Imperial Journal of Interdisciplinary Research (IJIR)
Vol.2, Issue-1 , 2016
ISSN : 2454-1362 , www.onlinejournal.in

Canny Mask Segmentation with Laplacian Sigmas In Content Based Image Retrieval System

ENSAF A. AL-ZURQA
Department of Information Technology
Faculty of Engineering and Information Technology, Taiz University
Taiz, Yemen

Abstract

Content-based image retrieval draws many of its methods from the field of image processing and is regarded by some as a subset of that field. It differs from these fields principally through its emphasis on the retrieval of images with desired characteristics from a collection of significant size. "Content-based" means that the search will analyze the actual contents of the image. The term 'content' in this context might refer colours, shapes, textures, or any other information that can be derived from the image itself. Without the ability to examine image content, searches must rely on metadata such as captions or keywords, which may be laborious or expensive to produce. In this paper, we work on specific "eggs" by using different laplacian sigmas and canny mask, to retrieve the images from the data base that are related to the original one. The main objectives of our work are classifying an image as belonging to a specific category and retrieving images that are similar to database by showing statistical results of the different sigmas values in our CBIR system.

Keywords: Enhancement procedures, image adjustment, canny mask segmentation, image analysis, laplacian sigmas .

1. Introduction

The earliest use of the term content-based image retrieval in the literature seems to have been by Kato [1992], to describe his experiments into automatic retrieval of images from a database by colour and shape feature. The term has since been widely used to describe the process of retrieving desired images from a large collection on the basis of features (such as colour, texture and shape) that can be automatically extracted from the images themselves. The features used for retrieval can be either primitive or semantic, but the extraction process must be predominantly automatic. Retrieval of images by manually-assigned keywords is definitely not CBIR as the term is generally understood – even if the keywords describe image content.

CBIR operates on a totally different principle from keyword indexing. Primitive features characterizing image content, such as colour, texture, and shape, are computed for both stored and query images, and used to identify (say) the 20 stored images most closely matching the query. Semantic features such as the type of object present in the image are harder to extract, though this remains an active research topic. Video retrieval is a topic of increasing importance – here, CBIR techniques are also used to break up long videos into individual shots, extract still keyframes summarizing the content of each shot, and search for video clips containing specified types of movement. Three commercial CBIR systems are now available – IBM's QBIC, Virage's VIR Image Engine, and Excalibur's Image Retrieval Ware. In addition, demonstration versions of numerous experimental systems can be viewed on the Web, including MIT's Photo book, Columbia University's WebSEEk, and Carnegie-Mellon University's Informedia. CBIR systems are beginning to find a foothold in the marketplace; prime application areas include crime prevention (fingerprint and face recognition), intellectual property (trademark registration), journalism and advertising (video asset management) and Web searching. Both the Alta Vista and Yahoo! Search engines now have CBIR facilities, courtesy of Virage and Excalibur respectively. The effectiveness of all current CBIR systems is inherently limited by the fact that they can operate only at the primitive feature level. None of them can search effectively for, say, a photo of a dog – though some semantic queries can be handled by specifying them in terms of primitives. A beach scene, for example, can be retrieved by specifying large areas of blue at the top of the image, and yellow at the bottom. There is evidence that combining primitive image features with text keywords or hyperlinks can overcome some of these problems, though little is known about how such features can best be combined for retrieval.

In this paper, we work on different images' content in; to retrieve the images from the data base that are related to the original one. The main objective is retrieving images that are similar in shape, texture and/or colour to an image that is done using different Laplacian sigmas and canny mask segmentation. We will discuss in details which technique is the best for our application, the deciding factor is applying the next steps until segmentation. The processing goes as we illustrate in following sections. Before processing the images, it needs to be enhanced in order to brighten eggs' colour, also images captured contains some noise that need to be removed or at least reduced for further correct processing on the image. After the image was enhanced and filtered, each egg must be extracted separately for further processing, so the next step here is detecting edges of eggs in each image, once the edges are detected from background, eggs are segmented (separated from one another) to make them ready for further processing. Extracting feature vectors needs image adjustments and pre-processing procedures. So, the rest of the paper is organized as follows: In Section 2, related works to CBIR are introduced. The pre-processing procedure (enhancement procedures) steps are illustrated in Section 3. In Section 4, image adjustment functions are presented as parts of feature extraction step. We present our segmentation work by canny mask, in section 5. In Section 6, we introduce the evaluation results for our work. Finally, in Section 7 concluded and future work is highlighted.

