Row-Centric Lossless Compression of Markov Images

Matthew G. Reyes
EECS Department, University of Michigan
mgreyes@umich.edu

David L. Neuhoff
EECS Department, University of Michigan
neuhoff@umich.edu

***Abstract*—Motivated by the question of whether the recently introduced Reduced Cutset Coding (RCC) [1], [2] offers rate-complexity performance benefits over conventional context-based conditional lossless coding for sources with two-dimensional Markov structure, this paper compares several row-centric coding strategies that vary in the amount of conditioning as well as whether a model or an empirical table is used in the encoding of blocks of rows. The conclusion is that, at least for sources exhibiting low-order correlations, 1-sided model-based conditional coding is superior to the method of RCC for a given constraint on complexity, and conventional context-based conditional coding is nearly as good as the 1-sided model-based coding.**

I. Introduction

Lossless coding of an image involves blocking (equivalently, grouping) and ordering the pixels in some way, and feeding them, together with a corresponding set of *coding distributions*, to an encoder, which without loss of optimality we can assume to be an *Arithmetic Encoder*. The coding distribution for a given pixel, or block of such, is conditioned on some subset of the pixels, referred to as its *context*, that have already been encoded. Moreover, the coding distribution can be computed based on a *statistical model* to which the image has been fit, or on *empirical statistics* collected from the image or a set of training images. This paper considers how choices of blocking the pixels, contexts for these blocks, and whether coding distributions are *model-based* or *empirical-based*, affect coding rate and complexity.

For simplicity we focus on bilevel images. To provide a well-founded testing ground with interesting correlation structure, we focus on images produced by a uniform Ising model [3] on a square grid graph $G = (V, E)$, whose nodes V are the sites of an $M \times W$ rectangular lattice and whose edges E are pairs of horizontally and vertically adjacent nodes. Letting the random variable X_i associated with each site i assume a value in the alphabet $\mathcal{X} = \{-1, 1\}$, a configuration $\mathbf{x} = (x_i : i \in V)$ has probability

$$p(\mathbf{x};\theta) \quad = \quad \exp\{\ \theta \sum_{\{i,j\}\in E} x_i x_j - \Phi(\theta)\}, \qquad (1)$$

where $\Phi(\theta)$ is the normalizing constant and $\theta > 0$ is the positive edge correlation parameter of the model. This model assigns probabilities that are monotone decreasing in the number of *odd bonds* in an image, which are adjacent pairs of sites with different values. This suggests that this could be a good model for complex bilevel images, such as landscapes and portraits, having numerous black and white regions with smooth or piecewise smooth boundaries between them. Indeed, in [4] a competitive lossy coder for bilevel images was based on this model. We assume M is so large that the model can and will be assumed to be row stationary.

We focus on what we call *row-centric* schemes, which are schemes in which rows are grouped into blocks of height N_b, and within each block, columns are sequentially encoded from left to right. By the Markov property, an optimal row-centric scheme encodes each column with context consisting of the column to the left and the pixels in the preceding row that are above and to the right of the column, with coding distribution equal to the conditional distribution of the column given the context. Due to the assumed stationarity, the rate of this scheme is equal to the entropy-rate $H_\infty = \frac{1}{W} H(\mathbf{X}_{r_2}|\mathbf{X}_{r_1})$, where $\mathbf{X}_{r_1}$ and $\mathbf{X}_{r_2}$ denote successive rows of the image, and H_∞ is the least possible rate for coding this source.

The above is an example of what we refer to as a *1-sided* scheme, in that a block is encoded conditioned on the pixels on one side of it. Computing either model- or empirical-based coding distributions for this scheme is prohibitively complex, the latter due to exponential storage of the empirical statistics and the former due to marginalizing over the pixels below the block, in addition to storage. There are, however, tractable, albeit approximate, model- and empirical-based 1-sided schemes. A tractable 1-sided model-based scheme is based on the idea of truncating the graph below the block, so that no marginalization is required. However, this truncation means that the computed coding distributions, while conditioned on the full 1-sided context, will not be the true conditional distribution for this context. As a result, there will be some redundancy in the coding relative to the optimum rate of H_∞. A tractable 1-sided empirical-based scheme involves limiting the size of the context for a block column so that the corresponding table of empirical statistics is manageable. The resulting coding distribution will be the true conditional distribution for the given context. However, since the context is smaller than the full 1-sided context, the achieved rate will again be greater than H_∞.

