Lossy Source Coding under a Maximum Distortion Constraint with Decoder Side-Information

Jayanth Nayak[1]
Univ. of California, Santa Barbara
e-mail: jayanth@ece.ucsb.edu

Ertem Tuncel
Univ. of California, Riverside
e-mail: ertem@ee.ucr.edu

Kenneth Rose
Univ. of California, Santa Barbara
e-mail: rose@ece.ucsb.edu

A basic problem in information theory is source coding under a distortion constraint when the decoder has side-information about the source [1]. Traditionally, the constraint imposed is that the expected distortion between the source and its reconstruction averaged over the block being coded not exceed a given value. In certain applications (e.g. medical imaging) however, this constraint is considered too weak: we require that with probability 1 the maximum samplewise distortion in a block not exceed a given value.

In this paper we shall focus on the problem of variable length coding of a memoryless source under a maximum distortion constraint when there is side-information solely at the decoder. Using a combinatorial approach, *we derive a non-single-letter expression for the minimum asymptotic average rate as well as single-letter bounds on this rate.* These quantities reduce to known ones when the reconstruction is required to be exact. *We also derive necessary and sufficient conditions for an upper and lower bound on the rate to coincide.*

The key elements of the problem setup are: a sequence of pairs of random variables $\{(X_i, Y_i) \in \mathcal{X} \times \mathcal{Y}\}$, drawn independently at each instant from some common distribution $P_{XY}(X,Y)$, a distortion measure $d\colon \mathcal{X} \times \mathcal{Z} \to [0,\infty)$, and a distortion level $D \geq 0$. $\mathcal{X}, \mathcal{Y}$ and $\mathcal{Z}$ are the source, side-information and reconstruction alphabets, respectively. These sets are assumed to be finite.

Unlike the exact reconstruction case, graph-theoretic tools are insufficient for an analysis of the rate and we need to work with general families of sets. Let $\mathcal{A}$ be the family of D-balls ($\{x \in \mathcal{X} : d(x,z) \leq D\}$) corresponding to the elements $z \in \mathcal{Z}$. Let $\mathcal{B}$ be the family of fan-out sets ($\{x \in \mathcal{X} : p(x,y) > 0\}$) of elements $y \in \mathcal{Y}$. Without loss of generality, we can assume that $\mathcal{B}$ and $\mathcal{A}$ are covering ($\cup_{d\in\mathcal{D}} d = \mathcal{X}$) independence systems ($s \in \mathcal{D}, s' \subset s \Rightarrow s' \in \mathcal{D}$). Let $x \in \mathcal{X}$ be observed at the encoder. At the decoder, the side-information restricts the possible observations to some element of $\mathcal{B}$ containing x. If we are required to reconstruct the source within distortion D, we can accomplish this by specifying a set that contains x and whose intersection with every element of $\mathcal{B}$ results in some element of $\mathcal{A}$. Let $\bar{\mathcal{B}}(\mathcal{A})$ denote the family of all sets whose intersection with every element of $\mathcal{B}$ results in some element of $\mathcal{A}$ (note that the family of sets whose intersection with every element of $\bar{\mathcal{B}}(\mathcal{A})$ is in $\mathcal{A}$, denoted $\bar{\bar{\mathcal{B}}}(\mathcal{A})$ strictly contains $\mathcal{B}$ in general). Therefore variable length scalar coding can be accomplished by first partitioning $\mathcal{X}$ with sets from $\bar{\mathcal{B}}(\mathcal{A})$ and then using a prefix-free code to encode the sets in the partition. For coding vectors of length n, $\overline{\mathcal{B}^n}(\mathcal{A}^n)$ are the sets available for forming the partition. This observation, along with some ideas from [2, 3] leads to the following:

[1]This work was supported in part by the NSF under grants EIA-9986057 and EIA-0080134, the UC MICRO program, Dolby Laboratories, Inc., Lucent Technologies, Inc., Microsoft Corporation, Mindspeed Technologies, and Qualcomm, Inc.