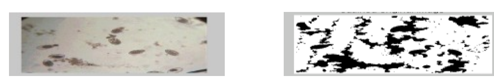
2. Related work

Research and development issues in CBIR cover a range of topics, many shared with mainstream image processing and information retrieval [1-8]. Some of the most important are: understanding image users' needs and information-seeking behaviour, identification of suitable ways of describing image content, extracting such features from raw images, providing compact storage for large image databases, matching query and stored images in a way that reflects human similarity judgements and efficiently accessing stored images by content, providing usable human interfaces to CBIR systems Key research issues in video retrieval include: automatic shot and scene detection, ways of combining video, text and sound for retrieval amd effective presentation of search output for the user. Enser [1995] reviews methods for providing subject access to pictorial data, developing a four-category framework to classify different approaches. He discusses the strengths and limitations both of conventional methods based on linguistic cues for both indexing and search, and experimental systems using visual cues for one or both of these. His conclusions are that, while there are serious limitations in current text-based techniques for subject access to image data, significant research advances will be needed before visually-based methods are adequate for this task. He also notes, as does Cawkell [1993] in an earlier study, that more dialogue between researchers into image analysis and information retrieval is needed. Aigrain et al [1996] discuss the main principles of automatic image similarity matching for database retrieval, emphasizing the difficulty of expressing this in terms of automatically generated features. They review a selection of current techniques for both still image retrieval and video data management, including video parsing, shot detection, keyframe extraction and video skimming. They conclude that the field is expanding rapidly, but that many major research challenges remain, including the difficulty of expressing semantic information in terms of primitive image features, and the need for significantly improved user interfaces. CBIR techniques are likely to be of most use in restricted subject domains, and where synergies with other types of data (particularly text and speech) can be exploited. Eakins [1996] proposes a framework for image retrieval, classifying image queries into a series of levels, and discussing the extent to which advances in technology are likely to meet users' needs at each level. His conclusion is that automatic CBIR techniques can already address many of users' requirements at level 1, and will be capable of making a significant contribution at level 2 if current research ideas can be successfully exploited. They are however most unlikely to make any impact at level 3 in the foreseeable future. In section 3, the first step of pre-processing procedures that are used in our works to enhance the images is illustrated.

3. Enhancement Procedures

The pre-processing procedure steps, that are used before the image analysis stages, are: Image Enhancement, Image adjustment, Noise Removal and Segmentation. Image Enhancement has many enhancement techniques, so we have put under study six different enhancement techniques. We will discuss in details which technique is the best for our application in such that the deciding factor is applying on the other stages until segmentation.

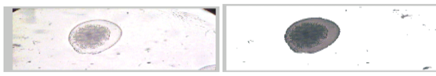
The cumulative histogram technique is one of enhancement techniques that is not suitable for our images since when applying the next steps, the final results were not sharp enough and contained a lot of noise as shown in figure (1).



(a) (b)

Figure (1): a) the output of the comulative histogram function
b)the output after detecting the objects from the background

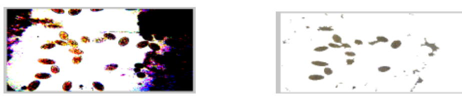
Applying the linear point operation to increase the contrast of the images gave us very good results in the enlarged images (zoomed x400). Different values of α>1 were tried to increase the contrast level. the optimal value of α is 1.9 which was obtained by trail and error.



