

Color Image Segmentation Applied to Medical Domain

Liana Stanescu¹, Dan Dumitru Burdescu¹, Cosmin Stoica¹

¹ University of Craiova, Faculty of Automation, Computers and Electronics, Craiova, Romania
{Stanescu, Burdescu_Dumitru, Stoica_Cosmin}@software.ucv.ro

Abstract. The article presents two practical ways of using the automated color image segmentation in the medical field: for content-based region query and for tracking the time evolution of the disease in patients following a certain treatment. A known technique was used for automated color medical image segmentation – the color set back-projection algorithm. Our previous work in extraction of color regions from a database of nature images using the same algorithm showed promising results. The images are transformed from RGB to HSV color space, quantized at 166 colors and processed by the color set back-projection algorithm that allows the color region detection. The algorithm is studied from two points of view: complexity and the retrieval quality. The experiments that were made on a database with color endoscopy images from digestive tract have shown satisfying results for both applications that are important in practical medical use and medical teaching.

Keywords: image processing, content-based region query, color set back-projection algorithm, HSV color space, color regions.

1 Introduction

Because digital images are produced and used in many domains, the process of retrieving visual information has become an area of intense research during the last years. The initial text-based methods used for finding similar images have been augmented by query-by-example methods that need only to establish a query image without any other helping information. This methodology takes into consideration both coarse granularity and fine granularity preprocessing of the image data. If in the first case, the image is considered as a whole, in the second case the image is processed and the relevant regions are extracted and used for query [10]. In this paper, a fine granularity approach to image indexing and retrieval is adopted.

Medicine is one of the domains where a large quantity of gray-scale or color images, produced by medical devices for establishing a more correct diagnosis, is accumulated. In the medical departments, thousands of images are collected including color endoscopy images, such as the ones used in this paper. In these digestive tract images, the abnormal aspect of the color and texture of the tissue is relevant in indicating a great number of diseases. A series of studies have been made and were created applications that permit medical image processing and the complex query

taking into consideration many criteria and characteristics [2], [3], [4], [5] and also automatic image segmentation [9], [10], [11].

The applications presented in this paper use a database that is populated by color images captured by an endoscope from a number of patients with stomach diseases. The color set back-projection algorithm is applied on each image for detecting the color regions. The region(s) that represents the sick zone are marked as relevant. The color, the minimum bounding rectangle and the number of pixels that roughly indicates the dimension of the sick region, are stored in the database.

The automated medical color image segmentation has two important uses:

1. in content-based region query
2. during the time evolution of the disease in patients following a certain treatment.

In content-based region query on medical images collections, the specialist chooses one or more of the detected regions for querying the database, with the purpose to retrieve images which are similar by color, texture or both; this can be useful for clarifying some uncertain diagnosis, for seeing the evolution and the treatment in patients with the same diagnosis or in medical teaching.

Concerning the second original use, by applying a certain drugs treatment to some patients, at certain time intervals that are strictly specified by the physician, the images are again collected by the endoscope from the same patients while the same algorithm is performed for detecting the color regions. The relevant region(s) that represents the sick tissue is marked. The comparison between the new and old regions detected as relevant for the same patient, taking into consideration the number of pixels, can help the physician to establish in what percentage the sick region is reduced because of the administrated drugs. This approach may lead to a more rapid and correct estimation of the percentage in which the medication has a good effect in the ulcer diagnosis. Based on the information stored in the database for each patient of the studied lot of sick persons, the statistics about treatment's efficiency can be generated. This may be of help for the patients, specialists, and drugs producers that may intend to test a new medical product.

Motivating factors for our work include:

- There are relatively few published studies for medical color images of the digestive tract.
- Our previous work in extraction of color regions from a database of nature images showed promising results.
- We expect the method to provide a useful, valuable capability for clarifying an uncertain diagnosis or observing the evolution of disease, or observing the regression of disease as a result of treatment; another use is medical teaching, to find additional examples of a particular visual phenomenon.

The paper is structured as follows: section 2 presents the algorithm used for detecting color regions from medical images. Section 3 presents the experiments made and the results obtained during the content-based region query process, and section 4 shows the results obtained in the process of time observation of the disease in patients that follow a certain treatment. Section 5 contains the conclusions.

2 The Color Set Back-Projection Algorithm

For detecting the color regions, it was chosen the color set back-projection algorithm, introduced initially by Swain and Ballard and then developed in the research projects at Columbia University [6]. This technique provides the automated extraction of regions and the representation of their color content. The extraction system for color regions has four steps [6]:

1. the image transformation, quantization and filtering (the transformation from the RGB color space to HSV color space and the quantization of the HSV color space at 166 colors)
2. back-projection of binary color sets
3. the labeling of regions
4. the extraction of the region features

The algorithm reduces the insignificant color information and makes evident the significant color regions, followed by the generation, automatically, of the regions for a single color, for two colors, for three colors.

