus, these results strongly prove that the proposed HA initialization scheme outperforms the randomly initialized, ACO, and AJNDH initialization schemes both qualitatively and quantitatively.

Meanwhile, another important finding that could be observed from the results tabulated in Tables 4–6 is that the other initialization schemes that have been put in comparison could produce good $F(I)$, $F'(I)$, and $Q(I)$ evaluation functions for only certain images. For example, the ACO initialization scheme produces the smallest $Q(I)$ value for image *Fire Fighter* but fails to produce the same good result (i.e. small values of $Q(I)$) for images *Birds* and *Capsicums*. The similar findings could be observed for other evaluation functions (i.e. $F(I)$ and $F'(I)$) on other images. Similarly, the ACO and AJNDH initialization schemes also suffer with the same problem. In addition, we also observe that the good performance achieved by these compared initialization schemes in the $F(I)$, $F'(I)$, and $Q(I)$ evaluation functions is not always consistent with the qualitative results as presented in Section 4.3.1. For example, although the randomly initialization scheme could produce second best $F(I)$, $F'(I)$, and $Q(I)$ values in image *Birds*, a significant classification error could be observed in the segmented image as the white feather of the birds is falsely clustered into the sky. Similar problems could be observed on the segmentation results of images *Swimmer*, *Moon*, *Nemo*, *Fire Fighter* produced by the ACO initialization scheme. In contrary, although the $F(I)$, $F'(I)$, and $Q(I)$ values produced by the proposed HA initialization scheme are also relatively small, it has successfully exhibited its robustness in preserving the salient features of the input color images as well as prevents the classification errors during the segmentation process.

Finally, it is also observed that the randomly initialized, ACO, and AJNDH initialization schemes produce inconsistent quantitative performance (i.e. based on the $F(I)$, $F'(I)$, and $Q(I)$ values obtained)

Table 6
Comparison of segmentation results based on the $Q(I)$ evaluation function.

Images	Initialization schemes			
	Randomly initialized (*1.0e+3)	ACO (*1.0e+3)	AJNDH (*1.0e+3)	HA (*1.0e+3)
House	0.7716	0.7056	***0.3495***	**0.2813**
Capsicum	0.5653	0.5740	***0.5347***	**0.4776**
Nemo	6.1464	**0.7001**	2.1969	***1.8929***
Fire Fighter	348.4360	**28.1336**	317.0170	***172.0360***
Birds	***0.3868***	2.0572	2.0572	**0.3803**
Swimmer	***0.5075***	0.5449	0.5813	**0.5073**
Moon	1.7599	***1.3058***	***1.3058***	**0.7068**

Table 7
Performance comparison of segmentation results based on average values of *MSE*, $F(I)$, $F'(I)$, and $Q(I)$ for 140 standard images.

Initialization schemes	Benchmark quantitative evaluation functions			
	MSE (*1.0e+2)	$F(I)$ (*1.0e+2)	$F'(I)$ (*1.0e+1)	$Q(I)$ (*1.0e+5)
Randomly initialized	3.1900	***7.3500***	***7.4400***	0.6890
ACO	3.5200	8.5300	8.6600	***0.5710***
AJNDH	***3.2200***	7.7900	7.8900	1.2400
HA	**2.9000**	**6.8400**	**6.9200**	**0.5690**

for the same image. For example, the ACO initialization scheme produces the best $Q(I)$ value for the image *Fire Fighter* but achieves the worst ranking in terms of the $F(I)$ and $F'(I)$ values. The same problems can also be observed for the randomly initialized and AJNDH initialization schemes. On the other hand, the proposed HA initialization scheme can perform consistently, by producing the relatively small $F(I)$, $F'(I)$, and $Q(I)$ evaluation functions for all images. This successfully proves the ability of the proposed HA initialization scheme in persisting consistent and good performance for any type of analyses and images.

