Nonlinear Analysis: Modelling and Control, Vilnius, IMI, 2000, No. 5

Quality of Colour Image Segmentation: the Measures

R. E. Blake[a] **and A. Juozapavičius**[b]

[a]Norwegian University of Science and Technology, Trondheim, Norway
[b]Vilnius University, Lithuania
richard.blake@idi.ntnu.no
algimantas.juozapavicius@maf.vu.lt

Received: 27.08.2000
Accepted: 27.09.2000

Abstract

Colour image segmentation is an important tool in many applications, like robotics, computer vision and data compression. Differing highly from grey-scale images, the colour segmentation usually has more complicated and time consuming algorithms and is controlled by a larger set of parameters. The measures of quality of colour segmentation are presented in the article, enabling users to evaluate the most efficient set of parameters for a procedure of colour segmentation.The evaluation of parameters are suggested to be provided by heuristical and statistical methods, also presenting relationship of such measures.

Keywords: colour image segmentation, quality of segmentation, measures, relationship of measures.

1 Introduction

Image segmentation is a very useful method in many applications, like computer vision and data compression. Various computer vision systems such as object recognition, industrial inspection, robotics and measurement, use image segmentation as an early step in the image processing [1]. In data

[1]This article was written when A. Juozapavičius was a visiting professor at the Norwegian University of Science and Technology, Trondheim, Norway

compression, the MPEG standard use efficient segmentation of images or groups of images to reduce data amount optimally, in conjunction with shape coding and texture coding [2]. In any case, image processing has to enable the robust decision making. These decisions used to occur at the lowest level of processing possible, referring to properties of pixels, or groups of pixels, and image segmentation is made therefore at the pixel level. In many cases this decision making relies upon segmentation process, and involves suitable measure of quality. This makes the importance of a quality measure widely recognized. The quality of segmentation affects the later steps in image processing. Once the quality is measured, it should be possible to tune parameters of the segmentation.

Colour image processing differs greatly from the monochrome one. Primarily, every pixel of colour image corresponds to a multidimensional point in suitable colour space, instead of being just scalar value as in a monochrome image. Secondary, the process (of colour segmentation) is controlled by more parameters then in grey-scale case. From the *dichotomies* in statistical pattern recognition [3], it may be stated that it's possible to segment colour image more reliably than a monochrome one. As a consequence, this situation leads to more complicated and more time consuming algorithms (see for example [4, 5, 6, 7]). Thus, it is even more important to devise methods of quality control that can help to tune parameters for optimal performance.

It is highly desirable therefore to have a procedure to monitor the quality of colour image segmentation and use it to help tune the parameters of the segmentation. This has been proposed before for range images, [8], and depends on defining some expected properties in the segmentation. In the next section we consider the general task of segmentation of a colour image and identify possible faults in the segmentation that will lead to a penalty cost.

There are many approaches in image segmentation algorithms suggested by numerious researchers [1, 3, 6, 7]. The segmentation of colour images can be based on some (usually luminance-only segmentation) or all of the image information. In the latter case, the segmentation is based on all of the colour components, either independently or combined. All these approaches vary very much in principles used, in algorithmics, design and implementation. The only feature in common of all of them is the presence of parameters for control of behaviour of segmentation procedure.

The subject of this paper is to introduce a measure of the quality of colour segmentation that can be used with any method of colour image segmentation. The segmentation procedure is thought of as a blackbox which takes as input a colour image and control parameters and gives as its output a segmentation. The measure of quality presented is analysed under conditions appropriate for applications.

2 Quality Criteria in Colour Image Segmentation

There are two main factors, guiding colour segmentation process: **colour** of a region, and **luminance**. These factors may be considered independently, and they are rarely conbined together in the same algorithm.

Using RGB space, the attributes of colour and luminance for each pixel may be expressed in a quantitative form: the colour is the ratio $r : g : b$ and the luminance is $\sqrt{(r^2 + g^2 + b^2)}$ (some applications use another expression for luminance: $r + g + b$). Physically these attributes are almost independent, they interact at low intensities only. This is because the colour ratio $r : g : b$ becomes seriously perturbed by quantisation and noise. Colour segmentation is the operation of finding, and labelling disjoint connected sets of pixels that have acceptable uniformity in the attributes of the members of each set.

Luminance. Luminance values are important. A single region can contain a wide range of luminance values. So long as these values are everywhere above a threshold, that makes the estimate of the colour ratio reasonably well conditioned. That is, the luminance represented by value $\sqrt{(r^2 + g^2 + b^2)}$, should be sufficiently large that $r : g : b$ is not seriously perturbed by noise.

