Multiresolution Color Patch Extraction

Pere Obrador
Imaging Systems Laboratory
HP Laboratories Palo Alto
HPL-2005-204
November 16, 2005*

color patches, image segmentation

Certain applications require the extraction of patches of color from an image, their size and location. These applications may be: color harmonization algorithms[1], non-photorealistic rendering[2], etc. These applications use not too big a palette of colors, and in both cases large areas of homogeneous color are favored along with high detail preserved in the smaller areas with a lot of color activity. The main problem this paper will tackle is to identify the underlying color in an image region, which will be referred to as its underlying color patch, and also try to protect as much as possible the high color activity detail areas. No perfect scene object segmentation is intended in this process, since different objects may be quantized to the same color, the result may be a merged color patch.

* Internal Accession Date Only
Published in SPIE Visual Communications and Image Processing, 15-19 January 2006, San Jose, CA, USA
Approved for External Publication

Multiresolution Color Patch Extraction

Pere Obrador

Hewlett-Packard Laboratories
Palo Alto, CA, USA
pere.obrador@hp.com

ABSTRACT

Certain applications require the extraction of patches of color from an image, their size and location. These applications may be: color harmonization algorithms[1], non-photorealistic rendering[2], etc. These applications use not too big a palette of colors, and in both cases large areas of homogeneous color are favored along with high detail preserved in the smaller areas with a lot of color activity. The main problem this paper will tackle is to identify the underlying color in an image region, which will be referred to as its underlying color patch, and also try to protect as much as possible the high color activity detail areas. No perfect scene object segmentation is intended in this process, since different objects may be quantized to the same color, the result may be a merged color patch.

Keywords: Coherent regions, Mathematical morphology, Jitter effect, Color harmonization, non-photorealistic rendering, toon shading

1. INTRODUCTION

Color harmony sets guidelines on how to create effective color combinations. There have been attempts at developing fully automatic color harmonization algorithms[1] (CH); in these it is of critical importance to identify areas with homogenous dominant color in which there is little color activity, and also of high importance is to identify smaller regions with high color activity (i.e., even if the color region is very small within that high color activity region, it may still be of high importance at harmonization time if the chroma of such region is significantly different from the rest of regions' chroma). Since the available palette of colors for harmonization is known before hand, the image (figure 10a.) is usually quantized using that palette, as seen in figure 10b. The quantized image is then manipulated in order to obtain a reduced number of regions that can be used for harmonization purposes (i.e., the extracted color patches may only be patches of colors that appear in the palette). In this application the location of the patch, i.e. whether it is touching one of the borders of the image, whether it is quite a central patch, etc., is of high importance too, hence the need of large homogenous regions.

In computer graphics, photorealistic rendering attempts to make artificial images of simulated 3d environments that look "just like the real world." So non-photorealistic rendering[2] (NPR) is then any technique that produces images of simulated 3d world in a style other than realism. Often these styles are reminiscent of paintings (painterly rendering), or of various other styles of artistic illustration (sketch, pen and ink, etching, lithograph, etc.) Of particular commercial interest are techniques that can render 3d scenes in styles which match the "look" of traditionally animated films. Often called 'toon shading, these techniques allow for seamless combination of these rendered elements with traditional animation. In painterly rendering and 'toon shading the palette that is used is quite limited too, and they also favor large homogenous regions with little color activity, and will usually favor the smaller high color activity areas to be leaved intact.

For NPR, and to a less extent in CH algorithms, it is of paramount importance not to modify the edges in the original image in the rendered output.

No perfect scene object segmentation is intended in this process, since different objects may be quantized to the same color, the result may be a merged color patch. CH really does not care about object boundaries; NPR has and extra step to add edges as brush-strokes or similar, which means that perfect object segmentation is not needed either.

(a) (b)

Figure 1. Examples of color patches (from the sample image in figure 10) and their underlying color patches at 2 different resolutions

Another aspect that should be noted is the multi-resolution nature of such a problem. Underlying color patches at different scales may look very different indeed. i.e., what may look as an underlying color patch at a very small scale, might just look as a non-underlying color at a very large scale (see figure 1a, where the vertical white stripe, reflection on nose in figure 10b, is removed at the larger scale in favor of the underlying color; on the other hand, at a smaller scale it is definitely an underlying color within its region as seen in figure 10c). And the other way around, large enough regions (given the scale/resolution) with high color activity will not have a clear underlying color patch and should be left alone.

Having these requirements in mind, a technique was developed to extract the underlying color patches in an image, quantized with a predetermined quantization table (palette). The current state of the art on similar approaches will be covered in the following section, with a detailed description of the proposed scheme, and finally multiple results will presented.

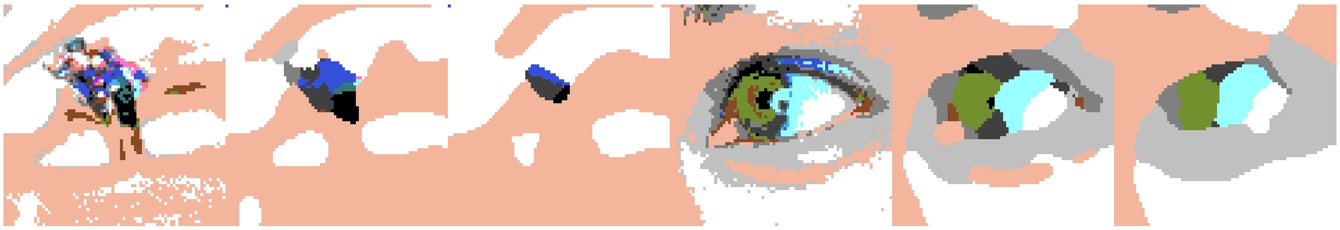
2. STATE OF THE ART

In order to determine the underlying color patch of an image, it's quite clear that the *mode(image,structuring_element)* operator is the first one to come up to mind (see figure 2). The operator picks the highest occurrence color within the structuring element coverage (throughout this paper constant circles of a certain specified diameter are used as structuring elements, i.e., disc21 specifies a circle or diameter 21 pixels). One great advantage of such a scheme is that it will never introduce colors that are not covered by the structuring element (that becomes a problem when using other schemes, as shown below). It is also a very fast operation to perform, but unfortunately, the results do not meet the requirements of the applications described in the introduction.