In addition to the tractable, approximate 1-sided schemes mentioned above, we also consider *0-sided* schemes in which no conditioning is used. For model-based coding distributions, a 0-sided scheme involves truncating the graph both above and below the block. This results in an even more severe approximation than the 1-sided model-based scheme because not only is the distribution approximate, but there is no conditioning. A 0-sided empirical-based scheme involves empirical statistics from the block alone, without any context from the preceding row, which results in greater redundancy.

While 0-sided coding suffers greater reduandancy compared to 1-sided coding it is interesting to consider, because combining it with *2-sided* coding yields a scheme we refer to as *0/2-sided* coding. In 2-sided coding, a block is encoded conditioned on the row above the block and also the row below the block. By the Markov property the conditional distribution so computed is equivalent to the conditional distribution of the block conditioned on the entire rest of the image. Obviously, the entire image cannot be encoded in this way, but if a set of blocks are initially encoded with 0-sided coding, then the remaining blocks can be encoded with 2-sided coding. This is the Reduced Cutset Coding algorithm introduced in [1], [2]. Though in principle one could devise a 0/2-sided empirical-based scheme, in this paper we consider only model-based 0/2-sided schemes.

This paper seeks to answer how combining the greater redundancy of 0-sided coding with optimal 2-sided coding compares with both 1-sided model-based and 1-sided empirical-based coding. Both model- and empirical-based schemes can be made to approach H_∞ by, respectively, letting block height N_b or context size s be very large. However, the computational complexity of both these approaches is exponential in the respective quantity. Therefore, we are interested in how these three schemes approach H_∞ as a function of the relevant parameter. We first compare 0/2-sided model-based coding with 1-sided model-based coding, and then 1-sided model-based coding with 1-sided empirical-based coding. 1-sided model-based coding has rate decreasing monotonically with N_b. For a given complexity, i.e., N_b, we show numerically that 1-sided model-based coding outperforms 0/2-sided model-based coding. Moreover, we demonstrate empirically that 1-sided model-based coding outperforms 1-sided empirical-based coding, though not by much. In summary, at least for Markov models exhibiting low-order correlations, there are both model-based and empirically-based 1-sided schemes with good performance and low complexity.

Other row-centric schemes include conventional context-based conditional coding, such as in [5], [6], [7], but not schemes where image pixels are coded in Hilbert scan order [8], [7]. We compare performance of our 1-sided model-based scheme with that of JBIG [5], a widely used standard, on a test set of *scenic* [4] bilevel images as well as bilevel images simulated according to the Ising model considered in this paper. We observe that 1-sided model-based coding outperforms JBIG for simulated and very simple bilevel images. However, for bilevel images with more complex structure, JBIG performs better, though 1-sided model-based is competitive. This suggests that a more complicated model is required for optimal coding of more complex scenic bilevel images.

The remainder of the paper is organized as follows. Section II discusses row-centric Arithmetic Encoding of an image, Section III discusses redundancies associated with 0-, 1-, and 0/2-sided row-centric coding, and Sections IV and V discuss numerical results on simulated and real bilevel images. Proofs are either omitted or merely sketched due to lack of space.

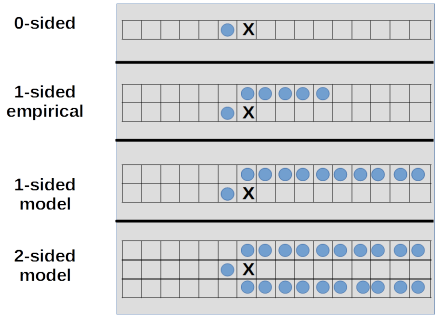

Fig. 1. Context sets in 0-sided, 1-sided empirical-based, 1-sided model-based, and 2-sided model-based coding. The pixel being encoded is indicated with an **X** while the context pixels are depicted with a blue circle.

