

Segmentation of color images based on the gravitational clustering concept

H. C. Yung

H. S. Lai

University of Hong Kong
Department of Electrical and Electronic
Engineering
Pokfulam Road
Hong Kong
E-mail: nyung@eee.hku.hk

Abstract. A new clustering algorithm derived from the Markovian model of the gravitational clustering concept is proposed that works in the RGB measurement space for color image. To enable the model to be applicable in image segmentation, the new algorithm imposes a clustering constraint at each clustering iteration to control and determine the formation of multiple clusters. Using such constraint to limit the attraction between clusters, a termination condition can be easily defined. The new clustering algorithm is evaluated objectively and subjectively on three different images against the K-means clustering algorithm, the recursive histogram clustering algorithm for color (also known as the multi-spectral thresholding), the Hedley-Yan algorithm, and the widely used seed-based region growing algorithm. From the evaluation, it is observed that the new algorithm exhibits the following characteristics: (1) its objective measurement figures are comparable with the best in this group of segmentation algorithms; (2) it generates smoother region boundaries; (3) the segmented boundaries align closely with the original boundaries; and (4) it forms a meaningful number of segmented regions. © 1998 Society of Photo-Optical Instrumentation Engineers. [S0091-3286(98)02803-7]

Subject terms: image segmentation; clustering; gravitational clustering; Markovian model; force effective function; RGB color space; objective evaluation; boundaries; segmented regions.

Paper 37027 received Feb. 27, 1997; revised manuscript received Jul. 11, 1997; accepted for publication Sep. 25, 1997.

1 Introduction

Image segmentation is one of the important steps often taken preceding image analysis. Its main purpose is to extract objects of interest contained in an image according to criteria such as similarity in pixel intensity, color values, texture, gradient, edges, histogram modes, or some statistical behavior. Pixels that satisfy one or more of these criteria are then grouped together where each of these may represent a complete or part of an object.¹⁻⁴ Over the past few years, different segmentation algorithms have been reported which can roughly be classified into three major categories: spatial segmentation,⁵⁻⁹ edge-based segmentation,¹⁰⁻¹² and segmentation by clustering in a measurement space.¹³⁻¹⁹ In general, these algorithms differ by how the pixel groups are formed and which criterion is being used.

Typical representatives of spatial segmentation are the seed-based algorithms^{3,5,6} and split-and-merge algorithms.⁶⁻⁹ The former algorithms are conceptually and computationally simple and their results are acceptable. Unfortunately, they are also rather sensitive to noise and have the tendency of forming a large number of small regions when real images are concerned. The latter algorithms are less sensitive to noise or sharp spatial variations, but the resulting object boundaries often appear undesirably blocky. For edge-based segmentation, the major problem lies with its sensitivity to noise and poor quality of the edges in general.³

On the other hand, segmentation by clustering in a measurement space such as a histogram space¹⁴⁻¹⁶ or one of the many color spaces¹⁷⁻¹⁹ has at times received considerable attention. For instance, an early work in the mid-1970s by Hartigan¹³ highlighted that clustering can be based on minimizing the total sum errors of K clusters representing all the pixels in the image, where individual sum error is calculated as the sum of the errors between the pixels in a cluster and the mean of these pixels (K-means algorithm). Around the same time, Ohlander¹⁴ proposed a recursive histogram-directed spatial clustering scheme that aimed to separate one mode of the histogram from another until all the segmented regions can each be represented by a single mode histogram. As real image histograms are spiky in nature, the results from this method tend to be noisy with numerous small regions. To alleviate this problem, the original histogram may be smoothed so that the resulting separation of the histogram modes is more accurate and less noise prone. Unavoidably, the smoothing operation partly determines the outcome of the segmentation. Stemmed from the same idea, Celenk¹⁵ proposed a 1-D histogram segmentation method in the perceptually uniform color space (LAB), of which the segmentation was heavily determined by how accurately the single mode histograms are extracted. A very similar algorithm known as the multi-spectral thresholding algorithm was described in Sonka et al.¹ for color images. Hedley and Yan¹⁶ also used histogram analysis but combined clustering with edge detection in which only low gradient pixels are subject to the analysis. Their results were compared with the K-means algo-

algorithm on a well-defined color road map where they claimed a better performance on detailed information and clustering time. However, the test image has very good contrast and sharp edges, and it is difficult to assess its performance against the K-means algorithm when it is not so.

From a different perspective, Lim and Lee¹⁷ used a variant of the RGB color space where thresholding was used for coarse segmentation and a fuzzy c-mean method was used for fine segmentation. Although numerous small regions resulted, the boundaries were fairly smooth. Similarly, Beni and Liu¹⁸ proposed a fuzzy clustering algorithm with minimal biases and used the maximum entropy principle to maximize the entropy of the centroids. Along this line, Mathews and Hearne¹⁹ proposed a non-metric fitness measurement which predicted the data properties by minimizing the conditional prediction errors where their algorithm was compared with the K-means algorithm and showed that they can achieve a smaller and more stable error rate on two Gaussian distributed data sets, though no real images were tested in this case.

In general, measurement space clustering has a number of merits. First, the transformation from the spatial domain is usually many-to-one, resulting in a reduced data set which has computing advantage. Second, although many of the clustering methods are inherently recursive or iterative, most of them generate reasonably smooth region boundaries and are less prone to noise and local boundary variations.^{16,20} For these reasons, clustering in a measurement space has been widely used for image segmentation, particularly when color images are concerned.

In this paper, a new clustering algorithm derived from the Markovian model of the gravitational clustering concept²¹ is proposed that works in the RGB measurement space for color image. To enable the model to be applicable in image segmentation, the new algorithm imposes a clustering constraint at each clustering iteration to control and determine the formation of multiple clusters. Using such constraint to limit the attraction between clusters, a termination condition can be easily defined, for example, when the attraction between clusters is zero. Conventionally, evaluation of segmentation results is never straightforward, but the new clustering algorithm is evaluated objectively and subjectively on three different images against the K-means clustering algorithm,¹³ the recursive histogram clustering algorithm for color (also known as the multi-spectral thresholding),^{1,3,14} the Hedley-Yan algorithm,¹⁶ and the widely used seed-based region growing algorithm.^{4,5} The Liu and Yang⁸ method is employed for the objective assessment, while three visual features are inspected in the subjective assessment. From the evaluation, it is observed that the new algorithm exhibits the following characteristics: (1) its objective measurement figures are comparable with the best in this group of segmentation algorithms; (2) it generates smoother region boundaries; (3) the segmented boundaries align closely with the original boundaries; and (4) it forms a meaningful number of segmented regions.

The organization of this paper is as follows: Sec. 2 presents a detailed description of the new algorithm; Sec. 3 briefly describes the evaluation conditions, test images and evaluation methods used; Sec. 4 provides details of the objective evaluation and the subjective inspection results; and

Sec. 5 concludes this paper with detailed discussions on the comparison.

