

Adaptive Region Growing Color Segmentation for Text Using Irregular Pyramid

Poh Kok Loo¹ and Chew Lim Tan²

¹ School of Design & the Environment,
Singapore Polytechnic,
Singapore 139651

² School of Computing,
National University of Singapore,
Singapore 117543

Abstract. This paper presents the result of an adaptive region growing segmentation technique for color document images using an irregular pyramid structure. The emphasis is in the segmentation of textual components for subsequence extraction in document analysis. The segmentation is done in the RGB color space. A simple color distance measurement and a category of color thresholds are derived. The proposed method utilizes a hybrid approach where color feature based clustering followed by detailed region based segmentation is performed. Clustering is done by merging image color points surrounding a color seed selected dynamically. The clustered regions are then put through a detailed segmentation process where an irregular pyramid structure is utilized. Dynamic and repeating selection of the most suitable seed region, fitting changing local condition during the segmentation, is implemented. The growing of regions is done through the use of multiple seeds growing concurrently. The algorithm is evaluated according to 2 factors and compared with an existing method. The result is encouraging and demonstrates the ability and efficiency of our algorithm in achieving the segmentation task.

1 Introduction

As compared to the binary and gray scale document images, color document images contain much richer set of information. The use of varying colors allows the subject textual area to be distinguishable from the background and non-subject regions. On the one hand, the color attribute provides an additional avenue for the extraction of textual components. On the other hand, it also introduces new complexity and difficulties. First is the variety of color spaces that can be used where each has its pros and cons. No single space is general enough for all uses. Second is the distance measurement problem. Till date there is yet a standard and precise way of measuring color distance, which is a crucial parameter for all segmentation tasks. Third is the number of unique color points. In a frequently used 24-bit true color image, the number can reach 16 millions. This will intensify the processing complexity. Nevertheless, there are various proposed methods attempting to overcome the problem and complexity in order to benefit from the advantage of using the color attribute in the segmentation process.

2 Existing Methods

There are many proposed techniques for color segmentation. As categorized by [1], color segmentation can be divided into feature-space based, image-domain based and physics based techniques. Feature-based methods focus their attention only on the color features where color similarity is the key and only criterion to segment image content. Spatial relationship among color is ignored. This has resulted in a problem where the segmented regions are usually fragmented. Extra and elaborate post-processing is required to retain the compactness of the regions. Image-domain based methods belong to a category of methods that take spatial factors into consideration. The technique utilizes both color and spatial factors in its homogeneity evaluation. Splitting/merging and region growing are two main techniques used in this category. The common processing steps are the selection of a seed region, the growing or splitting of regions from this seed point, the merging of homogenous regions and a stopping criterion for growing or splitting. By nature this is a sequential process where each pixel and all its neighbors have to be evaluated. The processing order becomes critical at points with the same homogeneity value. The selection of a suitable seed region is another problem where the initial selected seed region may dominate the growing or splitting process. Although many proposed methods attempt to solve this problem by making the best selection, the suitability of a region being a seed point does change during the segmentation process. These problems are reported in [1] and [2]. Physics based techniques are mainly used to process real scene images where the physical models of the reflections properties of materials are utilized.

Despite the large number of proposed color segmentation algorithms, only a handful of them have directly addressed the document image processing domain with the focus in text segmentation and extraction. The key requirement in this domain is not so much in attempting to find the best approximation in terms of the color features. Rather, the emphasis is on how well the segmentation process can achieve the retention of major document components (i.e. text or non-text) and at the same time realizes the compactness within each component. The challenge is to have just a sufficient number of unique colors for the former and minimizing the color uniqueness to attain the later. The method proposed in [3] processes colors in the HSI color space. Color clustering is applied only to the Hue component based on the concept of local maximum where a chain of pointers is constructed pointing to the neighboring Hue values with a larger pixel count. Regions are thus segmented by grouping pixels with Hue values pointing to a same local maximum. As seen in this method, a very intricate post-processing stage in performing connected component analysis is required to merge fragments of textual region. In [4], the same clustering concept is used in the RGB color space. It divides the RGB Cartesian space into multiple numbers of fixed size cubes where each cube will hold the occurrence of pixels having the color defined within the cube. A pointer chain is then constructed by analyzing the 26 potential neighboring cubes to locate the local maximum with the highest number of occurring pixels. It differs from [3] in its performance of clustering in a three dimensional space and the addition of a 4th dimension taking the spatial factor into consideration. The additional dimension is defined by dividing the image plan into horizontal strips where each strip will contain a fix number of image rows. Each bin in the 4th dimensional space will now contain the pixel occurrence of a specified color range located along a certain strip. Although it overcomes some of the problems faced due to the

