Iowa State University

From the SelectedWorks of Sarah A. Rajala

August 15, 1989

Image Segmentation Using Human Visual System Properties With Applications In Image Compression

Heidi A. Peterson, Purdue University
Sarah A. Rajala, North Carolina State University
Edward J. Delp, Purdue University

Available at: https://works.bepress.com/sarah_rajala/23/
IMAGE SEGMENTATION USING HUMAN VISUAL SYSTEM PROPERTIES WITH APPLICATIONS IN IMAGE COMPRESSION

Heidi A. Peterson†, Sarah A. Rajala*, and Edward J. Delp†

†Computer Vision and Image Processing Laboratory
School of Electrical Engineering, Purdue University
West Lafayette, Indiana 47907

*Department of Electrical and Computer Engineering, North Carolina State University
Raleigh, North Carolina 27695

ABSTRACT

Many image compression techniques involve segmentation of a gray level image. With such techniques, information is extracted that describes the regions in the segmented image, and this information is then used to form a coded version of the image. In this paper we present a region-growing-based segmentation technique that incorporates human visual system properties, and describe the use of this technique in image compression.

We also discuss the effect of requantizing a segmented image. Requantization of a segmented image is useful because it can lead to a reduction in the number of bits required to code the description of the regions in the segmented image. This results in a lower data rate. We show that the number of gray levels in a segmented image can be reduced by a factor of at least twelve, without noticeable degradation in the quality of the segmented image. This result is attributable to human visual system properties having to do with contrast sensitivity, and to the fact that requantization of a segmented image does not usually reduce significantly the number of distinct segments in the image. In addition, in this paper we explore the relationship between the number of segments in an image, and the extent of requantization possible before noticeable degradation occurs in the image.

Finally, we discuss the impact of the above results on image compression algorithms, and present some experimental results.

1. IMAGE SEGMENTATION

The segmentation technique we propose uses a variation of centroid-linkage region growing [1], and is based on the segmentation algorithm presented in [2, 3]. Before segmenting the image, as a pre-processing step the image is clamped. This clamping is justified by Weber's Law [4], which states that the contrast sensitivity of the eye decreases as the intensity of the visual stimulus moves away from the middle range of intensity values. Clamping allows a reduction in the dynamic range of the image, with only minor degradation in image quality.

After clamping, image segmentation is accomplished in two stages. In the first stage, a variation of centroid-linkage region growing is used to form an initial segmentation of the image. With centroid-linkage region growing, the image pixels are scanned in a raster fashion. At each pixel, regions (segments) neighboring the current pixel are compared to each other. If any neighboring regions have average intensities within a threshold of each other, these regions are merged. Then the current pixel is compared to each neighboring region. If the current pixel's gray level is within a threshold of the average intensity of a neighboring region, then the current pixel is merged with that region. If the current pixel matches more than one neighboring region, it is merged with the region it matches the best. If the current pixel does not match any of its neighboring regions, then a new region is started, with the current pixel as its first member.

An important feature of the segmentation algorithm described above is the threshold used to determine whether regions and pixels should be merged. We propose using a threshold based on human visual
system properties, specifically Weber's Law [3]. First consider

\[
\text{threshold}_1 = (0.1234 \times |128-A|) + 4, \tag{1}
\]

where A is the average gray level of the eight pixels neighboring the current pixel. This function, which is an approximation to Weber's Law, can be used as the merging threshold in the segmentation algorithm. The function is upward convex, therefore \text{threshold}_1 is largest in the highest and lowest intensity areas of the image. This means that as the average gray level of the pixels near the current pixel gets further from the middle of the gray level range, it becomes "easier" to merge pixels and regions.

A refinement of \text{threshold}_1 can be made based on the fact that Weber's Law does not hold at the very highest and lowest intensities. A second threshold can be defined as follows:

\[
\text{threshold}_2 = \begin{cases} 
\text{threshold}_1, & \text{threshold}_1 < T \\
\text{thmax}, & \text{threshold}_1 > T.
\end{cases} \tag{2}
\]

With this threshold, Weber's Law is no longer used in the highest and lowest intensity areas of the image, but rather \text{thmax} is used for the merging threshold in these areas. This second threshold is the merging threshold we use in our segmentation algorithm. After separating the image into regions using the above described algorithm, each region in the image is filled in with the gray level closest to the average intensity of that region.