2 Constrained Gravitational Clustering

2.1 Concept

The concept of this constrained gravitational clustering (CGC) algorithm is built upon the gravitational clustering method originally proposed by Wright²¹ for data analysis. In brief, gravitational clustering defines a finite system of particles in space, each with a specified initial location, a zero velocity, a given mass, and a negligible volume, that converge to the centroid of the system due to the gravitational attraction between the masses. The physical model of gravitational clustering includes complex calculations in velocity, accelerations, and collisions of particles, where Wright simplified it to a Markovian model in which movements of particles depends only on the locations and masses of the remaining particles and not on any past history. Thus the Markovian model of gravitational attraction between two particles p_k and p_i is defined as:

$$\mathbf{F}_{p_k p_i} = -G \frac{m_k \times m_i}{|\mathbf{s}_k - \mathbf{s}_i|^3} (\mathbf{s}_k - \mathbf{s}_i) \quad (1)$$

where G is the gravitational constant, m is the mass, and \mathbf{s} is the location vector. Due to this attraction, particles are clustered until a single mass is left in the system. The optimality of the clustering is determined in terms of the percentage of particles clustered versus the time taken to cluster them. As it is, this model cannot be applied directly to image segmentation because of the fact that optimum result can only be determined after all the particles are clustered into a single mass and the clustering time is calculated. Using clustering time as an indicator of optimality almost certainly has no correspondence to optimal image segmentation results, while the whole clustering process is potentially time-consuming for large digital images generally used in many applications nowadays.

The CGC algorithm resolves these two issues by first introducing a new constraint which makes the clustering meaningful with respect to the spatial property of images, and second, mapping the spatial image data into the RGB color space for data reduction. In the former case, a force effective function (FEF) is defined such that it governs the effectiveness of the attractive force between particles. In essence, when the net force acting on a particle is concerned, only those particles that satisfy the FEF constraint contribute to the calculation of the net force. According to this nonzero net force, the particle moves to a new location, and if more than one particle share the same location, these particles are merged to form a new particle. The clustering process iterates by determining the new net force on each particle until the net forces on all the particles are zero. With the FEF, the number of particles to be considered in each calculation is substantially reduced while the number of clusters formed corresponds to the number of segmented regions. For the latter case, the use of the RGB color space seems appropriate as true color images are in RGB, and the mapping itself is many-to-one. Thus, the resulting number of particles becomes smaller than the total number of image pixels and is independent of the image size. Moreover, scaling of the measurement space resolution is also possible

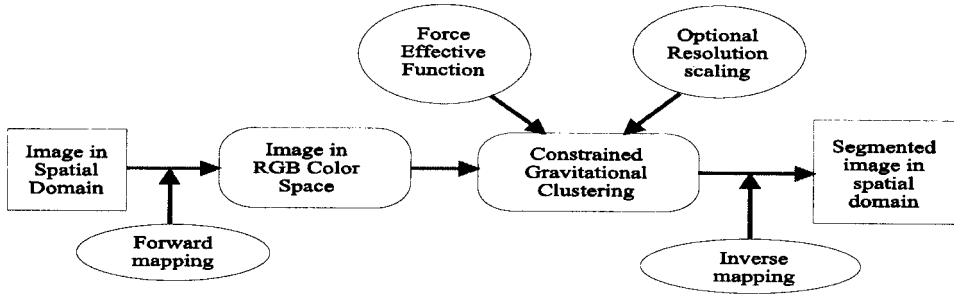


Fig. 1 Conceptual flow of the constrained gravitational clustering algorithm.

if further data reduction is desired. Figure 1 depicts the conceptual flow of the constrained gravitational clustering algorithm for image segmentation.

2.2 Algorithm

For a true color image g of size $M \times N$, each pixel is defined by three colors: red, green and blue. Therefore, the pixel value of g at (x,y) can be represented by

$$g(x,y) = \begin{pmatrix} g_R(x,y) \\ g_G(x,y) \\ g_B(x,y) \end{pmatrix}, \tag{2}$$

where $x=1,\dots,M$, $y=1,\dots,N$, and $g_R(x,y), g_G(x,y), g_B(x,y) \in [0,\dots,l-1]$, where l is the number of color levels. The mapping of the image pixels to the RGB color space can be illustrated in Fig. 2, where each pixel in the spatial domain is mapped to the measurement space according to their RGB values. (Note that four clusters are already noticeable after the mapping.) Each pixel represents a mass of 1. Whenever there is more than one pixel mapped to the same location in the RGB space, the mass of that particle before clustering equals to the total number of pixels mapped to it. After the mapping, each entry in the RGB space is regarded as a single particle.

For the clustering, let us define particle p_k in the RGB space to be characterized by two parameters: mass (m_k), and location vector (s_k), as given by:

$$p_k = \begin{pmatrix} s_k \\ m_k \end{pmatrix}, \tag{3}$$

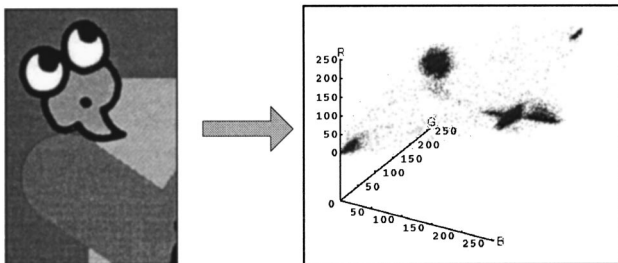


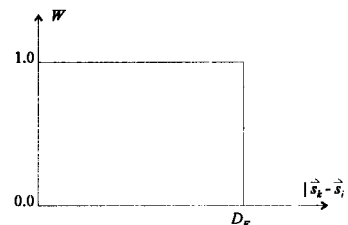
Fig. 2 Mapping of an image into the RGB color measurement space.

for $k=1,\dots,Q$ where $Q \ll M \times N$. The gravitational attraction, $F_{p_k p_i}$, between two particles p_k and p_i is defined by the Markovian model as given in Eq. (1). Therefore, the net force on particle, p_k due to all the other particles satisfying the FEF constraint in the RGB space can be written as follows:

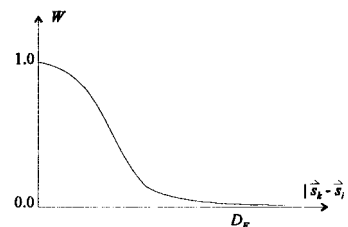
$$F_{p_k} = \sum_{\substack{k=1 \\ k \neq i}}^n F_{p_k p_i} \times W(p_k, p_i), \tag{4}$$

where $W(\cdot)$ is the FEF, which can take the form of a step function or a Gaussian function as illustrated in Fig. 3, where the value D_F is the deciding parameter.

In one extreme, if D_F is large enough to include all the particles in the RGB space, the CGC algorithm becomes the original gravitational clustering algorithm. In the other extreme, if $D_F=0$, each particle in the measurement space represents a cluster, hence no further clustering is possible. However, if D_F is in between the two extremes, it determines how the clustering is to be performed and the final number of clusters. To illustrate this point, Fig. 4 depicts the clustering results of the image in Fig. 2 using different values of D_F and shows how the number of clusters formed



(a) Step function



(b) Gaussian function

Fig. 3 Force effective functions.

