A Sampling-based Distributed Source Coding Scheme

H Mo\textsuperscript{1}, J Chen\textsuperscript{1}, J Li\textsuperscript{1} and M Zhao\textsuperscript{1}

\textsuperscript{1}Department of Electronic Engineering, Information School, Yunnan University, Kunming, Yunnan, 650091, China

Email: chenjh@ynu.edu.cn; www_mofeng58@163.com

Abstract. In many practical distributed source coding (DSC) schemes that apply channel coding and iterative decoding, correlation estimation provides the likelihood information about the source bits to initialize the decoding algorithms. Therefore, the decoding depends strongly on the accuracy of this correlation estimation information. In this study, a sampling-based scheme for DSC is proposed that takes advantage of the source redundancy without requiring prior knowledge of its statistics. At the sender, the source sequence is sampled to obtain the sampled and unsampled sub-sequences. Then, the un-sampled sub-sequence is compressed by an arithmetic coder. Meanwhile the syndromes of the sampled sub-sequence are also calculated. At the receiver, the correlation between the side information and the un-sampled sub-sequence is used to estimate the conditional marginal distribution of the sampled sub-sequence. Then, the estimated information is used to initialize the likelihood information of the decoding algorithms. Finally, the likelihood information is combined with the received syndromes to perform iterative decoding to recover the original sampled sub-sequence. Experiment results show that the proposed scheme achieves better rate-distortion performance compared with the existing DSC schemes.

Keywords: Distributed Source Coding (DSC), Correlation Estimation, Side Information, LDPC code

1. Introduction

Distributed source coding (DSC) addresses the problem of compressing several physically separated, but correlated sources which are unable to communicate with each other. The receiver exploits the correlation among the sources to perform joint decoding [1]. DSC can be applied in sensor networks. Recently, DSC has also been applied to the distributed video coding (DVC) by exploiting the temporal correlation between consecutive frames in a video sequence [2].

The Slepian-Wolf theorem [3] states that two correlated sources can be encoded separately without any loss of compression efficiency when they are jointly decoded, as long as the correlation among the sources is preserved throughout their transmission to the receiver. In this paper, we focus on the asymmetric Slepian-Wolf problems. Consider, for example, two discrete correlated sources X and Y, in which the source Y (called side information) is encoded at its entropy rate H(Y) while the source X is encoded at the conditional entropy rate H(X|Y). At the decoder side, the source X could be losslessly recovered with the aid of the correlated side information Y if enough compressed bits have been received. The current mainstream DSC schemes are based on channel coding and source coding. Blizard and Hellman [4] first proposed to apply channel codes to source coding problems. Wyner [5] pointed out the relationship between channel coding and source coding with side information. Pradhan and Ramchandran [6] revived this approach with their distributed source coding using syndrome
(DISCUS) framework. This approach has attracted much attention after being proved by various implementations of systems with low-density parity-check (LDPC) codes [7] and turbo codes [8]. All these schemes effectively utilize the correlation between the source X and side information Y to improve the overall performance of the system.

Entropy code-based distributed source coding schemes have also been used for asymmetric Slepian-Wolf problems. The idea behind using entropy codes for that issue is to exploit their high compression performance together with their capability to exploit the source memory. To this end, Grangetto et al. [9] proposed a distributed arithmetic coding scheme. The scheme based on quasi-arithmetic and overlapped arithmetic codes has also been designed in 0. These schemes reveal better performance than the approaches based on channel codes for short sequences.