(a) (b) (c)

Figure 2: (a) GIRL original quantized image, (b) *mode(GIRL,disc11)*, (c) *mode(GIRL,disc21)*

Figure 2 and 3 show the results for a disc11 and disc21, and it is clear that many of the areas with high detail have been wiped away, which reduces the overall number of regions greatly (see figure 9). The *mode* operator also does not maintain the edges on the rendered image. So, it is obvious that such a scheme could not be used in either NPR or CH.



(a) (b) (c) (d) (e) (f)

Figure 3: (a,d) Details of GIRL quantized original image, (b,e) *mode(GIRL,disc11)*, (c,f) *mode(GIRL,disc21)*

In the literature there have been multiple attempts at scene object segmentation, where multiple features are taken into account in order to segment out the objects, one of such features being color. So, one approach[3,4] could be to segment the image based on only the color feature: the image is usually analyzed, and an image-based palette be extracted from it, which would result in a much better scene object segmentation. Using a predetermined palette would yield much poorer results, and it is not clear how the detail regions of high color activity would be treated, since these segmentation algorithms are not designed with this in mind.

Now, if the color patch is looked at as being corrupted by noise, morphological filtering[5] becomes an interesting tool, since it has some good properties that could help in extracting underlying color patches (e.g. noise removal, edge preservation, etc.). Binary openings (closings) are useful for filtering images that have been corrupted by union (subtractive) noise. Analogous approaches have been taken in the gray scale. In particular, the gray-scale opening (closing) can be employed to filter maximum (minimum) noise because the noise to be filtered lies above (below) the signal. Typically noise is mixed, there being noise spikes both above and below the signal. From a statistical perspective, the problem with opening as a filter is that, unless the noisy image lies above the uncorrupted image, as with maximum noise, the filter will suffer bias because the opened noisy image will always lie beneath the noisy image. A dual comment applies to closing bias. As long as these noise spikes are sufficiently separated, they can be suppressed by application of either a composition of opening and closing or of closing and opening. This type of morphological filter may allow for the extraction of a color patch at a certain scale.

If the multi-resolution problem is considered, then a concatenation of such opening/closing with different structuring element may be applied. In the literature these filters are referred to as *alternating sequential filters*[5] which are a composition of alternating openings and closings with increasingly wide structuring elements. This type of filter will remove smaller sized noise spikes, that may be closer together (due to the smaller size of the structuring element), cleaning up those areas for the following filtering stage, that can deal with larger area noise spikes further apart, and so on. The main problem with such an approach is that you need to predetermine whether opening will be used first or second in the stage which generates a certain bias in the output region edges.

Other more sophisticated morphological filtering techniques, like the grayscale area open-close[6] (i.e., one type of opening by reconstruction), filters the flat zones of an image, where flat zones are connected regions of constant gray value. An area open-close keeps all connected components of the input of area larger than a certain limit, merging each of the smaller areas into one of the larger areas. This has a very interesting property for color patch extraction: the flat zones of the output image contain the flat zones of the input image (edges in output image, exist also in input image). This is usually implemented as a connected pyramid, i.e. *alternating sequential filters* of increasing sizes, allowing for multi-resolution flat zone merging. The main problem with this approach is that it will definitely merge the high detail color regions into larger flat zones, not fulfilling one of the main goals stated in the introduction.

Now, the morphological methods described above are applied to grayscale images. They need to be extended to color images in order to fulfill the goals stated in the introduction.

There has been a lot of research in recent years in how to extend morphological filtering to color images, some focusing on hue filtering[7], others[8,9,10] describe how to approach the filtering of a full color signal. In order to extend the concepts of grayscale morphological filtering, there is a need for an ordering of multivariate samples, some examples of such are presented in[11], even though there is no natural means for such an ordering. One such ordering, the marginal ordering in which ranking takes place within one or more of the marginal sets of samples, i.e., scalar ranking is performed within each channel, is the one this paper will focus on due to its simplicity. Note that the fact that each color channel may produce a different output near the region edge borders (edge jitter[8,9]) may generate colors that were not present in that same neighborhood of the input image. This should be avoided, since the generation of non-existent color patches would prove fatal in CH.

In order to overcome the limitations of the methods presented above the proposed method below will rely on a symmetric *alternating sequential filter* pyramid, with multiple structuring elements at each stage, and each of the morphological stages will be modified in order to prevent the edge jitter from happening.

3. PROPOSED METHOD

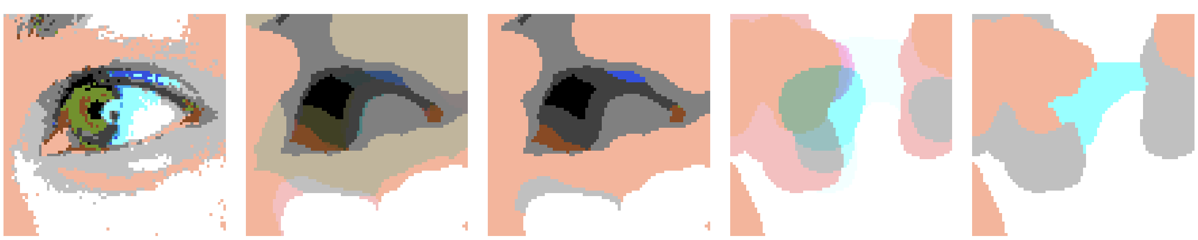
The proposed approach to extracting the color patches is based on an *alternating sequential filter* pyramid scheme introduced above. But some critical changes will be added in order to obtain the desired performance, i.e., extract the largest possible underlying color patches, keep the edges from the original image and keep detail in smaller regions with high color activity.