lack of spatial factor, the effectiveness is restricted to the size of the cube and the widths of the strips. The accuracy of the segmentation result will also depend on where the color space and the image plan are divided. Despite the problem it is an efficient method capitalizing on the efficiency and simplicity of using histogram and at the same time incorporating the spatial factor in the clustering progress. Another conceptually similar system proposed in [5] also attempts to incorporate spatial information into a feature-based type of color clustering.

In view of all the proposed methods, our contributions are in 4 areas. First is in the area of color measurement where a simple measurement method in the RGB color space is derived as described in section 3. Second is in the area of color quantization where an efficient method without the need of a color histogram is proposed in section 4.1. Third is in our region growing method where seeds are selected dynamically and repeatedly to suit the best local condition, which avoids the problem of having a fixed seed dominating the entire growing process. The problem of sequential processing encountered by the other region growing methods is also addressed by having multiple seeds to grow concurrently. The fourth area is in our use of the irregular pyramid structure which differs from the traditional pyramid in that it constructs the pyramid from an intermediate level instead of the original base level in pixel format. It has greatly enhanced the processing speed. The final contribution is described in section 4.2.

3 Color Space and Distance Measurement

In color segmentation the RGB color space is most commonly used where each color is represented by a triplet red, green and blue intensity. HSI is another common color space where a color is characterized by the degree of Hue, Saturation and Intensity variance. Another category of color space is based on the CIE color model. The main aim of this model is to provide a uniform color spacing that facilitates direct measurement of color distance. $L^*a^*b^*$ is one of such color space. While selecting a color space for image segmentation, the key consideration is the ability to have an accurate and efficient way to measure color distance. Color distance is used as a measurement of color similarity where pixels/regions satisfying a certain degree of color homogeneity are grouped to form a cluster. In this aspect the CIE $L^*a^*b^*$ color space seems to be the most promising where the color distance can be computed directly from the Euclidean distance of the Lab coordinates (i.e. ΔE). In spite of this, not many proposed methods make use of this color space. This may be due to the complexity of its conversion process from the RGB color space and also some controversy in its accuracy. In HSI color space, color distance is frequently measured along the individual axis separately. Although the Hue component alone can be used to measure color similarity as in [6], it is not sufficient for detailed segmentation. Both Saturation and Intensity value must also be utilized for finer segmentation results as in [3]. In addition to this requirement to analyze the three axes separately, a further complication exists when the Saturation value is low where all colors look almost the same despite varying Hue value. This is reported both in [1] and [7]. In view of these problems we have decided to use the RGB color space. It is efficient because no conversion is required. Although it also suffers from the non-uniformity problem where the same distance between two color points within the color space may be perceptually quite

different in different parts of the space, within a certain color threshold it is still definable in terms of color consistency.