(a) (b)

Figure (2): a) the output of the linear function
b)the output after detecting the objects from the background

But the nonlinear technique gave better results to the images zoomed x100 than the linear technique. The second type of image was more advantageous than the first type . The optimum values for α's were found by trial and error to be 0.9 and 0.7 for the first and second techniques respectively. When we test Rescaling technique on images, we find that it makes the mean of the image equals 0 and it's variance equal 1. Results of this technique are show in figure (3).



(a) (b)

Figure (3): a) the output of the rescaling technique
b)the output after detecting the objects from the background

To correct the non-uniform illumination in the pre-processed images, we first find a coarse estimate of the background illumination by determining the minimum of each 33-by-32 block in the image. A surf plot of the coarse estimate is shown in figure (4).

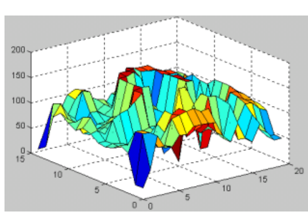


Figure (4): A surf plot of the coarse estimate

The coarse estimate is then expanded in size so that it is the same size as the original image. A bi-cubic interpolation is used to ensure that the data is smooth. Then computed illumination is then subtracted from the original image to correct non-uniformity.

After finishing image enhancement, we adjust the image. In our work, we used an adjustment function that is considered the second step in pre-processing procedure. This adjustment function is explained in the section 4.

4. Image Adjustment Function

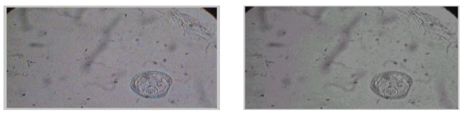
In our work, we used an adjustment function to rescale the image so that it covers the entire dynamic range ([0,1]). It gave good results except when it was used after the rescaling technique discussed previously. The adjusted images are shown in figure (5).

(a) (b)

Figure (5): a) The original image
b) The output of the adjustment function.

In our sample images, we found some images that have noise, which causes some problems in the programming. These problems were: This noise could be detected as a part of an egg. It may connect two or more eggs and this leads to detect them as one egg

Noise may be detected as an egg (this problem is solved in the last section). To remove these types of noise, as an important step in pre-processing procedure, we used the median filter. The median filter is an effective filter, when the noise pattern consists of strong, spike-like components and in the characteristic to be preserved is edge sharpness. We are using a filter of size 3 X 3, which gave us the best results by trial and error. The result is shown in figure (6).



(a) (b)
Figure (6): a) The original image.
b) The output of the filter.

As we mentioned in section 3, that the segmentation is one of the pre-processing steps which are used in image analysis stages. So in section 5, we introduce the segmentation analysis problems and then in subsection 5.1, we explain canny mask segmentation that is used in our work.

5. Segmentation Problems' analysis

To segment objects from each other, First we scan the out put image from the previous step to get the starting point, we add this point to the position matrix (this matrix contains the positions of the contour of each object), then we loop to find the remaining points of the contour. We take the first position we find without checking whether the other positions are valid or not. Every time we find a position we check whether it is already in the position matrix or not, if it is we stop to avoid entering an endless loop. Then we use the matrix of positions to get the segmented object from the original image by calculating the width and height of the object then putting it in a matrix with these dimensions, and then clearing it from the original image. This primary technique caused two problems:

The first problem was that the next position to be tested is out of the range of the image, i.e. the object to be segmented is at the border of the image. To solve this problem we pad the image with a white border before scanning it.

The second problem occurred because a position was reached that was already in the position matrix, but it was not the starting point of the contour this caused the search for more positions to stop. To solve this problem, we put a condition that if the position under test is in the position matrix, we check if it is the first position (i.e. the staring point) then we stop, otherwise we check the other positions. So we save the position found in some variables but not in the position matrix again, if no other new position is found we go back to the position saved and search around it for a new position. But sometimes there is more than one position around the position under test all of them are already in the position matrix, this caused the problem entering an endless loop in an object. The program kept looping around the second egg because of the pixel that joins between them.

To solve this, we use a function that returns the best position from the positions surrounding the position under test when all of them are already in the position matrix. Best position meaning a position has a valid next position around it. If no best position if found we use the last one found and try around it, this sometimes causes entering an endless loop, to avoid this we defined a maximum number of repetitions if exceeded we break using only the positions saved to define the new object segmented.

