REGION-BASED METHOD FOR VISUALLY LOSSLESS IMAGE COMPRESSION

Lenko Grigorov, grigorov@cs.queensu.ca
Purang Abolmaesumi, purang@cs.queensu.ca

School of Computing, Queen's University, Kingston, Ontario K7L 3N6, Canada

ABSTRACT

In this work a method for visually lossless image compression is proposed. Results in the field of Human Vision Modeling are used to select the visually significant regions from the image. Appropriate compression methods are applied to process regions depending on whether they are significant or not. The reconstructed image appears identical to the original, even though some of the information is discarded during compression. The approach is applicable to all types of color images: photographs, diagrams, text, and others. A reference implementation is developed and tested on a variety of examples. Promising results are obtained for images with mixed content, such as text and pictures, where standard methods perform unsatisfactorily.

1. INTRODUCTION

Computer networks and the internet have become a crucial factor for everyday life since they facilitate the exchange of ideas at the fast rate required by modern society. Digital images are widely used to communicate ideas, however, uncompressed digital images place a heavy burden on computer resources, both in terms of storage space and network bandwidth. While many different compression methods are employed, they can be separated into two categories: lossless and lossy. The lossless compression methods compact the image data using fully reversible transformations. Unfortunately, these methods generally do not achieve a high level of compression ratio. The lossy methods perform much better, however, they discard information during the process of compression.

The most widely used compression methods (such as JPEG) strive to achieve minimal loss of information based on mathematical measurements of difference. The particularities of the human vision system are often not taken into account and not used to the advantage. Thus, not only the image quality may be degraded noticeably, but also unwanted artifacts may be introduced. In this work we propose a new method for image compression which is developed in accordance to the properties of human vision. The images reconstructed from the compressed data appear identical to the originals. This is true for all types of images—including photographic pictures and pictures of diagrams and text.

Previous research in the area of image compression is summarized in [3] and [6]. In [4], the human vision model is used to enhance the JPEG2000 compression standard and outstanding compression results are achieved. The goal of similar approaches is to maximize the compression ratio while maintaining acceptable image degradation. In our work, the goal is to preserve the appearance of images and discard only visually insignificant data.

2. HUMAN VISION MODEL

The contemporary results about the human vision model, even though not yet decisive, are well summarized in [1] and [5]. A widely used method to measure the sensitivity of the eye is the use of special test images [4]. Stripes of alternating brightness or color are fit together (see Fig. 1). On one of the axes the frequency of the stripes per distance unit increases exponentially. On the other axis, the difference (of brightness or color) between neighboring stripes decreases exponentially. A subject looking at the test images would be able to tell immediately that below a certain level of difference between the neighboring stripes, they appear as a uniform background. Furthermore, this perception threshold varies with the overall brightness of the surrounding; i.e., the value is greatest for medium-brightness surrounding; with the increase or decrease of the overall brightness of the surrounding the eye is less sensitive to changes. Thus, we can define a threshold function which depends on the relative difference of the changes and on the overall brightness of the surrounding.

The physical preceptors of light in the eye respond to the wavelengths of the red, green and blue light. However, the signals are combined before they reach the brain. The signals which are received by the brain can be described as "what the brightness is", "how blue or how yellow it is" (blue chrominance), and "how red or how

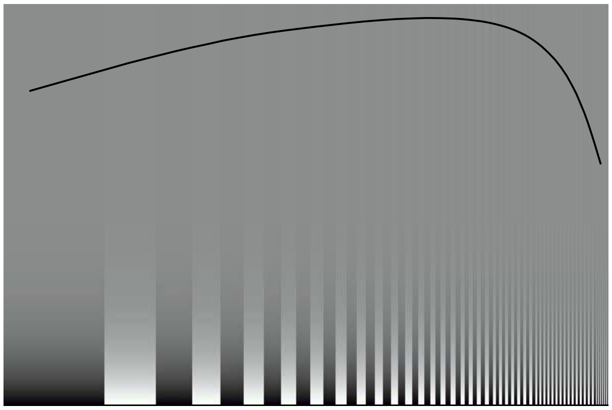

Figure 1 Perception test picture with threshold function

green it is" (red chrominance). Among the standards defined by [7] is the YCbCr color model (brightness, blue chrominance, red chrominance), which approximates how humans see. One of the most important facts to note is that this model is very suitable for compression purposes—it achieves very good decorrelation of the image bands (i.e., the value of a band cannot be predicted from the values of the other bands). This is the reason why the YCbCr model is employed for the compression method proposed in this work.