Uniformity. Most colour segmentation algorithms partition image into 4-connected (or 8-connected) regions, that correspond to approximately uniform colour ratios in the original image. That is, with each pixel represented by the red, green and blue triple, r, g, b, the colour ratio $r : g : b$ should be approximately uniform over the region.

Images met with in the real world often convey information through a change of luminance with the colour ratios remaining constant. These effects may be covered by monochrome segmentation methods. Also, luminance and uniformity of the colour are implemented more directly in some other colour models, like YIQ and HSV.

Segmentation procedure usually labels regions separated. We define four fault conditions in relating segment labels to colour regions:

- Fault 1: the same label is applied to pixels with different colours;

- Fault 2: colour pixels that are adjacent in the image and have compatible colours and luminance have different labels;
- Fault 3: pixels are left without a label;
- Fault 4: pixels with compatible colours and the same label have significantly different luminance.

The words *compatible* and *different* mean that the variance of attributes are respectively inside and outside of allowed bounds.

Fault 1 represents **confusion of colours**, in which different colours are collected under one region, or the variance of colours exceeds suitable limits. Such region should probably be split into several regions to improve the quality.

Fault 2 represents **over segmentation**, in which adjacent regions containing essentially the same colour have different labels. These should probably be merged into fewer larger regions.

Fault 3 represents **failure to segment**, in which significant pixels have been ignored by the segmentation algorithm. Such pixels should be labelled so that the segmented image represents the information in the original image most effectively.

Fault 4 represents **region and shadow**, i.e. there is too much luminance varying among pixels with the same label. The danger is to identify object without its shadowy parts, the situation which has to be corrected. Nevertheless this case may not not a fault at all. The appropriate sensitivity to variations in luminance, for compatible colours, depends very much on the application. This fault may need to be considered but we do not include it in our experiments.

The definition of the faults is independent of method used to obtain a segmentation. The procedure for measuring the quality is also independent of a method of obtaining the segmentation.

The formal definition of the faults leads to an expression of the quality of segmentation in a mathematical form. Denoting:

- N_{pair} - number of pairs of pixels in the image;
- N_{lc} - number of pairs of pixels with the same label and different colours;
- N_{adj-lc} - number of pairs of adjacent pixels with the same colour and different labels;

- N_{adj} - number of pairs of adjacent pixels;
- N_{nl} - number of pixels without label;
- N - number of pixels in the image;
- N_A - number of pixels in a region labeled by A;
- N_{A-T} - number of pixels in the region A, whose luminance exceeds the *mean* of the intensities of region's pixels by the given threshold T.

Then the measures of the quality of segmentation are represented by the formula:

$$C_1 = \frac{N_{lc}}{N_{pair}}, C_2 = \frac{N_{adj-lc}}{N_{adj}}, C_3 = \frac{N_{nl}}{N}, C_4 = \sum_A \frac{N_{A-T}}{N_A}$$

where summation is taken over all regions segmented.

The quality of the segmentation is measured by examining every pixel in every region of the image and accumulating costs for the severity of each fault condition that is found. Lower costs imply a more effective segmentation.

The pseudocodes for the cost functions are as follows:

```
Inputs: a colour image in RGB form, a segmentation mask,
        the region adjacency graph, a threshold value

Output: a triple of cost values C1, C2, C3 measuring the
        severity of the three faults

BEGIN

  Zeroise C1, C2, C3;

  FOR each region given in the segmentation mask DO

  BEGIN
     Compute the average R, G and B component values over
     the region, denoted by Rm, Gm and Bm;

     For each pixel in the region assign a cost to the
     discrepancy between R:G:B and Rm:Gm:Bm, and
```

```
   accumulate this sum as C1;
END;

FOR each region, a, given in the segmentation mask DO
BEGIN
       FOR each adjacent region, b, from the region
       adjacency graph DO
       BEGIN
             For each pair of pixels from the two regions
             assign a cost to the discrepancy
             between Ra:Ga:Ba and Rb:Gb:Bb,
             and add this to C2;
       END;
END;

FOR each pixel position where sqrt(R^2+G^2+B^2)>threshold DO

BEGIN
   If pixel does not belong to a region then increment C3;
END;

Express C3 as a fraction of pixels that are above threshold.
END;
```

The quality measures introduced reflect different aspects of a segmentation process, and may be treated separately from each other, according to the needs of application. Nevertheless, the definition of each of them shares some aspects, which are common to all of these measures. As a consequence, they are interrelated logically. It is expected that they are statistically interrelated too, because of common features of pixels used in the determination of the value of measures.