In order to support the abovementioned findings, the $F(I)$, $F'(I)$, and $Q(I)$ values of the randomly initialized, ACO, AJNDH, and HA initialization schemes for another 20 supplementary images (as depicted in Fig. A.1 in Appendix A) are tabulated in Table C.1 in Appendix C. Overall, notice from the quantitative results of $F(I)$, $F'(I)$, and $Q(I)$ values as shown in Table C.1, it is clearly shown the proposed HA initialization scheme produces relatively smaller values in these three evaluation benchmarks as compared to the other compared initialization schemes. Thus, the segmentation results produced by the proposed HA initialization scheme are more favored. Although both of the ACO and AJNDH initialization schemes could produce smaller values of $F(I)$, $F'(I)$, and $Q(I)$ than the proposed HA initialization scheme in certain images, the difference in these values is not significant, and furthermore the segmentation regions produced by the proposed HA initialization scheme is more homogenous and less distortion as compared to other initialization schemes when inspected visually.

Meanwhile, it is also worth to point out that the randomly initialization scheme could achieve slightly good performance in certain images as it is able to produce comparable $F(I)$, $F'(I)$, and $Q(I)$ values with the proposed HA initialization scheme. The excellence performance of the randomly initialization scheme in certain images is mainly due to the fact that both of the randomly initialized and HA initialization schemes are sharing the same cluster number during the segmentation process. However, the randomly initialization scheme fails to maintain its consistent performances for all images as the randomized mechanism that involved during the initialization process of the cluster center is unstable and has too much uncertainties involved. This could lead to less satisfactory initialization of cluster centers, which is then followed by a poor segmentation result. On the other hand, for the proposed HA initialization scheme, the abovementioned unstable nature and uncertainties issues during the initialization stage do not exist as it offers a more systematic procedure of splitting and merging modules to initialize the cluster center. Such systematic procedure can guarantee a more accurate and reasonable estimation of the initial cluster, and this ensures the proposed HA initialization scheme could persist its outstanding performance for all images.

To further evaluate the segmentation results produced by the proposed HA initialization scheme and the other state-of-art initialization schemes, the average value of *MSE*, $F(I)$, $F'(I)$, and $Q(I)$ for 140 natural images taken from the public segmentation database are given as tabulated in Table 7. As shown in Table 7, the results clearly prove that the proposed HA initialization scheme successfully

outperforms the randomly initialized, ACO, and AJNDH initialization schemes by producing the smallest average values of *MSE*, $F(I)$, $F'(I)$, and $Q(I)$. This is then followed by the randomly initialized and AJNDH initialization schemes. The ACO initialization scheme, on the other hand, produces the worst results by producing the largest values of *MSE*, $F(I)$, and $F'(I)$, and large value of $Q(I)$. Based on the average *MSE* values shown, we can conclude that the proposed HA initialization scheme is able to generate a more compact and stable cluster during the clustering process as compared to the other initialization schemes. Furthermore, the outstanding capabilities of the proposed HA initialization scheme in consistently producing the small values of $F(I)$, $F'(I)$, and $Q(I)$ during the segmentation process suggesting the potential of the proposed HA initialization scheme to be employed as a robust and excellence initialization scheme for the conventional FCM algorithm.

5. Conclusions

In this paper, the HA initialization scheme is proposed to overcome the sensitiveness of FCM algorithm to the initialization condition of clusters number and initial cluster centers as they have significant impacts on the segmentation quality. In other word, better segmentation result could be achieved if better initialization condition for FCM algorithm is provided. From the experimental results, it can be concluded that the proposed initialization scheme could produce better initialization condition for FCM algorithm than the randomly initialized, ACO and AJNDH initialization schemes by successfully reducing the classification errors and producing more homogeneous regions in the segmentation results. Meanwhile, the quantitative analysis also proves that the proposed HA initialization scheme has successfully produced better segmentation results. Thus, it is recommendable for this algorithm to be applied in the post image processing in consumer electronic products or machine vision systems, which are, for example, extensively used with the microscope in capturing microscopic images, especially in segmenting medical images. As future work, we will investigate the impact of different color conversion technique on the clustering results. More specifically, the newly proposed HSL color space such as those proposed in [25] will be employed in our HA module to determine whether the clustering results will change significantly when different variants of HSL color space is used. In addition, we also like to further investigate the capability of the HA initialization scheme to perform segmentation with other performance metrics such as the *F*-measure, variation of information, rand index as adopted by the literatures in [49,50].

