Distributed Compression in Camera Sensor Networks *

Nicolas Gehrig and Pier Luigi Dragotti
Communications and Signal Processing Group
Electrical and Electronic Engineering Department
Imperial College, Exhibition Road, London SW7 2AZ, United Kingdom
{nicolas.gehrig, p.dragotti}@imperial.ac.uk

***Abstract*— We address the problem of distributed compression in camera sensor networks. Our approach uses some geometrical information in order to estimate the correlation in the visual data. This correlation, which is related to the structure of the plenoptic function, can then be used to reduce the overall transmission rate from the sensors to a common central receiver. Our approach allows for a flexible allocation of the bit-rates amongst the encoders and can be made resilient to a fixed number of occlusions. Finally, we show that our distributed coding approach can be extended to general binary sources. The technique we propose uses linear channel codes and can achieve any point of the Slepian-Wolf achievable rate region.**

I. Introduction

The recent advent of sensor network technology is radically changing the way in which we process, sense and transport signals of interest. Sensor networks are given by a large number of low-power, smart devices with computational capabilities connected through wireless links. In this work, we focus on camera sensor networks, that is, we assume that each sensor is equipped with a digital camera and transmits the acquired visual data to a central receiver. The data acquired by different sensors is clearly highly correlated. If sensors could communicate amongst themselves, it would be easy to exploit this correlation in full. However, this collaboration is usually not feasible since it would consume most of the sensors' power. It is therefore necessary to develop separate compression algorithms that can still exploit this correlation without allowing cooperation amongst sensors.

This distributed compression problem has its theoretical foundation in two papers by Slepian and Wolf [1] and by Wyner and Ziv [2]. However, the theories developed in these papers are non constructive. The first constructive design of separate encoders was presented in [3]. Extensions of this encoding strategy have been subsequently presented in several papers (see [4], [5], [6] for example). These distributed compression schemes usually rely on the assumption that the correlation of the source is known a-priori. In this paper, we show how it is possible to predict the correlation structure of the acquired data by using a-priori knowledge of the locations of the cameras and of the objects of interests. We study in details some simplified geometric set-ups and present new distributed compression algorithms that exploit any geometrical information. Our algorithm allows for flexible allocation of the bit-rates amongst the encoders and can be made resilient to a fixed number of occlusions (Section II). We then show that our compression strategy can be generalized to any binary source and present ways to implement symmetric and asymmetric Slepian and Wolf codes by using linear channel codes (Section III).

*This work was supported in part by DIF-DTC project number 12.6.2.

II. Distributed Compression of Correlated Images

The structure of the visual information available at any view position and in any direction is given by the plenoptic function [7]. Therefore, if we can estimate the structure of this function, we can obtain some information about the correlation between the different views of the scene. In our approach, we use some geometrical information about the positions of the cameras and the objects of interest in order to estimate the plenoptic constraints.

The geometrical set-up of our camera sensor network is presented in Figure 1. We consider N cameras evenly placed on a horizontal line and all having the same orientation (perpendicular to the line of cameras). We assume that the observed scene is composed of simple objects with depths bound in $[z_{min}, z_{max}]$. According to the epipolar geometry principles, which are directly related to the plenoptic function, we know that the difference of an object's position on the images obtained from two consecutive cameras is given by: $\Delta = \frac{\alpha f}{z}$, where z is the depth of the object, α is the distance between the cameras and f is the focal length. Since this disparity Δ depends only on the object depth z, and since we know that there is a finite depth of field ($z \in [z_{min}, z_{max}]$), the range of possible disparities for any object is therefore bound in $[\frac{\alpha f}{z_{max}}, \frac{\alpha f}{z_{min}}]$. This result gives us some information about the correlation structure between the different views and can be used to design distributed compression algorithms.

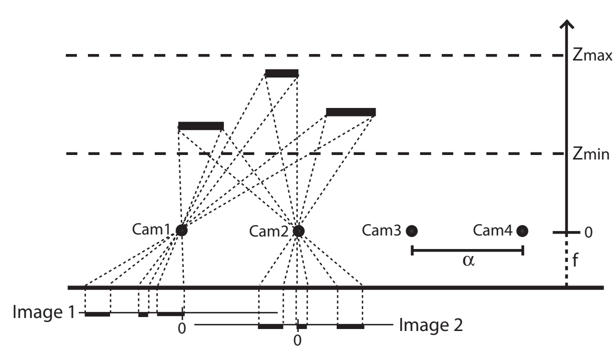

Fig. 1. Our camera sensor network configuration.