The second stage of our image segmentation algorithm involves removal of image segments that are so small or so weakly contrasted with neighboring segments that they are insignificant to the human viewer. A filter based on the modulation transfer function of the human visual system [3, 5] is applied to the image to remove these segments. Only regions with fewer than 16 pixels are examined as candidates for removal. The following window is applied at each pixel in the region under test:

\[
\begin{array}{ccc}
1/16 & -1/8 & 1/16 \\
-1/8 & 1/4 & -1/8 \\
1/16 & -1/8 & 1/16.
\end{array}
\]

Let \(E_i\) be the result of applying this window at pixel \(i\). Then the total energy of a segment is defined to be:

\[
\text{energy} = \frac{1}{N} \sum_i E_i^2, \tag{3}
\]

where \(N\) is the number of pixels in the region under test, and the summation is over all \(i\) such that pixel \(i\) is in the region under test. If the energy of the region under test is less than a predefined threshold, then that region is judged to be insignificant, and is merged with the neighboring region which has average intensity closest to the average intensity of the region under test. Since this filtering operation eliminates only those regions in the segmented image which are unimportant to the human viewer, the filtering operation should not degrade the quality of the segmented image. The final result of region growing is a gray level image composed of a number of regions, each with uniform gray level.

**2. SEGMENTED IMAGE REQUANTIZATION**

In order to requantize the segmented image, requantization thresholds and output levels must be specified. It would seem that this problem can be viewed as a problem of classical optimum quantizer design, for example a Max quantizer [6]. However, there are problems with applying the methods of Max to the requantization of our segmented image. The methods in [6] rely on a continuous probability density defined on the quantizer input. For the case at hand, the image pixels that are to be requantized take on discrete values and, therefore, do not have a continuous probability density. These methods also require an adequate distortion measure (see Section 4). Therefore, a different approach to the design of the requantizer for the segmented image is necessary.

The algorithm we propose for the selection of the requantizer parameters is also based on Weber's Law. Specifically, the spacing of the requantization thresholds will vary according to Weber's Law. The requantization thresholds will be densely spaced in the middle range of input gray levels, and as the
input gray level moves toward the extremes, the requantization thresholds will be further apart. This will result in relatively fine requantization for mid-range input gray level values, and more coarse requantization toward the extreme ends of the gray level range. The algorithm for the selection of the requantization thresholds is described below.

Suppose that the image to be requantized has gray levels in a range of width \( M \). In other words, the gray levels in the segmented image range from some gray level \( s \), to \( s + M \). Also, suppose that we desire the image to be requantized to \( N \) different gray levels. Let the unit requantization bin length, \( Q \), be the integer closest to \( M/N \) (\( Q \) has units of gray levels). Now, recall the approximation for Weber's Law of Equation (1). Notice that the value of this approximation ranges from 4.0 to 19.8, which is a ratio of approximately 4.75. To specify the equations for the lengths of the requantization bins, two numbers are needed whose ratio is close to 4.75 and whose sum is 2.0. The pair of numbers meeting both of those requirements is (0.35, 1.65). By weighting \( Q \) by a function of these two numbers, and \( M \) and \( N \), the requantization bins' lengths (measured in number of gray levels) can be varied according to the Weber's Law approximation of Equation (1).

There are two sets of equations which, together, accomplish the requantization threshold design described above, one set for \( N \) odd and one set for \( N \) even. To specify these equations we begin by numbering the requantization bins, starting with 1 for the bin corresponding to the lowest input gray level values, and up to \( N \) for the bin corresponding to the highest input gray level values. It should be mentioned that since the requantization bins must be integer in length, the values given by the equations below for requantization bin length are always rounded to the nearest integer.

We will first consider the case of \( N \) even. For this case, the middle two requantization bins are specified to be 0.35\( Q \) gray levels in length. In other words, the length of the bins numbered \( N/2 \) and \((N+2)/2 \) is 0.35\( Q \) gray levels each. For bins numbered between 1 and \((N-2)/2 \), the length of a requantization bin is given by:

\[
L_X = \frac{2.6Q}{2-N} X + (1.65 + \frac{2.6}{N-2})Q, \quad (4)
\]

where \( X \) is the bin number, and \( L_X \) is the length of bin \( X \), measured in number of gray levels. For bins numbered between \((N+4)/2 \) and \( N \), the length of a requantization bin is given by:

\[
L_X = \frac{2.6Q}{N-2} X + (0.35 + \frac{1.3(N+2)}{2-N})Q. \quad (5)
\]