Acknowledgements

The authors would like to express their sincere thanks to the associate editor and all reviewers who made great contributions to the improvement of the final paper. This research was supported by the Fundamental Research Grant Scheme (FRGS) entitled "Investigation of New Color Image Illumination Estimation Concept for Development of New Color Correction Techniques" and Universiti Sains Malaysia (USM) Postgraduate Fellowship Scheme.

Appendix A. Segmentation results of the 20 test images by the randomly initialized, ACO, AJNDH, and HA initialization schemes

Fig. A1.

Appendix B. Comparison of segmentation results by the randomly initialized, ACO, AJNDH, and HA initialization schemes based on the MSE evaluation function

Table B1.

Table B1
Comparison of clustering quality based on the *MSE* evaluation function.

Test	Image	Initialization schemes			
		Randomly initialized (*1.0e+2)	AFHA (*1.0e+2)	AJNJDH (*1.0e+2)	RFHA (*1.0e+2)
MSE	*Diver 1*	1.8738	6.7123	***1.6425***	**1.5621**
	Crown	**2.1728**	3.1039	3.0312	***2.2799***
	Car	**2.4405**	3.5353	**2.4405**	***2.5226***
	Cow	***3.0375***	***3.0375***	**2.2997**	***3.0375***
	Insect	***2.3378***	2.7968	4.8237	**2.2173**
	White Church	***1.8040***	2.4351	1.8225	**1.7999**
	Beach	***2.8764***	2.9097	**2.6842**	**2.6842**
	Red Church	***1.4772***	2.3447	1.7596	**1.4648**
	Pyramid	5.7642	4.7869	***3.9997***	**3.4171**
	Drum Players	0.4690	**0.2561**	***0.2843***	***0.2843***
	Statues	***1.2196***	2.2036	1.8679	**1.2064**
	Eagle	**3.0806**	***3.8248***	***3.8248***	**3.0806**
	White Boat	4.7628	4.8270	**3.2343**	***3.6140***
	Diver	3.0049	**2.6116**	***2.9407***	***2.9407***
	Island	***2.2623***	3.9973	***2.2623***	**2.1480**
	Cactus	***3.5109***	**3.1634**	3.7046	**3.1634**
	Horses	3.2572	**2.8643**	3.5025	***3.1542***
	Building	***3.6886***	4.8967	4.2599	**3.6699**
	Onion	2.6336	***2.5814***	3.5940	**2.5768**
	Pegion	**2.1085**	4.0238	2.4239	***2.2026***

Appendix C. Comparison of segmentation results by the randomly initialized, ACO, AJNDH, and HA initialization schemes based on the $F(I)$, $F'(I)$ and the $Q(I)$ evaluation functions

Table C1.