From the application point of view it is attractive to combine $C1, C2$ and $C3$ into a single cost value that gives a total ordering for the quality of a segmentation. This combination is expected to be independent from the specific segmentation algorithm. The section next present principles and results of experimentation in such direction.

3 Relationship of Measures

3.1 Heuristic Approach and Results of Experimentation

Preliminary computational experiments were produced in a way of a "common sense", combining measures in accordance to the basic steps of segmentation algorithms. At the very first attempts these experiments failed to find any combination method that gives an estimate of segmentation quality that agreed with the judgement of a skilled observer. This was true for different sets of images, even for a small set of test images, all presenting the same object, collected from different viewpoints. It is obvious that the more sophisticated approach is needed.

The problem is that $C1, C2$ and $C3$ represent more than one degree of freedom. Further experiments have shown that a search procedure, rather than a combination of $C1, C2$ and $C3$ in an expression, seems to capture the essence of the quality of the segmentation.

This means that the choice of optimum segmentation is based on a partial ordering. The following heuristic search procedure has been successful in finding effective segmentations:

```
Input: A list of segmentation parameters and the
       corresponding C1, C2 and C3 values.

Output: A selection of the optimum segmentation and its
       parameters.

 BEGIN
       Search the input list for local minima in C2,
       and list the local minima into L;

       Search L and locate the minimum of C1 + 10 C3^2,
       present this element of L as the result;
 END.
```

The input parameters belong to a multi-dimensional space. The local minima of $C2$ are found by searching the space along lines that hold all but one parameter constant.

Figure 1 shows a colour image of an object that has been used in tests. The image is 512*512 pixels and has red, green and blue values each of 8 bits.

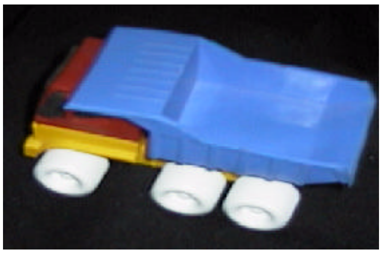


Fig.1. Original colour image (one of the images used for segmentation experiments). Images present an object (a truck) in different poses of rotation around vertical central axis.

The table below shows numerical results from some segmentations. We varied just two of the 6 parameters that control the segmentation. For the purposes of this paper we regard the segmentation method as a black box, but since $p1$ and $p2$ are actually percentages adjusting thresholds for action inside the segmentation algorithm, it was encouraging to observe that $C1, C2$ and $C3$ vary in a controlled way as $p1$ and $p2$ are adjusted.

The evaluation procedure searched the table and found 4 local minima in $C2$ as noted in the commments column. The procedure then identified the response $C1 = 0.644, C2 = -201.385$ and $C3 = 0.099$ as the optimum.

p1	p2	Fault 1 (C1) colour confusion	Fault 2 (C2) oversegmentation	Fault 3 (C3) missing labels	comments
20	12	0.614	-204.577	0.114	x
20	13	0.615	-204.566	0.114	x
20	14	0.616	-204.401	0.113	x
20	15	0.622	-204.502	0.110	min
20	16	0.625	-204.501	0.109	x
20	17	0.627	-204.309	0.107	x
22	12	0.628	-201.310	0.107	x
22	13	0.629	-201.316	0.107	x
22	14	0.634	-204.076	0.105	min
22	15	0.634	-204.063	0.105	x
22	16	0.635	-204.053	0.104	x
22	17	0.637	-204.054	0.103	x
24	12	0.641	-203.727	0.102	x
24	13	0.642	-203.881	0.101	min
24	14	0.643	-201.412	0.100	x
24	15	0.643	-201.230	0.099	x
24	16	0.644	-201.385	0.099	min, opt
24	17	0.645	-201.361	0.098	x

The results in the table are for segmentations reasonably close to the optimum.

It is important that the measure of segmentation should be well behaved at a larger distance from the optimum. Figure 2 shows a segmentation of figure 1 in which the main defect is that colours are confused: the segmentation has defined a region that covers red and yellow parts of the image.