Central to DSC is how to effectively use the correlation between the source and the side information at the decoder. This correlation information plays an important role in the practical DSC applications. However, none of the above-mentioned DSC schemes can simultaneously utilize the internal correlation of source and the correlation between the source and side information. Varodayan et al. 0 proposed an adaptive DSC scheme. The encoder sends, in addition to syndrome bits, a portion of the unencode “doping” bits to the decoder. The receiver selects the optimal side-information from different candidates of the side information and performs joint decoding to recover the source. He et al. 0 employed the “doping” idea of 0 and the “discrete universal denoising (DUDE)” idea of 0 that proposed a distributed source coding using improved side information (DSCUISI) scheme. At the encoder, the source sequence is sampled and divided into two parts, namely the sampled and unsampled sub-sequences. At the decoder, the correlation between the un-sampled sub-sequence and the side information is applied based on the DUDE algorithm to generate improved side information. In this work, inspired by 0, we propose a sampling-based distributed source coding (SDISC) scheme. The receiver exploits the correlation between the unsampled sub-sequence and side information to estimate the conditional marginal distribution of the sampled sub-sequence. Then, the estimated information is used to initialize the decoding algorithms by providing the likelihood estimates for the source bits. Compared with those schemes based on LDPC codes and DSCUISI, the proposed framework makes full use of both the internal correlation of the adjacent source symbols in X and the correlation between X and Y. Experiment results also demonstrate that the proposed scheme provides a better rate-distortion performance and faster convergence speed.

The remainder of the paper is organized as follows: Section II introduces the framework of the proposed SDISC scheme. In Section III, we compare the performance of the proposed scheme with that of LDPC codes of [7] and DSCUISI schemes. Finally, Section IV concludes this paper.

2. Description of the proposed SDISC schemes

To describe the correlation between the sources X and Y, a virtual channel is usually used in DSC researches. If we only consider the binary case, this correlation can be described using the binary symmetric channel (BSC) with a crossover probability p.

2.1. The Framework of SDISC

In this work, inspired by 0, we proposed the SDISC scheme to obtain a better rate-distortion performance. The proposed SDISC architecture is illustrated in figure 1. The entire SDISC architecture consists of three stages: (i) Sampling. Let the sources X and Y be correlated sequences \( \{X_i\}_{i=1}^\infty \) and \( \{Y_i\}_{i=1}^\infty \), respectively. For the source X, sampling every k symbols, it can be divided into a sampled X1 and an un-sampled X2 sub-sequences. (ii) Coding. The un-sampled sub-sequence X2 is encoded by applying an arithmetic encoder and the output codeword is transmitted to the decoder without loss. Meanwhile, the LDPC parity check matrix is used to calculated the corresponding syndromes of the sampled sub-sequence X1. For the source Y, in the asymmetric DSC applications, it is usually assumed to be transmitted to the decoder losslessly. (iii) Decoding. At the arithmetic decoder, by exploiting the received codeword, the un-sampled sub-sequence X2 is reconstructed losslessly. The reconstructed X2 and side information Y are used to estimate the conditional marginal...
distribution of the sampled sub-sequence (for ease exposition, the corresponding probability estimation is described in the sequel). Then, these probabilities and the received syndromes are used as the input of the soft-input decoder to recover the sampled sub-sequence $X_t$ (which is denoted as $\hat{X}_t$).

In general, a soft-input decoder can be viewed as a belief-propagation (BP) decoder that incorporates the additional soft information about the source symbols. In this work, we use the LLR BP decoding algorithm to perform joint decoding to recover the sampled sub-sequence. And the estimated posterior probabilities are used to initialize the likelihood information of the LLR BP decoding algorithm.

At the above-mentioned decoding stage, the side information $Y$ can be regarded as a sequence that obtained by letting the source $X$ through a noisy channel with the crossover probability $p$, then the problem of recovering the sampled symbols is similar to the denoising or error correction problem. In this case, the task of the decoder is to denoise the channel output knowing the channel characteristics but without prior knowledge of the source. To solve this problem, Weissman et al. 0 proposed the DUDE algorithm to estimate the source sequence when the statistical characteristics of the discrete memoryless channel (DMC) and the channel output are known, but the source statistics is unknown.

The DSCUIISI scheme proposed by 0 is based on the asymptotic penalty idea of 0 to quantify the minimum loss of the estimated symbols and applied to the actual DSC framework. Subsequently, Ordentlich et al. 0 proposed a simplified DUDE algorithm to incorporate a priori beliefs of the source symbols for denoising. This paper goes further along the idea of 0 and applied it to the proposed scheme. Unlike previous works on discrete universal denoising, in this work, we not only receive the side information $Y$, but the un-sampled sub-sequence is also known at the decoder.