3. SELECTION OF IMAGE REGIONS

Having approximated the threshold function described in the previous section, it can be applied to separate the regions of the image where the human eye would be able to perceive noticeable changes. There are three threshold functions for the Y, Cb and Cr bands, respectively, so they can be applied separately to the corresponding bands. The selection of regions is described for a single band only because the process is the same for all bands.

The image (band) is divided into two types of regions (see Fig. 2). The white regions describe pixels with visually significant information—such as edges. These regions are governed by the area above the perception threshold, which simply means that the observer will perceive a sharp change in the image (i.e., change of brightness or change of chrominance). The black regions describe the rest of the image. They are governed by the area below the perception threshold, which simply means that the observer might notice changes, but they will be very gradual and appear blurry. Clearly, the black regions can be compressed with loss of information while they would still appear identical to the original.

The approach used for the division of the image into black and white regions is based on the gradient of the image. Indeed, the higher the gradient, the bigger is the relative change in the image and thus the stronger the visual perception of the change. For each pixel of the Y band, the average value of the 3-by-3 neighborhood is calculated and used for the adjustment of the gradient according to the average brightness of the surrounding. After this, a flat threshold value can be applied across the whole image and the resulting division to black and white regions constitutes the desired extraction of visually significant regions.

a)

b)

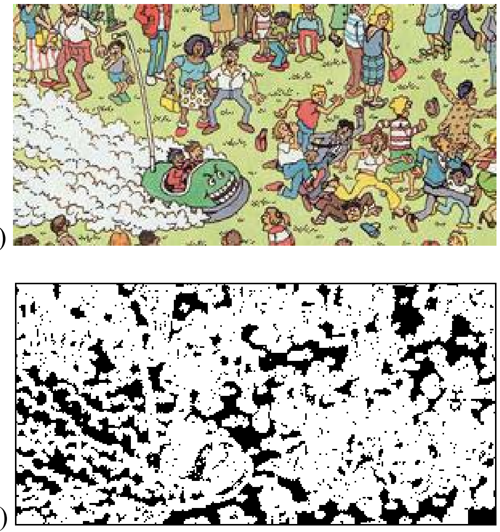

Figure 2 Black and white region selection. a) original image; b) regions for the Y band

4. COMPRESSION OF VISUALLY SIGNIFICANT REGIONS

The white regions contain visually significant information and thus it is desirable that any compression method applied to the pixels in these regions does not cause unwanted loss of detail. The white regions are selected based on the value of the gradient (i.e., the gradient has a large value). This implies that there are big differences in the values of the neighboring pixels. The human sight loses precision with the increase of stimulus. In other words, the eye cannot tell between small variations of large differences. The two least significant bits of pixels in the white regions for both the brightness and the chrominance bands do not play a role for the visual discrimination between the original and the compressed image. Thus, for all pixels in the white regions the two least significant bits are discarded.

As suggested in [3], the coding of each bitplane separately could offer a better level of compression. Additionally, pixel values are transformed using the gray code to increase the chance that neighboring pixels in the image will have same bits in a given bitplane and thus compressibility is improved.

As a summary, the following method for compression of the white regions is used: gray-code the pixel values and then output the six most significant bitplanes separately.