Next we consider the case of \( N \) odd. For this case, the middle three requantization bins are specified to be 0.35\( Q \) gray levels in length. In other words, the length of the bins numbered \((N-1)/2 \), \((N+1)/2 \), and \((N+3)/2 \) is 0.35\( Q \) gray levels each. For bins numbered between 1 and \((N-3)/2 \), the length of a requantization bin is given by:

\[
L_X = \frac{2.6Q}{3-N} X + (1.65 + \frac{2.6}{N-3})Q, \quad (6)
\]

again, where \( X \) is the bin number, and \( L_X \) is the length of bin \( X \), measured in number of gray levels. For bins numbered between \((N+5)/2 \) and \( N \), the length of a requantization bin is given by:

\[
L_X = \frac{2.6Q}{N-3} X + (0.35 + \frac{1.3(N+3)}{3-N})Q. \quad (7)
\]

With these equations defined, requantization thresholds are obtained by centering the middle requantization bins in the middle of the gray level range of the input image. For example, for \( N \) and \( M \) even, set the requantization threshold for the lower edge of requantization bin number \( N/2 \) to be gray level \( s + (M/2) - L_{N/2} \) and the requantization threshold for the upper edge of that requantization bin to be gray level \( s + (M/2) - 1 \). Similarly, set the requantization threshold for the lower edge of requantization bin number \((N+2)/2 \) to be gray level \( s + (M/2) \) and the requantization threshold for the upper edge of that requantization bin to be gray level \( s + (M/2) + L_{(N+2)/2} - 1 \). By working outward and adding the quantization bin lengths given by Equations (4), (5), (6), and (7) to the requantization thresholds that have already been determined, the remainder of the requantization thresholds can be specified. All that remains in the design of the requantizer is to specify an output level for each of the
The image segmentation and requantization techniques discussed in Sections 1 and 2 can be used for the purpose of image compression. Segmentation-based image compression involves extracting information describing the regions of the segmented image, and using that information to form a coded version of the image \([2,3,7]\). Since information must be encoded for each segment in the image, it is desirable for the image to be composed of as few segments as possible, and the descriptions of the segments to be coded with as few bits as possible. This would result in the best possible data rate. The description of the regions in the segmented image must include information about the interior of each region. Often the interior of a region is described simply by the average gray level of the pixels which are contained in that region. By requantizing the segmented image as described in the previous section, the number of different average gray levels used to describe the region interiors is reduced. Therefore fewer bits are required to encode the segments' average gray levels. So by requantization of the segmented image, the image can be coded with a lower data rate.

The data rate can be improved even further by a different application of the requantization mentioned in Section 2. With the image segments assigned one of only a few possible average gray level values, it is feasible for all segments with the same requantized average gray level to be grouped together for transmission. In this way, it is necessary to transmit an average gray level only at the beginning of each large group of segments, rather than for each individual segment.

Due to space limitations, we will not discuss any specific image compression algorithms. Interested readers should see \([2,3,7]\) for examples of coding algorithms that could benefit from the methods presented in Sections 1 and 2.

### 4. EXPERIMENTAL RESULTS

Experiments using various combinations of the ideas proposed above have been performed to determine for a given number of image segments, which yields the most visually pleasing segmented image. When we discuss "image quality" below, we are referring to how the images appear to us. Though various image quality metrics have been proposed \([8-10]\), the search for a meaningful, easily computable image quality metric has eluded the image coding field, and none of the proposed metrics have been widely accepted. A more controlled study of the quality of our images is planned in the future, when we will undoubtedly propose a new image quality metric to add to the confusion. The two images shown in Figure 1 were used as test images for this study. These images are 256 by 256 pixels, and have 256 gray levels.

The first experiment involves comparing the results of segmenting an image with and without the clamping operation suggested as a pre-processor. Figure 2b shows the result of clamping the image in Figure 2a to gray levels 31 to 158. These clamping thresholds were chosen in order to reduce the number of gray levels in the image by a factor of two, thereby reducing the number of bits required to represent the gray levels in the image by one. The specific values were subjectively chosen to be the best to achieve this. Figures 2c and 2d show the segmented versions of Figures 2a and 2b, respectively. The number of segments and gray levels in each image is summarized in Table 1. The images were segmented using the threshold of Equation (2) with \( \text{thmax} = T = 12 \). It can be seen from these images that...
the clamping operation does noticeably degrade the image, hence the segmented version of the clamped image is also of lower quality than the segmented version of the original image. A possible explanation for the loss in quality after clamping is the fact that Weber's Law, which was the supposed justification of the clamping operation, does not hold at the highest and lowest gray levels. Perhaps the clamping would be more successful if done in the range of gray levels where the human visual system is least contrast sensitive, rather than simply at the lowest and highest gray levels. However, as was discussed in Section 2, we propose requantization of the segmented image. Therefore it is not important to reduce the number of bits required to represent the gray levels in the image at this point, and we feel that clamping is not necessary.