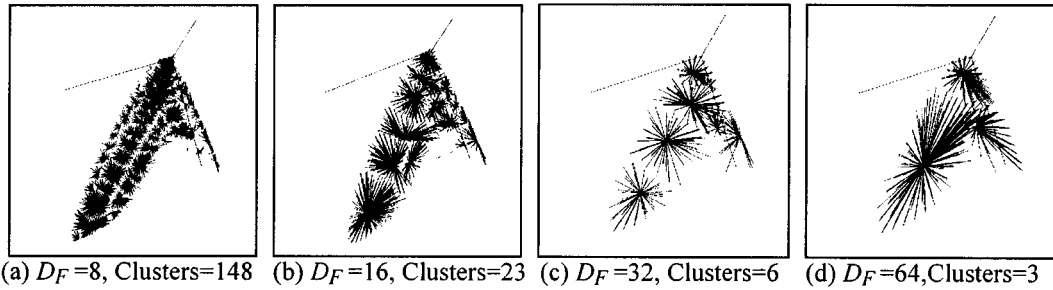


Fig. 4 Clustering at different values of D_F .

can be controlled. For small D_F , the number of particles considered for calculating the net force per particle is small and a larger number of clusters are formed as a result. As D_F increases, the number of particles considered in the net force calculation increases and the number of clusters formed decreases as a result. Of course, large D_F demands high computing overhead and vice versa.

With the above constraint, if the net force on p_k is non-zero, then it moves in the RGB space according to the direction and magnitude of the force. Let us define this movement to be $\Delta \mathbf{s}_k$ as a result of the net force on p_k , which is given by:

$$\Delta \mathbf{s}_k = \mathbf{u}_k t + \frac{1}{2} \mathbf{a}_k t^2, \tag{5}$$

and the acceleration to be given by:

$$\mathbf{a}_k = \frac{\mathbf{F}_{p_k}}{m_k}, \tag{6}$$

where \mathbf{u}_k is the velocity of particle p_k . According to the Markovian model, \mathbf{u}_k is assumed zero and t is normalized to 1. Therefore, the new location of p_k equals to $\mathbf{s}_k + \Delta \mathbf{s}_k$, or:

$$\mathbf{s}'_k = \mathbf{s}_k + \frac{1}{2} \frac{\mathbf{F}_{p_k}}{m_k} \tag{7}$$

where the mass of the particle remains unchanged [Fig. 5(a)] if there is no particles occupying its new location. If two particles p_k and p_j have moved to the same location as illustrated in Fig. 5(b), i.e. $\mathbf{s}'_k = \mathbf{s}'_j$, then they are merged to form a new particle p_{new} according to the following equation:

$$p_{new} = \begin{pmatrix} \mathbf{s}_{new} \\ m_{new} \end{pmatrix} = \begin{pmatrix} \mathbf{s}_{new} = \mathbf{s}'_k = \mathbf{s}'_j \\ m_k + m_j \end{pmatrix}. \tag{8}$$

After the new locations and masses for all the particles in the RGB space have been determined, the clustering process repeats by calculating the new net forces for the new particles in the new locations and new masses. This causes the particles to move yet again if any of these values are non-zero. The clustering process terminates when all the net forces in the system are zero, i.e. no movement is possible even if the process continues. The final number of particles formed in the RGB space corresponds to the number of clusters or regions in the segmented image.

During the clustering process, an internal table recording the movements and merging of particles is maintained. According to this table, a mapping of the input color intensities, (R, G, B) , to the output color intensities, (R', G', B') can be established. This mapping facilitates the inverse mapping from the RGB space back to the spatial domain. The flow of the algorithm is depicted in Fig. 6.

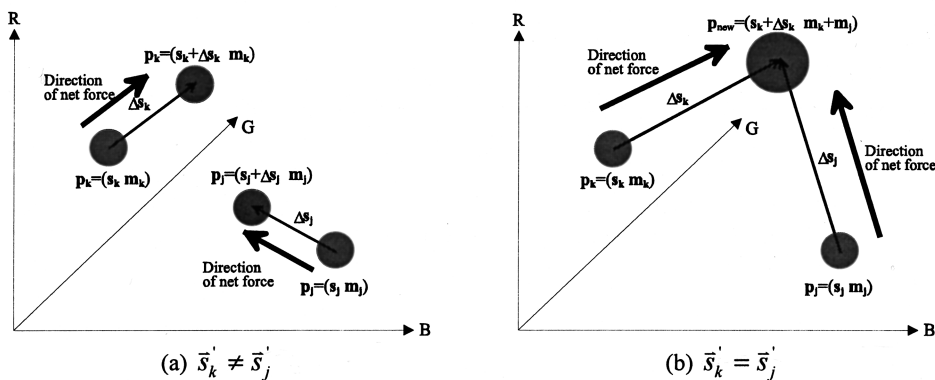


Fig. 5 Two possible cases of particle movement: (a) $\mathbf{S}'_k \neq \mathbf{S}'_j$, and (b) $\mathbf{S}'_k = \mathbf{S}'_j$.

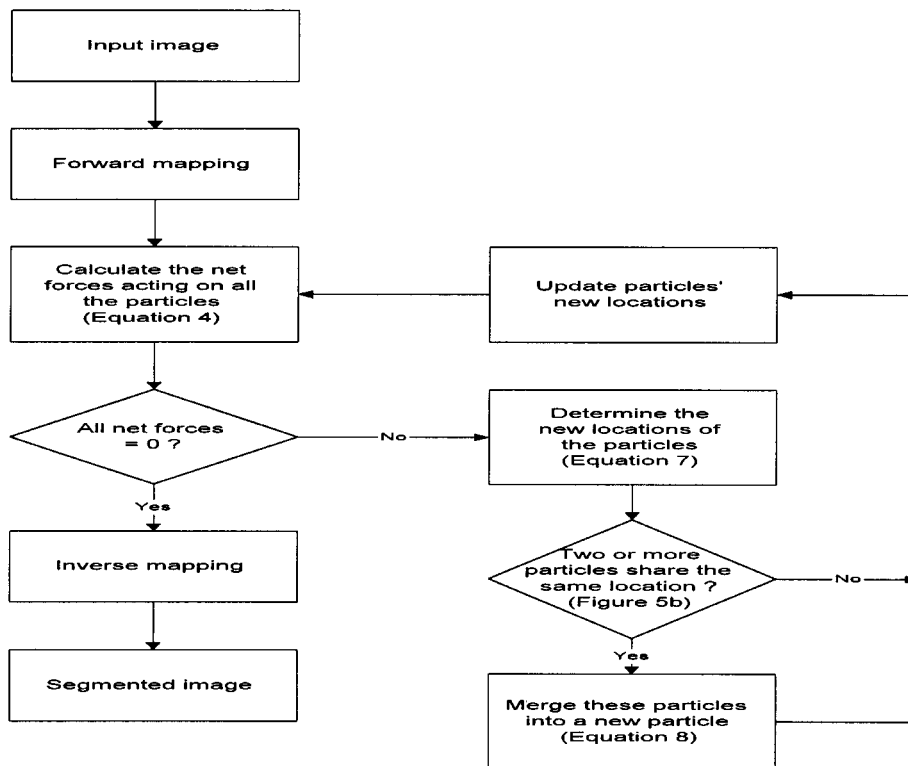


Fig. 6 Flow chart of the clustering process.