3.1. Edge Jitter removal

The edge jitter problem in color morphological filtering will have to be solved in order to obtain faithful underlying color patches. In figure 4 the effect of a large structuring element opening/closing on each individual color channel are shown. Multiple colors non-existent in the palette appear (see figures 4b and 4d) . In order to solve this, a statistical approach is proposed.

Opening and closing operations using constant structuring elements should never produce output values that were not present in the original image (in the area covered by the structuring element). The proposed solution is to remap the non-existing color to an existing color in that neighborhood, picking the maximum likelihood color from the quantized original image in that region.

A histogram is generated for all the pixels in the quantized original image, which basically shows how many pixels of each color in the palette are present in that particular structuring element area. Reordering the histogram from maximum occurrence to least occurrence yields a likelihood ordering of the colors in that area. Now, if the color of the pixel being considered is not present in the palette, it will be compared with the colors in the reordered histogram, starting with the maximum likelihood one, until at least two of the color planes match one of the palette colors (i.e., this would mean, since the palette is quite sparse, that with a high likelihood the two matching color planes contain the right filtered result, and the non-matching one needs to be re-mapped to its corresponding palette value) . If no matching 2-plane color is still found, then the process is repeated checking for only one color of the pixel matching the color in the reordered histogram colors (i.e., this would mean that the other two planes were incorrectly filtered, and they should be mapped to this is the maximum likelihood color with a common color plane). In this way, the most likely palette color (given the filtered output) will be chosen. Figure 4 shows the some results of such remapping.

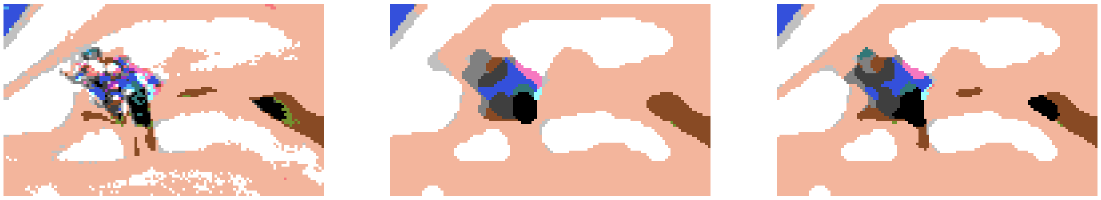


(a) (b) (c) (d) (e)

Figure 4: (a) Detail of GIRL quantized original, (b) opening(disc21) on each color channel, (c) proposed statistical color remapping after opening(disc21), (d) closing(disc21) on each color channel, (e) proposed statistical color remapping after closing(disc21)

3.2. Propose morphological filtering stage

Alternating sequential filters perform the opening and closing in a certain pyramid level with the same structuring element. This will usually reduce the amount of detail in smaller areas with a lot of color activity, as seen in figure 5.



(a) (b) (c)

Figure 5: (a) Details of GIRL quantized original, (b) $C_{disc11}O_{disc11}(GIRL)$, (c) $C_{disc3}O_{disc11}(GIRL)$, which retains greater detail in the high color activity areas

The way this problem is solved is by reducing the size of the second filter in each of the stages:

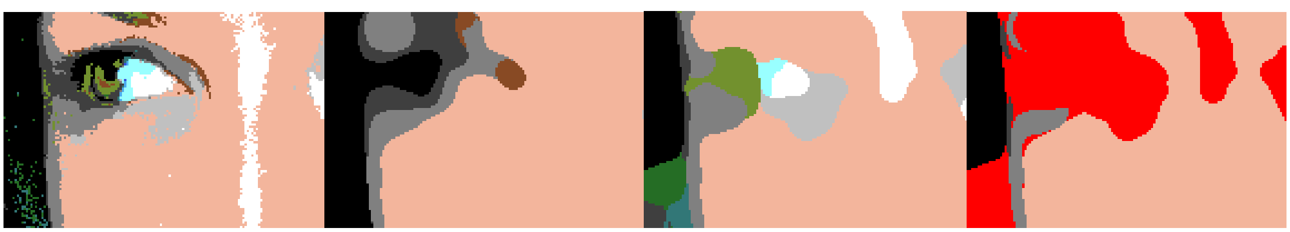
$$C_{g_2}O_{g_1}(f) = (f \circ g_1) \bullet g_2 \text{ with } g_2 < g_1 \qquad \textbf{eq. (1)}$$

Or

$$C_{g_2}O_{g_1}(f) = (f \circ g_1) \bullet g_2 \text{ with } g_2 < g_1 \qquad \textbf{eq. (2)}$$

which allows for finer detail to be retained in the output image.

Alternating sequential filters have a bias depending on whether the filtering stage is like **eq. (1)**, or whether it is like **eq. (2)**, as seen in figure 6. But there are large areas where both the output of these two equations are the same, which will be denoted as areas of agreement. In figure 6d those areas may be inspected, along with the rest of de areas of disagreement (in pure red). These areas of disagreement are usually detail rich high color content regions, which should be preserved close to the original.