5. COMPRESSION OF THE VISUALLY LESS SIGNIFICANT REGIONS

The compression of the black regions is a more sophisticated process than the compression of white regions, since there is greater freedom in discarding visually insignificant information. As discussed in Section 3, the gradient in the black regions is very small and the human eye perceives only smooth changes in such regions.

A method which can be successfully used in these circumstances is the prediction-based coding; pixel values can be approximated well from their neighborhoods. In order to increase the chance of successful prediction, the black regions can be additionally blurred before the start of compression. This is acceptable, since the observers already perceive the regions as blurred.

Since the black regions can have arbitrary shape, the definition of pixel neighborhoods has to be changed: a neighborhood consists of all pixels from the standard neighborhood which lie in the same black region. Both the convolution and the prediction-based coding use this modified notion of neighborhoods. Black regions are processed using the method for the traversal of arbitrarily shaped regions proposed in [2].

Testing shows that the Y band can be compressed more efficiently than the Cb and Cr bands for which even a single-bit inaccuracy leads to a perceptible “coloring” of the image in the black regions. On the other hand, the overall change of brightness is not detected so easily. The above facts lead to the selection of two different coding methods. For the coding of the Y band, the prediction error is quantized (with a step of 2) before writing to the output and a variable-length binary coding is used. For the coding of the Cb and Cr bands, the prediction error is not quantized and the variable-length binary coding is adjusted to larger expected prediction error.

As a summary, the following method for compression of the black regions is used: for each black region, initialize the neighborhood predictor with the first pixel in the region, then traverse the region and output the prediction errors for the pixels. If the Y band is compressed, the prediction errors are quantized. The prediction errors are coded with a variable-length binary code, which is different for the brightness band and the chrominance bands.

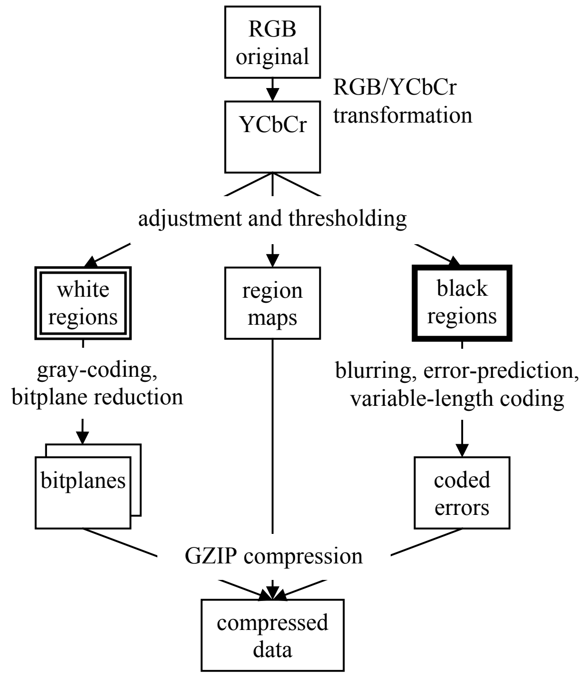

Figure 3 A diagram of the region-based visually lossless compression method

6. RESULTS

Different types of images were used to measure the performance of the region-based compression method: high-frequency images (photographs and illustrations), vector graphics, text and images with mixed content (snapshots of web pages). The output was additionally compacted using the standard GZIP compressor. A diagram of the algorithm is presented in Fig. 3 and the initial results are summarized in Table 1.

When used to compress photographic images, the method performs much better than LZW-TIFF and usually reduces the size of the images to less than 50% of the original size. It cannot outperform standard JPEG compression, however, JPEG-style compressors produce files of very small size at the expense of a much higher level of image degradation (in photographs this may be visually acceptable).

