Middle-East Journal of Scientific Research 19 (7): 980-984, 2014
ISSN 1990-9233

DOI: 10.5829/idosi.mejsr.2014.19.7.1491

An Efficient Clustering Algorithm for Image Segmentation

E. Fathima and K.P. Kaliyamoorthie

Department of Computer Science,
Bharath University, India

Abstract: Gray-level clustering is an important procedure inimage processing, which reduces the gray-level of an image. Inorder to display an image with high gray level in a screen withlower gray level, a good gray-level clustering algorithm isnecessary to complete this job. Based on the mean value andstandard deviation of histogram within a sub-interval, a novelrecursive algorithm for solving the gray-level reduction isproposed in this paper. It divides the sub-interval recursivelyuntil the difference between original image and clusteredimage within a given threshold. Experiments are carried outfor some samples with high gray level to demonstrate thecomputational advantage of the proposed method.

Key words: Gray-level clustering · Image processing · Clusteringalgorithm · Image segmentation · Histogram

INTRODUCTION

The main purpose of gray-level clustering is to reduce ahigh gray-level image into a lower gray-level image. Aclustered image can then be use in many applications. Forexample, satellite image clustering [1] is used toclassification of geology, the measure of area on forest, orscope of cultivated land. In image compression [2] application, a clustered image might reach a highercompression rate. The application for magnetic resonanceimaging [3], the injured area of apple is obtained afterclustered. In text classification [4] application, text is easy tobe extracted from a clustered image. Since the gray-level andresolution of a screen in a cell phone, MP4, or PDA isusually lower than that of a PC, to reduce the gray-level andresolution of an image is necessary for displaying on theseequipments.

There are two types of image clustering. The first type isbased on the gray-level. It includes single thres holding [5] and mult-thresholding [5-11] on gray-level. Singlethresholding just transforms original image to binary image. Mult-thresholding classifies the gray-level into severalintervals. Otsu's method [5] is the most classic method inmult-thresholding. However, it takes a lot of time tocomplete the mult-thresholding. Thus, Arora [11] proposeda quick mult-thresholding algorithm to solve this problem in2008.

The second type of image clustering is based on clusteringtechnology. K-means algorithm [12-13] is the classic methodin solving this problem. But, the number of clusters and theinitial set of each cluster center must be set by human. Final output set of each cluster center in K-means algorithmwill be the set of gray-level for each cluster. Since timeneeded for K-means algorithm is too high to be implemented, Dong proposed a fast K-means algorithm [13] in 2006 Based on two-layer pyramid data structure, Dong‘salgorithm finds the image cluster from the lower resolutionimage and then extends it to the higher resolution image.

With combining the methods of first type and second type, a precise and fast gray-level clustering algorithm is proposedin this paper. The first type method used in this paper isArora [11] algorithm, which is used to get the number ofclusters and the initial set of each cluster center. Then, Kmeansclustering algorithm is introduced to update the set of each cluster center until all centers are not changed. Therefore, a precise and fast gray-level clustering algorithmis created [14-15].

The rest of this paper is organized as follows. Theproposed algorithm of gray-level clustering is presented innext section. Section 4 applies the proposed algorithm andthe compared algorithms to several sample images. Concluding remarks and potential applications are providedin Section 5.

Proposed Algorithm: An $M \times N$ resolution digital image may be defined as atwo-dimensional function where x and y are spatial coordinates and the amplitude of F at any

Corresponding Author: E. Fathima, Department of Computer Science, Bharath University, India.

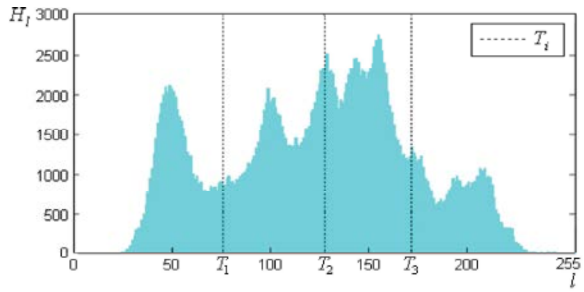


Fig. 1: Example of a histogram partitioned into 4 sub-intervals

pair ofcoordinates (x, y) is the gray level of the image at that point [18]. Thus, we use $F(x, y)$ to represent an original input image.