Theorem 1. *The minimum asymptotic average rate*

$$R^*(P_{XY}, D) = \lim_{n\to\infty} \frac{H_\chi(P_X^n, \overline{\mathcal{B}^n}(\mathcal{A}^n))}{n},$$

where $H_\chi(P,\mathcal{D}) \triangleq \min_c\{H(c(X)) : X \sim P,\ c$ *partitions* $\mathcal{X}$ *with sets from* $\mathcal{D}\}$ *for a set family* $\mathcal{D}$, *and* $H(\cdot)$ *is the entropy.*

Note that $R^*(P_{XY}, D)$ depends only on the marginal P_X and the set families $\mathcal{B}, \mathcal{A}$.

Using the results of Wyner-Ziv [1], the fact that $\overline{\mathcal{B}^n}(\mathcal{A}^n) \supset [\bar{\mathcal{B}}(\mathcal{A})]^n$, and that the rate is unchanged if we replace $\mathcal{B}$ by $\bar{\bar{\mathcal{B}}}(\mathcal{A})$, we obtain the following bounds on $R^*(P_{XY}, D)$:

$$H_\kappa(P_X, \mathcal{A}) - H_\kappa(P_X, \bar{\bar{\mathcal{B}}}(\mathcal{A})) \leq$$
$$\leq \max_{X\in Y\in\bar{\bar{\mathcal{B}}}(\mathcal{A})} \min_{\substack{X\in S\in\bar{\mathcal{B}}(\mathcal{A})\\ S\leftrightarrow X\leftrightarrow Y}} I(X;S|Y) \leq R^*(P_{XY}, D) \leq$$
$$\leq H_\kappa(P_X, \bar{\mathcal{B}}(\mathcal{A})),$$

where $H_\kappa(P,\mathcal{D}) \triangleq \min_{\substack{X\sim P\\ X\in T\in\mathcal{D}}} I(X;T)$ is the minimum asymptotic rate required to cover $\mathcal{X}^n$ with sets from $\mathcal{D}^n$. $X \in T \in \mathcal{D}$ denotes that the random variable T is defined over $\mathcal{D}$ such that $p(t|x) > 0 \Rightarrow t \ni x$. $\leftrightarrow$ relates elements of a Markov chain.

The next theorem presents necessary and sufficient conditions on the sets $\mathcal{B}$ and $\mathcal{A}$ such that the upper bound and the weaker lower bound coincide for some nowhere vanishing distribution P.

Theorem 2. *Given set families* $\mathcal{B}, \mathcal{A}$ *and* $\mathcal{C} \subset \bar{\mathcal{B}}(\mathcal{A})$,

$$H_\kappa(P,\mathcal{B}) + H_\kappa(P,\mathcal{C}) = H_\kappa(P,\mathcal{A}).$$

for some nowhere vanishing P *if and only if*

1. *There exist covering families* $\mathcal{B}_0 \subset \mathcal{B}$ *and* $\mathcal{C}_0 \subset \mathcal{C}$ *such that* $b \cap c \neq \emptyset, \forall b \in \mathcal{B}_0, c \in \mathcal{C}_0$. *Let* $\mathcal{A}_0 \triangleq \{b \cap c : b \in \mathcal{B}_0, c \in \mathcal{C}_0\}$.
2. *There exists a function* $r\colon \mathcal{X} \to (0,\infty)$ *such that*

$$\sum_{x\in a} r(x) = 1, \forall a \in \mathcal{A}_0$$
$$\sum_{x\in a} r(x) \leq 1, \forall a \notin \mathcal{A}_0$$

For the noiseless coding with side-information scenario, the first condition reduces to the condition for normality [4] of the confusability graph of the source with $\mathcal{B}$ as the set of cliques and $\mathcal{A}$ as the set of singletons along with the empty set. The second condition is trivially satisfied by the function $r(x) = 1, \forall x \in \mathcal{X}$.

References

[1] A. D. Wyner, J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. on Information Theory*, vol.22, no.1, Jan. 1976.

[2] N. Alon, A. Orlitsky, "Source coding and graph entropies," *IEEE Trans. on Information Theory*, vol.42, no.5, Sep. 1996.

[3] P. Koulgi, E. Tuncel, S. L. Regunathan, K. Rose, "On zero-error source coding with decoder side-information," *IEEE Trans. on Information Theory*, vol.49, no.1, Jan. 2003.

[4] J. Körner, G. Longo, "Two-step encoding of finite memoryless sources," *IEEE Trans. on Information Theory*, vol.19, no.6, Nov. 1973.

0-7803-8280-3/04/$20.00 ©2004 IEEE