(a) (b) (c) (d)

Figure 6: (a) Details of GIRL quantized original, (b) $C_{disc11}O_{disc21}(GIRL)$, (c) $O_{disc11}C_{disc21}(GIRL)$, (d) regions of agreement between (b) and (c), pure red (r=255,g=0,b=0, not part of palette) encodes disagreement areas

Figure 7 shows the actual proposed filtering stage, with a two parallel symmetrical a*lternating sequential filters* whose output are compared; the areas of agreement in both are selected as the filtering stage output. This filtering stage allows for a reduction in the overall number of regions (i.e., larger region color patches), but at the same time the areas with high detail are identified.

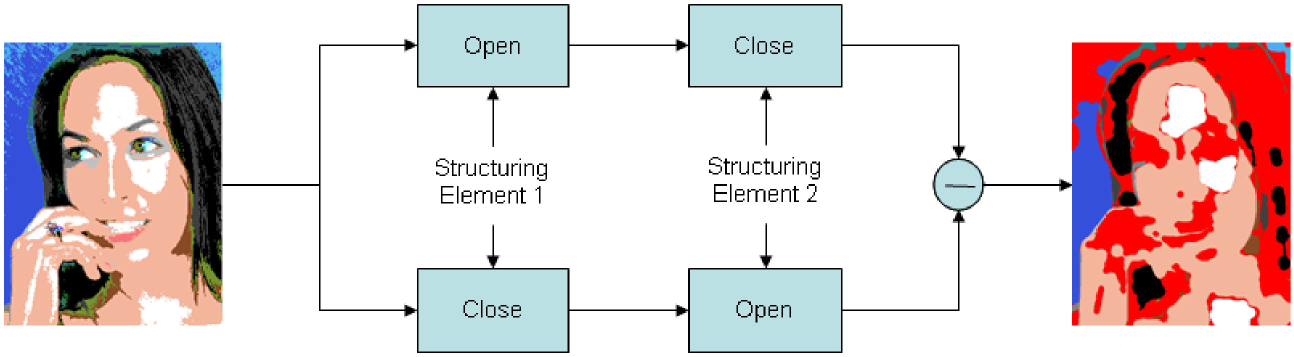


Figure 7: Graphical example of the proposed filtering stage, with structuring element 1 = disc-21, and structuring element 2 = disc-11

3.3. Proposed multi-resolution pyramid scheme

In order to extract the underlying color patches at multiple scales, a pyramid of such filtering stages is built, with increasingly larger structuring elements. The initial filtering stages clean up the "noise" in the order of the size of the structuring elements used, and that are apart that same structuring element size; this allows for the following stage to filter out that output with a larger structuring element, allowing for larger "noise" areas to be removed, and so on. The number of pyramid levels (filtering stages) and the structuring elements in each filtering stage need to be chosen in base of the smallest details that need to be preserved, and in base of the size of the largest underlying color patches to be extracted.

In order to keep the detail in the output, this should be incorporated from the input to the filtering stage, by, for instance, copying directly the input from the areas of disagreement into the output. Unfortunately this simple scheme would yield very poor color patches in those regions of disagreement, and since the next filtering stages use larger structuring elements, there would be very little hope to filter those out.

In the proposed scheme, the input is filtered using a modified $mode_{tempBased}(.)$ operator. The $mode_{tempBased}(.)$ is calculated taking into account only the pixels that lie within the areas of disagreement for that particular filtering stage (these areas of disagreement will be referred to as *template*, i.e., pure red areas in figures 6 and 7) in the input image, in this way no edges may be modified by such filtering (i.e., only the details within the areas of disagreement will be changed). This template based mode filtering is done with a small structuring element, which allows for a certain amount of underlying color patch extraction at that very small scale.

The template is generated by subtracting the outputs of the two parallel symmetrical a*lternating sequential filters*, i.e., any pixel that is not (r=0,g=0,b=0) is part of the template:

$$template = C_{g_2}O_{g_1}(input) - O_{g_2}C_{g_1}(input) \ , \quad \text{with } g_2 < g_1$$

The areas of agreement are then combined with a modified Union operator ($\cup_{tempBased}$) with the output of the $mode_{tempBased}(.)$ operator in the template areas (areas of disagreement):

$$output = \mathrm{mod}\,e_{tempBased}(template, input) \ \bigcup_{tempBased} \ [template] \qquad \textbf{eq. (3)}$$

This is the basic building block for the proposed pyramidal scheme as shown in figure 8. The pyramid can be composed of as many levels as the application may require, as long as the structuring elements keep increasing with the each new level. By starting with a small enough structuring element, the detail regions will be preserved, and by going to a large enough structuring element in stage N, the largest resulting color patches (if existing in the image) will be homogenized to that similar structuring element size. The usage of morphological filters along with the $mode_{tempBased}(.)$ allows the larger edges in the image to be preserved.

