G-Images : Towards Multilevel Unsupervised Image Segmentation

Harbir Singh
School of Computing Sciences,
University of East Anglia,
Norwich, England.
harbir.singh@uea.ac.uk

Reyer Zwiggelaar
Department of Computer Science,
University of Wales,
Aberystwyth, UK.
rrz@aber.ac.uk

Abstract

We present two novel approaches to initiating unsupervised segmentation of digital images using an algorithm that utilises the concept of information theory. The first approach uses Information Gain and the second is based on the Gini Index. In the two approaches, Information Gain and the Gini Index are calculated locally, at a pixel level, resulting in a G-image where high G value occurs at contrasting boundaries and zero G value within homogeneous regions. Subsequently, a multi-level thresholding approach based on the G-image is used to obtain the optimal segmentation results. The segmentation is guided by both local and global parametric constraints. Comparative, visual, evaluation on real and artificial images shows promising results.
Keywords*: Segmentation, Computer Vision, Pattern Recognition, Information Theory*

1. Introduction

Automated image segmentation is an important processing step with widespread applications in performing computer vision tasks such as pattern recognition and image retrieval. Image segmentation algorithms classify the picture elements of an image into different classes so that pixels corresponding to an object of interest belong to the same class [8]. Approaches to carrying out automated segmentation can be divided into two groups, namely supervised and unsupervised methods. Interpretation of objects of interest is often application dependent and in case of supervised segmentation priori information is used for image segmentation by incorporating properties of pixels in relation to its neighbouring pixels. Unsupervised approaches are undertaken when prior information of objects of interest is not available or difficult to obtain and this is often the case. Given the importance of unsupervised image segmentation, various methods are reported in the literature. Some of the methods for carrying out unsupervised image segmentation include the Bayesian approach [1], Markov trees and complex wavelets [13], histogram clustering [11], neuro-fuzzy systems [10], higher-order hidden Markov chains [3] and the use of entropy [15, 12]. Local homogeneity analysis based unsupervised segmentation is undertaken in [6] and a method for unsupervised segmentation of colour-texture regions in images is presented in [2]. Unsupervised segmentation using k-means and fuzzy c-means algorithm can be found in [7]. In this paper, we have used a novel approach to solve the task of unsupervised image segmentation using *Information Gain* [5] and the *Gini Index* [4].

2. Segmentation methods

We introduce the concept of a G-image, using *Information Gain* or the *Gini Index*, which forms the basis of the segmentation process. For a comparative evaluation we also discuss image segmentation based on J-images [2], H-images [6], and a standard k-means algorithm.

2.1 G-Image

Both the *Information Gain* and *Gini Index* are entropy based methods. We have developed these approaches because entropy based measures for homogeneous and random noise regions are low whilst at boundaries between regions these measures are significantly increased.

Our segmentation method is based on region growing using 8-neighbourhood pixel connectivity and incorporates *Information Gain* and the *Gini Index* as a heuristic at the pixel processing stage. However, derivation of G-images is an initial starting step for undertaking region segmentation. We consider a grayscale image, $I(x, y)$, of size $M \times N$ where $x\epsilon[0, M)$ and $y\epsilon[0, N)$. Each pixel is characterised by a grayscale value, which is restricted to one of L possible values $0, 1...L-1$, where maximum $L = 256$ gives a 8 bit quantisation scheme.

2.1.1 Gain-based segmentation

Let set S consist of n_s data points in the 8-neighbourhood of a candidate pixel, S_c. Considering 8-neighbourhood, our sample S, will consist of 9 points, which are labelled $j = 1..n_s$. This sample will contain points of various intensity values giving rise to randomness or degree of impurity in the sample. Our attempt is to classify this sample such that we can identify areas of high gain. The high-gain areas are taken to represent class boundary or edge pixels. In order to do so we need a prior number of classes. However, due to large randomness in intensity variation this number is not always available.