When compressing vector graphics, the results are the complete opposite. The method performs much better than JPEG, while it cannot outperform LZW-TIFF. The compression ratio is worse with low-frequency images, since in addition to the compressed data the method has to store the region bitmaps for the three bands. Furthermore,

Table 1 Average ratios of the region-based compression (RBC) and other compression methods

(average ratios)	RBC/ LZW-TIFF ratio	RBC/best-quality JPEG ratio	RBC/ original size ratio
Photographs	0.42	1.51	0.43
Vector graphics	1.45	0.65	0.11
Text (RGB image)	0.99	0.61	0.04
Mixed (web page)	0.87	0.89	0.18

TIFF works best when there are large, uniform-color regions—and vector graphics consist mostly of such regions. Still, region-based compression usually produces files which are not larger than twice the size of the LZW-TIFF output and are approximately 10% of the size of the original image. When compressing RGB images of text, the proposed method performs as well as LZW-TIFF.

The visually lossless compression method is designed to work with different types of images, including images where other lossy compressors are known to introduce unwanted artifacts. Since the method is based on the human vision model, it is completely versatile. This led to the conclusion that best performance should be expected with images with mixed content. Such images were obtained by taking snapshots of the homepages of different websites, mixing text, vector graphics and photographs. Indeed, for such input, the region-based compression outperforms both LZW-TIFF and JPEG. There is no visible image degradation (see Fig. 4), while the size of the compressed image is about 90% of the size of the output of the other compression methods.

The proof-of-concept implementation consists of a number of separate modules, some written in Java and others in Matlab. It was not our goal to produce optimized code and thus comparisons with other compression methods in terms of speed are irrelevant. We can predict, however, that this method would be slower than single-pass compressors, since it has to make a number of passes. On the other hand, most of the calculations involve very simple operations. The space complexity of the method also exceeds that of standard compressors, since the bitmap of the regions is required in addition to the image data.

7. CONCLUSION

The proposed method for region-based visually lossless compression combines results from different fields of research to offer a way to compress digital images without any loss of visually significant information. The properties of the human vision model are employed to take advantage of different ways to reduce the visually insignificant data in the compressed image. The YCbCr color model is used, both because of its suitability for compression and because of its resemblance of how people see. Each band is divided into regions with visually significant data and with visually less significant data. The regions are coded using the most appropriate coding method to achieve high compression performance. The method is applicable to all types of images, since it takes into account the properties of the human sight rather than mathematical definitions of closeness.

This work serves only as an introduction of this method and many possible improvements are not considered. Future work may involve adjustment of the perception threshold according to the frequency of changes in the image and thus better performance can be expected. Yet another path to explore is noise modeling.

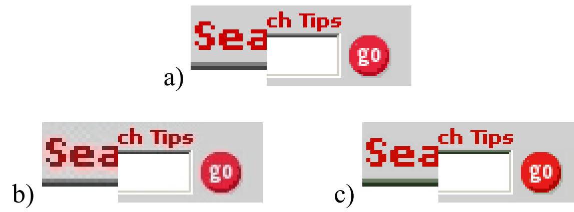

Figure 4 Comparison of quality. A zoomed part of each image is shown in the left. a) original image; b) image compressed with JPEG (best-quality settings); c) same image compressed with region-based method

8. REFERENCES

[1] S. Coren, L. Ward and J. Enns, *Sensation and Perception*, 5th ed, Harcourt Brace College Publishers, 1999.

[2] G. Fernandez i Ubiergo, “Lossless region-based multispectral image compression”, *Image Processing and Its Applications Conference*, Dublin, Ireland, vol. 1, pp. 64-68, 1997.

[3] R. Gonzalez and R. Woods, *Digital Image Processing*, 2nd ed, Prentice Hall, 2002.

[4] M. Nadenau, Integration of human color vision models into high quality image compression, Thesis #2296, Federal Technical University of Lausanne, 2000.

[5] C. Padgham and J. Saunders, *The Perception of Light and Color*, Academic Press Inc., New York, 1975.

[6] M. M. Reid, R. J. Millar and N. D. Black, “Second-generation image coding: an overview”, ACM Computing Surveys, 29(1), pp. 3-29, 1997.

[7] YCbCr Color Model, ITU-R BT.601 Recommendation, International Telecommunication Union, `http://www.itu.int`