M-015

Effect of Quality Control on Adaptive Distributed Source Coding for Multi-view Images

Mehrdad Panahpour Tehrani † Michael Droese† Toshiaki Fujii † Masayuki Tanimoto †

1. Introduction

Multi-view images of a scene can be used for several applications ranging from free viewpoint television (FTV) [1] to surveillance. Due to the enormous size of multi-view images, coding is one of the challenges to build such applications. In a scenario with limited processing and communication, work by Slepian and Wolf [2] suggest that it is possible to encode statistically dependent signals in a distributed manner to the same rate as with a system where the signals are jointly encoded. Therefore, distributed source coding of multi-view images is preferable if there is a major constraint on individual camera node performance (i.e., energy, which is consumed by sensing and communication operations). However, approaching the Slepian- Wolf bound is still an open issue for research.

Some work has been carried out [3] in designing a distributed source coding but the performance is not close to information theoretic bound. Aaron et al [4] proposed compression with side information using turbo codes. It approaches the theoretic bound, however it resembles our method in a different way.

We propose an adaptive distributed source coding method without inter-node communication for multi-view images; similar to the distributed source coding method proposed in [5] based on module-operation. To perform the decoding task, disparity estimation is employed to compensate the scene geometry [6] to provide the side information. In this research, we have examined the effect of the quality of the input image on the performance of the proposed coder by using conventional compression/decompression scheme. Experimental results show performance close to the limit of information theory for all qualities. Furthermore, the proposed architecture with adaptive scheme shows significant improvement over previous work.

2. Coding Method

In Fig. 1, we considered three nodes in a cluster (PN, CN and CN_s), which are statistically depended. A PN sends the whole image whereas a CN_s/CN only partially, using an adaptive coding at a rate close to theoretical bound - $H(CN_s|PN)/H(CN|PN,CN_s)$. CN_s sends sub-sampled image and encodes the rest of image, however CN encodes all image. In Fig. 1, *Rx* and *Ry* show the encoding rate, practically.

Fig. 2 shows the coding block diagram. Due to no inter-node communication amongst cameras, the CN/CN_s views are encoded independently at each node. The encoded data is transmitted to the joint decoder. At the joint decoder the side information from PN is provided by the scene geometry, which is obtained by an area-based matching method of [6].

† Department of Information Electronics,
Graduate School of Engineering, Nagoya University

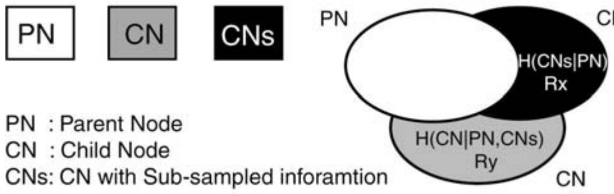

Fig.1: Three views coding architecture and entropies

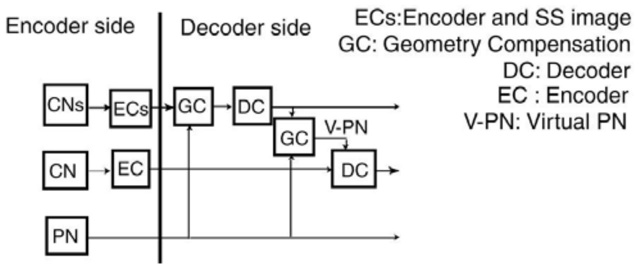

Fig.2: Three views coding block diagram

Before describing the encoding/decoding algorithms it is essential to note the variables used throughout the following sections. *"n x m"* describes a block to be encoded. *"D"* stands for the maximum gray level bound that is imposed on the multi-view image coding at each block. "Maximum disparity" stands for the number of pixels required to find all correspondences in a stereo setup. It also defines the size of a cluster.

2.1 Encoding

PN view is not encoded. However, the CN/CN_s view is divided to blocks. In CN_s, blocks are sub-sampled (i.e., syndrome image), and the rest of pixels (i.e., all pixels in CN) are encoded by using a module-operation with a "*D*" value. The encoded pixels generate an image called coset image. The adaptive value of the "*D*" is obtained by using the average absolute gradients of a block in vertical and horizontal directions (i.e., spatial frequency). It corresponds to the spatial frequency of the scene. In our adaptive coding scheme, the higher spatial frequency, the higher "*D*" value is used. Based on the range, where the measured average gradient is, the adaptive "*D*" to encode a block is obtained. Table 1 shows the look up table (LUT) to decide the adaptive "*D*" at each range. The performance of the decoded image can be control by changing the average "*D*" value used for an image at encoder side. Multiplying a linear weighting factor (i.e., ≥ 0) to the measured average gradient does the controlling procedure.

Furthermore, using the conventional images compression /decompression method such as JPEG, or JPEG2000 can control the input image quality to the encoder, as shown in Fig. 3.

Table 1: Look up table (LUT) for adaptive distributed source coding

Gradient Range	0	1	2 - 3	4 - 7	8 - 15	16 - 31	32 - 63	64 - 127	≥ 128
D	1	2	4	8	16	32	64	128	256

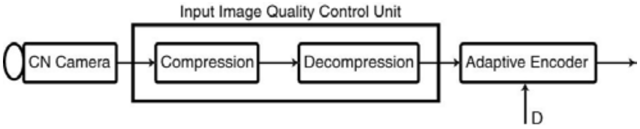

Fig. 3: Block diagram of Encoder with quality control unit

2.2 Decoding

It is using side information from the geometry compensation (GC) [6]. GC provides the corresponded block for CN_s decoding, whereas it generates a Virtual-PN (V-PN) view in the location CN for its decoding. The decoder applies an "inverse" module-operation, which is not unique. Therefore, that solution is chosen which minimizes the distance to the corresponding pixel of the other image (i.e., PN, V-PN).