II. Row-Centric Arithmetic Coding

As mentioned in the introduction, in row-centric coding, rows are grouped into $N_b \times W$ blocks, and within a block $\mathbf{X}_b$, columns of pixels are encoded from left to right. We use subscripts r_i to indicate the i-th row within a block and c_i to indicate the i-th column. Rows are enumerated top to bottom, and columns left to right. That is, $\mathbf{x}_{r_i}$ indicates the configuration on the i-th row of the block, where r_0 and r_{N_b+1} indicate, respectively, the row preceding and row succeeding the block. Likewise, $\mathbf{x}_{c_i}$ denotes a configuration on the i-th column of a block. The configuration on the pixels in row r_0 and columns c_i through c_j is denoted $\mathbf{x}_{r_0,c_i:c_j}$.

When coding column configuration $\mathbf{x}_{c_i}$, a *coding distribution* $p_K(\mathbf{x}'_{c_i})$ for all $\mathbf{x}'_{c_i}$ is passed, together with the configuration $\mathbf{x}_{c_i}$, to an Arithmetic Encoder. In 0-sided coding $p_K(\cdot)$ is conditioned only on $\mathbf{x}_{c_{i-1}}$. In 1-sided model-based coding, $p_K(\cdot)$ is conditioned on $\mathbf{x}_{c_{i-1}}$ and $\mathbf{x}_{r_0,c_i:c_W}$. In 1-sided empirical-based coding, $p_K(\cdot)$ is conditioned on $\mathbf{x}_{c_{i-1}}$ and $\mathbf{x}_{r_0,c_i:c_{i+s-2}}$, where s is the size of the context. In 2-sided model-based coding, $p_K(\cdot)$ is conditioned on $\mathbf{x}_{c_{i-1}}$, $\mathbf{x}_{r_0,c_i:c_W}$, and $\mathbf{x}_{r_{N_b+1},c_i:c_W}$. The contexts for these schemes can be visualized with Fig. 1.

The approximate number of bits produced by the AC encoder when encoding the i-th column is $-\log p_K(\mathbf{x}_{c_i})$. The *rate* R_{c_i} of encoding the i-th column of block $\mathbf{X}_b$ is the expected number of bits produced, divided by N_b. If $p(\mathbf{x}_{c_i}|\mathbf{x}_{K_{c_i}})$ is the true (conditional) distribution of column i given the context, then the rate of encoding the i-th column is

$$R_{c_i} = \frac{1}{N_b}\left[H(\mathbf{X}_{c_i}|\mathbf{X}_{K_{c_i}}) + D(p(\mathbf{x}_{c_i}|\mathbf{x}_{K_{c_i}})||p_K(\mathbf{x}_{c_i}))\right].$$

where D denote divergence. From this, the rate of encoding block b is

$$R_b = \frac{1}{WN_b}\left[\overline{H} + \overline{D}\right],$$

where $\overline{H}$ and $\overline{D}$ are the respective sums of the per-column entropies and divergences.

A. Model and empirical based coding distributions

For model-based methods, the coding distribution is computed by running Belief Propagation [2] on the Ising model

restricted to the subgraph induced by the block of rows, with a possibly modified edge correlation parameter. Messages are first passed from right to left on the resulting line-graph of superpixels (columns) in such a way that after the messages are received at the first column, encoding can proceed from left to right with the coding distributions being computed as they are needed. In the 0- and 1-sided cases, the edge correlation parameter is adjusted to mitigate the effect of truncating edges above or above and below the block. In the case of 1- or 2-sided coding, in which conditioning on either the upper, or both the upper and lower boundaries, is part of the coding distribution, this conditioning is incorporated by introducing *self correlation* [3] on the top row, or top and bottom rows, of the block that bias those sites toward the value of their boundary neighbor. Complexity is proportional to the number of column configurations, which grows exponentially with column height N_b.