In order to analyze the color distance measurement in the RGB color space for the definition of color similarity, we have conducted an experiment. The experiment starts with a pivot color. It will then randomly generate 620 non-repeating variation of color points with the same distance from the pivot color computed by using the same distance function (i.e. Euclidean or Manhattan). The color point is a 20x20pixels square which is about the size of a 12point character. All colors are then visually inspected by 10 human subjects to determine its similarity. Each observer will vote for one of the 7 categories as shown in table 1. This process is then repeated for color distance, in the range of 10 to 500 by a step of 10, computed by different distance measurements. The final result is obtained by taking the majority vote. The result of the experiment reveals that the Manhattan distance is a better distance measurement where the generated color points exhibit a more stable visual color similarity. In contrast, the Euclidean distance measurement will produce a wider variation of color perception with the same color distance. This finding shows that color is formed by the additive of the varying red, green and blue intensity and not so much of the physical Euclidean distance between the color points. The various categories of threshold limit obtained through the experiment are shown in table 1. It is categorized into 4 main groups. The first group belongs to those below 71 where the same color is observed with a very low intensity variance. The second group ranges from 71 to 120 where the color appears to be from the same color series (e.g. dark/light brown) with varying degree of intensity. The third group ranges from 121 to 190 where different colors are observed with varying color ranges. Color above 190 becomes quite random and thus is considered as undefined and cannot be interpreted.

Table 1. Categories of color threshold limits.

Threshold	Visual inspection result
10 to 30	Same color.
31 to 70	Same color, low intensity variance.
71 to 90	Same color series.
91 to 120	Same color series, low intensity variance.
121 to 150	Difference color, small color range.
151 to 190	Difference color, wider color range.
Above 190	Very random occurring color.

Based on this experimental result, the following color distance function is derived. The function will compute the total absolute variation of the respective RGB values between 2 color vectors (i.e. C_i and C_j). The further additive factor σ is to discriminate between well distributed color variance among all RGB values and those with un-even variance distribution. The former reflects better color consistency than the later. If the distance is within the threshold T_l then the 2 colors are considered “close”. Otherwise they are treated as 2 unique colors. Although the use of a single threshold to determine the “closeness” between two colors may not be the most precise way of color measurement in the RGB color space, in our context for text segmentation it is more than sufficient. In [8], the authors also make use of a human perception evaluation of color differences to guide the color clustering process.

$$\begin{aligned}
 \text{dist}(\overline{C}_i, \overline{C}_j) &= r' + g' + b' + \sigma \\
 r' &= |C_i^r - C_j^r|, \quad g' = |C_i^g - C_j^g|, \quad b' = |C_i^b - C_j^b| \\
 \sigma &= (|r' - g'| + |r' - b'| + |g' - b'|) / 3 \\
 \text{close}(\overline{C}_i, \overline{C}_j, T_1) &= \begin{cases} \text{true,} & \text{if } \text{dist}(\overline{C}_i, \overline{C}_j) < T_1 \\ \text{false,} & \text{otherwise} \end{cases}
 \end{aligned}$$

4 Proposed Method

Our proposed method is a combination of color feature based and region based color segmentation process. The algorithm utilizes a color feature based technique to perform fast segmentation of regions with very close colors in a pre-processing stage. Based on the result, region-based growing method is then employed to perform detailed segmentation of the remaining regions taking both color and spatial factors into consideration.

4.1 Pre-processing Stage

In a 24-bit true color input image, the number of unique colors will frequently exceed half of the image size. Most of these colors are perceptually close and cannot be differentiated by human beings. In our study of the RGB color space, a color variance of 30 and below will fall into this category. As a result the pre-processing stage will attempt to aggregate colors within a boundary of $T_1=15$ surrounding the pivot color as a cluster. This will ensure the maximum color distance among all colors inside the cluster is within 30. Due to the usually large number of color points, we employ a simple and yet efficient way of clustering. The process will loop through the entire image. As it moves, pivot color points are identified. A color point with variance exceeding 2 times the T_1 limit is inserted as a new pivot color. Those within the color limit are clustered with the closest pivot color. The efficiency and accuracy of this process lie in between fixed partitioning of the color space as in bit-dropping technique and the selection of a color seed by giving preference to a bigger region [8]. Our proposed method is more efficient without the pre-requisite to build a histogram for the selection of a suitable color seed. It is also more accurate than the bit-dropping technique by building the pivot color list dynamically as it loops through the image. As compared to bit-dropping where a fixed partition is used regardless of the actual color distribution, this process will avoid non-existent color points. The final output from this stage is a group of pixel clusters having color similarity within a limit of 30.