3 Evaluation Conditions

To evaluate the performance of the CGC algorithm, four other image segmentation algorithms were also implemented and tested. These algorithms are the K-means clustering (K-means),¹³ recursive histogram clustering (RHC),^{1,14} seed-based region growing (SRG),⁵ and the Hedley and Yan algorithm (SCS).¹⁶ The K-means clustering is chosen here because it has been widely used in data analysis as well as image segmentation. It has also been compared with other clustering algorithms in the past.^{16,19} For the recursive histogram-directed spatial clustering algorithm, it represents a large class of histogram clustering techniques. In our evaluation, the RHC algorithm adopts a histogram smoothing filter for reducing noise content in the histogram and employs the multispectral thresholding approach for handling color images. The SRG algorithm is chosen because it has been widely used too. It works on a simple principle and produces good results. There are two reasons for choosing the SCS algorithm. The first is that it represents a combined histogram analysis and edge detection approach. The second is that it has been compared with the K-means algorithm using the image of a road map.¹⁶ Although the edges and contrast of road map image tested are sharp, giving the SCS algorithm a certain advantage, their comparison would serve as a useful reference for the evaluation in this research.

In the evaluation, for all five algorithms and three images, two sets of results were obtained. The first set of results was taken by applying the algorithms as they are, no preprocessing or postprocessing. The second set of results were taken with a region merging algorithm acting as a postprocessor to help remove some of the spatial disconti-

nuity common in measurement space methods, and to reduce the number of region altogether.

3.1 Test Images

The three images used in the evaluation were “toon,” “cells” and “bus” (Figs. 7–9). These images belong to part of a larger picture or photograph digitized by a scanner. “Toon” is an image of a small cartoon figure surrounded by some colorful but well-defined regions. The boundaries consist of straight lines (horizontal, vertical and diagonal) and curves, and the color contrast between regions is high. In contrast, “cells” is an image of some purple-dyed white cells taken from a slide under a microscope. The image is slightly out of focus and its contrast is normal. The dyed white cells are light purple and have a dark purple nucleus each, whereas the other cells are dark



Fig. 7 “toon.”

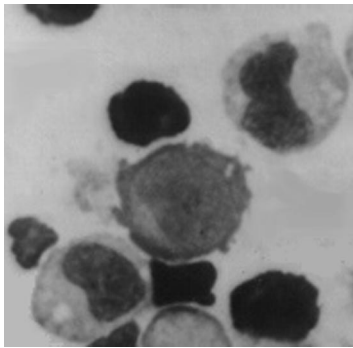


Fig. 8 “cells.”

brown. The image presents a higher degree of segmentation difficulty than the “toon” image. The “bus” image is part of a photo taken from an overhead camera looking down at a busy road. The double-decker bus is mainly white with a faint pattern on the side. The boundaries look fuzzy, but the front right side and three wheels are clearly visible. The background includes part of a road where construction work can be seen at the top of the image and broken road markings near the bottom. This image is considered to be the most difficult amongst the three. The jagged boundaries and the busy background are expected to cause noticeable boundary errors and incorrect segmentation.

3.2 Evaluation Methods

Both objective measurement and subjective inspection were used to evaluate the performance of the algorithms concerned. For the objective measurement there exist many different evaluation methods.^{7,8,22–24} Most of these methods require a reference image for the evaluation, which is not always available. Of all these methods, the one proposed by Liu and Yang⁸ is adopted in this research because it can be used to evaluate real images locally and globally without needing a reference image. However, it should also be noted that this method does not provide an exact grading of the segmentation algorithms. It merely gives a broad and general indication of how well the algorithm performs. The evaluation function L is defined as:

$$L[h(x,y)] = \frac{\sqrt{R}}{1000A_T} \sum_{i=1}^R \frac{e_i^2}{\sqrt{A_i}}, \quad (9)$$

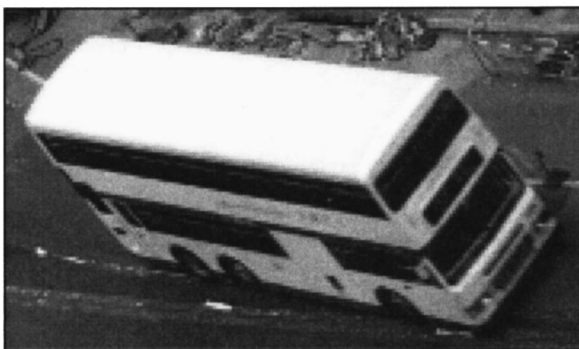


Fig. 9 “bus.”

Table 1 Results according to smallest L without region merging.

		K-means	RHC	SRG	SCS	CGC
“Toon”	Parameter	$K=5$	$W=11$	$T=40$	$T_g=0.75$	$D_F=16$
	L	0.92	61.23	0.21	37.58	0.19
	R	261	7695	1094	6011	1190
“Cells”	Parameter	$K=4$	$W=5$	$T=10$	$T_g=0.85$	$D_F=16$
	L	0.45	1.68	0.20	4.76	0.28
	R	273	1893	3966	830	487
“Bus”	Parameter	$K=3$	$W=11$	$T=10$	$T_g=0.8$	$D_F=8$
	L	2.09	4.22	0.33	5.17	1.03
	R	322	1212	7602	1152	4336

where $h(x,y)$ is the segmented image; R is the total number of regions resulted in $h(x,y)$; A_i is the number of pixels in the i 'th region and A_T is the total number of pixels in the image; e_i^2 is the color error of region i which is defined as the sum of the euclidean distance of the color vectors between the input image and the segmented image for each pixel in the region. As the color error is squared and the number of regions square-rooted, it seems fair to assume that $L[h(x,y)]$ is biased toward the effect of color errors, rather than the correct number of segmented regions.

The subjective inspection is based on three criteria: (1) smoothness of the boundaries, (2) boundary correctness, and (3) the complexity of the segmented image. In this case, smoothness refers to the compactness of the edges, correctness refers to how close the segmented boundaries align with the original, and complexity refers to the number of meaningful regions formed.

4 Results

4.1 Objective Evaluation

The first set of tests was conducted on all five algorithms with the three test images (without region merging). For each algorithm, its parameter, e.g. D_F in CGC, was varied over a range where their segmented results and L for all three test images were determined. The best L value, the corresponding number of segmented regions (R), and the parameter to achieve this value for each algorithm are given in Table 1.

From Table 1, a number of observations can be made with respect to L . First, the performances of both the RHC and SCS algorithms are poor. For the “toon” image, their L values are much higher than the other three. For the other two images, their L values are high, but not exceptionally high. Second, both the SRG and CGC algorithms perform similarly for all three images and are substantially better than the others. Compared with the SRG, the CGC seems to be better on the “toon” image, slightly worse on the “cells” image and 3 times worse on the “bus” image. In this regard, the SRG algorithm is the best and most consistent, and the CGC algorithm follows closely behind. Third, the K-means algorithm performs reasonably in third place as its L varies from ~ 6 times worse than the best case on the “bus” image to only 2 times worse than the “cells”

Table 2 Relative ranking according to L without region merging.

	K-means	RHC	SRG	SCS	CGC
"Toon"	3	5	2	4	1
"Cells"	3	4	1	5	2
"Bus"	3	4	1	5	2

image. For clarity, the relative ranking of these algorithms according to L is depicted in Table 2. As it shows, the SRG algorithm is consistently better than the rest, followed by the CGC algorithm, then the K-means. Between the RHC and SCS algorithms, the RHC seems to perform slightly better than the SCS algorithm.

