ON THE NON-OPTIMALITY OF FOUR COLOR CODING OF IMAGE PARTITIONS

Sameer Agarwal and Serge Belongie

Department of Computer Science and Engineering
University of California, San Diego
La Jolla, CA 92093
{sagarwal,sjb}@cs.ucsd.edu

ABSTRACT

Recent interest in region based image coding has given rise to graph coloring based partition encoding methods. These methods are based on the four color theorem for planar graphs, and assume that a coloring for a graph with the minimum possible number of colors will result in the most compressible representation. In this paper we show that this assumption is wrong. We show that there exist graphs with chromatic number k that can be colored with $k+1$ colors resulting in bitmaps representing image partitions which are more compressible than the corresponding bitmaps generated using any k coloring of the same graph. We conclude with some conjectures on optimal coloring of weighted graphs.

1. INTRODUCTION

Region based image coding is an area that has seen tremendous interest in recent years. One of the tools required for a successful region based image encoding is an efficient image partition encoding method. The classical method for solving is problem is based on using chain codes [1, 2]. However as the number of regions in an image partition increases, the performance of the chain coder goes down rapidly. An interesting alternative approach to the partition encoding problem is based on using lossless compression methods on image bitmaps created by associating a single color with all the pixels in a segment [3, 4].

The simplest way of constructing such a bitmap is to assign to each pixel a color corresponding to the index of the region to which it belongs. This piecewise constant bitmap can be compressed using any of the conventional image compression methods. The method while simple has an obvious flaw. The average number of bits per pixel required to store the image grows with the number of segments in the image.

A closer look at the problem shows that we do not need individual labels for each region, all that is really needed is the ability to discriminate between adjacent regions in the image. So perhaps we can do with fewer colors than the number of distinct regions. This is the well known map coloring problem in graph theory. In our case the maps are restricted to a plane and the solution to this problem is the well known Four Color Theorem [5].

The Four Color Theorem states that any planar graph can be colored using four colors such that no two adjacent vertices have the same color. Hence there exists a four color image corresponding to the image partition which will allow us to discriminate between any two regions. This representation places a hard upper bound of 2 bits per pixel on the amount of space required to store it.

The existing work on using graph colorings for encoding image partitions assumes that the minimum number of colors required to efficiently color and compress a graph is also the number of colors required to get the smallest representation of the partition. In the following we shall show that this is not the case. Graph theoretic minimality does not correspond to information theoretic minimality.

The rest of the paper is organized as follows. In section 2 we formally define our problem, define the notion of *entropy of a graph coloring* and use it to contrast our problem with the standard graph coloring problem as studied in graph theory. We then state and prove our main result. Section 3 concludes with a discussion of our results and some conjectures on optimal colorings of weighted graphs.

2. GRAPH COLORING AND ENTROPY

Formally the map coloring problem is formulated as a vertex coloring problem on an unweighted graph G, where individual vertices of a graph represent regions in the map, and edges between the vertices represent adjacency relations between the region. The minimum number of colors required to color the vertices of a graph G such that no two adjacent vertices have the same color is called the chromatic number of the graph and is denoted by $\chi(G)$.

Now consider the weighted graph G, such that with each vertex v_i we associate a weight p_i, which denotes the proportion of image pixels in the corresponding region. Since we are interested in using this coloring for compressing the

Fig. 1. A segmented image [6] and its four colored version.

bitmap, we are not interested in finding just any coloring of the graph. We are interested in finding that $k-$coloring which will result in the most compressible $k-$colored bitmap. The entropy of the image is a measure of its compressibility. Hence the entropy of a graph coloring is defined as:

$$H = -\sum_{i=1}^{k} C_i \log C_i \quad (1)$$

where

$$C_i = \sum_{j=i}^{N} p_j \delta_{\phi(j)i} \quad (2)$$

and $\phi(j)$ is the function that maps the set of integers $\{1, \ldots, N\}$ representing the segments to the set of colors $\{1, \ldots, k\}$, and δ_{ij} is the Kronecker Delta. The lower the color entropy the higher compression we will be able to get. Hence the aim is to construct the coloring with the lowest possible entropy. This is in contrast to the classical graph coloring problem where we are interested in constructing an admissible coloring with the fewest colors.