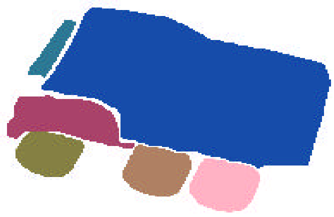


Fig.2. Segmentation with colour confusion
($C1 = 2.088, C2 = -184.840, C3 = 0.053$)

The three fault scores are also shown and it can be seen that the first fault score is larger than those in the table, 2.088 compared with values between 0.614 and 0.645. This the response of the first fault score to a region that contains a mixture of colours. The second fault score is slightly larger (ie. less negative) than those in the table, showing a small increase in over segmentation. The third fault score is slightly smaller than those in the table showing that a few more significant pixels have received labels.

Figure 3 shows a segmentation of the image of figure 1 where there is obviously more oversegmentation. The second fault score has increased and the figure shows that parts of the image with uniform colours have been split up.

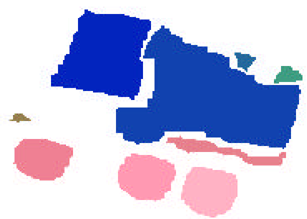


Fig.3. Segmentation with fragmented regions and missing labels (oversegmentation).
($C1 = 0.262, C2 = -39.127, C3 = 0.402$)

The second fault score has responded to the oversegmentation and has become larger (ie. less negative) than that for figure 2.

There are also many significant pixels that have not been allocated a label and the third fault score has increased.

Figure 4 shows the segmentation that is produced for the case marked as optimum in the table. It can be seen that the segmentation is a good compromise. The gaps between regions are a property of this particular segmentation method which is designed to find good candidate regions that will later be filled out with region growing.

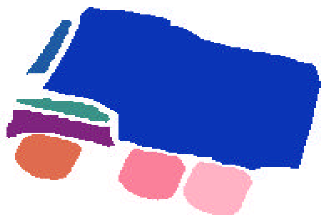


Fig.4. Segmentation using optimal parameters.
$(C1 = 0.644, C2 = -201.385, C3 = 0.099)$

3.2 Statistical Approach and Results

The quality measures do not give immediate information on parameters which are essential to segmentation procedures. These measures it appears to contain epitomized data on types of segmentation methods. In addition to this, many attributes of pixels it has to be noted to be shared in calculation by all of quality measures. All these aspects listed make a statistical approach to the analysis of a relationship of quality measures highly preferable.

Interpreting quality measures as empirical data therefore, we are considering two disjoint sets of dependent and independent variables. We choose appropriate independent variables, as the most universal ones, guiding as many segmentation algorithms as possible. After looking over segmentation algorithms [3, 5, 6, 7] we defined the set of independent statistical parameters used for statistical analysis as follows:

- iterations, smoothing without crossing edges,
- iterations, smoothing reasonably flat regions,
- level of region separation, controlling greater accumulation or greater separation of regions,
- absolute value of variance inside a masking window,

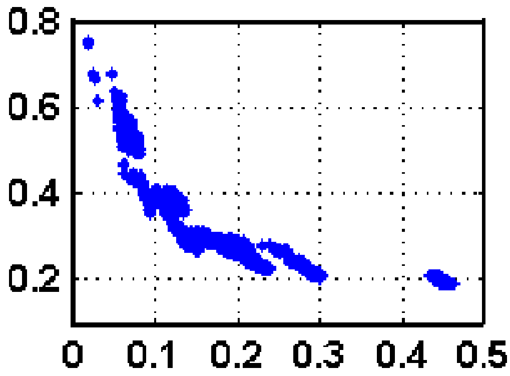
- ratio between the variance inside smaller masking window and the variance inside bigger masking window,

The dependent parameters are, of course, the quality measures C_1, C_2, C_3. To have input data for the calculation of empirical relations between these measures we used a series of images. Figure 5 presents an original image, one of the series of 200 images, segmented according to all of the possible combinations of values of independent variables and then the values of the measures computed. These images are with $128*128$ resolution, and reflect an object (block with a cylinder), rotating around vertical axis.

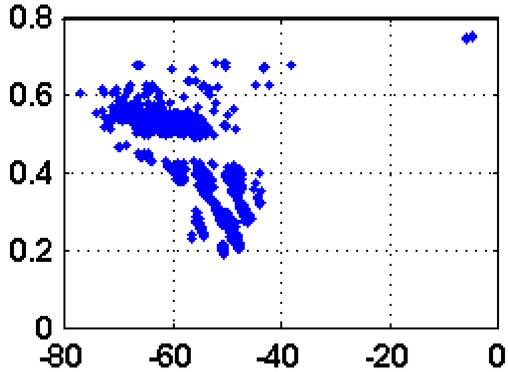
Fig.5. Original image, reflecting block-cylinder in a rotating position.