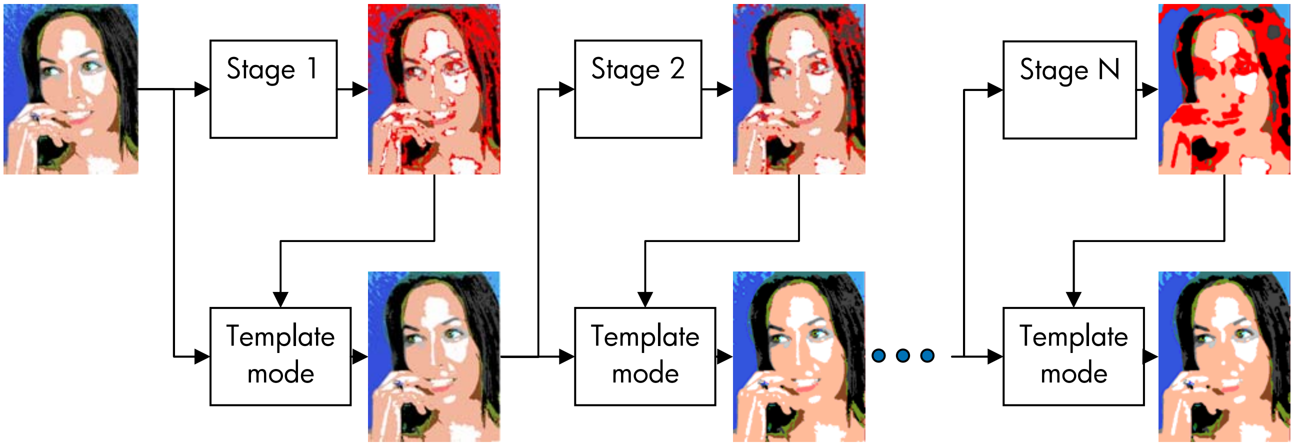


Figure 8. The proposed pyramidal method to extract underlying color patches, where "Template mode" incorporates the operation $output = \mathrm{mod}\,e_{tempBased}(temp, input) \cup_{tempBased} temp$. See blown up examples in the results section below.

4. RESULTS

The presented scheme has been implemented in a 4 level pyramid, with the following structuring elements:
Stage 1: g_1 = disc-5, g_2 = disc-3; Stage 2: g_1 = disc-7, g_2 = disc-5;
Stage 3: g_1 = disc-11, g_2 = disc-7; Stage 4: g_1 = disc-21, g_2 = disc-11
The $mode_{tempBased}(.)$ structuring element is set to disc-3, in lower levels of the pyramid (stages 1&2), and set to disc-5 in higher levels (stages 3&4).

The sizes of the tested images in the above examples, as well as the results in this section are: GIRL 419 x542 pixels, portrait image, FIRECRACKER 387x518 pixels, HAT 640x433 both highly texturized. Figures 10, 12 and 13 show the actual tests with original, quantized to 25 colors, and 4 different outputs from the proposed scheme, one at each pyramid level.

Figure 9 shows how the overall number of color patches is consistently reduced by increasing the number of pyramid levels, even though after 3-4 stages it starts flattening out. What is not clear from this graph is the quality of these color patches, i.e., whether they are homogenous patches with no smaller patches contained in them.

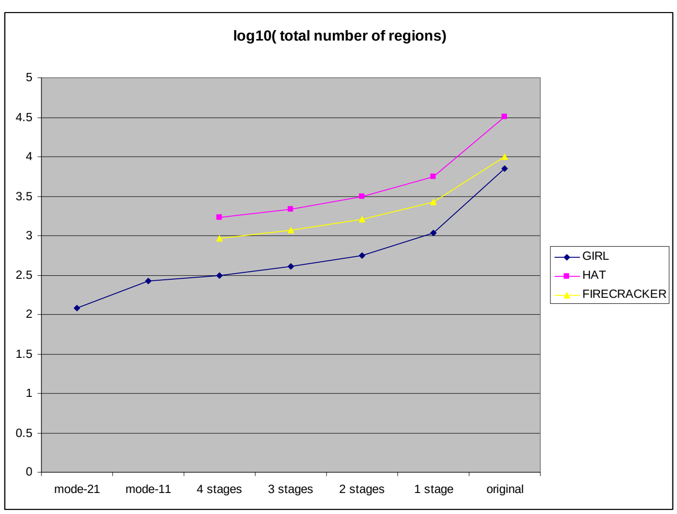
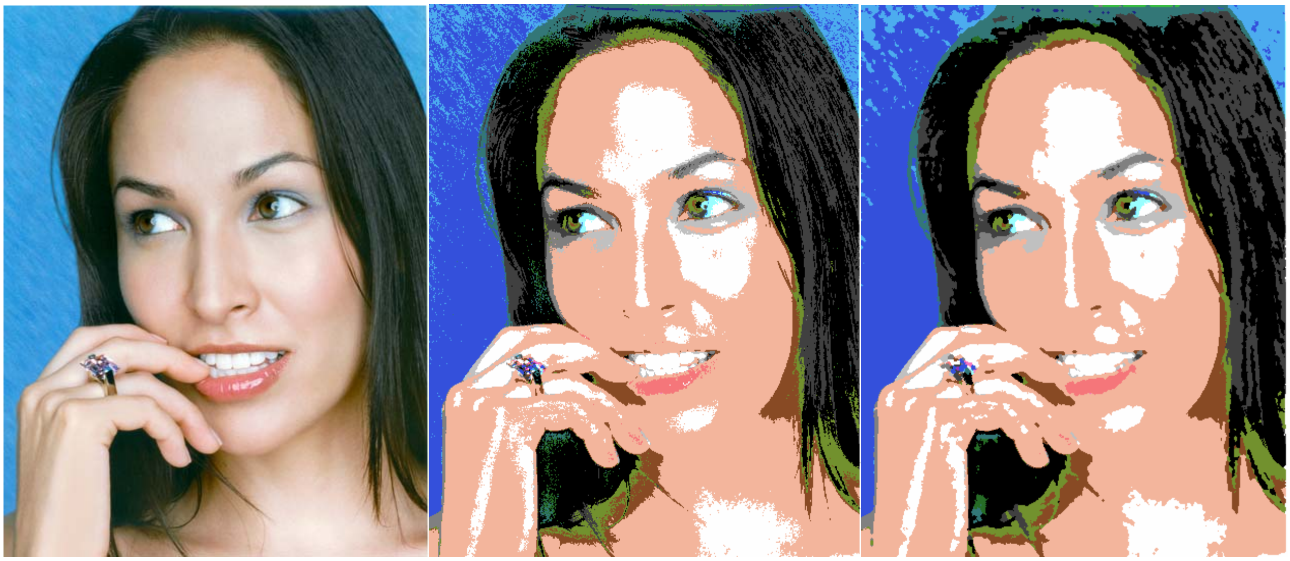
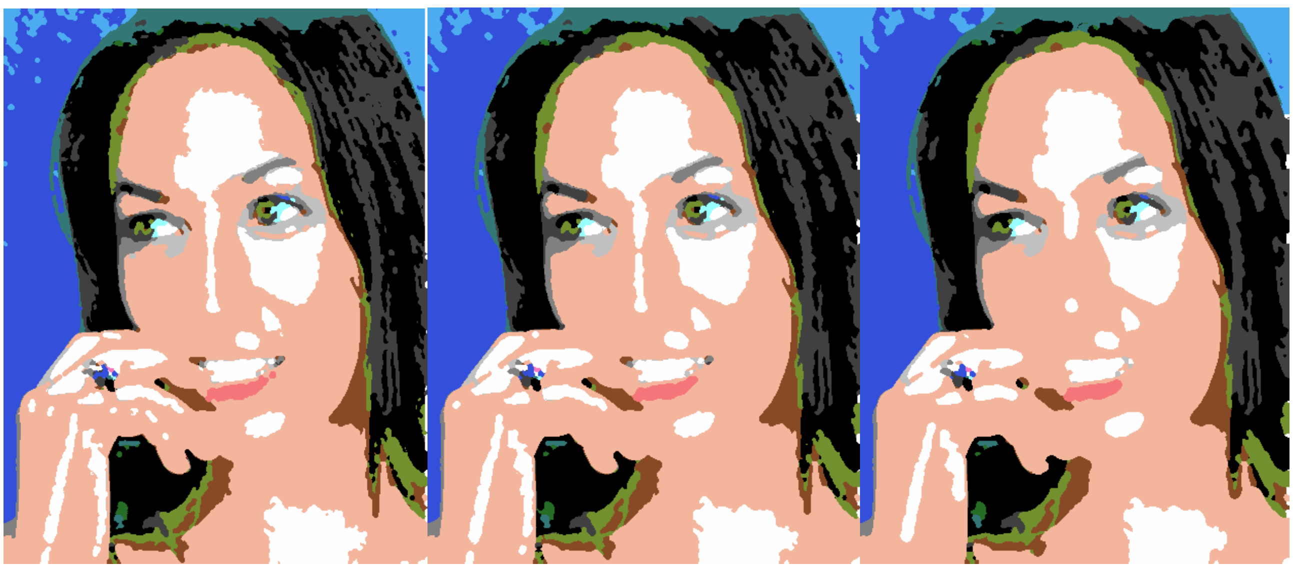