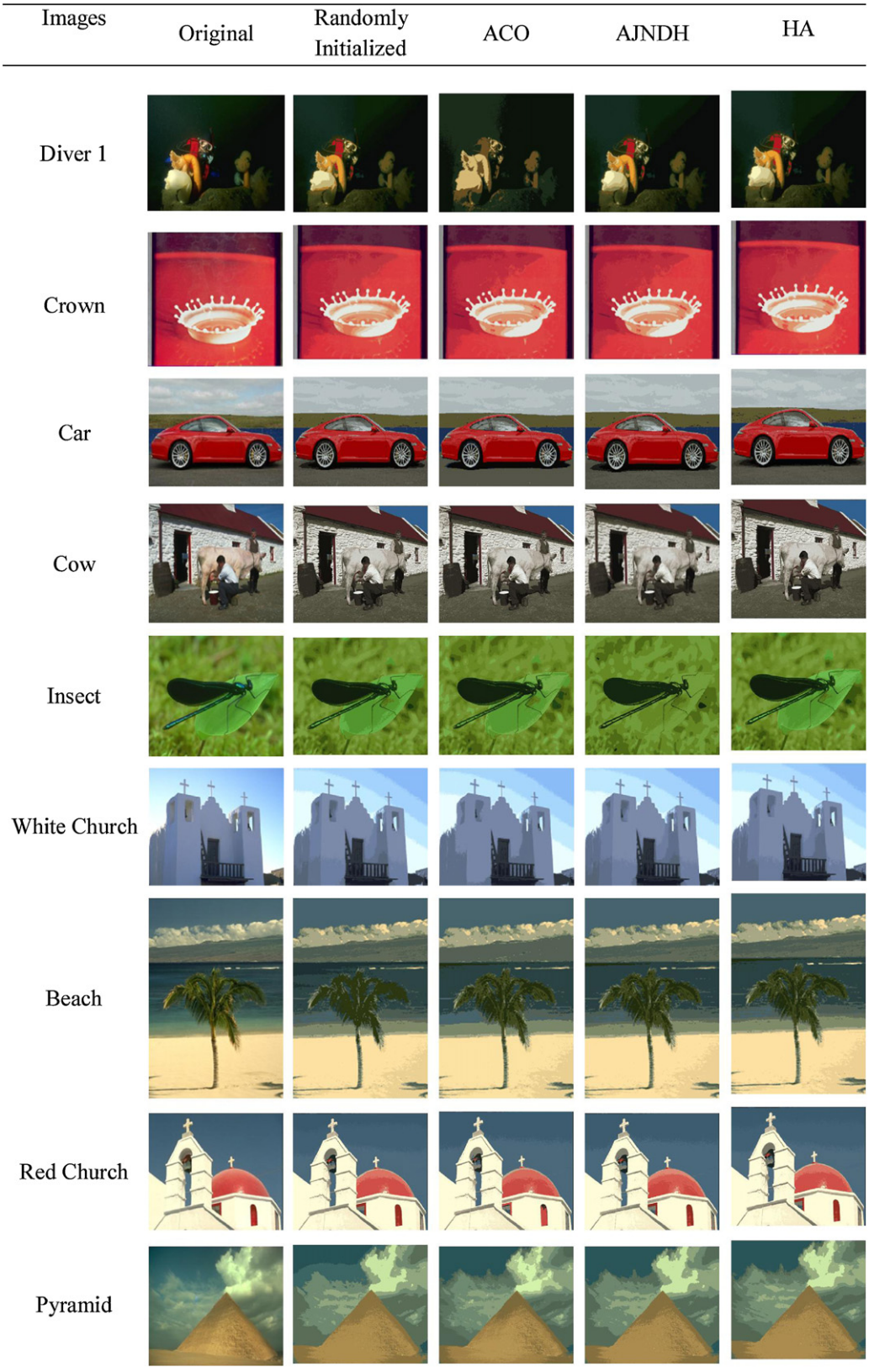

Fig. A1. Image segmentation results. First column: images' name. Second column: the original image. Third column: the randomly initialization scheme segmentation. Fourth column: the ACO initialization scheme segmentation. Fifth column: the AJNDH initialization scheme segmentation. Sixth column: the proposed HA initilization scheme segmentation.

Fig. A1. (*Continued*).

Table C1
Comparison of segmentation results based on the $F(I)$, $F'(I)$, and $Q(I)$ evaluation functions.