Input data to calculate statistical relationships, were presented in rectangular matrices (one matrix for each image), each having 3990 lines (the number equal to the one of combinations of values of independent parameters). These matrices were computed according to the principles and algorithms described above.

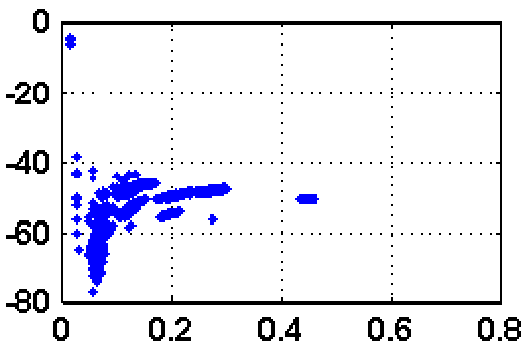
The relationships of C_1, C_2, C_3, at the first step, may be evaluated by visualizing the distribution of pairwise combinations of values of these measures. Fig. 6 gives an example of visual distributions of corresponding values of measures taken pairwise to each other. These distributions are almost the same to all the images analysed.



The distribution of values C1 versus C3



The distribution of values C2 versus C3



The distribution of values C1 versus C2

Fig.6. Plot of pairwise distribution of measures.

Plots of distribution mentioned produced very similar results for most of the test images. The standard statistical characteristics (like residuals, mean, standard deviation, residual standard error, $t-$values, $F-$statistics) and tests (simple hypothesis of statistical dependence or independence) calculated following these plots lead to following conclusions:

- quality measures C_1 - C_2 and C_2 - C_3 are statistical dependent insignificantly; this is because of the range of values of C_2 is more than 10 times greater than the range of values of C_1 and C_3 - so even the major changes in values of measures C_1 or C_3 are not able to make sensitive influence to the values of C_2
- quality measures of C_1 - C_3 are interrelated statistically in a significant way (this follows from $p-$value); so the hypothesis about their statistical independence can't be rejected
- statistical dependence of measures C_1 and C_3 can't be approximated by simple linear regression with appropriate error, at least second degree polynomial regression has to be used (like $C_3 = aC_1^2+bC_1+c$); nevertheless neither of regression equations reflect the situation that there are gaps in ranges of pairwise distribution of these measures, as seen in fig. 6.

4 Conclusions

The paper gives an analysis of the procedure to assess the quality of colour image segmentation. The quality measures suggested are based on an essential use of a triple value of pixels, which is suitable in colour images only. Two approaches used, a heuristic approach, and a statistical one reflect the character of relationship between these measures.

In a heuristic approach, a tabulation of segmentation parameters and the triples can be interpreted by a two step search procedure that finds the optimum parameters. The statistical analysis of a relationship of quality measures C_1, C_2, C_3 confirms the heuristic approach. Sample results have been presented that illustrate these methods on a colour image.

Further research has to be done in order to implement these results into praxis of many colour image segmentation procedures.

References

1. Sadjadi, Firooz A. (editor). "Automatic Object Recognition". *SPIE Proceedings*, **v.1471**, ISBN: 0-8194-0580-9, 486 p., 1991

2. Ahanger, G. "Techniques for Automatic Digital Video Composition". *Ph.D. Dissertation*, Dept. of Electrical and Computer Engineering, Boston University, December 11, 1998

3. Gonzalez, R.C., and Woods, R.E. *Digital Image Processing.* Addison-Wesley Publishing Company, 1992

4. Zugaj, Didier and Lattuati, Vincent. "A new approach of color images segmentation based on fusing region and edge segmentation outputs". *Pattern Recognition*, **31** (2), pp. 105-113, 1998

5. Park, Sang Ho, and Lee, Sang Uk. "Color image segmentation based on 3-D clustering". *Pattern Recognition*, **31** (8), pp. 1061-1076, 1998

6. Ji, S. and Park, H.W. "Image segmentation of color image based on region coherency". *IEEE Trans. PAMI*, **v. 20**, pp. 80-83, 1998

7. Pujas, Ph., and Aldon, M.J. "Robust colour image segmentation". *Proc. International Conference on Advanced Robotics*, pp.145-155, 1995

8. Hoover, Adam, Jean-Baptiste, Jiang, Xiaoyi, Flynn, Patrick J., Bunke, Horst, Goldgof Dmitry B., Bowyer, Kevin, Eggert, David W., Fitzgibbon, Andrew, Fisher, Robert B. "An experimental comparison of range image segmentation algorithms". *IEEE Trans. PAMI*, **v. 18**, No. 7, July 1996