To conclude with, the second step of the color set back-projection algorithm is the following [6]:

1. Detection of single color regions

1.1 Making the image histogram, $H[m]$, all the values $m'=m$ for which $H[m] \geq p_0$ are detected.

1.2 For each m' the color set c having the property $c[k]=1$ for $k=m$ and $c[k]=0$ in other cases is found. On the image $R[m,n]$ the back-projection algorithm for each color set c is applied and the color regions are found. For each region n the local histogram $L_n[m]$ is stored.

1.3 The residue histogram $H_r[m]=H[m]-\sum_n L_n[m]$ is computed.

2. Detection of two colors regions, detection of three colors regions...

For each detected regions the color set that generated it, the area and the localization are stored. The region localization is given by the minimal bounding rectangle (MBR). The number of color pixels represents the region area.

For the implementation of the color set back-projection algorithm, the initial image is transformed in HSV format and quantized. At the end of this process, the color set of the image is obtained [6]. On the matrix that memorizes only the quantized colors from 0 to 165 it is applied a 5x5 median filter, which has the role of eliminating the isolated points. Having the HSV quantized matrix it is possible to begin the process of regions extraction presented above. It may be observed that this process is in fact a depth – first traversal, described in pseudo-code in the following way [1], [12]:

```
FindRegions (Image I, colorset C) is:
InitStack(S)
Visited =  $\emptyset$ 
for *each node P in the I do
    if *color of P is in C then
        PUSH(P)
        Visited  $\leftarrow$  Visited  $\cup$  {P}
        while not Empty(S) do
            CrtPoint  $\leftarrow$  POP()
```

```

Visited ← Visited ∪ {CrtPoint}
For *each unvisited neighbor S of
CrtPoint do
    if *color of S is in C then
        Visited ← Visited ∪ {S}
        PUSH(S)
    end if
end for
end while
* Output detected region
end if
end for

```

The total running time of a call of the procedure FindRegions (Image I, colorset C) is $O(m^2 \cdot n^2)$, where “m” is the width and “n” is the height of image.

Proof:

Recall that the number of pixels of the image is $m \cdot n$, where “m” is the width and “n” is the height of the image. As it is observed next the first loop FOR of the algorithm is executed once at most for each pixel P in the image. Hence, the total time spent in this loop is $O(n \cdot m)$. The WHILE loop processes the stack S for each pixel that has the same color of its neighbor. The inner loop FOR processes the pixels of unvisited neighbor. So, the total time spent in these loops is $O(m \cdot n)$, because all pixels of the image are processed once at most. From previous statements it is inferred that the total running time of this procedure is $O(m^2 \cdot n^2)$.

3 Content-Based Region Query - Experiments and Results

In content-based region query, the user selects one or several query regions, trying to find in the database the images containing similar regions from the point of view of color, size or spatial extent.

Taking into account that the color information of each region is stored as a color binary set, the color similitude between two regions may be computed with the quadratic distance between color sets [6].

Other two important distances are taken into consideration: the distance in area between two regions and the distance in MBR width (w) and height (h) between two regions q and t [6]. The single region distance is given by the weighted sum of the color feature $d_{q,t}^f$, area $d_{q,t}^a$ and spatial extent $d_{q,t}^s$ distances. The user may also assign a relative weight α to each attribute. For example, the user may weigh the size parameter more heavily than feature value and location in the query. The overall single region query distance between regions q and t is given by [6]:

$$D_{tot} = \alpha_a \cdot d_{q,t}^a + \alpha_s \cdot d_{q,t}^s + \alpha_f \cdot d_{q,t}^f. \quad (1)$$

Figure 1 shows an image representing gastric cancer diagnosis. The sick part in the image is the prominent oval zone presenting bleedings. The color regions detected by

the color set back-projection algorithm are presented in figure 2. From the medical point of view the region 5 and 6 are important.

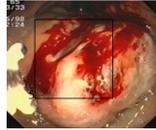


Fig. 1. An image representing gastric cancer diagnosis.

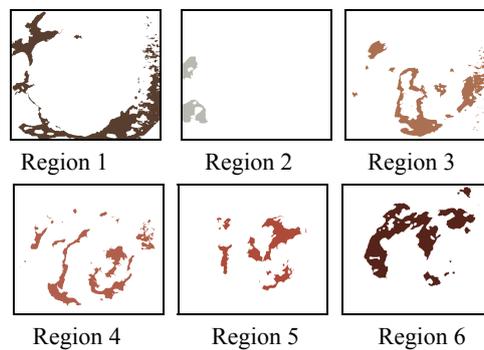


Fig. 2. Color regions detected by the algorithm.