4.2 Detailed Segmentation Stage

This stage will perform a detailed analysis of the resulting clustered regions from the pre-processing stage and continue to merge regions having a larger color variance. Region growing is used as a means to perform clustering where an irregular pyramid structure [9] is used. A pyramid is a data structure holding image data points in suc-

cessive levels of reduced resolution. The lowest level is the original input image at full resolution. Each successive higher pyramid level L_{i+1} will hold a smaller representative data set $R_{i+1,k}$ of the lower level L_i . As a result L_{i+1} is a proper subset of L_i where the number of data points on level $i+1$ (i.e. N_{i+1}) is less than level i (i.e. N_i).

$$L_i = \{R_{i,j}\} \text{ where } j = 1 \text{ to } N_i$$

$$L_{i+1} = \{R_{i+1,k}\} \text{ where } \begin{cases} L_{i+1} \subset L_i \\ k = 1 \text{ to } N_{i+1} \\ N_{i+1} < N_i \end{cases}$$

The creation of each pyramid level will follow 4 main stages. The 1st stage is the formation of a pyramid level with the selected surviving regions from the lower level. The 2nd stage is the analysis of neighboring relationship among regions where two regions are neighbors if their children on the lower level are neighbors. The 3rd is the selection of a new survivor which forms the candidate for the next higher pyramid level. The survivor selection process is based on the determination of local maxima. A region survives if it has the highest surviving value among its immediate neighborhoods. The final stage is the selection of children where each survivor will claim a set of neighboring non-survivors as its children. The process will then repeat for the next higher pyramid level. The detailed pyramid formation process is shown in our previous paper [10]. In the current context we can view the pyramid building process as a way to perform image segmentation by growing seed regions. The selection of a survivor is equivalent to the selection of a seed region. The claiming of non-survivors by the survivor as its children is comparable to the growing of the seed. In contrast to a regular pyramid, this effect can be achieved through the use of irregular pyramid structure because of its flexibility in survivor selection and the ability to perform selective claiming of suitable neighboring non-survivors. Intended segmentation result can thus be obtained through a suitable definition of the seed selection criteria and the appropriate designation of growing or claiming rules. In our algorithm the selecting and the growing of seed regions will follow three basic eligibility criteria. A region is eligible to participate in the selection and the growth processes if it satisfies all three eligibility criteria as shown below. The selection and the growth of a seed region will only occur among regions that exhibit spatial adjacency, color closeness and size inferiority. An arbitrary region X is eligible to participate in the process initiated by a pivot region S , if it is adjacent to the region S (i.e. neighbor). A neighboring region will only be considered if the color variance between itself and the pivot region S falls within a color threshold T_2 where the superscript 'c' denotes color. Finally only smaller neighbor in terms of size is evaluated where the superscript 'a' represents size.

$$eligible(S, X) = \begin{cases} true, & \text{if } adjacent(S, X) \wedge \\ & close(S^c, X^c, T_2(S)) \wedge \\ & S^a > X^a \\ false, & \text{Otherwise} \end{cases}$$

Unlike the traditional pyramid structure where the base pyramid level is the original input image in pixel format, our proposed algorithm will begin the pyramid in a