To circumvent this lack of knowledge, we introduce a global threshold T which we can vary to influence the calculation of expected information (Eq. 1) of the sample and thereby influence calculation of gain. In a way, setting T involves prior knowledge of the class we want to segment out. For example, we know that in a grayscale high resolution computed tomography (HRCT) lung image, bone has a threshold higher than 240. Therefore, by setting a T value close to 240 helps us to extract areas of high gain which pertain to the bone class (see Fig. 6 h).

Based on the global threshold value we partition the sample data into 2 classes such that if the intensity of a sample point (pixel) is less than the global threshold T, then that pixel is assigned a class label of include and if the intensity of a pixel is more that the global threshold T, then that pixel is assigned a class label of exclude. In this case the label of include means the pixel is part of a homogeneous region, whilst the label of exclude means it is likely that the pixel is belongs to the boundary between two homogeneous regions. T is selected incrementally spanning the range of the grayscale values (0-255).

Given that our class label C has two possible values, (C=include or C=exclude), let n_{s_i} be the number of pixels belonging to class C_i in our sample. The expected information for the whole sample S, is obtained as

$$EI(S) = -\sum_{i=1}^{m} p_i log_2(p_i) \tag{1}$$

where p_i is the probability of occurrence of class C_i in the sample which is given by ns_i/ns. m is the total number of class labels, which is 2 in this case, and $ns = \sum_i ns_i$.

To estimate local constraint, let I_j be the intensity at each pixel j, μ be the mean of all pixels in the 8-neighbourhood of pixel j and σ be the standard deviation of all pixels in the neighbourhood. The pixels at which the following condition

$$\sqrt{(I_j - \mu)^2} < \sigma \tag{2}$$

is satisfied are assigned a class label of include. In other words if at a pixel Eq. 2 is satisfied, then that pixel is assigned a class label include else it is assigned a class label exclude. The pixels which are assigned a class label of include, are taken to represent the least randomness or impurity at the candidate pixel and hence are likely to be part of a homogeneous region.

The entropy E, at the candidate pixel S_c is given as

$$E(S_c) = p_c * EI(S_c) \tag{3}$$

where p_c is the probability that S_c belongs to class c, where $c \epsilon 1, 2$.

Total gain is defined as

$$G = EI(S) - E(S_c) \tag{4}$$

Based on the G values, a G-image is derived. This is a grayscale image whose pixel values are the G values calculated over local windows centred on those pixels. The areas with a high G value in the G-image represent region boundaries while areas with low or zero values represent homogeneous patterns.

2.1.2 Gini-based segmentation

The *Gini Index* (GI) as a measure of diversity is mostly used in economics [14]. It is used in the study of inequality in the distribution of a given variable in a population such as inequality in the distribution of income, distribution of wealth, production, health and education, etc. The GI ranges between 0, where there is no concentration (perfect equality), and 1, where there is total concentration (perfect inequality). It has been used in non-economic area such as in lung nodule diagnosis with the purpose of characterising lung nodules as malignant or benign by analysing the degree of voxel density concentration in lung nodules [14]. It also finds application in machine learning for example in the construction of decision tree classifiers [4]. We have aimed to show its applicability in the field of unsupervised image segmentation where low Gini values occur in regions of homogeneity and high Gini values occur in border areas, which are regions of high inequality in the sense of homogeneity of regions.

Modelling the GI as an impurity function and proceeding on lines similar to our Gain-based approach, let our sample be represented by S points in the 8-neighbourhood of a candidate pixel. This sample will contain a variation of intensity values and therefore represent many possible classes. Based on a global threshold T, we assign a class label include to all points whose intensity is less than T and a class label exclude to other points. In order to model the sample variance, we calculate GI for each class and sum over all classes in the sample to obtain the final GI at the candidate pixel, as shown in Eq. 5.