We detect noise by two consecutive ways:

- First we check the size of the segmented object by comparing it to the maximum and minimum expected values of the height and width of the parasite, if it is within range we assume that it is a valid egg of a parasite. But sometimes some matrices containing noise are within range, so we pad the matrix to the maximum expected values of height and width, then we calculate the ratio of colour that is not white to all colours including the white background. If this ratio is less than some threshold we assume that this matrix contains noise.

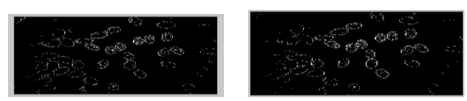
After noise removal, we need to perform the last step in pre-processing phase 'segmentation' for two reasons

- First we need to separate the eggs from the background.
- In images containing more than one egg, we need to get each egg alone.

The image used in detecting objects from background is the output image from the previous steps, in these images the objects to be segmented from the background differ greatly in contrast from the background, changes in the background can be detected by operators that calculate the gradient of an image. In the following subsection 5.1, we illustrate why we choose canny mask in our segmentation work and then explained the mask itself.

5.1 Canny Mask Segmentation

We have several ways to sement the objects from background and calculate the gradient of an image: The Sobel method finds edges using the Sobel approximation to the derivative. It returns edges at those points where the gradient of the image is maximum. In the Prewitt method finds edges using the Prewitt approximation to the derivative. It returns edges at those points where the gradient of the image is maximum. The Roberts method finds edges using the Roberts approximation to the derivative. It returns edges at those points where the gradient of the image is maximum. To find edges by looking for zero crossings after filtering the image, we use a Laplacian of Gaussian filter. The zero-cross method finds edges by looking for zero crossings after filtering the image with a filter we specify. And finally, the Canny method finds edges by looking for local maxima of the gradient of the image. The gradient is calculated using the derivative of a Gaussian filter. The method uses two thresholds, to detect strong and weak edges, and includes the weak edges in the output only if they are connected to strong edges. This method is therefore less likely than the others to be "fooled" by noise, and more likely to detect true weak edges. These techniques were tried on different images. We concluded that applying the sobel mask on images enhanced using the linear technique gave the best results in images zoomed x400, while applying the canny mask after the correction of non-uniform illumination technique gave best results on images zoomed x100. Figure (7) shows the output of the two masks.

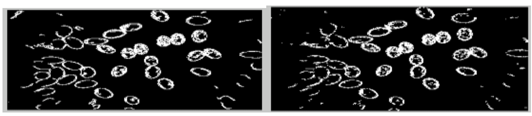


(a) (b)
Figure (7): a) output of the canny mask
b) output of the sobel mask

The binary gradient mask shows the lines of high contrast in the image, these lines do not quite delineate the outline of the object of interest, compared to the original image the appears gaps in the lines surrounding the objects in the gradient mask, these gaps will disappear if the output image is dilated using linear structuring elements, we are using structuring elements in these forms

$$S1=\begin{bmatrix}1&1&1\\1&1&1\\1&1&1\end{bmatrix}\quad S2=\begin{bmatrix}0&1&0\\0&1&0\\0&1&0\end{bmatrix}\quad S3=\begin{bmatrix}0&0&0\\1&1&1\\0&0&0\end{bmatrix}$$

The binary gradient mask is dilated using the vertical structuring element (S2) followed by the horizontal structuring element (S3), the out put is shown in figure (8).



(a) (b)
Figure (8): the outputs of the two masks after dilation

The dilated gradient mask show the outline of the eggs quite nicely, but there are still holes in the interiors of the eggs, to fill these holes we use a filling function. Then we smooth the object by eroding the image twice with the diamond structuring element (S3). After smoothing objects, we use the original image to get the colours of the segmented object.

After we show how canny mask work and why it is the best for our work, we test different Laplacian sigmas and so the statistical results of using them.