region format (i.e. group of pixels). After the initial pre-processing stage, regions with small color variance are formed. Connected component analysis is then used to identify individual regions from the respective colors. These extracted regions will form the base of the pyramid. This change in using the intermediate result avoids the processing of the first few pyramid levels which are the most time consuming. After the formation of a pyramid base and the determination of neighborhood relationship, the survivor/seed selection process will begin. The selection process is based on a surviving value. The value is computed by allowing each region to vote for its closest neighbor satisfying the 3 eligibility criteria as defined before. The process will then determine a local maximum having the largest number of neighborhood's vote. Region that has the largest number of 'suitable' neighbors is a good seed candidate that will enable maximum "healthy" growing of the seed region into its surrounding. With the selected seed, the growing of region will begin. In a traditional irregular pyramid construction process [10], this is the child selection stage where each survivor will claim a set of suitable surrounding non-survivors as its children. In our context, we will treat this stage as the growing of the selected seed regions into the eligible surrounding neighbors. Instead of having the seed actively growing into its neighbors, our algorithm allows the neighbors to take an active role in searching for the most suitable seed. A region X will become part of a seed S , if S is the nearest to the region X among all eligible seed regions α of X . Region X will evaluate all its surrounding seeds α , which it is eligible to evaluate, and merge with one having the lowest color variance. This will maximize the color closeness among regions that are merged. After this stage the entire process to construct another new pyramid level will repeat. A newly formed region $R_{i+1,k}$ on pyramid level $i+1$ will encapsulate a seed region $R_{i,j}$, and a group of non-surviving regions β_r on the lower level i . All the non-survivors β_r are guaranteed to be the nearest eligible neighboring regions.

$$\begin{aligned}
 \text{nearest}(S, X) &= \left\{ S = \alpha \left| \begin{array}{l} \min(\text{dist}(X^c, \alpha^c)) \forall \alpha \text{ eligible}(\alpha, X) \\ \alpha \in \text{neighbouring seeds / survivors of } X \end{array} \right. \right\} \\
 R_{i+1,k} = \{R_{i,j} \cup \beta_r\} \text{ where } &\left\{ \begin{array}{l} R_{i,j} \in L_i \cap L_{i+1} \\ \beta_r \in L_i \\ r = 1 \text{ to } |R_{i+1,k}| - 1 \\ \text{eligible}(R_{i,j}, \beta_r) \\ \text{nearest}(R_{i,j}, \beta_r) \end{array} \right.
 \end{aligned}$$

The pyramid construction process will stop when the reduction rate is below 0.1. Reduction rate is defined as the decreasing rate in the number of newly created regions on the next higher pyramid level. Through experiment we observe that 0.1 is the break-even point where any continuation in the growing process beyond this stage will not yield any noticeable improvement in the segmentation result. The next stage is the text extraction process where the connected components on each respective color layers are extracted and their textual status is verified. The verification is done through a simple component's size, width and height consistency check to determine its textual identity.

4.3 Threshold Derivation

This section will describe the derivation of the threshold T_2 which is used in the eligibility function to determine the color closeness of two regions. Two regions are “close”, if their color differences fall below this threshold. It is different from T_1 used in the pre-processing stage (i.e. section 4.1) which is an empirical value. The threshold $T_2(R_{i,j})$ is dynamically and locally determined by taking the average of all color variances between the pivot region $R_{i,j}$ and all its neighbors $N_m(R_{i,j})$. This will enable a good adaptation to the varying color contrast conditions across image regions. The design of the formula is based on the assumption that regions belong to the same image object will have lower color contrast than those in different image objects. During the initial growing stage of a region, the majority of the neighboring regions may belong to the same object. The threshold will provide an unrestrictive and yet steady growth of the regions. As the growing reaches its maturity (i.e. reaching the boundary) the majority of the regions with homogenous background will stop at this threshold value (i.e. average variance). For the remaining regions with a complex background, an upper color limit of T_3 is required to avoid excessive over segmentation of region. In order to have a suitable upper bound value that can apply to all situations T_3 is determined globally. As shown below T_3 is defined as the overall average color variance among all pixels P for the entire image plus the standard deviation among all the variances. In order to have a good estimate for T_3 , flat regions with zero variance and regions whose variance are beyond the 190 threshold as stated in table 1 are ignored.

$$T_2(R_{i,j}) = \min \left(\frac{\sum_m \text{dist}(R_{i,j}^c, N_m^c(R_{i,j}))}{|N(R_{i,j})|}, T_3 \right)$$

$$T_3 = \frac{\sum_q^{\text{imgSize}} \sum_m^{8cc} \{ v \mid v = \text{dist}(P_q^c - P_m^c) \text{ where } 190 \geq v > 0 \}}{\text{total number of } v} + \text{stdev}(\forall v)$$