Test	Image	Initialization schemes			
		Randomly initialized (*1.0e+3)	AFHA (*1.0e+3)	AJNJDH (*1.0e+3)	RFHA (*1.0e+3)
$F(I)$	*Diver 1*	***0.2543***	1.5562	0.2899	**0.2279**
	Crown	***0.9769***	1.4722	1.4175	**0.9768**
	Car	0.2244	0.2942	***0.2112***	**0.1896**
	Cow	***0.8379***	***0.8379***	**0.5235**	***0.8379***
	Insect	***0.2892***	0.4287	0.8857	**0.2753**
	White Church	***0.3746***	0.6051	0.3890	**0.3522**
	Beach	0.2445	***0.2366***	**0.2216**	**0.2216**
	Red Church	***0.3295***	0.4609	0.3659	**0.3192**
	Pyramid	1.0677	**0.8503**	1.2122	***1.0512***
	Drum Players	0.9319	**0.5079**	***0.5763***	***0.5763***
	Statues	**0.1979**	0.7893	0.3427	***0.2024***
	Eagle	***0.2683***	0.5713	0.4073	**0.2598**
	White Boat	**0.9049**	***1.1161***	***1.1161***	**0.9049**
	Diver 2	1.12299	1.3849	**0.6832**	***0.7903***
	Island	0.8632	**0.8196**	0.8457	***0.8456***
	Cactus	0.3576	0.7225	***0.3415***	**0.2950**
	Horses	***1.00617***	**0.7023**	1.0401	**0.7023**
	Building	1.0406	**0.7814**	1.1187	***0.9126***
	Onion	**0.1613**	0.1979	0.1859	***0.1723***
	Pegion	0.5044	***0.5036***	0.8647	**0.4767**

Test	Image	Initialization schemes			
		Randomly initialized (*1.0e+2)	AFHA (*1.0e+2)	AJNJDH (*1.0e+2)	RFHA (*1.0e+2)
$F'(I)$	*Diver 1*	***0.2579***	1.6031	0.2947	**0.2313**
	Crown	**0.9870**	1.4918	***1.4367***	**0.9870**
	Car	0.2286	0.3028	***0.2158***	**0.1934**
	Cow	***0.8463***	***0.8463***	**0.5272**	***0.8463***
	Insect	***0.2973***	0.4397	0.9079	**0.2814**
	White Church	***0.3818***	0.6176	0.3961	**0.3581**
	Beach	0.2510	***0.2443***	**0.2280**	**0.2280**
	Red Church	***0.3336***	0.4673	0.3713	**0.3230**
	Pyramid	1.2559	1.0062	***0.9935***	**0.5532**
	Drum Players	0.9368	**0.5111**	***0.5806***	***0.5806***
	Statues	***0.2714***	0.5800	0.4125	**0.2628**
	Eagle	**0.9179**	***1.1322***	***1.1322***	**0.9179**
	White Boat	1.1287	1.3957	**0.6867**	***0.7944***
	Diver 2	0.8814	**0.8344**	0.8634	***0.8632***
	Island	0.3599	0.7337	***0.3442***	**0.2974**
	Cactus	1.0190	***0.7129***	1.0555	**0.7127**
	Horses	1.0534	**0.7906**	1.1340	***0.9238***
	Building	**0.1672**	0.2061	0.1935	***0.1784***
	Onion	0.5138	***0.5134***	0.8879	**0.4860**
	Pegion	***0.7412***	1.3732	0.8211	**0.6590**

Test	Image	Initialization schemes			
		Randomly initialized (*1.0e+3)	AFHA (*1.0e+3)	AJNJDH (*1.0e+3)	RFHA (*1.0e+3)
$Q(I)$	*Diver 1*	1.0157	3.8594	***0.8117***	**0.7672**
	Crown	***3.7284***	4.0840	3.9176	**2.8727**
	Car	0.3986	0.3948	***0.3238***	**0.3109**
	Cow	**10.9058**	**10.9058**	***47.7409***	**10.9058**
	Insect	**0.4826**	0.6261	1.5019	***0.5788***
	White Church	***0.7637***	1.5202	0.8091	**0.6961**
	Beach	0.2959	***0.2847***	**0.2645**	**0.2645**
	Red Church	0.9074	1.0013	***0.9044***	**0.8755**
	Pyramid	3.5723	***1.6213***	3.3046	**1.4464**
	Drum Players	218.3890	121.7550	**65.2754**	***65.3106***
	Statues	1.7088	1.7838	**1.1837**	***1.6753***
	Eagle	2.4176	**2.3296**	**2.3296**	***2.4164***
	White Boat	284.5980	**34.5471**	256.8220	***22.31620***
	Diver 2	1.6661	1.6824	***1.6557***	**1.6552**
	Island	134.7760	**2.5992**	61.2981	***52.8101***
	Cactus	6.4033	**2.9239**	***3.7440***	**2.9239**
	Horses	3.4558	3.0736	***2.9547***	**2.7117**
	Building	**0.1156**	0.1523	0.1467	***0.1459***
	Onion	1.5045	***1.4099***	1.7085	**1.3253**
	Pegion	4.634.8	2.9134	***2.8578***	**2.6959**