The results of the query based on two regions, the Region 5 and 6, are presented in figure 3.

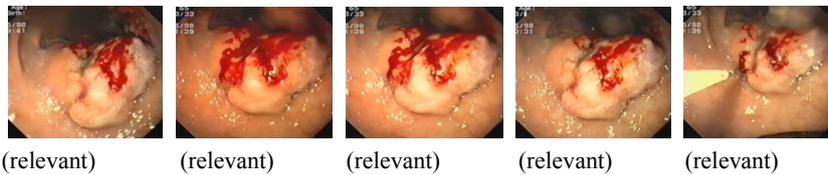


Fig.3. The results of the content-based query based on the regions Region 5 and Region 6.

The query finds 5 relevant images in the first 5 returned images.

The quality of the content-based region query process was tested in the following conditions:

- in the database there are 960 images representing: polyps, colitis, ulcer, ulcerous tumor and esophagitis
- for each image the color regions are detected using the color set back-projection algorithm
- the information about regions (color, area, MBR, position) is stored in the database
- content-based region queries were made on one or many color regions significant from the medical point of view. For each query, the images from the databases were inspected and relevance was assigned to them (1 – relevant, 0 – irrelevant).

In table 1, there are presented some of the results of the queries made. For each experiment the following information is given: the number of query regions, the weights for the distances (color, area and MBR) and the number of relevant images found in the first five retrieved images.

Table 1. The experimental results.

Nr. of Query Reg.	Color	Area	MBR	Nr. of relevant images
1	100	0	0	4
1	100	0	0	4
1	100	100	0	3
1	100	100	100	2
2	100	0	0	4
2	100	100	0	3
2	100	1000	100	2

4 Tracking the Time Evolution of the Disease - Experiments and Results

In order to evaluate the efficiency of the color set back-projection algorithm in tracking the time evolution of the disease, it was tested on a gastroenterologic imagistic database at Filantropia University Hospital of Craiova. A double blind, double dummy, placebo controlled study was performed in order to evaluate the efficacy and safety of a new proton pump inhibitor in the treatment of peptic ulcer. Two hundred and two patients were included and allocated in two groups: a placebo group (102) and a treatment group (100). One endoscopic image was taken at inclusion for each patient, in order to evaluate the cicatrization process under treatment. During the 6 weeks of follow-up, another endoscopic image of the ulcer was taken after the first 3 weeks and another one at the end of the treatment (6 weeks). Both images were stored in the database.

The endpoint of the study was to estimate the percentage of patients in which the ulcer size diminished with more than 50% after three weeks of treatment and with more than 80% at the end of the treatment. A randomization number was allocated to each patient. Each image stored at randomization was labeled XXXT₀, in which XXX was the randomization number and T₀ designed the initial aspect of the ulcer, before treatment. Using the same algorithm, the image stored at the second visit (three weeks of treatment) was labeled XXXT₁ and XXXT₂ at the end of the treatment. Only the endoscopist made the initial evaluation, i.e. the diagnosis of peptic ulcer at T₀. In order to evaluate the colour set back-projection algorithm, at the following visits both the endoscopist and the computer performed the search in the database, and an independent expert validated the results. The validated results were quantified as „Identified” or ”Non-Identified” both for the observational-based method and the computer-based one and included in statistics. The duration of the retrieval process was also taken into account and included in the statistics. In figure 4 are presented the important regions detected by the algorithm at the moment T₀ and T₁ for the selected image. At the moment T₁, after a period of treatment the area of the region is smaller

The percentage of correct diagnosis attended by using the classic observational method is compared with the results obtained with the computer-based one in terms of reliability and reproducibility of this method. Thus it is evaluated the advantage of using a computer-based retrieval system in terms of reducing the time spent on such an operation. For this purpose a statistical computation was used, performing the non-parametric two-sample Wilcoxon test for independent data and the analysis of variance.

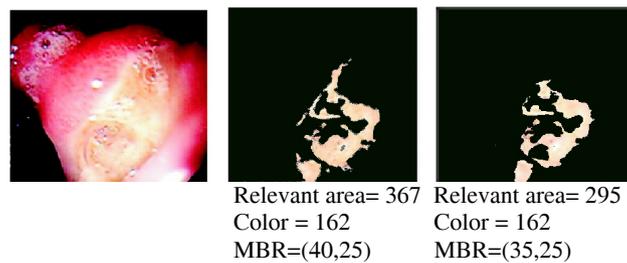


Fig. 4. Image with relevant color regions at the moment T_0 and T_1 .