5 Experimental Results

Evaluation of color segmentation is not an easy task. There exists no general methodology to evaluate the correctness of a color segmentation result. The validity of the resulting colors and the segmented regions will vary according to the human perception and the original intent. As a result the measurement is more qualitative rather than quantitative. The existence of the wide variety of color space representations and the utilized segmentation techniques also add on to the evaluation complexity. This has also led to the difficulty in comparative study among algorithms. Since the ultimate aim of our proposed algorithm is to achieve color document image segmentation in preserving both the major graphical and textual components with the later as the focus, we will evaluate the result based on 2 factors. The most important factor is to evaluate how well the algorithm can ensure spatial compactness in the segmented

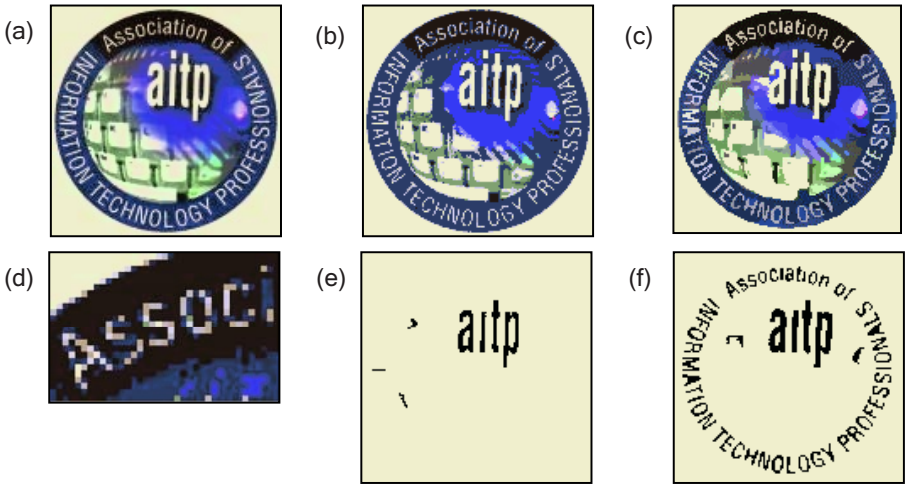


Fig. 1. Test sample : logo (a) original image, (b) pointer segmentation, (c) pyramid segmentation, (d) zoom-in of image b, (e) pointer text extraction, (f) pyramid text extraction.



Fig. 2. Test sample : onions advertisement (a) original image, (b) pointer segmentation, (c) pyramid segmentation (d), (e) pointer segmentation color layer 1 and 2, (f) pyramid text extraction.

textual region. The 2nd factor will assess the effectiveness in the retention of the original image content (i.e. non-text) after segmentation. The evaluation method for the 1st factor is by counting the number of correctly extracted textual component from the respective color layers that is visually recognizable. The 2nd factor is measured by visually counting the number of retained major features in the image. For doing a comparative study we will use the ‘pointer’ method [4] as described above. Figure 1 shows the test sample of a logo extracted on the web. After color segmentation and text extraction, all recognizable text regions are detected with minor over segmentation for the character ‘A’ (i.e. 1c, 1f). Figure 1b shows the segmentation result of the ‘pointer’ method where only the centre bigger texts (i.e. aitp) are extracted (i.e. figure 1e). All smaller texts along the circular path are classified as noise. As shown in figure 1d, the ‘pointer’ method has failed to group pixels within the character as a single color cluster. As a result, pixels belonging to the same character are fragmented into multiple color layers. Without further connected component analysis, each fragment of the character will be too small in size to be classified as text. The lack in the color