The next issue investigated is the use of the post-segmentation filter described in Section 2. We were interested in the effect the filter has on the quality of the segmented image. Figures 3a and 3b show the segmented versions of the images in Figure 1, and Figures 3c and 3d show the result of post-segmentation filtering the images in Figures 3a and 3b. These images were also segmented using Equation (2) with \( th_{max} = T = 12 \), and with the energy of Equation (3) thresholded at 14. For comparison purposes, the number of gray levels in each image and the number of segments in each image is summarized in Table 1. From these images it can be seen that the post-segmentation filter does not degrade the visual quality of the segmented images.

Another experiment deals with the extent of requantization possible before the segmented image suffers visible degradation. Figures 4b through 4d and 4f through 4h show the segmented images of Figures 4a and 4e requantized to eight, twelve and fourteen gray levels. The requantization was performed using the algorithm described in Section 3, with the output gray level for each bin set to the mean gray level of the pixels in that bin. As an example of the results of the requantizer design algorithm, Table 2 presents the thresholds and output levels for requantization of the image in Figure 4a, for the case of 12 requantization bins (This requantizer is used to produce the image of Figure 4c). Table 2 also shows the lengths of the requantization bins which were calculated from Equations (4) and (5). Again, the number of segments and gray levels in each image is given in Table 1.

From the images in Figure 4, it can be seen that requantization to fourteen gray levels causes very little degradation in the quality of the segmented images. This is due to at least two factors. First, requantization of a segmented images does not reduce the number of segments in the image significantly (Note the number of segments in each image before and after requantization). Because most neighboring segments in an image are in different requantization bins, not many neighboring segments get merged during requantization. By preserving a large number of the image segments, the relative quality of the segmented image is unchanged. Notice that as the number of requantization bins is decreased so that the number of segments in the image begins to decrease significantly, the degradation in the requantized image then begins to be more noticeable. The second factor has to do with the contrast sensitivity of the eye. Though requantization affects the gray level of each segment, the contrast between segments is preserved well enough that the eye does not detect a change in contrast before and after requantization. So, by requantization it is possible to reduce the number of gray levels in a segmented image, without degrading the quality of the image. For example, in the case of the image in Figure 4e, requantization to 14 gray levels reduces the number of gray levels in the segmented image from 239 to 14 (a factor of 17), with little degradation in image quality.

In connection with the post-segmentation filtering, another question to be answered is whether such filtering is necessary in our algorithm. It is possible that requantization of a segmented image would accomplish the task of eliminating insignificant image segments, making post-segmentation filtering unnecessary. Figures 5a and 5b show the result of requantizing the segmented images of Figures 3a and 3b (Recall that Figures 3a and 3b have not been post-segmentation filtered). Refer to Table 1 for the number of segments and gray levels in these images. Comparing Figures 5a and 5b to the post-segmentation filtered images in Figures 5c and 5d, it appears that by including post-segmentation filtering, the number of segments in the image can be reduced without significant degradation in the segmented image.

A final question to be addressed is whether the number of requantization levels necessary to avoid visible degradation in the requantized image is related to the number of segments in the image. We experimented with requantizing segmented versions of the two test images composed of between 200 and...
8000 segments. From these experiments it seems that the number of requantization levels required is more a function of the particular image, than of the number of segments in the image. The image of Figure 1a consistently required more requantization levels than the image of Figure 1b, regardless of the number of segments in the segmented image being requantized. For the image in Figure 1b, approximately eight to ten requantization bins were necessary to maintain the segmented image's quality. For the image in Figure 1a, approximately fourteen requantization levels were required. Additional testing with other test images is planned to explore this question further.

5. ACKNOWLEDGMENT

This research was supported in part by the National Science Foundation under grant RII-8620257.

6. REFERENCES


Figure 1. The test images. Each image is 256 × 256 pixels, and composed of 256 gray levels.
Figure 2. (a) The original test image. (b) The image in Figure 2a, clamped to gray levels between 31 and 158. (c) and (d) The images of Figures 2a and 2b, respectively, after segmentation.

Figure 3. (a) and (b) The images of Figures 1a and 1b, respectively, after segmentation. (c) and (d) The images of Figures 3a and 3b, respectively, after post-segmentation filtering.
Figure 4. (a) The image of Figure 3a, after post-segmentation filtering. (b), (c) and (d) The image of Figure 4a requantized to 14, 12, and 8 gray levels, respectively. (e) The image of Figure 3b, after post-segmentation filtering. (f), (g) and (h) The image of Figure 4e requantized to 14, 12, and 8 gray levels, respectively.
Table 1. Summary of the number of segments and gray levels in all the images.