Let X and Y be the horizontal positions of a specific object on the images obtained from two consecutive cameras. Assume that the image width is made of 2^R pixels. For a specific X, we know that $Y \in [X + \frac{\alpha f}{z_{max}}, X + \frac{\alpha f}{z_{min}}]$. Our first coding approach consists in sending X perfectly from the first encoder and then, modulo encode Y as $Y' = Y \bmod \lceil \alpha f(\frac{1}{z_{min}} - \frac{1}{z_{max}}) \rceil$ and send it from the second encoder. At the decoder, the original Y is recovered as the only possible position belonging to $[X + \frac{\alpha f}{z_{max}}, X + \frac{\alpha f}{z_{min}}]$ and corresponding to Y'.

This simple approach takes full advantage of the geometrical information to minimize the global transmission bit-rate, however, its asymmetrical structure may be problematic for some practical applications. We propose thus a distributed coding scheme that allows for a flexible allocation of the bit-rates amongst the encoders. Let $\widetilde{Y}$ correspond to $Y - \lceil \frac{\alpha f}{z_{max}} \rceil$. We can thus see that the difference $(\widetilde{Y} - X)$ is contained in $\{0, 1, \ldots, \delta\}$, where $\delta = \lceil \alpha f(\frac{1}{z_{min}} - \frac{1}{z_{max}}) \rceil$. Looking at the binary representations of X and $\widetilde{Y}$, we can notice that the difference between them can be computed using only their last R_{min} bits where $R_{min} = \lceil \log_2(\delta + 1) \rceil$. Our distributed coding approach is presented in Figure 2 and can be described as follows: For each object's position, send the last R_{min} bits from both sources and send only complementary subsets for the first $(R - R_{min})$ bits. This simple technique is very powerful since it allows for a completely flexible allocation of the bit-rates amongst the two encoders.

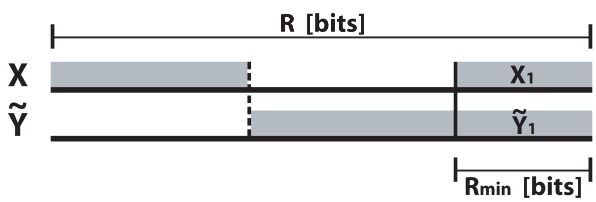

Fig. 2. Our distributed compression approach for two correlated sources. The last R_{min} bits are sent from the two sources but only complementary subsets of the first $(R - R_{min})$ bits are necessary at the receiver for a perfect reconstruction of X and Y.

According to Slepian and Wolf [1], the minimum information that must be sent from each encoder is given by the conditional entropy $H(X|Y)$ and $H(Y|X)$ respectively. Assuming that the difference between X and Y is uniformly distributed and that $\delta + 1$ is a power of 2, we can state that $H(X|Y) = H(Y|X) = R_{min}$. The remaining information that has to be sent in order to allow for a perfect reconstruction is related to the mutual information $I(X;Y)$ and is by definition available at both encoders. If X is uniformly distributed, we have that $H(X) = R$ and $I(X;Y) = H(X) - H(Y|X) = R - R_{min}$. Therefore, our coding approach is, in this case, optimal.

Our coding approach can be generalized to any number of sources and can be made resilient to a fixed number of occlusions. Consider a system with N cameras as depicted in Figure 1. Assume that any object of the observed scene can be occluded in at most $M \leq N - 2$ views. The following distributed coding strategy is sufficient to allow for a perfect reconstruction of these N views at the decoder and to interpolate any new view: Send the last R_{min} bits of the objects' positions from only the first $(M+2)$ sources, with $R_{min} = \lceil \log_2((M+1)\delta) \rceil$ and $\delta = \lceil \alpha f(\frac{1}{z_{min}} - \frac{1}{z_{max}}) \rceil$. Then, for each of the N sources, send only a subset of its first $(R - R_{min})$ bits such that each particular bit position is sent from exactly $(M+1)$ sources. Since no more than M views are occluded, this strategy ensure the reception of the last R_{min} bits from at least two sources and the reception of each particular bit position from at least one source. Therefore, a full reconstruction of all the views is possible.