When the number of segmented regions is concerned, the K-means algorithm always produces the smallest R as compared with the rather large numbers by the other four algorithms in some cases. For instance, although SRG consistently gives the best L for all three images, its R in each case is large, particularly for the "bus" image. The same applies to the CGC, RHC, and SCS algorithms, at a smaller scale. The particular point can be explained by Eq. (9) that as L weighs more heavily on the total color error than R , algorithms like the K-means that produce small R but large color error would have a large L . On the other hand, algorithms such as the SRG that produce a large R but small color error would have a small L . From the segmentation point of view, it is perhaps desirable to have a small R and a small color error. Both the K-means and CGC algorithms seem to exhibit this property to some extent, but not the others. In the case of the RHC and SCS algorithms, due to the inherent noisy nature of histograms and the problem of using edge detection to differentiate high and low gradient

Table 3 Results according to smallest L with region merging.

		K-means	RHC	SRG	SCS	CGC
"Toon"	parameter	$K=7$	$W=3$	$T=40$	$T_g=0.8$	$D_F=16$
	L	0.25	0.89	0.22	1.79	0.21
	R	167	233	168	693	161
"Cells"	parameter	$K=5$	$W=5$	$T=50$	$T_g=0.85$	$D_F=16$
	L	0.28	0.34	0.25	0.57	0.24
	R	277	455	264	367	244
"Bus"	parameter	$K=3$	$W=11$	$T=80$	$T_g=0.9$	$D_F=24$
	L	1.06	1.26	0.95	1.52	0.99
	R	254	340	253	381	283

pixels, they tend to generate large R and errors at the same time, and hence large L .

To further illustrate this point, let us inspect the "cells" results, which are depicted in Fig. 10. Broadly, both the K-means and CGC results appear to be neat with sufficient details to represent the various regions in the original image. Their boundaries are smooth and reasonably correct, and their backgrounds are correctly segmented. The SRG result is fine apart from the fuzzy boundaries and incorrectly segmented background. Both the RHC and SCS results appear to have numerous small regions and noisy boundaries, with a rather untidy background as well. The severe distortion of the SCS result is clearly visible.

The second set of tests in our evaluation was to perform segmentation with region merging. The evaluation conditions and criteria are the same as before, and the corresponding results are given in Table 3, with the ranking in Table 4. From these tables, observations similar to the first test set can be made. First, the objective performance of the

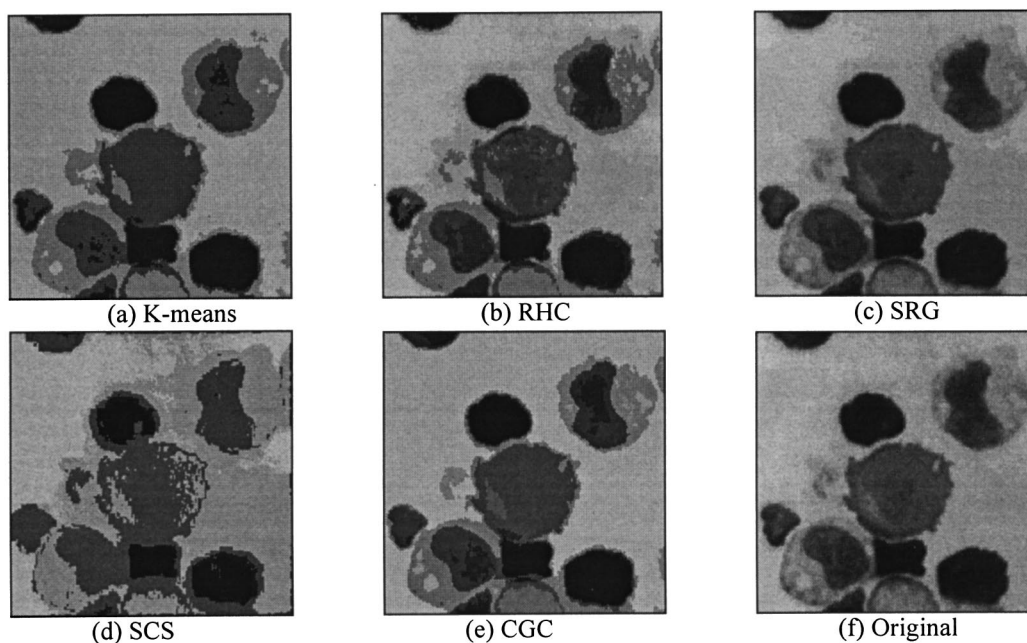
**Fig. 10** Resulting images according to smallest L without region merging.

Table 4 Relative ranking according to L with region merging.

	K-means	RHC	SRG	SCS	CGC
"Toon"	3	4	2	5	1
"Cells"	3	4	2	5	1
"Bus"	3	4	1	5	2

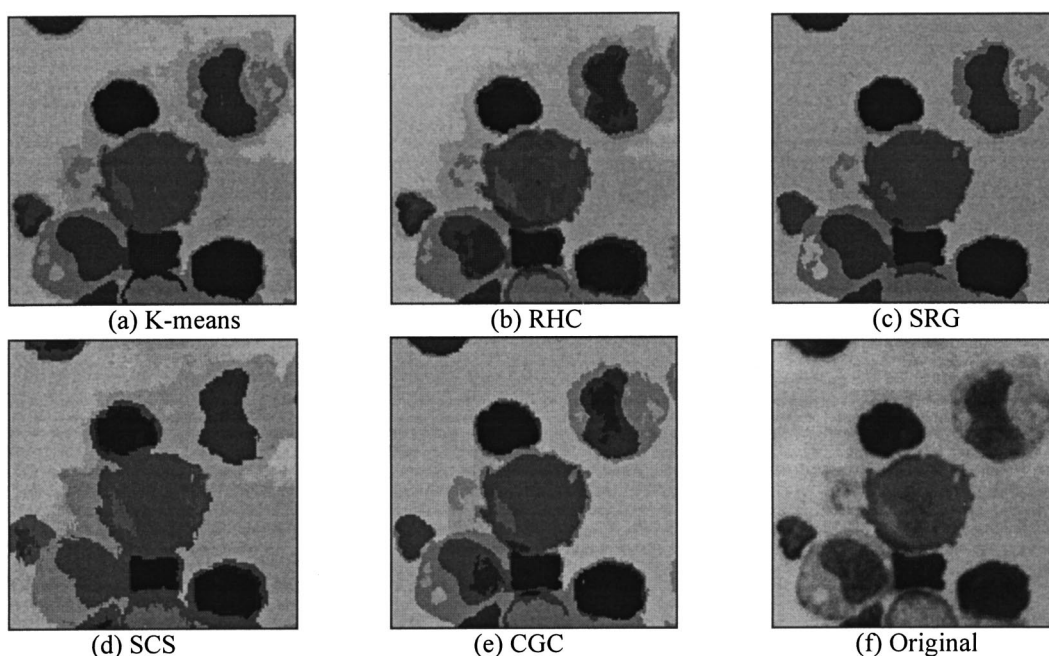
SRG and CGC algorithms are almost the same. The difference is minor and the trend is consistent for all three images. Second, the other three algorithms show a marked improvement in L overall. The K-means values are much closer to the SRG and CGC, being at worst 0.07 higher. In the case of the RHC and SCS algorithms, the L values are much more acceptable than before, although they are still poor compared with the K-means values.