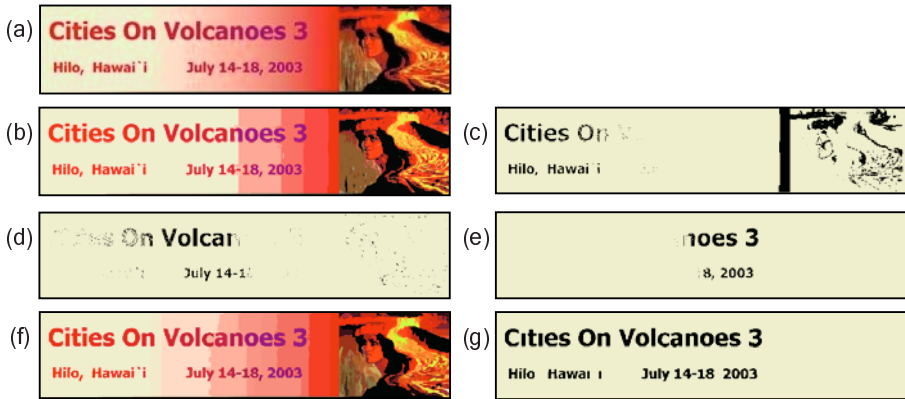


Fig. 3. Test sample : volcanoes advertisement (a) original image, (b) pointer segmentation, (c),(d),(e) pointer segmentation color layers, (f) pyramid segmentation, (g) pyramid text extraction.

value continuity for pointer chasing may be the main cause for the failure. Figure 2 is a test sample of an advertisement on the web for onions. As shown in figure 2f all text regions are extracted correctly by our method. In contrast, the ‘pointer’ method produces results as illustrated in figures 2d and 2e showing components on 2 different color layers holding the main bulk of the textual contents. In order to better demonstrate the degree of fragmentation, these images are used instead of the image after text extraction where most of the components are removed as graphic or noise. As seen in these 2 images, the textual regions are split into 2 main color clusters. Although visually the representative colors for the 2 layers are close, both co-exist as the color peaks in their own local region in the feature space and thus they are clustered as 2 separate color layers. Figure 3 demonstrates the same effect for another test sample where the word ‘On’ and the character ‘n’ in ‘volcanoes’ are segmented into multiple color layers as shown in figures 3c, 3d and 3e. Our method has proven again to be effective with the correct and full extraction as shown in figure 3g. The test sample: weather service logo in figure 4e demonstrates the ability of our algorithm in preserving major regional contents (i.e. woman and tower) whereas the ‘pointer’ method has resulted in both regions almost being absorbed into the background area (i.e. figure 4c). This may be due to the problem of using histogram where the color pixel count for both regions is small. The possibility of such a weak color becoming a local maximum is very low. As such it is absorbed by a neighboring stronger color in terms of the area coverage. The last sample in figure 4b shows an image having colors that are perceptually very close. This has stretched the ‘pointer’ method to its extreme where the result is a single cluster of color (i.e. figure 4d). In contrast, our method (i.e. figure 4f) performs satisfactorily. The results after evaluating 38 images according to the 2 factors are shown in table 2 which confirms that our algorithm achieved the intended task. For textual components, our method achieved a 84% identification rate (i.e. 128/152) as compared to the 70% attained by the ‘pointer’ method. The majority of the results are satisfactory except for images with very high color variance among the textual fragments where the aim of compactness cannot be obtained.