References

[1] N.R. Pal, S.K. Pal, A review on image segmentation techniques, Pattern Recognition 26 (1993) 1277–1294.
[2] H.D. Cheng, X.H. Jiang, Y. Sun, J. Wang, Color image segmentation: advances and prospects, Pattern Recognition 34 (2001) 2259–2281.
[3] B. Bhanu, J. Peng, Adaptive integrated image segmentation and object recognition, IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews 30 (2000) 427–441.
[4] S. Jin, W. Kim, J. Jeong, Fine directional de-interlacing algorithm using modified sobel operation, IEEE Transactions Consumer Electronics 54 (2008) 587–862.
[5] S. Yang, Y. Kim, J. Jeong, Fine edge-preserving technique for display devices, IEEE Transactions Consumer Electronics 54 (2008) 1761–1769.
[6] J. Leimer, Design factors in the development of an optical character recognition machine, IRE Transactions on Information Theory 8 (1962) 167–171.
[7] Y.-K. Chen, Y.-F. Chang, Global automatic thresholding with edge information and moving average on histogram, in: IEEE International Symposium on Signal Processing, Information Technology, 2005, pp. 731–736.
[8] D.-S. Kim, I.-J. Jeon, S.-Y. Lee, P.-K. Rhee, D.-J. Chung, Embedded face recognition based on fast genetic algorithm for intelligent digital photography, IEEE Transactions on Consumer Electronics 52 (2006) 726–734.
[9] B. Hiew, A. Teoh, D. Ngo, Pre-processing of fingerprint images captured with a digital camera, in: International Conference on Control Automation, Robotics and Vision, 2006, pp. 1–6.
[10] W. Cui, G. Wu, R. Hua, H. Yang, The research of edge detection algorithm for Fingerprint images, in: World Automation Congress, 2008, pp. 1–5.
[11] S.E. Umbaugh, R.H. Moss, W.V. Stoecker, G.A. Hance, Automatic color segmentation algorithms-with application to skin tumor feature identification, IEEE Engineering in Medicine and Biology Magazine 12 (1993) 75–82.
[12] G.A. Hance, S.E. Umbaugh, R.H. Moss, W.V. Stoecker, Unsupervised color image segmentation: with application to skin tumor borders, IEEE Engineering in Medicine and Biology Magazine 15 (1996) 104–111.
[13] R. Leitch, M. Gallanti, Task classification for knowledge-based systems in industrial automation, IEEE Transactions on Systems, Man and Cybernetics 22 (1992) 142–152.
[14] V.P. Dinesh Kumar, T. Thomas, Clustering of invariance improved legendre moment descriptor for content based image retrieval, in: International Conference on Signal Processing, Communications and Networking, 2008, pp. 323–327.
[15] G.C. Karmakar, L.S. Dooley, A generic fuzzy rule based image segmentation algorithm, Pattern Recognition Letters 23 (2002) 1215–1227.
[16] E.H. Ruspini, Numerical methods for fuzzy clustering, Information Sciences 2 (1970) 319–350.
[17] J.C. Dunn, A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters, Journal of Cybernetics 3 (1973) 32–57.
[18] J.C. Bezdek, Pattern Recognition with Fuzzy Objective Function Algorithms, Kluwer Academic Publishers, Norwell, Massachusetts, 1981.
[19] D.-W. Kim, K.H. Lee, D. Lee, A novel initialization scheme for the fuzzy C-means algorithm for color clustering, Pattern Recognition Letters 25 (2004) 227–237.
[20] J.C. Bezdek, M.R. Pal, J. Keller, R. Krisnapuram, Fuzzy Models and Algorithms for Pattern Recognition and Image Processing, Kluwer Academic Publishers, Norwell, Massachusetts, 1999.
[21] S. Khan, Cluster center initialization algorithm for K-means clustering, Pattern Recognition Letters 25 (2004) 1293–1302.
[22] Z.-B. Wang, R.-H. Lu, A new algorithm for image segmentation based on fast fuzzy C-means clustering, in: Proceedings of the 2008 International Conference on Computer Science and Software Engineering IEEE Computer Society, 2008, pp. 14–17.
[23] Y. Guo, H.D. Cheng, W. Zhao, Y. Zhang, A novel image segmentation algorithm based on fuzzy C-means algorithm and neutrosophic set, in: Proceedings of the 11th Joint Conference on Information Science, 2008, pp. 1–6.
[24] F. Smarandache, A Unifying Field in Logics: Neutrosophic Logic. Neutrosophy, Neutrosophic Set, Neutrosophic Probability, 4th ed., American Research Press, Rehoboth, 2005.
[25] V. Patrascu, New fuzzy color clustering algorithm based on HSL similarity, in: IFSA-EUSFLAT, 2009, pp. 48–52.
[26] K. Bhoyar, O. Kakde, Color image segmentation based on the JND color histogram, International Journal of Image Processing (IJIP) 3 (2011) 283–292.
[27] Z. Yu, O.C. Au, R. Zou, W. Yu, J. Tian, An adaptive unsupervised approach toward pixel clustering and color image segmentation, Pattern Recognition 43 (2010) 1889–1906.
[28] K. Chen, Y. Ma, L. Jun, Segmentation by fusion of self-adaptive SFCM cluster in multi-color space components, International Journal of Image Processing (IJIP) 6 (2012) 157–166.
[29] K.S. Chuang, H.L. Tzeng, S. Chen, J. Wu, T.J. Chen, Fuzzy C-means clustering with spatial information for image segmentation, Computerized Medical Imaging and Graphics 30 (2006) 9–15.
[30] S.F. Bahght, S. Aljahdali, E.A. Zanaty, A.S. Ghiduk, A. Afifi, A new validity index for fuzzy C-means for automatic medical image clustering, International Journal of Computer Applications 38 (2012) 1–8.
[31] B. Sowmya, B. Sheela Rani, Colour image segmentation using fuzzy clustering techniques and competitive neural network, Applied Soft Computing 11 (2011) 3170–3178.
[32] R. Krishnapuram, J.M. Keller, The possibilistic C-means algorithm: insights and recommendations, IEEE Transactions on Fuzzy Systems 4 (1996) 385–393.