Then the array of an original input image $F(x, y)$ is:

(1,0) (1,1) (1, 1)
(1,0) (1,1) (1, 1)
(0,0) (0,1) (0, 1)

A set of gray-level is clustered by the proposed algorithmbased on the histogram. That is, a histogram is partitioned byproposed algorithm into n sub-intervals [16]. According to theintensity of gray-level from small to large, sub-intervals arenumbered as 0, 1, 2, ..., n-1. Figure 1 shows an example ofhistogram partitioned into 4 sub-intervals, i.e., n=4. Let Ti bemult-threshold between sub-intervals, where i = 0, 1, 2,...,n.

The values of mult-threshold $T0$, $T1$, $T2$, $T3$ and $T4$ are 0, 76,128, 172 and 256, respectively, shown in Figure 1. The semult-thre sholds partition gray-level into 4 sub-intervals, i.e., sub-intervals [0, 76), [76, 128), [128, 172) and [172, 256), respectively. Let Vi be the mean value of gray-level in sub-2009 interval$i$, where i = 0, 1, 2,..., n-1. There are 4 mean values$V0$, $V1$, $V2$ and $V3$ in Figure 1, which are 54, 106, 149 and 195, respectively.

In order to measure the visual quality of clustered image,peak signal to noise ratio (*PSNR*) is introduced in theproposed algorithm. *PSNR* [11] is expressed as:

$$PSNR = 20 \log 10 \, (255) / RMSE$$

The unit of *PSNR* is dB. The higher the *PSNR* is, the moresimilar between the clustered image and original image. We use Arora [11] algorithm to derive the number of clusters and the initial set of each cluster center. And then, K-means clustering algorithm [12-13] is introduced to updatethe set of each cluster center until all centers are not changed.

With combining the Arora algorithm and K-means algorithm, a precise and fast gray-level clustering algorithm is proposedin this paper. That is, Effective Multilevel Thresholding Algorithm.

Input: Original image $F(x, y)$
Output: Thresholded image $G(x, y)$
set$n$ = 2, a = 0, b = 256, $PSNR$ = 0
calculateμ and σ of histogram in interval [a, b)
set$T0$ = 0, $T1$ = μ, $T2$ = 256

Loop-1:
$a = Tn/2 - \sigma$ and $b = Tn/2 + \sigma$
$n = n + 2$
calculateμ and σ of histogram in interval [a, b)
for$i$ = n down to n/2+2 do Ti = Ti-2
setTn/2 = μ, Tn/2-1 = a and Tn/2+1 = b
for$i$ = 0 to n-1 do
calculateμ of histogram in interval [Ti, Ti+1)
setVi = μ
calculate*RMSE* and *PSNR*
repeat Loop-1 until the increment of $PSNR$ < 0.1
Loop-2:
for$i$ = 0 to n-2 do
calculateμ of histogram in interval [Vi, Vi+1)
setTi+1 = μ
for$i$ = 0 to n-1 do
calculateμ of histogram in interval [Ti, Ti+1)
setVi = μ
repeat Loop-2 until all Vi are not changed
For each pixel (x, y)
$G(x, y) = Vi$, if $F(x, y)$ lies in interval [Ti, Ti+1)

Figure 2(a) is original image of Lena. Its gray-level is 256 and its resolution is 512 by 512. Table 1 lists the values of n,Ti, Vi and *PSNR* for each loop during processing in theproposed algorithm with the image of Figure 2(a).

(a) (b)

Fig. 2: Lena images of (a) original image with 256 graylevel
(b) clustered image with 12 gray-level

Table 1: Processing for Lena

*NVi*and *Ti PSNR*
Loop-1:
2 *Vi*: 81, 161
Ti : 0, 124, 25619.66
4*Vi* : 54, 106, 149, 195
Ti : 0, 76, 128, 172, 25625.95
6 *Vi* : 54, 92, 118, 142, 161, 195
Ti : 0, 76, 103, 129, 153,
172, 25628.71
8*Vi* : 54, 92, 109, 123, 137,
149, 161, 195
Ti : 0, 76, 103, 115, 129,
143, 153, 172, 25629.53
10*Vi* : 54, 92, 109, 119, 126,
132, 140, 149, 161, 195
Ti : 0, 76, 103, 115, 122,
128,136, 143, 153, 172, 25629.70
12 *Vi* : 54, 92, 109, 119, 124, 126,
130, 134, 140, 149, 161, 195
Ti : 0, 76, 103, 115, 122, 124,
127,132, 136, 143, 153,
172, 25629.72
Loop-2:
12*Vi* : 47, 70, 95, 111, 124, 134,
144, 154, 164, 174, 195, 211
Ti : 0, 56, 84, 103, 118, 129, 139,
149, 158, 170, 186, 203, 25634.28