Let p_i be relative frequency of the class labels (C=include or C=exclude) in the dataset. The GI for splitting the sample S into S_1 and S_2 is defined as

$$gini(S_1, S_2) = \frac{|S_1|}{|S|} gini(S_1) + \frac{|S_2|}{|S|} gini(S_2) \quad (5)$$

where

$$gini(S) = 1 - \sum_{i=1}^{m} p_i^2 \quad (6)$$

and $|S|$ denotes the total number of elements in set S and m is the total number of class labels which is equal to 2 in our case. S_1 is the set of pixel points in the sample S which are assigned a class label of include and S_2 is the set of pixel points in the sample S which are assigned a class label of exclude.

In the Gini-based approach class assignment is done only once and additional steps of calculating expected information both for the sample points and at the candidate pixel level are avoided. In Figs 1- 4 (b,d), the G-images resulting from Gini-based approach show a very close similarity to the Gain-based G-images. Given the close match between the resultant G-images we may conclude that the two approaches are identical and either of the G-images could provide suitable starting stage for subsequent segmentation.

2.1.3 Algorithm outline

The steps of the algorithm are outlined below.

1. For each global threshold level ranging from $T = x_s$ to $T = x_e$ perform the following steps

   2.1 Using 8-neighbourhood pixel connectivity, for all pixels with intensity below global threshold assign class label of include

   2.2 Using 8-neighbourhood pixel connectivity, for all pixels with intensity above global threshold assign class label of exclude

   – *For Gain-based method*

   2.3 Calculate expected information of pixels in 8-neighbourhood using Eq. 1

   2.4 At the current pixel, using 8-neighbourhood pixel connectivity, assign class to current pixel according to criteria in Eq. 2

   2.5 Calculate expected information at the current pixel using Eq. 3

   2.6 Calculate gain at the current pixel using Eq. 4

   – *For Gini-based method*

   2.3 Calculate Gini index at current pixel incorporating 8-neighbourhood connectivity using Eq. 5

2.2 J-Image [2]

A region growing method based on image quantisation called JSEG is proposed in [2]. The image pixels are first replaced by quantised values forming a class-map of the image where a criterion for good segmentation (J) is defined similar to Fisher's multi-class linear discriminant. Applying the criterion to local windows in the class-map results in the J-image, in which high and low values correspond to possible region boundaries and region centres. Finally, a region growing method is used to segment the image based on the J-image. The developed approach does not involve the complexity of a class map and image quantisation.

2.3 H-Image [6]

Jing et al. proposed a similar method to JSEG but with a different segmentation criteria. To quantise the homogeneity of a pattern an H-image was derived with each pixel value being replaced by the calculated H value. The pixels of an image were viewed as a set of spatial data points located in a 2D plane with the top left corner being the origin. Based on the H value a H-image was derived whose pixel values were the H values calculated over local windows centred on those pixels. The dark and bright areas in the H-image which represented the region centres and region boundaries were used in carrying out region growing based on local homogeneity analysis. Our proposed approach is inspired by this method, however, introduces a different way of undertaking unsupervised segmentation using both global/external and local constraints.

3. Results and discussion

We tested our method on a variety of images as shown in Figs 1- 4. Our method performs equally well when compared to existing approaches. Considering Fig. 1(b), the Gain image shows sufficient detail when compared to the H-image in Fig. 1(c) and the method is successful in identifying features such as the eyes, hat and face in the image. The result is equally matched by the Gini image in Fig. 1(d).

In Fig. 2(b,d), the Gain and Gini images show the intricate pattern of spots detected. Minute details such as the whiskers were also picked up. The performance closely matches the result of the H-image, Fig. 2(c).

Considering Fig. 3, results based on lung computed tomography (CT) data are shown. Our method is successful in identifying the two lung regions along with the trachea in the centre and a few structures of interest within the two lung regions, Fig. 3(b,d). This is again comparable to the brightest lines in the H-image, Fig. 3(c).