2.2. Estimation of Posterior Distribution of Sampled Sub-sequence

Letting $A$ be a finite alphabet of cardinality $|A| = M$, taking $A = \{\alpha_1, \alpha_2, \ldots, \alpha_M\}$. For a vector $u$, let $u[i]$ denote the $i$-th element of $u$. We assume a given DMC and its transition probability matrix $\Pi = \{\Pi(b, d)\}_{b,d\in A}$ are known at the decoder. $\Pi(b, d)$ denotes the probability of the channel output symbol $d$ when the input is $b$. As mentioned above, the correlation between the sources can be described by a virtual channel, so the transition probability matrix of the virtual channel is also the matrix $\Pi$. Suppose a noiseless sequence $x^n \in A^n$ is transmitted over the channel, and a noisy observation sequence $y^n \in A^n$ is received. At the decoder, a part of the source $X$ is known. Suppose the sampled symbol at time $t$ is $x_t$, we utilize the known adjacent symbols $x_{t-2}, x_{t-1}, x_{t+1}, x_{t+2}, \ldots$ and $y_{t-2}, y_{t-1}, y_{t+1}, y_{t+2}, \ldots$ as the context to estimate the conditional marginal distribution of $x_t$. After the symbol at time $t$ is discarded by sampling, the resulting un-sampled sub-sequence can be denoted as $x^{n\setminus t}$. Therefore, the corresponding posterior distribution of each clean source symbol given the corresponding context is estimated, which is denoted as $P_{x_t|y^n, x^{n\setminus t}}$. According to 0, the formula for estimating the corresponding conditional distribution is given by:

$$P_{x_t|y^n, x^{n\setminus t}} \cdot \Pr(Y^n = y^n, X^{n\setminus t} = x^{n\setminus t}) = P_{x_t, y^n, x^{n\setminus t}}$$  \hspace{1cm} (1)$$

where $\Pr(Y^n = y^n, X^{n\setminus t} = x^{n\setminus t})$ is the joint probability of the known sequences, namely, it is a constant.

Since DMC is memoryless and known, the output at time $t$ is independent of the output at other times. In addition, although the inputs at other times are not independent of their respective outputs, the

![Figure 1](image-url)
inputs at other times and the output at time $t$ are independent of each other. Therefore, according to the conditional independence formula, we have:

$$
P(Y_t = y_t, Y^{n\setminus t} = y^{n\setminus t}, X^{n\setminus t} = x^{n\setminus t} | X_t = x_t) = P(Y^{n\setminus t} = y^{n\setminus t}, X^{n\setminus t} = x^{n\setminus t} | X_t = x_t) \times P(Y^t = y^t | X^t = x^t)$$

(2)

Then,

$$
P(Y_t = y_t, Y^{n\setminus t} = y^{n\setminus t}, X^{n\setminus t} = x^{n\setminus t} | X_t = x_t) = P(Y_t = y_t | X_t = x_t)$$

(3)

Now we define a vector operation, denoted as $\alpha \odot \beta$. And $\alpha \odot \beta$ represents that two vectors $\alpha$ and $\beta$ with equal dimension are multiplied by corresponding elements to obtain a new vector $\gamma$, whose dimension is the same as $\alpha$ and $\beta$. According to the vector operation, we have:

$$(\alpha \odot \beta)[i] = \alpha[i] \cdot \beta[i], \quad (\alpha \odot \beta)^T \cdot \mu = \alpha^T \cdot (\beta \odot \mu)$$

(4)

where $\mu$ indicates a vector whose dimension equals to $\alpha$. We convert equation (3) into a vector form, and use the above vector operations, then

$$
P_{x_t,y^{n\setminus t},y^t,x^{n\setminus t}} = P_{x_t,y^{n\setminus t},y^t} \odot \pi_{y^t}$$