Figure 9. Overall number of color patches (regions) extracted at different levels of the pyramid, in log10 scale. Also shown is the number of regions extracted by simply applying a *mode(.)* operator with disc-21 and disc-11.



(a) (b) (c)



(d) (e) (f)

Figure 10. (a) Original GIRL image, (b) Original GIRL image quantized to 25 palette colors, (c) output of stage 1 of the proposed scheme in figure 8, (d) output of stage 2 in figure 8, (e) output of stage 3 in figure 8, (f) output of stage 4 in figure 8

A way to measure how good the extracted color patches are is to measure how many small patches are left behind. The better larger color patches means smaller color patches that might have been embedded within them. Remember that the scheme will not remove all small patches, as they are needed in the detailed high color activity areas.

Figure 11 shows, for each image, an integral function of the size of the color patches in number of pixels, after having reordered the patches from the larger to smaller. They show the percentage of area covered by the larger 20 color patches in the image. For certain images, as GIRL and FIRECRACKER, in the first 20 or so largest regions, it is not clear that the 4-stage pyramid yields the best result in number of patches, since smaller structuring elements sometimes will have a better chance of merging very large regions. After the first 20 largest regions, consistently the results are better for the 4-stage pyramid, followed by the 3-stage pyramid and so on.

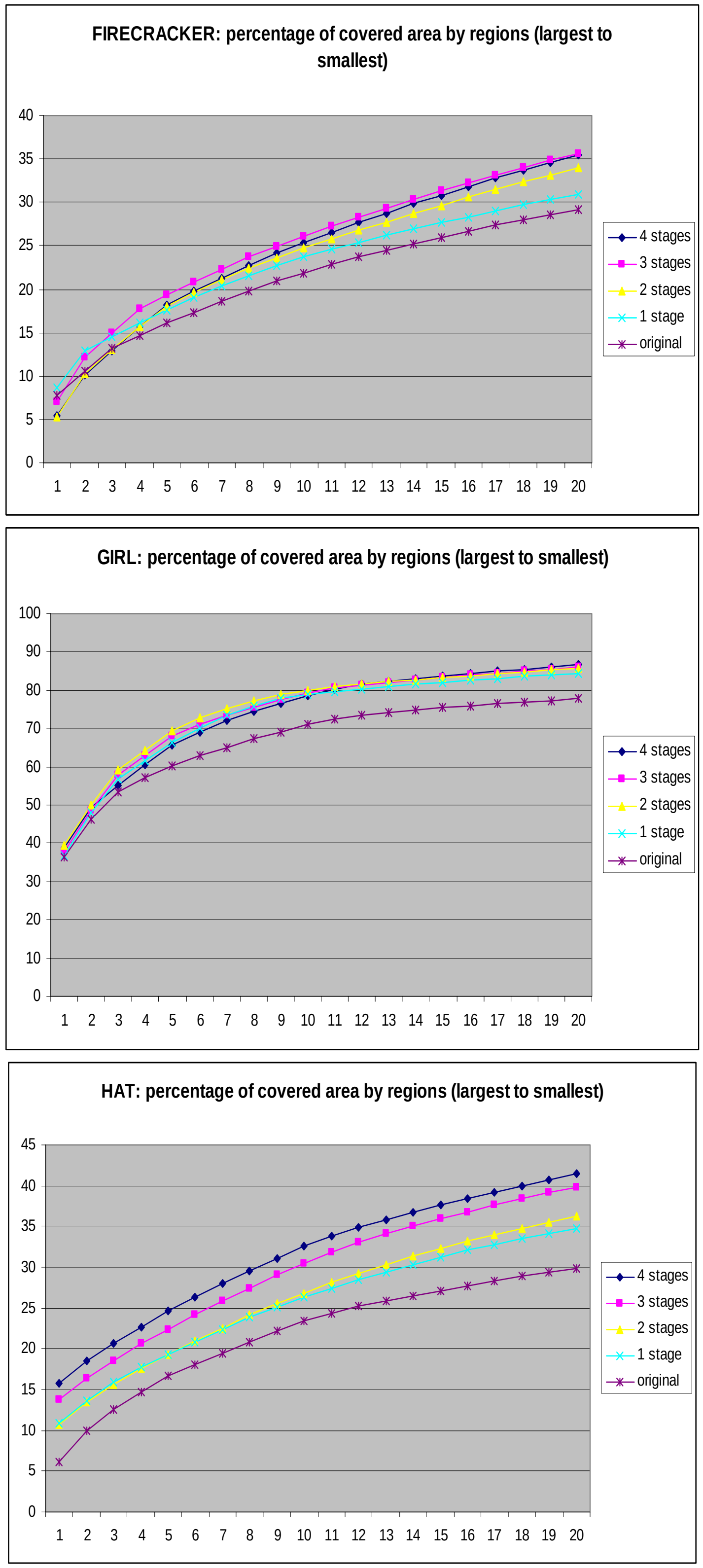


Figure 11. Integral plots in area percentage for each of the tested images. Values on x-axis are the reordered color patches based on their area (largest to smallest). Highly texturized images grow slower since the largest regions are usually much smaller.

Figure 12. (a) Original FIRECRACKER image, (b) Original image quantized to 25 palette colors, (c) output of stage 1 of the proposed scheme in figure 8, (d) output of stage 2 in figure 8, (e) output of stage 3 in figure 8, (f) output of stage 4 in figure 8

The graphs in figure 11 show how large the largest regions of the image are, and how fast they add up in area.

The way to measure how good the extracted patches are in the smaller size regions is to measure how fast the graphs in figure 11 reach a certain high percentage. Figure 14a shows the number of regions needed for each image and each stage output, it takes to reach the 90% mark or covered area. The plot is shown in log10 scale in order to appreciate better the results. The figure shows a consistent decrease in number of regions needed to accomplish that 90% mark the more levels in the pyramid are selected.

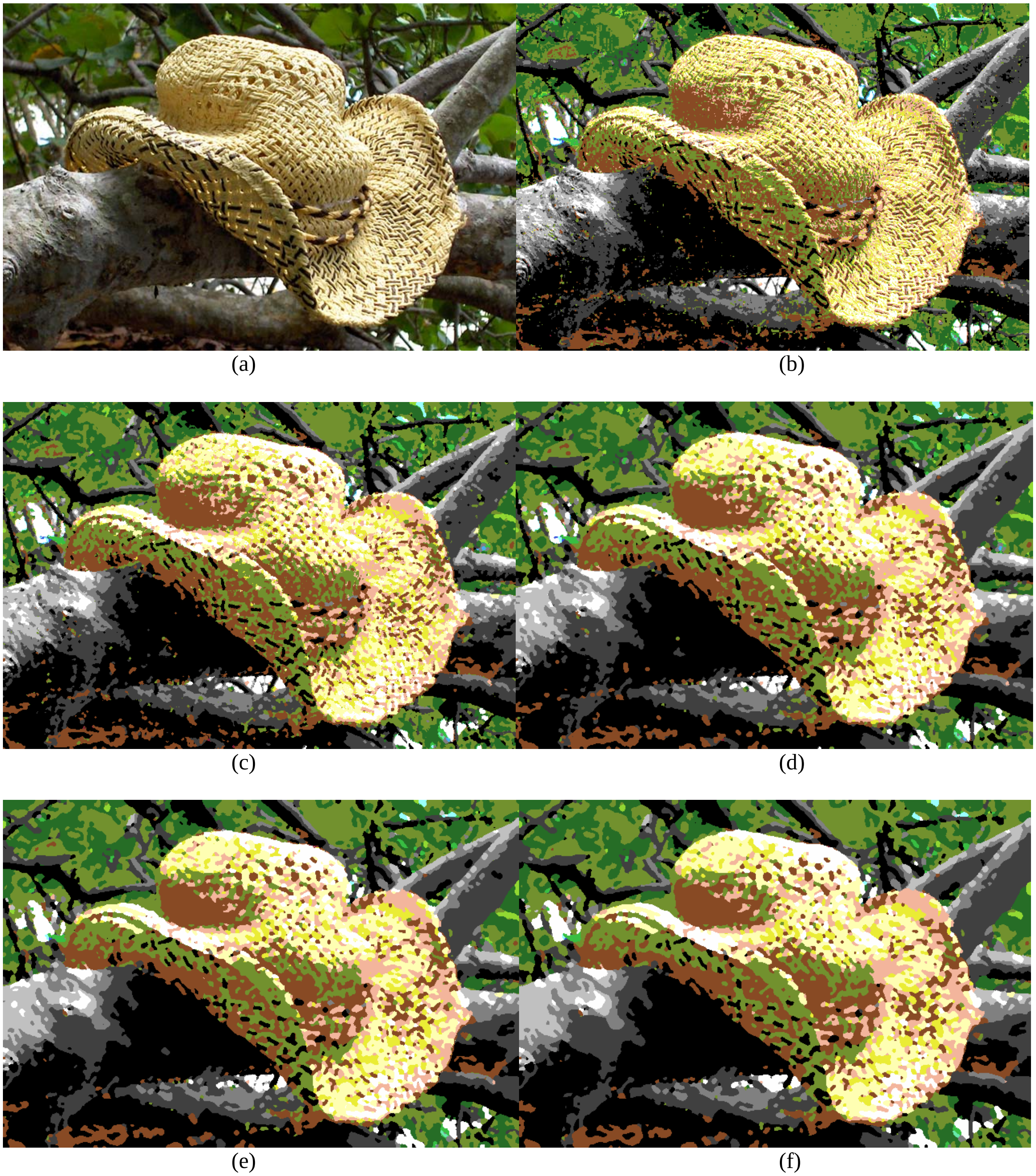
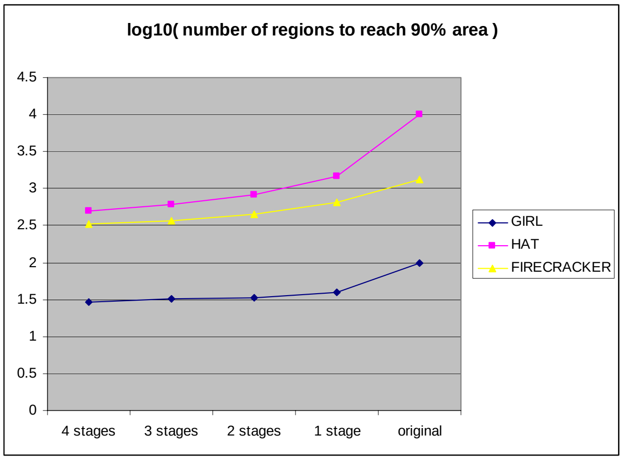
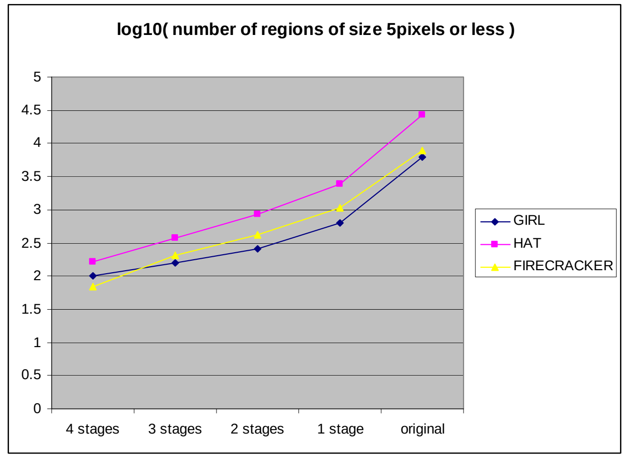