The K-means, SRG, and CGC algorithms produce similar R , whereas both the RHC and SCS algorithms still produce large R . For all the algorithms, the merging of regions seems to have some positive effect on the final R and L . For instance, in the case of the CGC algorithm on the "toon" image, for the same D_F , R and L with region merging are 161 and 0.21, respectively, and without region merging are 1190 and 0.19. The difference in L is rather small, yet their R s are vastly different. This can be explained: if the algorithms tend to generate a large R before region merging, then many of these regions would be merged after the postprocessing and result in a smaller R . On the other hand, merging regions would likely result in an increase in color errors. So if R dominates before the merging, then merging would give a smaller L (from 1.68 to 0.34 in RHC at $W=5$), otherwise a larger L may result (from 0.19 to 0.21 in CGC at $D_F=16$).

To inspect the effect of region merging, the "cells" images in this test set are depicted in Fig. 11. Broadly, the CGC results appear to be clean with smooth and correct boundaries, as well as a correctly segmented background. For the SRG result, region merging helped to remove the boundary fuzziness, but distortion also becomes apparent, particularly for the top right hand cell. Other than that, the background and other cells have been segmented reasonably well. Both the K-means and RHC results are similar. The K-means result looks smoother with less distortion, but the background is incorrectly segmented and so is the top right hand cell. The RHC result still appears noisy even after the region merging. For the SCS result, region merging helped to merge many smaller regions, but as a result, the image appears overmerged and boundary distortion appears to be extensive.

4.2 Subjective Inspections

Figures 12 to 14 depict the visually best segmented images for each algorithm on all three images with or without region merging. The inspection criteria are based on those mentioned in Sec. 3.2. When comparing the images in Fig. 12, we can observe that the RHC and SCS results depicted in (b) and (d) are poorer than the other three. Boundaries are obviously noisy and jagged in both cases. Of the two, the RHC result shows severe errors along the diagonal and curved boundaries. In addition, the small black region on the bottom right hand corner of the image disappeared altogether. The SCS result is slightly better but the errors on the boundaries are still obvious. For the other three algorithms, on close inspection, the K-means result seems to have the sharpest straight line boundaries, which are slightly fuzzy in the SRG and CGC cases. When curve boundaries are concerned, the K-means result seems to be slightly more jagged and the boundary errors are a little

**Fig. 11** Resulting images according to smallest L with region merging.

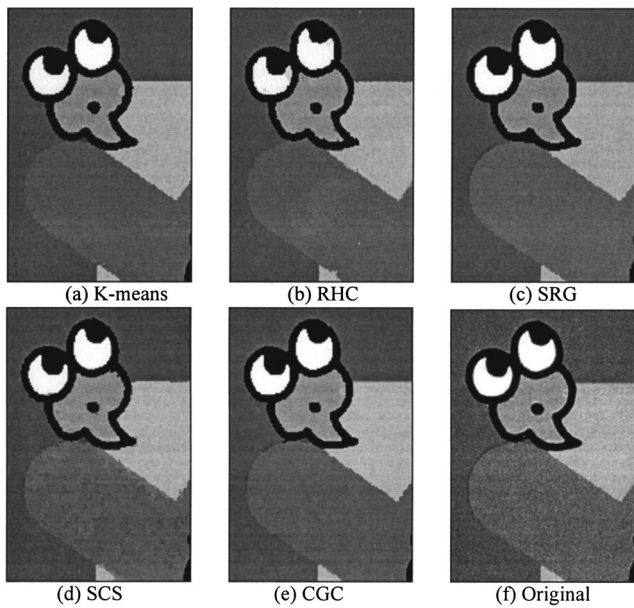


Fig. 12 Segmented results of "toon" image.

more apparent, although the overall appearance of these boundaries is still smooth and acceptable. In this case, the SRG and CGC results resemble the original reasonably closely. The only difference between these two results is that the SRG result has an overmerging problem under the small cartoon figure, just above the semicircular boundary, and the CGC has a similar problem for the boundary slightly toward the left. Taking everything into account, the CGC, K-means, and SRG results are all very similar and acceptable. The RHC and SCS results are poor and not acceptable.

When comparing the images in Fig. 13, we can observe that the worst performer in this case is the SCS algorithm. Its segmented regions suffer from a serious overmerging effect in which the middle and bottom few cells have all been merged together or with the background. Its appearance is also worsened by the erratic boundaries. A better result is obtained by the SRG algorithm. Although the SRG has the best L in the objective evaluation, its inability to correctly segment the cell at the top right hand corner, the bottom center cell, and some of the nuclei, degrades its visual quality. On the other hand, the RHC result appears to be much better than what its L value indicates. For instance, most of the regions have been correctly segmented and the whole image looks simple with reasonably correct boundaries. However, some boundaries appear to be noisy and jagged, and its background has also been segmented into a number of smaller regions. Of the remaining two, the K-means result appears to have a smaller number of regions and a correct background. The boundaries are smooth and correct, except for a mild hint of overmerging, which is noticeable around the rectangular cell at the bottom of the image. For the CGC result, it appears to be the best as all the regions have been correctly segmented with smooth and correct boundaries, and a correct background. The rectangular cell is clearly isolated and vital details in these cells have been appropriately retained. The only problem perhaps is the incomplete boundary of the bottom center cell with a dark outline and light internal region. In this case, only the K-means algorithm is able to segment this cell correctly.

When comparing the images in Fig. 14, a number of points can be noted. First, the K-means result is poor as part of the bus is merged with the background, and the boundaries of the big white bus top appear to be jagged. Second, for the RHC result, the boundaries of the white top

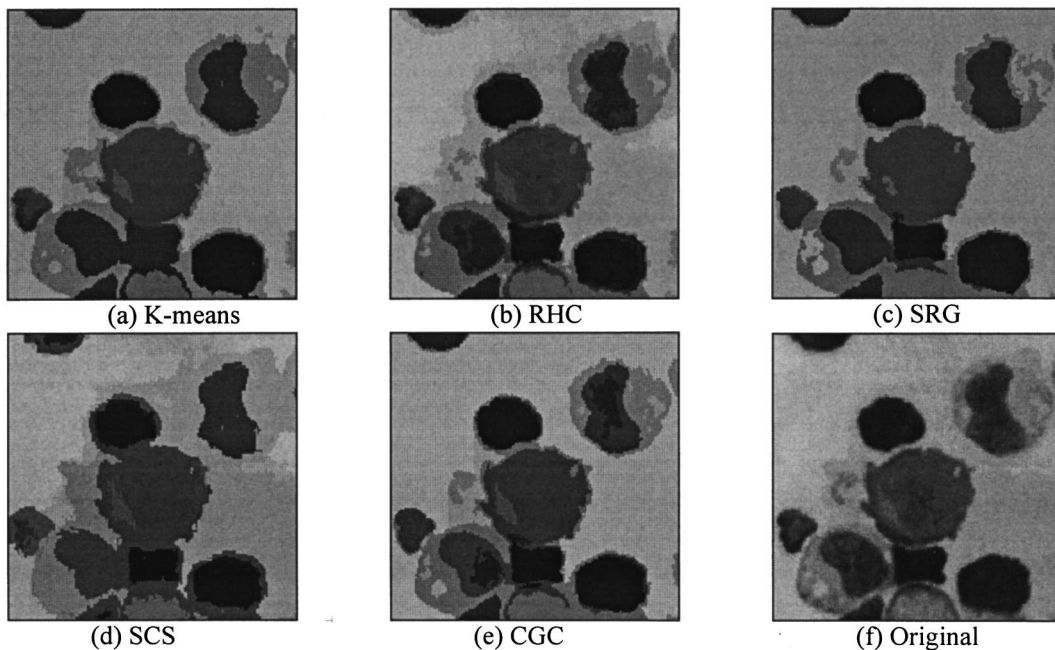


Fig. 13 Segmented results of "cells" image.