(5)

where $\pi_t$ denote the $i$-th column of $\Pi$. Sum $x_t$ on both sides of equation (3), we have:

$$
P(Y_t = y_t, Y^{n\setminus t} = y^{n\setminus t}, X^{n\setminus t} = x^{n\setminus t}) = \sum_{x_t} P(Y_t = y_t | X_t = x_t) \cdot P(X_t = x_t, Y^{n\setminus t} = y^{n\setminus t}, X^{n\setminus t} = x^{n\setminus t})$$

(6)

Dimension expansion is performed on both sides of equation (6) to expand it into a $M$-dimensional vector, and then

$$
P_{y_t,y^{n\setminus t},y^t,x^{n\setminus t}} = \Pi^T P_{x_t,y^{n\setminus t},y^t} \odot \pi_{y^t}$$

(7)

From equation (5) and equation (7), we can get:

$$
P_{x_t,y^{n\setminus t},y^t} = P_{x_t,y^{n\setminus t},y^t} \odot \pi_{y^t} = \Pi^T P_{y_t,y^{n\setminus t},y^t} \odot \pi_{y^t}$$

(8)

Equation (8) is commonly used to estimate the corresponding posterior distribution of each clean source symbol $x_t$ based on the known $x_2$ and $Y$ at the decoder.

According to the equation (8), we first need to calculate the empirical conditional distribution of each observed symbol $y_t$, given a context of adjacent previous and subsequent symbols. In this work, the symbols $x_{t-1}, x_{t+1}, y_{t-1}, y_{t+1}, y_t$ are selected as the context template to estimate the empirical conditional distribution of $y_t$. Therefore, there are $2^4$ set of binary conditional distributions. The context template is used as a sliding window to traverse the sequence from the start point to the end in the statistical process. Every time a specific context combination appears, the corresponding count value is increased by 1. For example, suppose that the second point $x_{t+1}$ in the context template is sampled, expressed as $x$, and the other condition points are known, the data group is denoted as 00101. Then the number of occurrences of 00101 and 01101 are respectively added by 1. The statistical process is detailed in 0 and 0. After the statistical process is completed, the corresponding posterior distribution of each clean source symbol $x_t$ can be estimated by applying equation (8).

3. Experimental Results

To demonstrate the improved effectiveness of the proposed approach over DSCUISI 0 and LDPC codes of [7], we describe the results of experiment conducted on a bit-plane of a standard grayscale image. For the source sequence $X$, we use the 7th bit-plane of a 512×512 grayscale image. The side information $Y$ is obtained by sending $X$ through a BSC with a crossover probability $p$. According to figure 1, the proposed scheme needs to transmit the code stream including the syndromes and the bits generated by the arithmetic coder. Considered, for example, the case of the sampling every 10 symbols ($k=10$), the source sequence $X$ is divided into the sampled $X_1$ (which contains 26214 bits) and unsampled $X_2$ (235930 bits). Suppose that the code length of the syndromes of $X_1$ and the compressed $X_2$ are expressed as $s$ and $I$, respectively, the total rate of the scheme is obtained by $(s + I)/262144$.

In this work, to obtain a trade-off between accuracy and computational efficiency, the size of the context is selected as 5. The code rate $R$ of the regular LDPC codes used includes 5/8 (3072x8192) and 6/8 (2048x8192). The LLR BP with a maximum number of 100 iterations is used as the decoding algorithm. Each result is the average value of 50 trials.
Taking the 7th bit-plane of Lena as an example, we simulate the proposed scheme for different values of the parameters $p$ and $R$. The results are also compared with those based on LDPC codes and DSCUISI. Figure 2 shows the bit error rate (BER) performance of the different DSC schemes under the different $R$ and $p$. From figure 2, it is clearly shown that the proposed scheme obtains the best BER performance among the schemes of interest. In addition, with the increase of $R$ and $p$, the performance of the proposed scheme is significantly better than the DSCUISI scheme. One of the main reasons is that the proposed scheme can make full use of both the internal correlation of the adjacent source symbols in $X$ and the correlation between $X$ and $Y$, and then provide a more accurate log-likelihood ratio for decoding. The BER performance vs. the number of iterations for different DSC schemes are also compared, as presented in figure 3. It is shown that the proposed framework converges faster than any other schemes within about 30 iterations. From these figures, it is clearly observed that the proposed scheme has better BER performance and faster convergence speed than any other schemes.