Lastly, we tested the method on artificial image data of different object shapes with the same contrast, Fig. 4. Once

Figure 1. Results: (a) original image, (b) Gain-image at optimum threshold, (c) H-image, (d) Gini-image at optimum threshold, (e) J-image and (f) k-means.

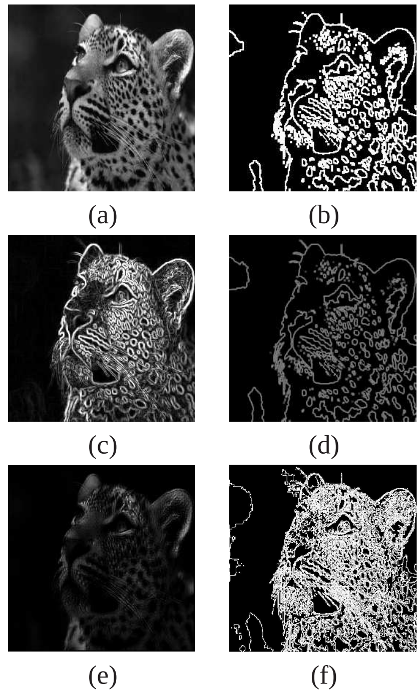


Figure 2. Results: (a) original image, (b) Gain-image at optimum threshold, (c) H-image, (d) Gini-image at optimum threshold, (e) J-image and (f) k-means.

again the performance of our method is comparable to the other methods in identifying the object regions. These results indicate a clear boundary detection between classes for both the H and G images and as such provide an appropriate starting point for image segmentation. It seems that the H images seem to represent a noisy version of the G images.

When comparing the result with the J-images, Figs 1-4(e), these seem less well defined and as such might provide a poor starting point for a segmentation process. In our implementation we used quantisation of grayscale images into 64 bins. The original algorithm deals with colour images and uses a more complex quantisation method based on peer-group filtering [2]. Deng and Manjunath mention that even though JSEG can be applied on grayscale images, the result are reasonable to an extent but not as good as colour images because intensity alone is not as discriminative as colour is.

In addition, when comparing with the k-means resultant images (see Figs 1- 4(f)), it is clear that the obtained class boundaries in the H and G images are a subset of the detected edges. On the other hand, the k-means images provide more detail as weak edges which do not represent class boundaries are also highlighted. In this case the value of k was set equal to six and the results indicate a large number of small noisy regions.

Accurate medical image segmentation to extract relevant parts of the anatomy is a crucial precursor for diagnosis and quantitative analysis. Some CT lung image results are shown in Fig. 6. A CT lung slice for the mid-thoracic region was segmented at different global threshold values, T. This shows that depending on the value of T various anatomical structures are extracted, e.g. at high T values the rib bones are found whilst at lower values soft tissue class boundaries are enhanced.

To further ascertain the utility of our method we show in Fig. 7 attempts at object selection dependent on contrast. Fig. 7(a) shows contrasting objects. Based on different global threshold values, we were able to select the objects as shown in the subsequent images. This would be difficult using methods which merely detect edges of objects and added steps of region labelling and selection based on region labels would be required.

Although not covered here, the extension of the developed G images can easily be extended to G volumes and as such can be used for anatomical segmentation of volumetric medical data, such as CT or MRI [16]. Preliminary results can be seen in Fig. 5, which shows a 3D display of lung cavity generated from a segmented 2D image stack of rib bones obtained as in Fig. 6(f). Extension to colour images is also planned where an approach similar to [6] could be applied