Let θ^*_{0,N_b} denote the adjusted parameter used for encoding a block with 0-sided coding and let $p_0(\mathbf{x}_b;\theta^*_{0,N_b})$ indicate the corresponding distribution on the block. The coding distribution for column $\mathbf{X}_{c_i}$ is denoted $p_0(\mathbf{x}_{c_i}|\mathbf{x}_{c_{i-1}};\theta^*_0)$.

For 1-sided coding, θ^*_{1,N_b} denotes the adjusted parameter accounting for truncation of edges below the block, $p_1(\mathbf{x}_b|\mathbf{x}_{r_0};\theta^*_{1,N_b})$ is the distribution on the block, and $p_1(\mathbf{x}_{c_i}|\mathbf{x}_{c_{i-1}},\mathbf{x}_{r_0,c_i:c_W};\theta^*_{1,N_b})$ is the coding distribution for column $\mathbf{X}_{c_i}$.

For 2-sided coding, the block is encoded using coding distribution $p(\mathbf{x}_{c_i}|\mathbf{x}_{c_{i-1}},\mathbf{x}_{r_0,c_i:c_W},\mathbf{x}_{r_{N_b+1},c_i:c_W};\theta)$ computed with respect to the block model $p(\mathbf{x}_b|\mathbf{x}_{r_0},\mathbf{x}_{r_{N_b+1}};\theta)$ based on the original parameter θ.

Empirical coding distributions are based on a table of the frequencies of different configurations of a column for all possible configurations of the context. Letting $\mathbf{x}_T$ denote the configuration being encoded and $\mathbf{x}_K$ denote the configuration of the context, the table consists of values of the form $p^*(\mathbf{x}_T,\mathbf{x}_K)$, from which the coding distribution $p^*(\mathbf{x}_{c_i}|\mathbf{x}_{c_{i-1}},\mathbf{x}_{r_0,c_i:c_i+s-2})$ can be computed, where again s is the size of the context.

There are 1-pass and 2-pass empirical coding methods. In this paper we consider only the 2-pass method in which the relevant frequencies are collected from a set of training images, and then, in a second pass, the rows of the image are encoded using the collected frequencies as coding distributions.

B. Reduced Cutset Coding [1], [2]

Reduced Cutset Coding (RCC) is a 0/2-sided method introduced in [1] and further analyzed in [2]. In it, an image is divided into alternating blocks of rows $\mathbf{X}_L$ and $\mathbf{X}_S$ of sizes $N_L \times W$ and $N_S \times W$, called *lines* and *strips*, respectively. Lines are encoded first in a 0-sided manner, i.e., with no conditioning. The parameter θ^*_{0,N_L} used for the coding distributions of columns is chosen to be the one that minimizes divergence with the true distribution of lines.

III. Row-Centric Coding Redundancy

In this section, we return to the problem, posed in Section I, of attaining rate as close as possible to the entropy rate $H_\infty = H(\mathbf{X}_{r_2}|\mathbf{X}_{r_1})$. We give formulas for the rates of each coding strategy, and we discuss the resulting redundancies associated with the different coding strategies considered in this paper. While we cannot analytically evaluate the rate of decrease of the redundancies, by performing numerical experiments as in the next section, we gain a sense of the relative rates of decrease.

We let R^{0E} and R^{0M} denote the rates for coding N_b rows with 0-sided empirical- and model-based coding, respectively, where dependence on N_b is supressed. Likewise for R^{1E}, R^{1M}, and R^{2M}.