In section 6, we will show some statistical results of the different techniques we used in the different stages of our system. We tested five different image enhancement techniques, before we apply the edge detection function using a gradient mask, we have six different gradient masks, so we needed to know which enhancement technique works better with which mask. We tried the different combinations on a data set containing of two images from each type of parasite, two from each zooming level (we have two zooming levels). We note that the Laplacian mask and the Zero-crossing mask gave very similar results. So we will only show the results of the Laplacian mask.

6. Results using different Laplacian Sigmas

In the classification our work part, we tested different values of sigma and we get the following results. (The values inside the cell represent are the percentage of successful classification).

6.1 Eccentricity Sigma

The following percentage results occurs

Parasite Type	**sigma**	0.1	0.2	0.3	0.4	0.5
Moniezia		98%	98%	98.2%	98.3%	98.5
Paramphistomum		97%	97.2 %	97.4%	97.7%	98

Table (1)

We found that all parasite types are classified correct for all values of eccentricity sigma. So we can choose any value of sigma.

6.2 Hu-moment Sigma

The following results occurs

Parasite Type	**sigma**	0.1	0.2	0.3	0.4	0.5
Moniezia		97%	97%	97%	97%	97%
Paramphistomum		97%	75%	50%	0%	0%

Table(2)

We found that the best value of sigma that successfully classifies all parasite types is 0.1 .

6.3 Co-occurrence Sigma

The following results occur:

Parasite Type	sigma	0.1	0.2	0.3	0.4	0.5
Moniezia		98%	98%	99%	99%	99%
Paramphistomum		97%	98%	98%	99%	99%

Table(3)

We found that the worst value of sigma that successfully classifies all parasite types is 0.1.

6.4 Histogram Sigma

The following results occur:

Parasit Type	Sigma	0.1	0.2	0.3	0.4	0.5
Moniezia		96%	96%	96.2%	96.5%	96.8
Paramphistomum		97%	97.2%	97.4%	97.8%	98%

Table(4)

From all the previous results, So we conclude that the combining between the Sobel mask and the linear technique gives the best results, those results were also sharp and the eggs were not distorted, it helped in very good segmentation.

7. Conclusion

Content-based image retrieval (CBIR), is the application of computer vision to the image retrieval problem, that is, the problem of searching for digital images in large databases. CBIR draws many of its methods from the field of image processing and computer vision, and is regarded by some as a subset of that field. It differs from these fields principally through its emphasis on the retrieval of images with desired characteristics from a collection of significant size.

The extent to which CBIR technology is currently in routine use is clearly has great impact on the more general applications of image searching. However, the main drawback of current CBIR systems is more fundamental. It is that the only retrieval cues they can exploit are primitive features. Hence current CBIR systems are likely to be of significant use only for applications at level 1. Retrieval does not always yield images that have clearly the same feature, unless the database contains many images with a dominant one. Searching by laplacian sigmas and using canny mask give often surprising results. Apparently the features used for matching are not the most effective ones.

There are the notions of precision (the ratio of relevant images to the total number of images retrieved) and recall (the percentage of relevant images among all possible relevant images) as we notice in our work. We conclude from all the previous results for ratio of relevant images that the combining between the cumulative histogram technique, the canny mask and different sigmas give the best results, but these results were based on the noise ratio.

References

[1] E. Loupias and s. Bres. Key point-based indexing for pre-attentive similarities: The kiwi system. *Pattern Analysis and Applications*, 4(2/3):200–214, 2011.

[2] Timothy J. Mills, David Pye, David Sinclair, and Kenneth R. Wood. Shoebox: A digital photo management system. Technical Report Technical Report 2000.10, AT&T Research,http://ftp.uk.research.att.com:/pub/docs/att/tr.2000.10.pdf, 2000.

[3] Sven Siggelkow and Hans Burkhardt. Fast invariant feature extraction for image retrieval. In Remco C. Veltkamp, Hans Burkhardt, and Hans-Peter Kriegel,

editors, *State-of-the-Art in Content-Based Image and Video Retrieval*. Kluwer, 2009.