[33] S. Behnke, N.B. Karayiannis, Competitive neural trees for pattern classification, IEEE Transactions on Neural Networks 9 (1998) 1352–1369.
[34] S. Ray, R.H. Turi, Determination of number of clusters in K-means clustering and application in colour image segmentation, in: 4th International Conference on Advances in Pattern Recognition and Digital Techniques (ICAPRDT'99), 1999, pp. 137–143.
[35] J.C. Bezdek, Pattern Recognition with Fuzzy Objective Functions, Plenum Press, New York, 1981.
[36] V. Patrascu, A generalization of fuzzy C-means algorithm using a new dissimilarity function, in: The 11-th International Fuzzy Systems Association World Congress, Beijing, China, 2005, pp. 591–596.
[37] V. Pătraşcu, Fuzzy image segmentation based on triangular function and its *n*-dimensional extension, in: M. Nachtegael, D. Van der Weken, E. Kerre, W. Philips (Eds.), Soft Computing in Image Processing, Springer, Berlin Heidelberg, 2007, pp. 187–207.
[38] G.W. Wyszecki, S.W. Stiles, Color Science: Concepts and Methods, Quantitative Data and Formulas, Wiley, New York, 1982.
[39] G. Paschos, Perceptually uniform color spaces for color texture analysis: an empirical evaluation, IEEE Transactions on Image Processing 10 (2001) 932–937.
[40] K.Y. Song, J. Kittler, M. Petrou, Defect detection in random colour textures, Image and Vision Computing 14 (1996) 667–683.
[41] C.-C. Chang, L.-L. Wang, Color texture segmentation for clothing in a computer-aided fashion design system, Image and Vision Computing 14 (1996) 685–702.
[42] Y. Gong, G. Proietti, C. Faloutsos, Image indexing and retrieval based on human perceptual color clustering, in: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1998, pp. 578–583.
[43] A. Hanbury, Constructing cylindrical coordinate colour spaces, Pattern Recognition Letters 29 (2008) 494–500.
[44] H.-D. Cheng, Y. Sun, A hierarchical approach to color image segmentation using homogeneity, IEEE Transactions on Image Processing 9 (2000) 2071–2082.
[45] P. Loo, C. Tan, Adaptive region growing color segmentation for text using irregular pyramid, in: S. Marinai, A. Dengel (Eds.), Document Analysis Systems VI, Springer, Berlin Heidelberg, 2004, pp. 264–275.
[46] S.N. Sulaiman, N.A.M. Isa, Adaptive fuzzy-K-means clustering algorithm for image segmentation, IEEE Transactions on Consumer Electronics 56 (2010) 2661–2668.
[47] J. Liu, Y.-H. Yang, Multiresolution color image segmentation, IEEE Transactions on Pattern Analysis and Machine Intelligence 16 (1994) 689–700.
[48] M. Borsotti, P. Campadelli, R. Schettini, Quantitative evaluation of color image segmentation results, Pattern Recognition Letters 19 (1998) 741–747.
[49] Arbela, X.P. ez, M. Maire, C. Fowlkes, J. Malik, Contour detection and hierarchical image segmentation, IEEE Transactions on Pattern Analysis and Machine Intelligence 33 (2011) 898–916.
[50] S. Alpert, M. Galun, R. Basri, A. Brandt, Image segmentation by probabilistic bottom-up aggregation and cue integration, in: IEEE Conference on Computer Vision and Pattern Recognition, 2007, pp. 1–8.