It can be shown that the rate for encoding a block with 0-sided model-based coding is

$$R^{0M} = H_\infty + \frac{1}{N_b W}\left[\overline{D}^{0M} + I(\mathbf{X}_{r_1};\mathbf{X}_{r_0})\right],$$

where $I(\mathbf{X}_{r_1};\mathbf{X}_{r_0})$ is the information between rows r_1 and r_0, and $\overline{D}^{0M}$ is the sum of divergences between $p(\mathbf{x}_{c_i}|\mathbf{x}_{c_{i-1}};\theta)$ and $p_0(\mathbf{x}_{c_i}|\mathbf{x}_{c_{i-1}};\theta^*_{0,N_b})$ over all columns c_i. Likewise, it can be shown that the rate for coding with 0-sided empirical-based coding is

$$R^{0E} = H_\infty + \frac{1}{N_b W}\left[I(\mathbf{X}_{r_1};\mathbf{X}_{r_0})\right].$$

Note that both 0-sided methods suffer the information penalty for independently encoding rows of the image. However, we do not include a divergence term in the above because given enough training data, the empirical coding distribution $p^*(\mathbf{x}_{c_i}|\mathbf{x}_{c_{i-1}})$ for the i-th column will well-approximate the true distribution $p(\mathbf{x}_{c_i}|\mathbf{x}_{c_{i-1}};\theta)$.

As for 2-sided coding, it can be shown that

$$R^{2M} \;=\; H_\infty - \frac{1}{N_b W} I(\mathbf{X}_{r_{N_b}};\mathbf{X}_{r_{N_b+1}}|\mathbf{X}_{r_0}),$$

which, of course, is not an actual coding rate. However, one can code every other block with 0-sided coding, then code the remaining blocks with 2-sided coding, and achieve rate

$$\frac{1}{2}\left[R^{0M}+R^{2M}\right] = H_\infty + \frac{1}{2N_b W}\left[\bar{D}^{0M} + I(\mathbf{X}_{r_{N_b+1}};\mathbf{X}_{r_0})\right].$$

This can be shown by substituting for R^{0M} and R^{2M} and applying standard information relations. In fact, this is the rate of RCC [2] when lines and strips have height $N_L = N_S = N_b$.

Looking now at 1-sided coding, it can be shown that 1-sided model-based coding achieves rate

$$R^{1M} = H_\infty + \frac{1}{N_b W}\overline{D}^{1M},$$

where $\overline{D}^{1M}$ is the sum of divergences between $p(\mathbf{x}_{c_i}|\mathbf{x}_{c_{i-1}},\mathbf{x}_{r_0,c_i:c_W};\theta)$ and $p_1(\mathbf{x}_{c_i}|\mathbf{x}_{c_{i-1}},\mathbf{x}_{r_0,c_i:c_W};\theta^*_{1,N_b})$ over all columns. Similarly, the rate of encoding a block with 1-sided empirical-based coding is

$$R^{1E} = H_\infty + \frac{1}{N_b W}\overline{D}^{1E},$$

where $\overline{D}^{1E}$ is the sum of divergences between $p(\mathbf{x}_{c_i}|\mathbf{x}_{c_{i-1}},\mathbf{x}_{r_0,c_i:c_W};\theta)$ and $p^*(\mathbf{x}_{c_i}|\mathbf{x}_{c_{i-1}},\mathbf{x}_{r_0,c_i:c_{i+s-2}})$

over all columns. Note that the two 1-sided coding schemes do not suffer an explicit information penalty because there is conditioning on the previous row. On the other hand, if the context size c could be chosen as $c = W + 2 - i$ for each column i, then the divergence term $\overline{D}^{1E}$ would vanish. Thus $\overline{D}^{1E}$ is really a sum of conditional information terms. However, both $\overline{D}^{1M}$ and $\overline{D}^{1E}$ are less than $\overline{D}^{0M}$, so it is of interest how these smaller divergences on all blocks compare with the 0/2-sided scheme of RCC in which half the blocks have a larger divergence, plus an information penalty, while the other half actually receive a coding rate reduction.

We have the following proposition relating the coding rates achieved by 0-, 1-, and 2-sided model-based coding, in which we introduce the subscript N_b on the rates $R^{0M}_{N_b}$, $R^{1M}_{N_b}$, and $R^{2M}_{N_b}$ corresponding to a block of N_b rows.