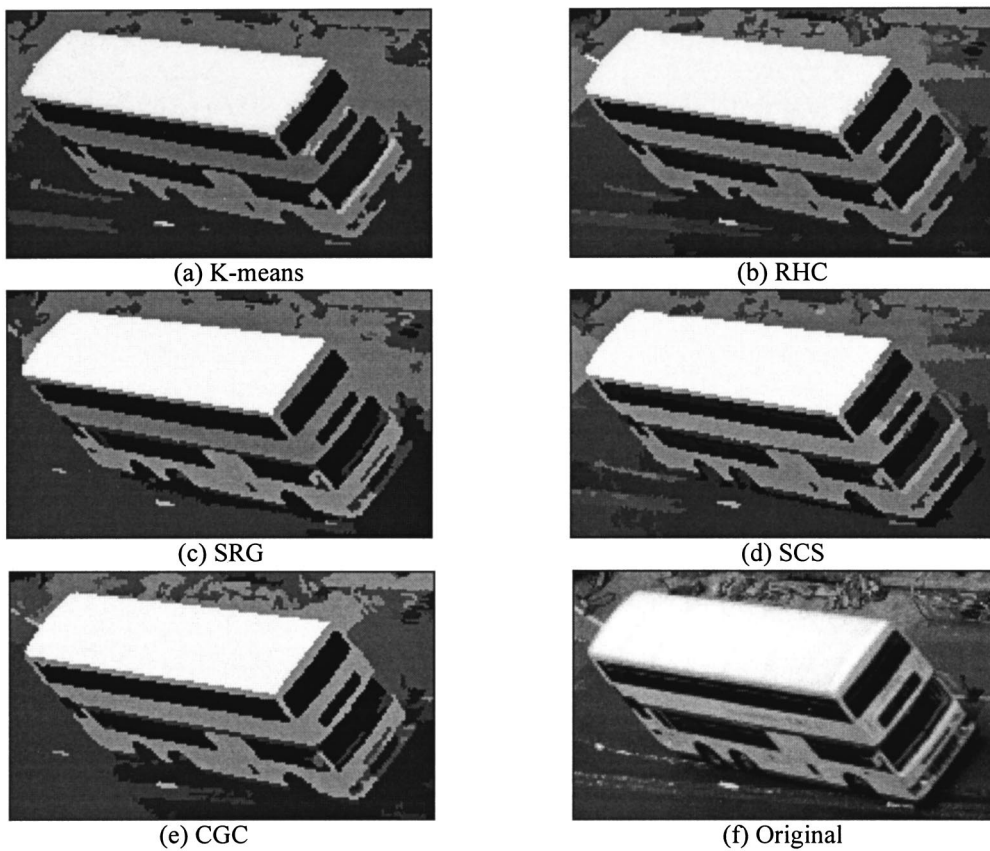


Fig. 14 Segmented results of "bus" image.

also appears to be jagged, and the background is incorrectly segmented. Third, the boundaries of the SRG result appear to be smoother, but merging of bus regions with the background is also evident. On the positive side, its background is clean and most of the bus has been correctly segmented. Fourth, the SCS result has too many details retained. For example, the front of the bus is segmented into numerous unwanted regions. This probably accounts for the large number of segmented regions remaining according to the objective measurement. As a whole, the CGC result ap-

pears to be the best with the bus and background correctly segmented. Boundaries are smooth and correct, yet without too many small regions.

In order to further study the results obtained, the results in Fig. 14 were enlarged, and the same part in each case is shown in Fig. 15. When the raw images before edge detection are compared, it can be seen that the CGC result has the best appearance with the simplest number of regions, as well as smooth and correct boundaries. The SRG result also looks quite good if not for the merged background. The

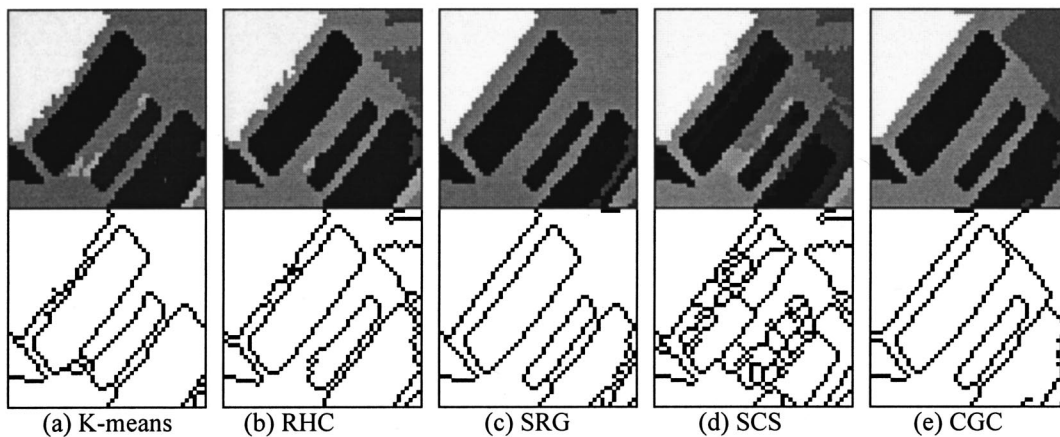


Fig. 15 Edge-detected results of portion of the "bus" image.

RHC result appears to be clean except for a few noisy boundaries. Similarly, the K-means boundaries suffer from the noise problem and the background is merged with the bus. The worst is again the SCS result which shows incorrect segmentation and boundary distortion. Their edge-detected results highlight the above points further.

In summary, the new clustering algorithm performs extremely well on both objective and subjective terms. Objectively, its evaluation values are among the smallest. Subjectively, the visual quality of its segmented results on the three test images have been consistently high where their boundaries are smooth and correct, and the results contain a meaningful number of regions in all cases. The performance of the K-means and SRG algorithms are very close. Although the SRG algorithm is better objectively, the K-means results give a slightly higher visual quality and much better number of segmented regions without region merging. With region merging, the SRG results appear slightly overmerged, and therefore not as good as the K-means in terms of boundary correctness. The RHC algorithm has the problem of jagged boundaries and a large number of segmented regions even with region merging. The SCS algorithm also suffers from a large number of segmented regions and jagged boundaries, where its boundary correctness is the worst amongst the five algorithms compared. Although this result differs from the Hedley and Yan¹⁶ results, it should be noted that their test image was very well-defined, whereas the test images considered here are not. Since they did not compare the number of segmented regions and the smoothness of boundaries, it would be difficult to make a direct comparison. Besides, our implementation of their algorithm may not be exactly the same as the one used in their test. However, due to the nature of the SCS algorithm, the results obtained here are not entirely surprising.