In all, both the observer and the computer software correctly identified the healing staging at T_1 (three weeks visit) in 170 patients, i.e. 84.15% while the attended statistical value was 168 patients, and in 158 patients at T_2 (end-of-treatment visit) i.e. 78.21% while the attended statistical value was 155. At T_1 , the human observer correctly staged the ulcer in 188 patients (93.07%), missing 14, while the computer software staged correctly 181 (89.60%) missing 21. At the last visit (T_2) the human observer staged correctly 186 patients (92.07%) missing 16 and the computer provided a good staging in 169 (83.66%) missing 33.

There were slight differences between human observer and computer system in the appreciating the healing staging in the way that the endoscopist missed the correct staging in certain cases and the software in others, but in all the inter-observer agreement was 84.15% at T_1 and 78.21% at T_2 as seen in table 2.

Table 2. Evaluation of inter-observer agreement in human- and computer-based query in the database.

Patients: 202		Correctly staged		Missed by both computer and observer		Missed by computer but not the observer		Missed by observer but not the computer	
		Nr	%	Nr	%	Nr	%	Nr	%
T1	HO	188	93.0	3	1.48	11	5.44	8	3.96
	CS	181	89.6	3	1.48	15	7.42	18	8.91
	Both	170	84.1	3	1.48	18	8.91	11	5.44
T2	HO	186	92.0	5	2.47	11	5.44	6	2.97
	CS	169	83.6	5	2.47	23	11.38	28	13.86
	Both	158	78.2	5	2.47	28	13.86	11	5.44

Despite this, no significant statistical difference between human observer and computer system was noticed by using the Chi-square test and the Wilcoxon nonparametric test, arguing for a good liability of the method (table 3).

Table 3. Statistic analysis of differences between human- and computer-based querying in the database.

	Observer-based evaluation of staging vs. computer-based evaluation at T1 (3 weeks)	Observer-based evaluation of staging vs. computer-based evaluation at T2 (end-of-treatment)
N	202	202
Missing	0	0
DOL	1	1
Chi2	1.966	2.828
p (Chi2)	0.1609	0.0927
G-square	1.589	2.396
p (G-square)	0.2075	0.1217
Contingence	0.098	0.117
Phi	0.099	0.118
corrected Chi2	0.899	1.767
corrected p	0.341	0.1838
Fisher's exact test	0.1661	0.1482
Mann-Whitney U	19695.00	18685.00
U'	21109.00	22119.00
Wilcoxon T	40198.00	39188.00
Significance level for one-sided test	11.11%	0.52%
Significance level for two-sided test	22.22%	1.03%

Table 4. Descriptive statistics of time spent by human observer and computer software to retrieve and evaluate a patient's record.

Time for retrieval (seconds)	Human observer, visit T1	Computer software, T1	Human observer, visit T2	Computer software, T2
N	202	202	202	202
Mean	37.520	22.149	37.807	18.881
Std. Deviation	16.962	8.878	14.499	9.914
Std. Error	1.193	0.625	1.020	0.698
Min.	9.000	5.000	6.000	7.000
Max	76.000	50.000	76.000	54.000
Missing	0	0	0	0

The speed of the retrieval process was also tested, comparing the time spent by the observer and the computer to find each patient's record in the database. This process was electronically measured and stored in the computer for statistics. The result is that the software has a significantly higher speed than human observer with no significant

decrease of the retrieval quality. In average, at T_1 visit, the observer spent 37.52 seconds for each patient to find and evaluate the record, while the computer needed only 22.14. At visit T_2 the human used 37.80 seconds to retrieve and evaluate the record, and the computer only 18.88 seconds (table 4). These differences had statistical significance as showed the t and Z test with a 95% confidence interval (table 5).

Table 5. Evaluation of statistical significance of differences observed in terms of retrieval speed between human observer and computer software

Timing comparisons: human observer vs. computer software	Mean difference	DOL	t	p	95% inf.	95%sup
T1	15.371	201	11.937	<0.0001	12.832	17.910
T2	18.926	201	15.094	<0.0001	16.453	21.398

5 Conclusions

This article presents the experimental results obtained in two medical applications based on automated image segmentation realized with the help of the color set back-projection algorithm. Our goal was to evaluate this algorithm for detection of diseased tissue regions, since in previous work we had observed satisfactory results for region detection in nature images. We hypothesized that the endoscopic abnormalities are detectable in these images by color change.