We begin with the following lemma.

Lemma 1. *For all integers $n > 1$ there exists a graph G with chromatic number $\chi(G) = n$ and maximum independent set $I(G)$, such that the subgraph induced by $G - I(G)$ has chromatic number n, i.e. $\chi(G - I(G)) = n$.*

Proof. The proof is trivial and can be demonstrated by constructing an example. Consider the complete graph K_n. We know that $\chi(K_n) = n$. Consider now a new graph G' constructed as follows :

1. Add n^2 vertices to K_n.
2. Connect vertices $\{i * n, \ldots, (i+1) * n - 1\}$ to the i^{th} vertex of K_n.

The chromatic number of G' is still n. It can now be trivially shown that the maximum independent set for the graph G' is the set of n^2 vertices in $G' - K_n$, which in turn proves the lemma. □

This leads to our main result,

Theorem 1. *For every integer $k > 1$ there exist graphs with chromatic number k, such that the minimum entropy coloring uses more than k colors.*

Proof. By Lemma 1 we know there exists a graph such that removing the maximum independent set from it does not

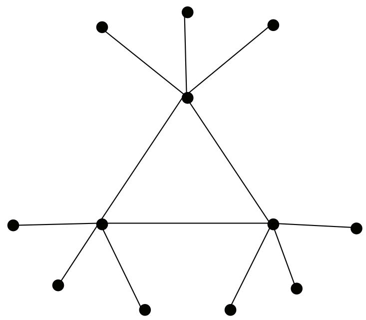

Fig. 2. The graph G' for $k = 3$

reduce its chromatic number. Let G be such an unweighted graph. Let $I(G)$ denote the maximum independent set of G. We will now associate weights with the vertices of G and show that there exists a $k+1$-coloring that has entropy lower than the best (minimum entropy) k-coloring of G. We shall do this by showing that the entropy of the $k+1-$coloring is bounded above by a lower bound on the entropy of the minimum entropy $k-$coloring.

For an arbitrary $0 < \epsilon < 1$, label each vertex in $I(G)$ with weight $1 - \epsilon/|I(G)|$, label the remaining vertices with the weight $\epsilon/|G - I(G)|$. Now color all vertices in $I(G)$ with the color $k+1$. The rest of the vertices can be colored using k colors independent of the vertices in $I(G)$. Now in the worst case (entropy wise), all the vertices will be equally distributed amongst the remaining k colors, resulting in the maximum entropy. Let $H_{k+1}(G)$ denote the entropy of this $k+1$ coloring.

$$\begin{aligned} H_{k+1}(G) &= -[(1-\epsilon)\log(1-\epsilon) + \sum_{i=1}^{k} \frac{\epsilon}{k} \log(\frac{\epsilon}{k})] \\ &= -[(1-\epsilon)\log(1-\epsilon) + \epsilon \log(\frac{\epsilon}{k})] \qquad (3) \end{aligned} \tag{4}$$

Now we consider the minimum entropy k-coloring. Since the removal of the set $I(G)$ does not reduce the chromatic number, any k-coloring of the graph G will have at least one vertex $x \in I(G)$ with color different from the rest of the vertices in $I(G)$. Hence the lowest entropy $k-$coloring will result when $|I(G)| - 1$ vertices are colored 1, the rest except for $k-2$ vertices are colored 2. The remaining $k-2$ colors are assigned one vertex each. In the following $H_k(G)$ denotes the color entropy of this $k-$coloring. α denotes the size of the independent set $|I(G)|$ and β denotes the size of the induced subgraph $|G - I(G)|$.