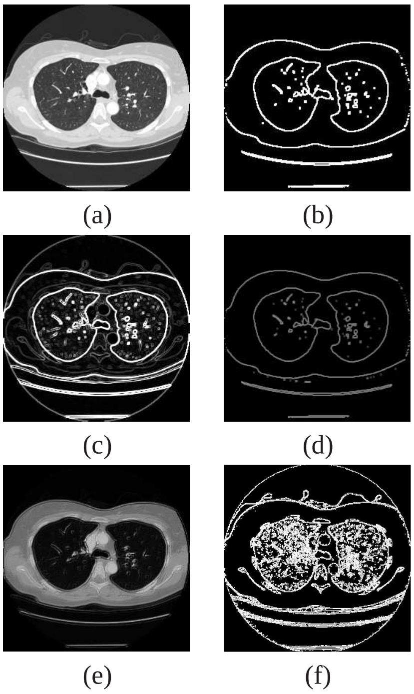


Figure 3. Results: (a) original image, (b) Gain-image at optimum threshold, (c) H-image, (d) Gini-image at optimum threshold, (e) J-image and (f) k-means.

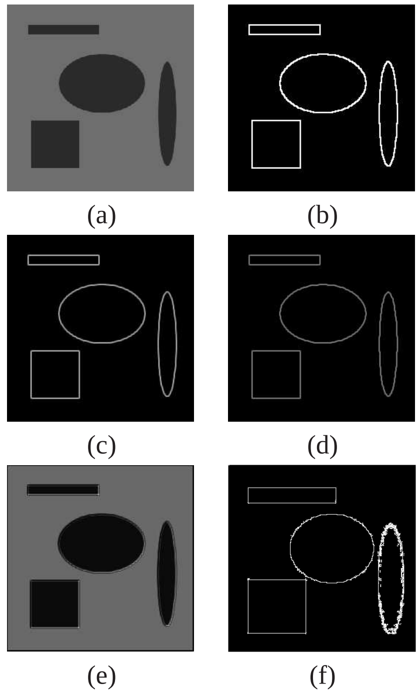
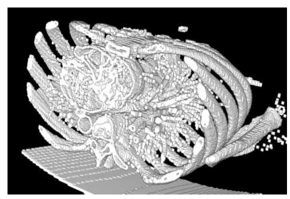


Figure 4. Results: (a) original image, (b) Gain-image at optimum threshold, (c) H-image, (d) Gini-image at optimum threshold, (e) J-image and (f) k-means.



(a)

Figure 5. 3D display of lung cavity.

to the three RGB colour values and the results combined by taking the norm of the RGB component results. Initial segmentation success with a diverse array of images highlights the robustness of the approach.

In future we intend to extend our analysis for carrying out unsupervised segmentation to 3D volumes and do further analysis in the region growing and merging area. We also wish to undertake a modified approach capable of performing well in the presence of noise and uncertainty [9]. In addition, a full comparison with existing techniques based on root-mean-square-distances to truth data, including full contrast and noise aspects, is being developed.

4. Conclusions

We have presented a novel approach to initiating segmentation when little prior knowledge is known about the scene. In addition we have also compared our method with existing techniques, highlighting the uniqueness of our method.

References

[1] L. Cheng and T. Caelli. Unsupervised image segmentation: a bayesian approach. *The 16th International Conference on Vision Interface, Halifax, Canada*, 7(2), June 2003.

[2] Y. Deng and B. Manjunath. Unsupervised segementation of color-texture regions in images and video. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(8):800–810, 2001.

[3] C. Derrode, C. Carincotte, and S. Bourennane. Unsupervised image segmentation based on high-order hidden markov chains. *IEEE International Conference on Acoustics, Speech, and Signal Processing, Montreal, Canada*, May 2004.

[4] M. Dunham. *Data Mining Introductory and advanced topics*. Prentice Hall, New Jersey, 2003.