However, decoding of the coset image is not possible, if the "*D*" value for each block is not known. To solve this problem, there are three ways as follow:

(1) Sending "*D*" value for each block from encoder to decoder.

(2) Estimating "*D*" value by using V-PN image (i.e., side information). Table 1 is used for V-PN image at decoder to estimate the "*D*".

(3) Estimating "*D*" value by using CN image (i.e., coset image). The maximum coset value of each block refers to the range and then the "*D*" is decided using Table 1.

The first method due to overhead on the transmission rate is not preferable. Therefore, we would like to estimate the "*D*" value of each block to decode the coset image. Experimental results on different block sizes and image scenes show that the third method performance is nearly the same as the first method. Hence, we proposed to use the third method for decoding.

3. Experiment

The data set consists of 3 views of 320x240 pixels per view. The camera interval is 15mm with 30cm distance to object. The block size in adaptive coder is 4x4. The performance of the adaptive coder is compared with a fixed coder, which is using the same "*D*". In Fig. 4 the decoded CN quality has compared with V-PN quality. It shows that the adaptive coder has gained over V-PN with lower value of "*D*" (i.e., in average) in comparison with fixed coder. Table 2 shows that the proposed coding can satisfy the Slepain-Wolf bound. Table 3 shows comparison of the bound obtained experimentally for different quality of input image. JPEG compression/decompression change the input image quality.

As it has been shown in equation (1), the tables show the ratio of *Rx*, Ry and the ideal rate $H(CN_s|PN)$, $H(CN|PN,CN_s)$ with *R1*, *R2*, and the ratio of the combined rate $Rx+Ry+H(PN)$ and the ideal combined rate $H(PN,CN_s,CN)$ with *R3,* respectively.

$$R1=\frac{Rx}{H(CN_S\,|\,PN)},\ R2=\frac{Ry}{H(CN\,|\,PN,CN_S)},\ R3=\frac{Rx+Ry+H(PN)}{H(PN,CN_S,CN)} \quad (1)$$

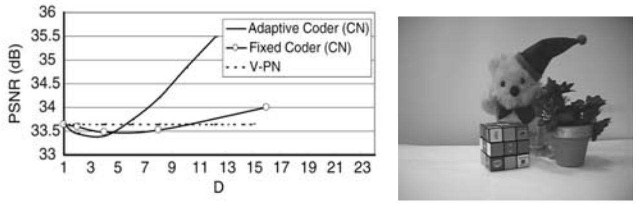

Fig.4: (*left*) PSNR vs. "*D*" for adaptive and fixed coders in comparison with V-PN (*right*) 36.37dB for average *D=16*

Table 2: Rates *Rx*, *Ry* as compared to Slepian-Wolf bound for different coding quality

	D = 16			***D = 23***			***D = 32***		
	R1	***R2***	***R3***	***R1***	***R2***	***R3***	***R1***	***R2***	***R3***
Fixed Coder	1.25	1.45	1.17	1.37	1.62	1.24	1.5	1.63	1.32
Adaptive Coder	0.78	0.85	0.91	0.94	0.98	1.01	1.23	1.26	1.17

Table 4: Rate ***Rx*, *Ry*** achieved by the adaptive scheme as compared to Slepian-Wolf bound for different input image quality

QUALITY	***R1***	***R2***	***R3***
10	1.81	1.31	1.25
20	1.51	1	1.11
40	1.24	0.89	1.03
60	1.14	0.88	1.01
80	1.1	0.88	1
100	1.3	1.25	1.13

4. Conclusion

We proposed an adaptive asymmetric distributed source coding of three correlated views close to Slepian-Wolf bound. The proposed coding performance can approach to the Slepian-Wolf bound by controlling the input image or coding quality. Furthermore, it outperforms the conventional coding scheme. In our future research, we want to develop a suitable compression scheme for coset images statistic.

References

[1] P. Na Bangchang, T. Fujii, M. Tanimoto, 'Experimental System of free viewpoint television', *Proc. of IS&T/ SPIE Symposium on Electronic Imaging*, Santa Clara, CA, USA, Vol. 5006-66, pp. 554-563, Jan 2003.

[2] D. Slepian and J.K. Wolf, "Noiseless Coding of Correlated Information Sources," IEEE Transaction on Information. Theory, vol. IT-19, pp. 471-480, July 1973.

[3] J. Kusuma, L. Doherty, K. Ramchandran, 'Distributed Compression for Sensor Networks', *IEEE Signal Processing Society Conference*, ICIP 2001.

[4] A. Aaron and B. Girod, "Compression with side information using turbo codes," *Proc. IEEE Data Compression Conference, DCC-2002*, Snowbird, UT, April 2002.

[5] M. P. Tehrani, T. Fujii, M. Tanimoto, 'Distributed source coding of multiview images", *Proc. of IS&T/ SPIE Symposium on Electronic Imaging*, VCIP 2004, San Jose, CA, USA, Vol. 5308, No. 31, pp. 300-309, Jan. 2004.

[6] M. Droese, T. Fujii, M. Tanimoto, "Ray-Space Interpolation based on Filtering in Disparity Domain", *Proc. 3D Image Conference 2004*, Tokyo, Japan, June 29-30 2004.