$$\begin{aligned} H_k(G) = &-\Bigg[\left((1-\epsilon)(1-\frac{1}{\alpha})\right)\log\left((1-\epsilon)(1-\frac{1}{\alpha})\right) + \\ &\left(\frac{1-\epsilon}{\alpha} + \epsilon - \frac{\epsilon(k-2)}{\beta}\right)\log\left(\frac{1-\epsilon}{\alpha} + \epsilon - \frac{\epsilon(k-2)}{\beta}\right) + \\ &\sum_{i=1}^{k-2} \frac{\epsilon}{\beta}\log\left(\frac{\epsilon}{\beta}\right)\Bigg] \\ = &-\Bigg[\left((1-\epsilon)(1-\frac{1}{\alpha})\right)\log\left((1-\epsilon)(1-\frac{1}{\alpha})\right) + \\ &\left(\frac{1-\epsilon}{\alpha} + \epsilon - \frac{\epsilon(k-2)}{\beta}\right)\log\left(\frac{1-\epsilon}{\alpha} + \epsilon - \frac{\epsilon(k-2)}{\beta}\right) + \\ &\frac{(k-2)\epsilon}{\beta}\log\left(\frac{\epsilon}{\beta}\right)\Bigg] \\ \geq &-\Bigg[\left((1-\epsilon)(1-\frac{1}{\alpha})\right)\log\left((1-\epsilon)(1-\frac{1}{\alpha})\right) + \\ &\frac{(k-2)\epsilon}{\beta}\log\left(\frac{\epsilon}{\beta}\right)\Bigg] \end{aligned} \tag{5}$$

The left hand side takes extremal values for $\epsilon = 0, 1$, hence

$$H_k(G) \geq max\left\{-\left(1-\frac{1}{\alpha}\right)\log\left(1-\frac{1}{\alpha}\right), -\frac{k-2}{\beta}\log\left(\frac{1}{\beta}\right)\right\} \tag{6}$$

Since ϵ was chosen arbitrarily we can always choose ϵ to be sufficiently small so that H_{k+1} is smaller than the lower bound on H_k. Hence the graph G with the above weight structure has a lower entropy with $k+1$ colors than the best k-coloring. □

For instance, for the graph in figure 2 for values of $\epsilon \leq 0.1$ a four colouring as described above will give a lower entropy than the best three coloring.

3. DISCUSSION

Theorem 1 proves that even though the chromatic number of an unweighted graph is k, the minimum number of colors required to achieve the lowest entropy maybe more than k. This has some interesting implication and raises some interesting questions.

1. In our proof we have only shown that there exist graphs for which the optimal number of colors that minimizes entropy is larger than the chromatic number of the graph. The issue of determining the optimal optimal number of colors which will minimize the entropy remains open.

2. We are not aware of any work on finding colorings that will explicitly reduce the color entropy. We conjecture that this problem is NP-Hard. To this end we conjecture that the following algorithm will find the minimum entropy coloring for the graph G.

   index=1
   while $V(G) \neq \phi$ **do**
   X = MaxWeightedIndependentSet(G)
   Color(X) = index
   index = index+1
   $G = G - X$
   end while

   Indeed if the above above greedy algorithm is optimal, the problem will be trivially NP-Hard, since identifying the Maximum Independent Set on an unweighted graph is an NP-Complete problem [7].

We believe further work in using map coloring for partition compression should focus on settling the above two questions.

4. ACKNOWLEDGEMENTS

The authors would like to thank Gary Cottrell for raising the question ”Why not use five ?” and Pamela Cosman for useful discussions.

5. REFERENCES

[1] H. Freeman, “Boundary encoding and processing,” in *Picture Processing and Psychopictorics*, B. S. Lipkin and A. Rosenfeld, Eds., pp. 241–266. Academic Press, 1970.

[2] B. Jähne, *Digital Image Processing*, Springer Verlag, 1998.

[3] M. Accame, F.G.B. De Natale, and F. Granelli, “Efficient labeling procedures for image partition encoding,” *Signal Processing*, vol. 80, no. 6, pp. 1127–31, June 2000.

[4] Armando J. Pinho, “Representing image partitions with four colors,” in *Advanced Concepts for Intelligent Vision Systems - Proc. of the 2nd Int. Symposium on Intelligent Vision Systems, ACIVS 2000*, August 2000, pp. 87–91.

[5] K. Appel, W. Haken, and K. Koch, “Every planar map is four colorable. part II. reducibility.,” *Illinois J. Math.*, vol. 27, pp. 491–567, 1977.

[6] D. Martin, C. Fowlkes, D. Tal, and J. Malik, “A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics,” in *ICCV*, July 2001.

[7] Michael Sipser, *Introduction to the Theory of Computation*, PWS Publishing Company, 1996.