Proposition 3.1: For all N_0 and N_2,

$$R^{0M}_{N_0+1} < R^{0M}_{N_0}, \qquad R^{2M}_{N_2+1} > R^{2M}_{N_2}, \qquad R^{0M}_{N_0} > R^{2M}_{N_2},$$

and for all N_b and N_2,

$$R^{1M}_{N_b+1} < R^{1M}_{N_b} \qquad R^{1M}_{N_b} < R^{0M}_{N_b} \qquad R^{1M}_{N_b} > R^{2M}_{N_2}$$

The first line of inequalities is a restatement of results from [2]. To see that $R^{1M}_{N_b+1} < R^{1M}_{N_b}$, consider blocks B_1 and B_2 of heights N_b and N_b+1, respectively. The main idea is that for a given parameter, the number of bits required to encode the last N_b rows of of B_2 is the same as for all rows of B_1. Moreover, the number of bits required to encode the first row of B_2 is less than the average number of bits required for the N_b rows of B_1. Therefore, the average number of bits over the $N_b + 1$ rows of B_2 is smaller than the average number of bits over the N_b rows of B_1. In particular, this inequality holds with coding parameter θ^*_{1,N_b} chosen to minimize the rate for encoding B_1 averaged over all configurations of r_0. The average rate for encoding B_2 can be further reduced with coding parameter $\theta^*_{N_b+1}$ chosen to minimize the rate of encoding B_2.

To see that $R^{1M}_{N_b} < R^{0M}_{N_b}$, we note that the rate of encoding a block of N_b rows decreases if we add conditioning on $\mathbf{x}_{r_0}$, and it can be further decreased by using the coding parameter θ^*_{1,N_b} that optimizes coding rate for a block of N_b rows conditioning on the previous row.

To see that $R^{1M}_{N_b} > R^{2M}_{N_2}$, we note that the rate of encoding a block conditioned on row r_0 can be decreased by condtioning on row r_2 as well, and this rate can be decreased by using coding parameter θ^*_{2,N_b} that minimizes coding rate for a block conditioned on both the previous and next rows.

IV. Numerical Results: Typical Ising Image

Using Gibbs sampling, we generated 17 configurations of a 200×200 image modeled by an Ising MRF with $\theta = 0.4$. On this dataset we tested three strategies: 0/2-sided model-based coding, 1-sided model-based coding, and 1-sided empirical-based coding. The estimates θ^*_{0,N_b} and θ^*_{1,N_b} were found as in [2] and are shown in Table I. The original parameter θ was used for 2-sided coding.

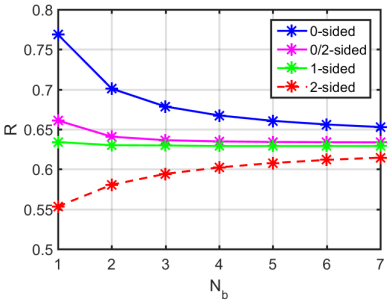

Fig. 2. 0-, 0/2-, 1-, and 2-sided coding rates for model-based methods.

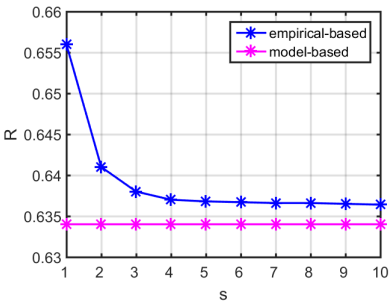

Fig. 3. Empirical- and model-based coding rates for 1-sided coding.

Figures 2 and 3 show the rates attained by the various row-centric coding schemes considered in this paper, as a function of block height N_b. These rates were computed by averaging the negative logarithm of the coding distributions evaluated at the actual pixel/super-pixel values. In [2] we observed that for a given complexity, i.e., the maximum of N_L and N_S, the best performance of 0/2-sided coding was found when lines and strips have the same size, i.e., $N_L = N_S = N_b$. Thus in the model-based comparison, our 0/2-sided method uses lines and strips of equal height.