III. Slepian-Wolf Codes using Linear Channel Codes

The coding strategy shown in Figure 2 can be generalized to any binary sources. In this section, we first present a simple comprehensive example based on the Hamming $(7,4)$ code, and then we propose a general constructive approach using any linear channel code.

A. A simple example with the Hamming $(7,4)$ code

We consider the Hamming $(7,4)$ code $\mathcal{C}$ whose parity check matrix is given by:

$$\mathbf{H} = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}. \tag{1}$$

We know that a 7-bit codeword x belongs to the Hamming code $\mathcal{C}$ if and only if its syndrome is equal to zero:

$$s_x = \mathbf{H}x^T = 0 \iff x \in \mathcal{C}. \tag{2}$$

The minimum distance between any two of the 16 codewords of the Hamming code is three. This code is therefore able to correct up to one bit error per codeword. Assume e_i is the error pattern corresponding to an error at bit position i (e_i has six 0 and one 1 at position i). We define $y = x \oplus e_i$ where $\oplus$ corresponds to the binary addition. This codeword y does clearly not belong to $\mathcal{C}$ since its distance from x is equal to one. Its syndrome is given by:

$$s_y = \mathbf{H}y^T = \mathbf{H}(x \oplus e_i)^T = \mathbf{H}x^T \oplus \mathbf{H}e_i^T = \mathbf{H}e_i^T. \tag{3}$$

We can therefore see that the syndrome of an erroneous codeword does not depend on the original codeword but only on the error pattern. This means that if we change the i^{th} bit of all the codewords of $\mathcal{C}$, this produces 16 new codewords, all having syndrome $\mathbf{H}e_i^T$. This new set of codewords is called *coset number* i and has the same properties than C (coset 0), that is, the minimum distance between any two codewords is still three. All the 2^7 possible 7-bit blocks are thus distributed in 8 distinct cosets. Notice that this Hamming $(7,4)$ code has a particular structure such that the syndrome of an erroneous codeword gives the binary representation of the error position, or similarly, the coset number. This property allows for a straightforward decoding.

Consider now two discrete memoryless uniformly distributed 7-bit binary random variables x and y, correlated such that their Hamming distance is at most one ($d_H(x, y) \leq 1$). Assume that x and y belongs to cosets i and j respectively. The difference between x and y is given by the error pattern $e_k = x \oplus y$ (x and y differs at position k). We know that the syndromes of x and y are given by $s_x = \mathbf{H}x^T = \mathbf{H}e_i^T$ and $s_y = \mathbf{H}y^T = \mathbf{H}e_j^T$ respectively. We can see that:

$$s_k = \mathbf{H}e_k^T = \mathbf{H}(x \oplus y)^T = \mathbf{H}x^T \oplus \mathbf{H}y^T = s_x \oplus s_y. \quad (4)$$

This result shows that knowing only the syndromes of x and y, we can retrieve the syndrome of their difference pattern and therefore, the bit position where they differ.

Our coding technique can now be presented as follows: Assume the following block representations for x, y and $\mathbf{H}$:

$$x = [x_a \quad x_b] \qquad y = [y_a \quad y_b] \qquad \mathbf{H} = [\mathbf{H}_a \quad \mathbf{H}_b] \quad (5)$$

where the first and the second blocks are of length 4 and 3 respectively. The syndromes of x and y are computed at their respective encoders as: $s_x = \mathbf{H}x^T = \mathbf{H}_a x_a^T \oplus \mathbf{H}_b x_b^T$ and $s_y = \mathbf{H}y^T = \mathbf{H}_a y_a^T \oplus \mathbf{H}_b y_b^T$. Encoder 1 transmits s_x together with a subset of x_a. Encoder 2 transmits s_y together with the subset of y_a which is complementary to the one chosen by the first encoder. For example, as presented in Figure 3, the encoder 1 could send $[x_1 \quad x_2 \quad s_x^T]$ and encoder 2 $[y_3 \quad y_4 \quad s_y^T]$.

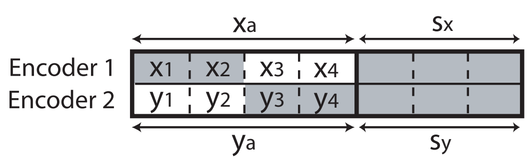

Fig. 3. Example of distributed source coding of two correlated 7-bit blocks using the Hamming $(7, 4)$ code. A total of only 10 bits are sent to the receiver (gray squares). The two blocks can be perfectly recovered at the receiver if their Hamming distance is at most one.