5 Conclusions

In this paper, a new clustering algorithm derived from the Markovian model of the gravitational clustering method was described. By introducing a constrained force effective function on the attractive force between the particles in the RGB color space, the formation of multiple clusters can be better controlled and determined, where the clustering process terminates when the attractive force between particles converge to zero. This new clustering algorithm was evaluated and compared objectively against four other algorithms using the Liu and Yang evaluation method, and subjectively based on the boundary smoothness, correctness, and the complexity of the segmented image expressed in the number of segmented regions. On the objective evaluation without region merging, the new clustering algorithm is on a par with the seed-based region growing method which is the best amongst the five algorithms. The K-means algorithm follows in third place comfortably with the remaining two algorithms giving rather poor evaluated values. When region merging is applied to the segmented images indiscriminately, the evaluated values and the standard deviation become smaller, but their relative performance has little change. Subjectively, the new clustering algorithm consistently gives smooth and correct boundaries for all the test images and a meaningful number of segmented regions when compared with the original. The K-means and seed-

based region growing algorithm both perform well when visually inspected, although the seed-based algorithm tends to generate a substantially larger number of segmented regions, and the K-means results appear to be mildly under-segmented. Weighing all these, the K-means algorithm is considered to have a slightly better performance than the SRG algorithm. The remaining two algorithms perform poorly in this category due to the nature of the histogram thresholding and the classification of high and low gradient pixels using edge detection. In conclusion, the new clustering algorithm is considered to be the best of the five segmentation algorithms in objective and subjective terms.

Future work will be carried out as follows. First, different force effective functions will be attempted where the relationship between D_F and the number of clusters formed will be studied. Second, the effect of scaling the color space will be considered as it will have further impact on the data reduction and computing requirement of the algorithm. Finally, other representative segmentation algorithms will be included in future comparisons, where other practical images will also be considered.

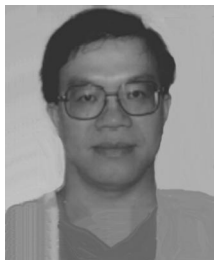
Acknowledgments

The authors gratefully acknowledge the financial support of the Committee of Research and Conference Grant of the University of Hong Kong under grants 337/062/0016 and 337/062/0036 and the Li Po Chun scholarship.

References

1. M. Sonka, V. Hlavac, and R. Boyle, *Image Processing, Analysis and Machine Vision*, Chapman & Hall Computer, University Press, Cambridge (1994).
2. R. J. Schalkoff, *Digital Image Processing and Computer Vision*, John Wiley and Sons Inc. (1989).
3. R. M. Haralick and L. G. Shapiro, "Survey: Image Segmentation Techniques," *Comp. Vision, Graphics, Image Process.* **29**, 100–132 (1985).
4. N. R. Pal and S. K. Pal, "A review on image segmentation techniques," *Pat. Recog.* **26**(9), 1277–1294 (1993).
5. R. Adams and L. Bischof, "Seeded region growing," *IEEE Trans. PAMI* **16**(6), 641–647 (1994).
6. J. Gauch and C. W. Hsia, "A comparison of three color image segmentation algorithms in four color spaces," in *Visual Comm. and Image Proc.* '92, *Proc. SPIE* **1818**, 1168–1181 (1992).
7. Y. J. Zhang, "Segmentation evaluation and comparison: a study of various algorithms," in *Visual Comm. and Image Proc.* '93, *Proc. SPIE* **2094**, 801–812 (1993).
8. J. Liu and Y. H. Yang, "Multiresolution color image segmentation," *IEEE Trans. PAMI* **16**(7), 689–700 (1994).
9. P. Bertolino and A. Montanvert, "Multiresolution segmentation using the irregular pyramid," *Proc. Int. Conf. Image Proc.* '96, **1**, 257–260, IEEE (1996).
10. L. G. Roberts, "Machine perception of three-dimensional solid," in *Opt. and Electro-Opt. Information Process.*, J. T. Tippell, Ed., MIT Press, Boston, pp. 159–197 (1965).
11. Y. Rui, A. C. She, and T. S. Huang, "Automated region segmentation using attraction-based grouping in spatial-color-texture space," in *Proc. Int. Conf. Image Process.* '96, **1**, 53–56, IEEE (1996).
12. G. Healey, "Segmenting images using normalized color," *IEEE Trans. Systems, Man and Cybernetics*, **22**(1), pp. 64–73 (1992).
13. J. A. Hartigan, *Clustering Algorithms*, John Wiley & Sons, Inc., Canada (1975).
14. R. Ohlander, "Analysis of natural scenes," PhD Thesis, Carnegie-Mellon University, Pittsburgh, Pennsylvania (1975).
15. M. Celenk, "A color clustering technique for image segmentation," *Comp. Vision, Graphics, and Image Process.* **52**, pp. 145–170 (1990).
16. M. Hedley and H. Yan, "Segmentation of color images using spatial and color space information," *J. Electronic Imaging* **1**(4), pp. 374–380 (1992).
17. Y. W. Lim and S. U. Lee, "On the color image segmentation algorithm based on the thresholding and the fuzzy c-means techniques," *Pat. Recog.* **23**(6), pp. 935–952 (1990).
18. G. Beni and X. Liu, "A least biased fuzzy clustering method," *IEEE Trans. on PAMI* **16**(9), pp. 954–960 (1994).

19. G. Matthews and J. Hearne, "Clustering without a metric," *IEEE Trans. on PAMI* **13**(2), pp. 175-184 (1991).
20. H. S. Lai and H. C. Yung, "An effective image segmentation algorithm in the RGB color space and its performance analysis," Research Report, Dept. of Electrical and Electronic Engr., The University of Hong Kong (Jan. 1997).
21. W. E. Wright, "Gravitational clustering," *Pattern Recog.* **9**, pp. 151-166 (1977).
22. Y. J. Zhang and J. J. Gerbrands, "Segmentation evaluation using ultimate measurement accuracy," *Image Process. Algorithms and Techniques III Proc. SPIE* **1657**, 449-460 (1992).
23. W. A. Yasnoff, J. K. Mui, and J. W. Bacus, "Error measures for scene segmentation," *Patt. Recog.* **9**, pp. 217-231 (1977).
24. L. Yang, F. Albrechtsen, T. Lonnestad, and P. Grottum, "A supervised approach to the evaluation of image segmentation methods," Research report, Department of Informatics, University of Oslo, Norway (1995).



H. C. Yung received his BSc and PhD degrees from the University of Newcastle-Upon-Tyne, England, in 1982 and 1985, respectively. He held the position of lecturer at the same university from 1985 to 1990, where he developed teaching and research activities in digital image processing and VLSI. From 1991 to 1993, he was employed as senior research scientist by the Defence Science and Technology Organization, Australia, where he headed

a research team on the algorithmic development of military-grade

signal analysis systems. He joined the University of Hong Kong in late 1993 where he is now an associate professor. Dr. Yung has co-authored a computer vision book and has published extensively in the areas of digital signal/image processing, parallel and multiprocessing methodology, and VLSI in books, journals, and conference proceedings. He is a chartered electrical engineer, a member of the HKIE, IEE, and a senior member of the IEEE.



H. S. Lai received the BEng degree in computer engineering from the University of Hong Kong in 1994 and MSc degree from the University of Surrey, United Kingdom, and was awarded the Cable and Wireless Prize for being the best overall student in the course in 1995. He is currently a PhD candidate and a Li Po Chun Scholar in the Department of Electrical & Electronic Engineering, University of Hong

Kong. His research interests are digital image processing and intelligent transportation system.