<table>
<thead>
<tr>
<th>Image</th>
<th>Gray Level Range</th>
<th>Number of Gray Levels</th>
<th>Number of Segments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 a</td>
<td>18-221</td>
<td>203</td>
<td>N/A</td>
</tr>
<tr>
<td>b</td>
<td>0-249</td>
<td>250</td>
<td>N/A</td>
</tr>
<tr>
<td>2 a</td>
<td>18-221</td>
<td>203</td>
<td>N/A</td>
</tr>
<tr>
<td>b</td>
<td>31-158</td>
<td>128</td>
<td>N/A</td>
</tr>
<tr>
<td>c</td>
<td>23-215</td>
<td>150</td>
<td>1727</td>
</tr>
<tr>
<td>d</td>
<td>31-158</td>
<td>122</td>
<td>1639</td>
</tr>
<tr>
<td>3 a</td>
<td>23-215</td>
<td>150</td>
<td>1727</td>
</tr>
<tr>
<td>b</td>
<td>1-245</td>
<td>239</td>
<td>8638</td>
</tr>
<tr>
<td>c</td>
<td>24-205</td>
<td>131</td>
<td>490</td>
</tr>
<tr>
<td>d</td>
<td>1-246</td>
<td>239</td>
<td>6948</td>
</tr>
<tr>
<td>4 a</td>
<td>24-205</td>
<td>131</td>
<td>490</td>
</tr>
<tr>
<td>b</td>
<td>28-194</td>
<td>14</td>
<td>412</td>
</tr>
<tr>
<td>c</td>
<td>28-190</td>
<td>12</td>
<td>382</td>
</tr>
<tr>
<td>d</td>
<td>32-176</td>
<td>8</td>
<td>274</td>
</tr>
<tr>
<td>e</td>
<td>1-246</td>
<td>239</td>
<td>6948</td>
</tr>
<tr>
<td>f</td>
<td>9-225</td>
<td>14</td>
<td>5664</td>
</tr>
<tr>
<td>g</td>
<td>11-224</td>
<td>12</td>
<td>5111</td>
</tr>
<tr>
<td>h</td>
<td>22-215</td>
<td>8</td>
<td>3660</td>
</tr>
<tr>
<td>5 a</td>
<td>27-195</td>
<td>14</td>
<td>1418</td>
</tr>
<tr>
<td>b</td>
<td>9-226</td>
<td>14</td>
<td>7022</td>
</tr>
<tr>
<td>c</td>
<td>28-194</td>
<td>14</td>
<td>412</td>
</tr>
<tr>
<td>d</td>
<td>9-225</td>
<td>14</td>
<td>5664</td>
</tr>
</tbody>
</table>

Figure 5. The images of Figures 3a, b, c, d, respectively, requantized to 14 graylevels.

Table 2. The specifications of the requantizer designed for the image of Figure 4a, using the algorithm outlined in Section 2.

<table>
<thead>
<tr>
<th>Bin Number</th>
<th>Bin Length</th>
<th>Gray Level Range</th>
<th>Output Gray Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>24</td>
<td>24-47</td>
<td>28</td>
</tr>
<tr>
<td>2</td>
<td>21</td>
<td>48-68</td>
<td>57</td>
</tr>
<tr>
<td>3</td>
<td>17</td>
<td>69-85</td>
<td>77</td>
</tr>
<tr>
<td>4</td>
<td>13</td>
<td>86-98</td>
<td>91</td>
</tr>
<tr>
<td>5</td>
<td>9</td>
<td>99-107</td>
<td>105</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>108-113</td>
<td>111</td>
</tr>
<tr>
<td>7</td>
<td>6</td>
<td>114-119</td>
<td>117</td>
</tr>
<tr>
<td>8</td>
<td>9</td>
<td>120-128</td>
<td>125</td>
</tr>
<tr>
<td>9</td>
<td>13</td>
<td>129-141</td>
<td>135</td>
</tr>
<tr>
<td>10</td>
<td>17</td>
<td>142-158</td>
<td>148</td>
</tr>
<tr>
<td>11</td>
<td>21</td>
<td>159-179</td>
<td>168</td>
</tr>
<tr>
<td>12</td>
<td>25</td>
<td>180-205</td>
<td>190</td>
</tr>
</tbody>
</table>