![Figure 2. BER vs. $p$ performance of LDPC codes, DSCUISI and the proposed scheme.](image1)

![Figure 3. BER vs. the number of iterations performance of DSCUISI and proposed scheme.](image2)

4. Conclusion

In this paper, the SDISC scheme is proposed to obtain lower BER for DSC applications. The proposed scheme is first attempt to utilize the un-sampled sub-sequence and the side information to estimate the conditional marginal distribution of the sampled sub-sequence to achieve the purpose of making full use of both the internal correlation of the adjacent source symbols in $X$ and the correlation between $X$ and $Y$. Experiment results demonstrate that the proposed scheme outperforms the existing channel code-based DSC schemes in terms of BER performance and the convergence speed. In future work, we will extend this scheme to two-dimensional sampling to further exploit the correlation within the source to improve the accuracy of decoding.

5. References

[1]. Xiong, Z. (2004). Distributed source coding for sensor networks. *IEEE Signal Process. Mag.*, 21(5), 80-94.
[2]. Yang, H., Qing, L., He, X., Ou, X., & Liu, X. (2018). Robust distributed video coding for wireless multimedia sensor networks. *Multimedia Tools and Applications*, 77(4), 4453-4475.
[3]. Slepian, D., & Wolf, J. (1973). Noiseless coding of correlated information sources. *IEEE Trans Inf. Theory*, 19(4), 471-480.
[4]. Blizard, R. B. (1969). Convolutional coding for data compression. Martin Marietta Corp., Denver Div., Rep.
[5]. Wyner, A., & Ziv, J. (1976). The rate-distortion function for source coding with side information at the decoder. *IEEE Trans. Inform. Theory*, 22(1), 1-10.
[6]. Pradhan, S. S., & Ramchandran, K. (2003). Distributed source coding using syndromes (DISCUS): Design and construction. *IEEE Trans. Inform. Theory*, 49(3), 626-643.
[7]. Liveris, A. D., Xiong, Z., & Georgiades, C. N. (2002). Compression of binary sources with side information at the decoder using LDPC codes. *IEEE Commun. Lett.*, 6(10), 440-442.
[8]. Aaron, A., & Girod, B. (2002, April). Compression with side information using turbo codes. *In Proceedings DCC 2002. Data Compression Conference* (pp. 252-261). IEEE.

[9]. Grangetto, M., Magli, E. (2007). Distributed arithmetic coding. *IEEE Commun. Lett.*, 11(11), 883-885.

[10]. Artigas, X., Malinowski, S., Guillemot, C. (2007). Overlapped quasi-arithmetic codes for distributed video coding. In 2007 IEEE International Conference on Image Processing (Vol. 2, pp. II-9). IEEE.

[11]. Varodayan, D. Adaptive distributed source coding. *IEEE Trans Image Process*, 21(5), 2630-2640.

[12]. He Zhiyuan, Chen Jianhua, Wang Zong. (2020). Distributed Source Coding Using Improved Side Information. *Journal of Electronics & Information Technology*, 42(7), 1-8.

[13]. Weissman, T. (2005). Universal discrete denoising: Known channel. *IEEE Trans. Inf. Theory*, 51(1), 5-28.

[14]. Ning, H., & Hua, Q. (2009). Research of LDPC decoding based on modified LLR BP algorithm. *Journal of southwest university (Natural science edition)*, 119-124.

[15]. Ordentlich, E., Seroussi, G., Verdú, S., & Viswanathan, K. (2008). Universal algorithms for channel decoding of uncompressed sources. *IEEE Trans. Inf. Theory*, 54(5), 2243-2262.

**Acknowledgments**

This work is supported by the National Natural Science Foundation of China (61861045).