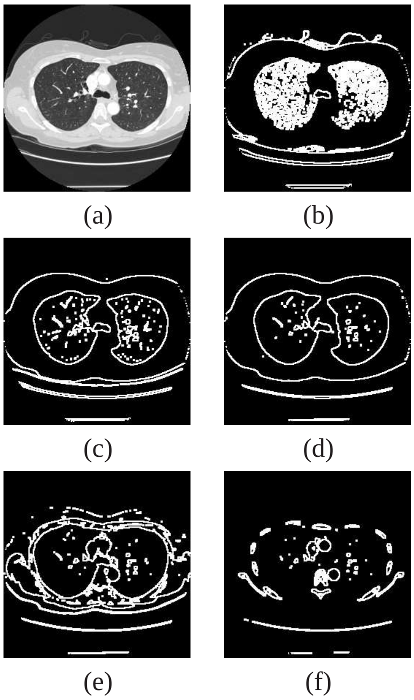


Figure 6. Sequence of CT lung images at different global threshold levels: (a) original image, (b) T=60, (d) T=120, (f) T=180, (g) T=210, (h) T=243.

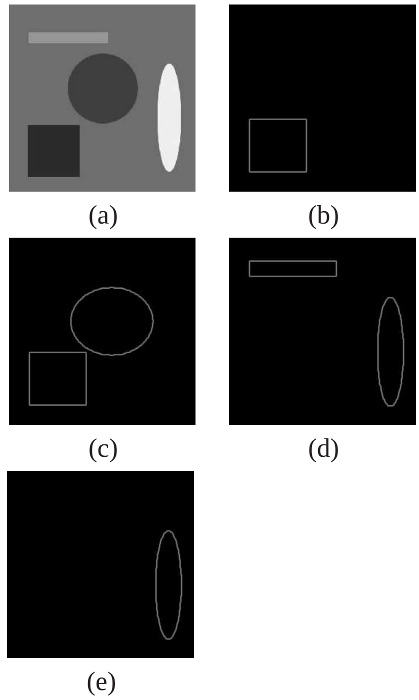


Figure 7. Effect of contrast: (a) original image, (b) T=50, (c) T=70, (d) T=120, (e) T=160.

[5] J. Han and M. Kamber. *Data Mining Concepts and Techniques*. Morgan Kaufmann, San Francisco CA, 2001.

[6] F. Jing, M. Li, H. Zang, and B. Zang. Unsupervised image segmentation using homogeneity analysis. *Proc. IEEE International Symposium on Circuits and Systems*, 2:456–459, May 2003.

[7] M. Khalighi, H. Soltanian-Zadeh, and C. Lucas. Unsupervised mri segmenatation with spatial connectivity. In *SPIE International Symposium on Medical Imaging, San Diego, CA*, pages 23–28, 2002.

[8] D. Marr. *Vision*. Freeman, San Francisco CA, 1982.

[9] S. Mitra and T. Acharya. *Data Mining Multimedia, Soft Computing and Bioinformatics*. Wiley-Interscience, USA, 2003.

[10] H. Muhammed. Unsupervised image segmentation using neuro-fuzzy systems. *31st Applied Image Pattern Recognition Workshop, Washington, D.C., USA*, pages 171–177, 2002.

[11] J. Puzicha, T. Hofmann, and J. Buhmann. Histogram clustering for unsupervised image segmentation. *Pattern Recognition Letters*, 20(9):899–909, 1999.

[12] P. Saha and J. Udupa. Optimum image thresholding via class uncertainity and region homogeniety. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(7):689–706, 2001.

[13] C. Shaffrey, N. Kingsbury, and I. Jermyn. Unsupervised image segmentation via markov trees and complex wavelets. *IEEE International Conference on Image Processing*, 3:801–804, Sept. 2002.

[14] A. Silva, P. Carvalho, and M. Gattass. Diagnosis of lung nodule using gini coefficient and skeletonization in ct images. In *ACM Symposium on Applied Computing*, March 2004.

[15] C. Sin and C. Leung. Image segmentation by edge pixel classification with maximum entropy. *International Symposium on Intelligent Multimedia, Video and Speech Processing*, pages 283–286, May 2001.

[16] H. Singh, M. Crawford, J. Curtins, and R. Zwiggelaar. Automated 3d segmentation of lung airway tree using gain based region growing approach. In *LNCS Medical Image Computing and Computer Assisted Intervention, Saint-Malo, France*, Sept. 2004.