[4] David McG. Squire, Wolfgang M¨uller, Henning M¨uller, and Thierry Pun. Content-based query of image databases: inspirations from text retrieval. *Pattern Recognition Letters*, 21:1193–1198, 2000.

[5] Jules Vleugels and Remco C. Veltkamp. Efficient image retrieval through vantage objects. *Pattern Recognition*, 35(1):69–80, 2006.

[6] C. T. Yu Y. A. Aslandogan. Multiple evidence combination in image retrieval: Diogenes searches for people on the web. In *Proceddings of 23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2000)*, pages 88–95, 2000.

[7] Andrew Berman and Linda Shapiro. A flexible image database system for content-based retrieval. *Computer Vision and Image Understanding*, 75(1/2):175–195, 1999.

[8] R. Brunelli and O. Mich. Image retrieval by examples. *IEEE Transactions on Multimedia*, 2(3):164–171, September 2000.

[9] D. Tao, X. Tang, X. Li, and X. Wu. Asymmetric bagging and random subspace for support vector machines-based relevance feedback in image retrieval. IEEE Trans. on Pattern Analysis and Machine Intelligence, 28(7):1088–1099, 2006.

[10] Z.N. Li and M. Drew. Fundamentals of Multimedia. Prentice Hall, 2004.

[11] J. Fournier, M. Cord, and S. Philipp-Foliguet. Retin: A content-based image indexing and retrieval system. *Pattern Analysis and Applications*, 4(2/3):153–173, 2001.

[12] Theo Gevers and Arnold Smeulders. Pictoseek: Combining color and shape invariant features for image retrieval. *IEEE Transactions on Image Processing*, 9(1):102–119, January 2000.

[13] ByoungChul Ko, Jing Peng, and Hyeran Byun. Region-based image retrieval using probabilistic feature relevance learning. *Pattern Analysis and Applications*, 4(2/3):174–184, 2001.

[14] Jorma Laaksonen, Markus Koskela, Sami Laakso, and Erkki Oja. Picsom – content-based image retrieval with self-organizing maps. *Pattern Recognition Letters*, 21:1199–1207, 2000.

[15] T.-S. Lai. *CHROMA: a Photographic Image Retrieval System*. PhD thesis, School of Computing, Engineering and Technology, University of Sunderland, UK, January 2000. `http://osiris.sund.ac.uk/˜cs0sla/thesis/`.

[16] Ingemar J. Cox, Matthew L. Miller, Thomas P. Minka, Thomas Papathomas, and Peter N. Yianilos.The Bayesian image retrieval system, PicHunter: Theory, implementation and psychophysical experiments. *IEEE Transactions on Image Processing (to appear)*, 2000.

[17] B. S. Manjunath, Jens-Rainer Ohm, Vinod V. Vasudevan, and Akio Yamada, "Color and Texture Descriptors". In: IEEETransactions on Circuits and Systems for Video Technology, Vol. 11, No. 6, June 2001,pp. 70-715

[18] Alberto Amato, Vincenzo Di Lecce, "Edge Detection Techniques in Image Retrieval: The Semantic Meaning of Edge", 4th EURASIP Conference on Video/Image Processing and Multimedia Communications, Zagreb, Croatia. pp. 143-148.

[19] J.Huang, S.R.Kumar, M.Mitra, W.J.Zhu and R.Zabih, ”Image Indexing using color corrolograms”,Proc. IEEE conf. on
Computer vision and Patternn REcongnition,PP 762-768, June 1997.

[20] Gupta, R. Jain, Visual information retrieval, Communications of the ACM 40 (5) (1997) 70-79.

[21] Y. Rui, T.S.Huang, S.-F.Chang, Image retrieval Past, present and future, in: M. Liao(Ed.),Proceedings of the International Symposiumon Multimedia Information Processing Taipei,Taiwan,1997.

[22] Theo Gevers and Arnold Smeulders. Pictoseek: Combining color and shape invariant features for image retrieval. IEEE Transactions on Image Processing, 9(1):102–119, January 2000.