Khang Siang Tan received his B. Eng. degree in Mechatronic Engineering from Universiti Sains Malaysia (USM), Malaysia in the year 2009. In year 2011, he received his M.Sc. degree in Electrical and Electronic Engineering from the same university. During his time in pursuing his M.Sc. degree, he was attached to the Imaging and Intelligent Systems Research Team (ISRT), School of Electrical and Electronic Engineering, USM. His research interests include image enhancement and image segmentation.

Wei Hong Lim received his B. Eng. degree in Mechatronic Engineering with First Class Honors from USM, Malaysia in the year 2011. He is currently pursuing his Ph.D. degree in Electrical and Electronic Engineering and is attached to the Imaging and Intelligent System Research Team (ISRT), School of Electrical and Electronic Engineering, USM. His research interests include digital image processing and artificial intelligent.

Nor Ashidi Mat Isa received the B. Eng. degree in Electrical and Electronic Engineering with First Class Honors from USM in 1999. In 2003, he went on to receive his Ph.D. degree in Electronic Engineering (majoring in Image Processing and Artificial Neural Network). He is currently an Associate Professor and lecturing at the School of Electrical and Electronic Engineering, USM. His research interests include intelligent systems, image processing, neural network, biomedical engineering, intelligent diagnostic systems and algorithms. As of now, he has led his ISRT research group to publish at both national and international arena. Their contributions can be found in numerous journals, chapters in books and proceedings.