Figure 13. (a) Original HAT image, (b) Original image quantized to 25 palette colors, (c) output of stage 1 of the proposed scheme in figure 8, (d) output of stage 2 in figure 8, (e) output of stage 3 in figure 8, (f) output of stage 4 in figure 8

Another way to measure this is to count how many regions smaller than a certain area are left in the image after the proposed processing at each stage. Figure 14b shows exactly that for regions of 5 pixels and less. This result is even more spectacular than the previous one, since it shows a very steep decline with the number of pyramid levels, even in log10 scale.



(a)



(b)

Figure 14. (a) Overall number of color patches, ordered from larger to smaller, it takes to reach 90% of the image area, (b) overall number of color patches of size 5pixels or less after processing

5. CONCLUSIONS AND FUTURE WORK

This paper presents two main contributions: a parallel symmetrical a*lternating sequential filters* scheme which allows for color patch extraction, while maintaining edges and detail regions, and also, a maximum likelihood scheme to fix edge jitter in color morphological filters applied to sparsely quantized images. The obtained results are good enough to be used in the two main applications presented above: CH and NPR.

Future work will look at extending some of these concepts to full color morphological filtering, and also look at possibilities of using such scheme to aid in image compression.

The color version of this paper: please refer to the HP-Labs web-site (http://www.hpl.hp.com/techreports/).

REFERENCES

1. P. Obrador, “Automatic color scheme picker for document templates based on image analysis and dual problem”, submitted to EI126, proceedings of Digital Publishing Conference, 2006.
2. Gooch & Gooch, “Non-Photorealistic Rendering”, A.K. Peters Ltd, Publishers, 2001.
3. Y. Deng; B.S. Manjunath; H. Shin; “Color image segmentation”, Computer Vision and Pattern Recognition, 1999.
4. J. J. Corso and G. D. Hager, “Coherent regions for concise and stable image description”, Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference Vol 2, 20-26 June 2005 Page(s):184 – 190 IEEE Computer Society Conference on. Volume 2
5. E.R. Dougherty and R.A. Lotufo, “Hands-on morphological image processing”, SPIE press, 2003.
6. P. Salembier and J. Serra, “Flat zones filtering, connected operators, and filters by reconstruction”, IEEE Transactions on Image Processing, vol. 4, No. 8, pp 1153-1160, August 1995.
7. A. G. Hanbury and J. Serra, “Morphological Operators on the Unit Circle”, IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 10, NO. 12, DECEMBER 2001
8. R. Peters , “Mathematical Morphology for Angle -Valued images”, *in Proc. SPIE* 3026, 1997, pp.84 - 94.
9. M. L. Comer and E. J. Delp, “Morphological operations for color image processing,” *J. Electron. Imag.*, vol. 8, pp. 279–289, July 1999.
10. G. Louverdis, M.I. Vardavoulia, I. Andreadis, Ph. Tsalides, “A new approach to morphological image processing”, pattern recognition 35, 1733-1741, 2002
11. V. Barnett, “The ordering of multivariate data," Journal of the Royal Statistical Society A, vol. 139, pp. 318-355, 1976.