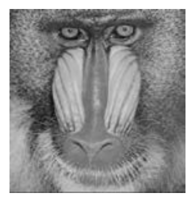
From Table 1, it is easy to see that the final *n* is 12 and *PSNR* is34.28 dB. Figure 2(b) shows the clustered image derived bythe proposed algorithm. Although the gray-level in Figure2(b) is 12, it is really similar to original image. Therefore, theproposed algorithm can be used to the equipment with lowergray-level displaying screen.

RESULTS

In order to prove the usefulness of the proposed algorithm, four different images with 512×512 resolution are used forexperiments. Figures 3(a), 3(b), 3(c) and 3(d) are originalimages named Baboon, Jet, Peppers and Washbowl, respectively. The compared methods used in this paper are K-means algorithm [12], Dong's algorithm [13] and Arora'salgorithm [11]. Since K-means algorithm and Dong'salgorithm need the initial *n* and the initial set of each clustercenter, we use the outputs of n and *Vi*from Arora's algorithm.Figures 4(a), 4(b) and 4(c) are the clustered imagesmaking by K-means algorithm [12], Dong's algorithm [13], 260 and Arora's algorithm [11], respectively. Compared them toFigure 2(b), it is found that there are blocks in the clusteredimages of compared methods. For examples, block imagesappeared on Lena's shoulder, forehead and hat.

Table 2: Methods Compared

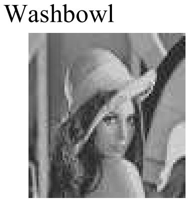
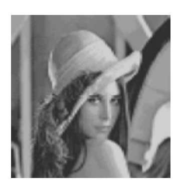
Image Method	*PSNR*	Time(ms)
Lena K-means algorithm	33.14	4340
Dong's algorithm	33.57	3304
Arora's algorithm	29.72	316
Proposed algorithm	34.28	317
Baboon K-means algorithm	33.27	5722
Dong's algorithm	33.43	3066
Arora's algorithm	29.03	322
Proposed algorithm	33.67	323
Jet K-means algorithm	30.79	5207
Dong's algorithm	30.79	3005
Arora's algorithm	26.29	281
Proposed algorithm	31.16	282
Peppers K-means algorithm	31.78	3156
Dong's algorithm	32.35	2719
Arora's algorithm	28.25	318
Proposed algorithm	32.45	320
Washbowl K-means algorithm	32.96	5205
Dong's algorithm	32.81	2989
Arora's algorithm	28.72	307
Proposed algorithm	33.43	308




(a) (b)


(c) (d)

Fig. 3: Original images (a) Baboon (b) Jet (c) Peppers (d) Washbowl

(a) (b)

(c)

Fig. 4: Clustered images using (a) K-means algorithm (b) Dong'salgorithm (c) Arora's algorithm

Both of the compared algorithms and the proposed algorithm are implemented with Matlab and run on a personal computer with a AMD K8-3.2G processor, 1 GBRAM and 512 KB cache. Table 2 lists the values of *PSNR*and times needed of each method for the testing images Lena, Baboon, Jet, Peppers and Washbowl. From Table 2, basedon the same clustered number, it is obvious to see that theproposed algorithm builds the clustered images with thehighest values of *PSNR* among the other methods. The*PSNR*s of the proposed algorithm for images Lena, Baboon, Jet, Peppers and Washbowl are 34.28, 33.67, 31.16, 32.45 and 33.43, respectively. Although the time needed of Arora'salgorithm [11] is the lowest over other methods, its *PSNR*sare the lowest in all testing images [17]. The speed of theproposed algorithm is also extremely close to that of Arora'salgorithm [11], which is faster than K-means algorithm [12] and Dong's algorithm [13].

CONCLUSIONS

Based on the mean value, standard deviation and *PSNR*, anew gray-level clustering algorithm is proposed in this paperfor image segmentation. Evaluation results show theproposed algorithm succeeds in deriving clustered images.The clustered image built by the proposed algorithm isextremely close to that of original image [18-20]. The algorithm alsoreduces the *RMSE* much more than other schemes, whichmakes it better suited for real equipment with lower grayleveldisplaying screen.