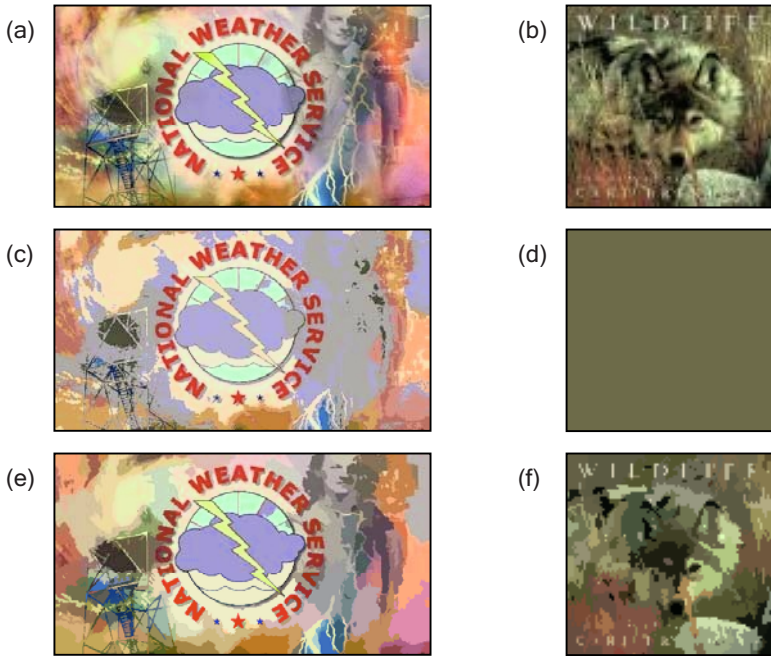


Fig. 4. Two test samples: Weather service logo, Wildlife magazine (a), (b) original image, (c), (d) pointer segmentation, (e), (f) pyramid segmentation.

Table 2. Evaluation result.

Factors	Original	Pyramid	Pointer
Text	152	128 (84%)	107 (70%)
Non-text	93	86 (92%)	72 (77%)

6 Conclusions

We have proposed a novel color segmentation technique to be used for text extraction. The segmentation is done in the RGB color space. A simple color distance measurement (i.e. Manhattan) and category of color thresholds are derived. The proposed algorithm is divided into two stages. The initial pre-processing stage utilizes a dynamic seed selection and clustering process with very close color range. The clustered regions are then used as an input to the detailed segmentation process (i.e. 2nd stage) where both color and spatial factors are considered. The segmentation is based on a region growing technique. Irregular pyramid is used for the segmentation. It differs from other region-growing methods in its selection of seeds and the way regions are grown. The algorithm is evaluated according to 2 factors and compared with the ‘pointer’ method [4]. The results have demonstrated the ability of our algorithm in achieving segmentation for text extraction.

References

1. L. Lucchese and S.K. Mitra, "Color Image Segmentation: A State-of-the-Art Survey" (invited paper), *Image Processing, Vision, and Pattern Recognition, Proc. of the Indian National Science Academy*, Vol. 67, A, No. 2, March 2001, pp. 207-221.
2. H.D.Cheng, X.H.Jiang, Y.Sun and Jingli Wang, "Color Image segmentation: advances and prospects", Vol. 34, *Pattern Recognition*, 2001, pp. 2259-2281.
3. H. Wang, "Automatic Character Location and Segmentation in Color Scene Images", *Proc. of the 11th International Conf. on Image Analysis and Processing*, 2001.
4. T. Perroud, K. Sobottka and H. Bunke, "Text Extraction from Color Documents – Clustering Approaches in Three and Four Dimensions", *6th International Conference on Document Analysis and Recognition*, Sep 2001
5. D. Lopresti and J.Zhou, "Locating and Recognizing Text in WWW Images", *Information Retrieval* 2, 2000, pp.177-206.
6. S. Lefevre, L. Mercier, V. Tiberghien and N. Vincent, "Multiresolution Color Image Segmentation Applied to Background Extraction in Outdoor Images", *IS&T European Conf. on Color in Graphics, Image and Vis.*, 2002, pp.363-367.
7. J. Liu and Y-H. Yang, "Multiresolution Color Image Segmentation", *IEEE Transactions on PAMI*, Vol.16, 1994.
8. Y. Gong, G. Prietti and C. Faloutsos, "Image Indexing and Retrieval Based on Human Perceptual Color Clustering", *Computer Vision and Pattern Recognition*, 1998.
9. P. Meer, "Stochastic image pyramids", *Comp. Vision, Graphics and Image Proc*, Vol. 45, No. 3, 1989, pp. 269-294.
10. P.K.Loo and C.L.Tan "Using Irregular Pyramid for Text segmentation and Binarization of Gray Scale images", *Proceedings of the 7th International Conference on Document Analysis and Recognition*, Vol. 1, Aug 2003, pp. 594-598.