In the first application we used these regions to query our database by image content, and inspected the retrieved images for similarity to the query image. We obtained an average of 3 relevant images in the first 5 returned by the system, when we used color and area as search parameters. The results obtained are promising, but a larger and more extensive investigation is required to properly evaluate this method for practical medical use. A more careful evaluation is needed in order to establish correctly the algorithms quality and the usefulness of the application in clinical environment. Currently, our software tool for automatic image segmentation on color and content-based region query is used in medical teaching.

Concerning the second original application, endoscopic measurement of lesions is of great importance in the design and performance of clinical trials, as, in this case of peptic ulcer disease. Although endoscopes are constructed with wide-angle lenses that significantly distort the image by creating a relative compression of points in its periphery, the presented computerized method showed a remarkable capacity to retrieve similar images in a quite large database giving an acceptable reproducibility for independent evaluations. Some authors [7] have already developed a computer program to correct the distortion of the wide-angle lens, which permits reducing errors to 1.8% +/- 2.2% ($p < 0.05$). On the other hand, evaluations of ulcer healing by measuring video-endoscopy images as a dynamic process could not be realized up to now. The documentation of the dynamic healing process shattered either on the patient's compliance or on the inconstancy of the image cut due to wobbling [8]. In

this work, instead of presenting a continuous film, instant takes of ulcer healing were processed. The inter-observer disagreement between human observer and the computer software was significantly low and the speed of the computerized method was higher, proving benefits in terms of time and cost efficiency. In the future, the computerized method will be developed in order to increase the capability of the system to recognize more complex lesions than ulcers and to diminish the percentage of errors.

References

1. Stanescu, L., Burdescu, D., Mocanu, M.: Detecting Color Regions and Content-based Region Query in databases with Medical Images. *Periodica Politecnica, Transactions on Automatic Control and Computer Science*, vol. 49(63) Timisoara (2004)
2. Müller, H., et al.: Comparing feature sets for content-based medical information retrieval. *SPIE Medical Imaging*, San Diego, CA, USA (2004)
3. Müller, H., et al.: A review of content-based image retrieval systems in medicine-clinical benefits and future directions. *International Journal of Medical Informatics*, vol. 73, 1--23 (2004)
4. Lehmann, T.M., et al.: The IRMA Reference Database and Its Use for Content-Based Image Retrieval in Medical Applications. Ammenwerth E, Gaus W, Haux R, Lovis C, Pfeiffer KP, Tilg B, Wichmann HE (Hrsg): *GMDS 2004 - Kooperative Versorgung - Vernetzte Forschung - Ubiquitäre Information*. Verlag videel OHG, Niebüll, 251--253 (2004)
5. Shyu, C. et al.: ASSERT, A physician-in-the-loop content-based image retrieval system for HRCT image databases. *Computer Vision and Image Understanding*, vol. 75, Nos. 1/2, 111--132 (1999)
6. Smith, J.R.: *Integrated Spatial and Feature Image Systems: Retrieval, Compression and Analysis*. Ph.D. thesis Graduate School of Arts and Sciences, Columbia University (1997)
7. Vakil, N., Smith, W., Bourgeois, K.: Endoscopic measurement of lesion size: improved accuracy with image processing. *Gastrointest Endosc.* Nov-Dec; 40(6) 779--780 (1994)
8. Jaspersen, D., Keerl, R., Weber, R., Huppmann, A., Hammar, C.H., Draf, W.: Dynamic imaging of gastric ulcer healing using the most modern Morph-Software. *Bildgebung*. Jun; 63 (2) 101-4 (1996)
9. Henning, M., Samuel, M., Gilles, C., Pierre-Alexandre, P., Christian, L., Antoine, G.: Automatic abnormal region detection in lung CT images for visual retrieval, *Swiss Medical Informatics*, vol. 57 2--6 (2005)
10. Mezaris V., et al: A Test-Bed for Region-Based Image Retrieval Using Multiple Segmentation Algorithms and the MPEG-7 eXperimentation Model: The Schema Reference System. In: *Proceedings of CIVR 2004, Lecture Notes in Computer Science 3115 Springer ISBN 3-540-22539-0* (2004)
11. Gordon, S., Zimmerman, G., Greenspan H.: Image segmentation of uterine cervix images for indexing in PACS. In: *Proceedings of the 17th IEEE Symposium on Computer-Based Medical Systems, CBMS 2004*. Bethesda, MD (2004)
12. Burdescu, D., D., Stanescu, L.: A New Algorithm For Content-Based Region Query In Databases With Medical Images. *Studies In Health Technology and Informatics Vol. 114 Medical and care Compunetics 2*, IOS Press Amsterdam ISBN 1 58603 520 7, 132--139 (2005)