REFERENCES

1. Baraldi, A. and F. Parmiggiani, 1995. A neural network for unsupervisedcategorization of multivalued input patterns: An application tosatellite image clustering, IEEE Transactions on Geoscience RemoteSensing, pp: 305-316.
2. Biswas, S. and N.R. Pal, 2000. On hierarchical segmentation for imagecompression, Pattern Recognition Letters, pp: 131-144.
3. McCarthy, M.J., B. Zion, P. Chen, S. Ablett, A.H. Darke and P.J. Lillford, 1995. Diamagnetic susceptibility changes in apple tissue afterbruising," Journal of the Science of Food and Agriculture, pp: 13-20.
4. Song, Y., A. Liu, L. Pang, S. Lin, Y. Zhang and S. Tang, 2008. A Novel Image Text Extraction Method Based on K-means Clustering, Seventh IEEE/ACIS International Conference on Computer and Information Science, 14-16, pp: 185-190, 261 Authorized licensed use limited to: LA TROBE UNIVERSITY. Downloaded on August 10, 2009 at 01:49 from IEEE Xplore. Restrictions apply.
5. Otsu, N., 1979. A threshold selection method from gray-level histograms, IEEE Transactions on Systems, Man and Cybernetics, pp: 62-69.
6. Sahoo, P.K., S. Soltani, A.K.C. Wong and Y.C. Chen, 1988. A surveyof thresholding techniques, Computer Vision, Graphics and ImageProcessing, pp: 233-260.
7. O'Gorman, L., 1994. Binarization and multi thre sholding of documentimages using connectivity, CVGIP: Graphical Models and Image Processing, pp: 494-506.
8. Yan, H., 1996. Unified formulation of a class of image thresholdingtechniques, Pattern Recognition, pp: 2025-2032.
9. Geraud, T., P.Y. Strub and J. Darbon, 2001. Color image segmentationbased on automatic morphological clustering, Proc. IEEEInternational Conference on Image Processing, pp: 70-73.
10. Sezgin, M. and B. Sankur, 2004. Survey over image thresholdingtechniques and quantitative performance evaluation, Journal of Electronic Imaging, pp: 146-165.
11. Arora, S., J. Acharya, A. Verma and P.K. Panigrahi, 2008. Multilevelthresholding for image segmentation through a fast statisticalrecursive algorithm, Pattern Recognition Letters, pp: 119-125.
12. Bezdek, J.C., 1981. Pattern recognition with fuzzy objective functionalgorithms, Plenum Pub Corp.
13. Dong, L., P. Ogunbona, W. Li, G. Yu, L. Fan and G. Zheng, 2006. A fastalgorithm for color image segmentation, Proc. the First International Conference on Innovative Computing, Information and Control, pp: 685-688.
14. Kumaravel, B. Anatha Barathi, 2013. Personalized image search using query expansion, Middle-East Journal of Scientific Research, ISSN: 1990-9233, 15(12): 1736-1739.
15. Kumaravel, A. and R. Udayakumar, 2013. Web Portal Visits Patterns Predicted by Intuitionistic Fuzzy Approach, Indian Journal of Science and Technology, ISSN: 0974-6846, 6(5S): 4549-4553.
16. Kumaravel, A. and K. Rangarajan, 2013. Algorithm for Automation Specification for Exploring Dynamic Labyrinths, Indian Journal of Science and Technology, ISSN: 0974-6846, 6(5S): 4554-4559.

17. Kumaravel, A. and Oinam Nickson Meetei, 2013. An Application of Non-uniform Cellular Automata for Efficient Crytography, Indian Journal of Science and Technology, ISSN: 0974-6846, 6(5S): 4560-4566.
18. Pattanayak, Monalisa and P.L. Nayak, 2013. Green Synthesis of Gold Nanoparticles Using Elettaria cardamomum (ELAICHI) Aqueous Extract World Journal of Nano Science & Technology, 2(1): 01-05.
19. Chahataray, Rajashree and P.L. Nayak, 2013. Synthesis and Characterization of Conducting Polymers Multi Walled Carbon Nanotube-Chitosan Composites Coupled with Poly (P-Aminophenol) World Journal of Nano Science & Technology, 2(1): 18-25.
20. Parida, Umesh Kumar, S.K. Biswal, P.L. Nayak and B.K. Bindhani, 2013. Gold Nano Particles for Biomedical Applications World Journal of Nano Science & Technology, 2(1): 47-57.