As predicted by Proposition 3.1, Figure 2 shows that $R^{1M}_{N_b}$ is decreasing in N_b, $R^{1M}_{N_b} < R^{0M}_{N_b}$ and $R^{1M}_{N_b} > R^{2M}_{N'_b}$ for all N_b and N'_b. Also in Figure 2, we observe that for a given block size N_b, 1-sided model-based coding achieves lower rate than 0/2-sided model-based coding. Indeed, 1-sided model-based coding with $N_b = 1$ nearly as good as 0/2-sided coding with $N_b = 7$. Moreover, using the 2-sided coding rate as a lower bound for H_∞, we can say that with $N_b = 3$, 1-sided model-based coding comes to within 3.5% of H_∞.

Figure 3 shows the rate of 1-sided model-based coding with $N_b = 1$, and 1-sided empirical-based coding with $N_b = 1$ for varying sizes of context. Note that context size $s = 1$ actually corresponds to 0-sided empirical-based coding, since in this scheme, only the pixel to the left is used as context. We observe that 1-sided model-based coding with $N_b = 1$ achieves lower rate than 1-sided empirically-based coding with all context sizes we considered. The difference between the rates of 1-sided model-based and 1-sided empirical-based coding shrinks with context size and when the context size is 5, the difference is about 0.0025 bpp or 0.4%. Improvements

TABLE I
CODING PARAMETERS θ^*_{0,N_b} AND θ^*_{1,N_b}

N_b	1	2	3	4	5	6	7
θ^*_0	0.62	0.48	0.45	0.44	0.43	0.43	0.42
θ^*_1	0.45	0.43	0.43	0.42	0.42	0.42	0.41

after that are very slow. Again using the rate of 2-sided model-based coding as a lower bound for H_∞, we observe that 1-sided empirical-based coding with context size 5 comes within 4% of entropy-rate.

Another interesting observation is made by recalling from the previous section that while both 0-sided model-based and 0-sided empirical-based coding methods suffer an information penalty, the model-based scheme suffers an additional divergence penalty $\overline{D}^{0M}$. Therefore, by comparing the $N_b = 1$ point on the 0-sided rate curve of Figure 2 with the $s = 1$ point on the empirical-based rate curve of Figure 3, we can estimate that the normalized divergence between $p(\mathbf{x}_b; \theta)$ and $p(\mathbf{x}_b; \theta^*_{0,1})$ for a single row is about 0.1 bits per pixel. Moreover, by again using the 2-sided model-based rate curve as a lower bound for H_∞, we can upper bound the normalized information $I(\mathbf{X}_2; \mathbf{X}_1)$ between successive rows by 0.041 bits per pixel.

V. APPLICATION TO BILEVEL IMAGES

We compared 1-sided model-based coding to JBIG on a test set of bilevel images as well as typical images for Ising models with different correlation parameter values. The test images are shown in Figure 4 and the coding rates are displayed in Table II. As expected, 1-sided model-based coding outperforms JBIG for images drawn from an Ising model. We observe that as the value of the correlation increases, the gap between 1-sided model-based coding and JBIG decreases. As one can see, the scenic bilevel images are smoother than the typical Ising images, and the trend of improving JBIG performance continues as JBIG outperforms 1-sided model-based coding on these more realistic images. The different relative performances of these two algorithms suggests, not surprisingly, that a more complicated model than the (uniform, first-order) Ising model is required to faithfully capture the dependencies in realistic bilevel images. However, given such a model, we would expect an appropriately adapted version of row-centric model-based coding to do well.

VI. CONCLUDING REMARKS

In this paper we posed the problem of considering different approaches to what are called row-centric coding. We presented the problem in the context of a standard MRF image model in order to provide a well-founded testing ground in which model-based and empirical-based approaches can be compared, and moreover, 1-sided coding can be compared to the tradeoffs in 0/2-sided coding.

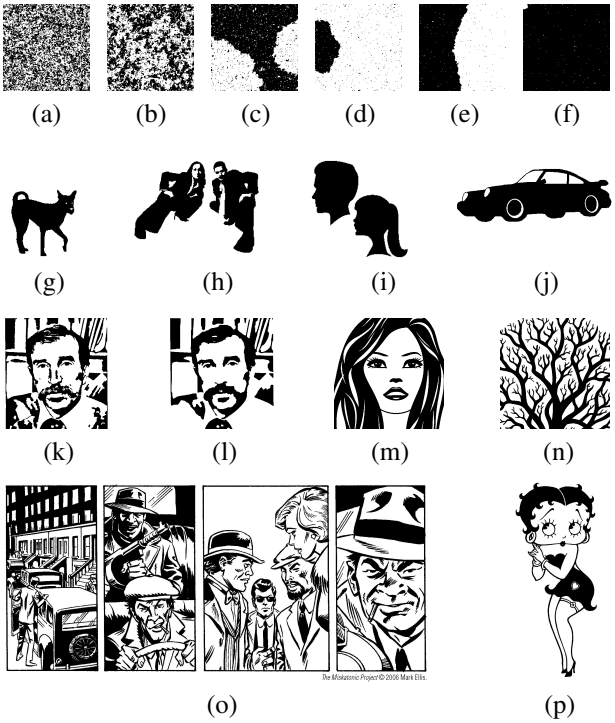

Fig. 4. Bilevel test images.

TABLE II
CODING RATES FOR TEST IMAGES USING JBIG AND 1-SIDED MODEL-BASED CODING WITH $N_b = 1$

Test Image	JBIG Rate	R^{1M}	$\theta^*_{1,1}$
(a)	0.94	**0.84**	0.32
(b)	0.71	**0.63**	0.45
(c)	0.30	**0.26**	0.66
(d)	0.15	**0.11**	0.82
(e)	0.09	**0.05**	0.98
(f)	0.06	**0.02**	1.15
(g)	0.10	**0.07**	1.23
(h)	0.08	**0.07**	0.92
(i)	0.06	**0.03**	0.94
(j)	**0.02**	0.03	0.91
(k)	**0.09**	0.097	0.86
(l)	**0.05**	0.06	0.91
(m)	**0.07**	0.08	0.94
(n)	**0.11**	0.13	0.91
(o)	**0.11**	0.14	0.76
(p)	**.03**	.03	0.83

REFERENCES

[1] M.G. Reyes and D.L. Neuhoff, "Lossless Reduced Cutset Coding of Markov Random Fields", DCC, Snowbird, UT, 2010.

[2] M. G. Reyes and D. L. Neuhoff, "Cutset Width and Spacing for Reduced Cutset Coding of Markov Random Fields," ISIT 2016, July 2016.

[3] R.J. Baxter, *Exactly Solved Models in Statistical Mechanics*, New York: Academic, 1982.

[4] M. G. Reyes, D. L. Neuhoff, T. N. Pappas, "Lossy Cutset Coding of Bilevel Images Based on Markov Random Fields," *IEEE Trans. Img. Proc.*, vol. 23, pp. 1652-1665, April 2014.

[5] "Progressive bi-level image compression, ISO/IEC Int. Std. 11544, 1993.

[6] N. Memon and X. Wu, "Recent developments in context-based predictive techniques for lossless image compression," *The Computer J.*, vol. 40, no. 2, pp. 127-136, 1997.

[7] N. Memon, D.L. Neuhoff, and S. Shende, "An analysis of some common scanning techniques for lossless image coding," *IEEE Trans. Image Proc.*, vol. 9, no. 11, pp. 1837-1848, 2000.

[8] A. Lempel and J. Ziv, "Compression of two-dimensional data, *IEEE Trans. Inform. Theory*, vol. IT-32, no. 1, pp. 18, 1986.