At the decoder, the syndrome of the difference pattern between x and y is obtained by computing the sum of the two syndromes $s_x \oplus s_y$. Using the corresponding error pattern (recovering the difference position), the missing bits of x_a and y_a can easily be retrieved. Finally, x_b and y_b are obtained as: $x_b^T = \mathbf{H}_b^{-1}(s_x \oplus \mathbf{H}_a x_a^T)$ and $y_b^T = \mathbf{H}_b^{-1}(s_y \oplus \mathbf{H}_a y_a^T)$, where the inverse matrix of $\mathbf{H}_b$ is given by:

$$\mathbf{H}_b^{-1} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}. \quad (6)$$

Since x and y are uniformly distributed, we have: $H(x) = H(y) = 7$ bits. We know that y can take 8 different equiprobable values for a specific x. Hence, $H(y|x) = H(x|y) = 3$ bits. The joint entropy of x and y is therefore equal to $H(x, y) = H(x) + H(y|x) = 10$ bits. Our coding scheme uses a total of 6 bits to send the two syndromes and a total of 4 bits to send the two complementary subsets of x_a and y_a and is therefore optimal.

B. Constructive approach using any linear channel code

Assume we have an (n, k) binary linear code $\mathcal{C}$ with parity check matrix $\mathbf{H}$ in its reduced form such that: $\mathbf{H} = [\mathbf{H}_1 \quad \mathbf{H}_2]$, where $\mathbf{H}_1$ and $\mathbf{H}_2$ are of size $(n-k \times k)$ and $(n-k \times n-k)$ respectively. Assume without loss of generality that $\mathbf{H}_2$ is non-singular. Notice that if the code is systematic, $\mathbf{H}_2$ is simply the identity matrix.

We know that such a code is capable of correcting 2^{n-k} different error patterns. Assume $\mathcal{C}$ is able to correct perfectly up to M errors per n-bit code block. We know that the following relation must hold:

$$2^{n-k} \geq \sum_{j=0}^{M} \binom{n}{j} \qquad \text{(sphere packing bound).} \quad (7)$$

This code $\mathcal{C}$ generates 2^{n-k} cosets each containing 2^k codewords of length n.

Let x_i be a binary block of length n represented as: $x_i = [a_i \quad b_i \quad q_i]$, where a_i, b_i and q_i are of length k_1, k_2 and $n-k$ respectively (k_1 and k_2 are chosen such that their sum is equal to k). The syndrome of x_i is defined as: $s_i = \mathbf{H}_1[a_i \quad b_i]^T \oplus \mathbf{H}_2 q_i^T$. We know that x_i belongs to the coset number k, if its syndrome is given by: $s_i = \mathbf{H}x_i^T = \mathbf{H}e_k^T$ (e_k is the coset leader of coset number k, i.e., the codeword with minimum weight).

Consider now two n-bit blocks x_1 and x_2, correlated such that their Hamming distance $d_H(x_1, x_2)$ is at most m. Assume that the channel code $\mathcal{C}$ is able to correct up to $M \geq m$ errors. Our distributed coding strategy consists in sending only $[a_1 \quad s_1^T]$ and $[b_2 \quad s_2^T]$ from the encoders 1 and 2 respectively. The transmission bit-rates are therefore given by: $R_1 = n - k_2$ bits and $R_2 = n - k_1$ bits, corresponding to a total of $R_1 + R_2 = 2n - k$ bits.

At the receiver, we let e_d correspond to the "difference pattern" between x_1 and x_2 as: $e_d = x_1 \oplus x_2$. We know that the syndrome of e_d is given by $s_d = \mathbf{H}e_d^T = \mathbf{H}(x_1^T \oplus x_2^T) = s_1 \oplus s_2$. We can now retrieve the error pattern e_d corresponding to this syndrome s_d using one of the following techniques: If the code is not too large, a simple lookup table storing the corresponding pattern error for each possible syndrome can be used. For larger code, an iterative method has to be used. Using an iterative decoding scheme such as the one proposed in [6], we can recover e_d as the closest codeword to the all zero sequence satisfying the syndrome s_d. Notice that this iterative decoding approach is particularly suited for LDPC codes which are amongst the best block codes known for memoryless channels [8].

Knowing the difference pattern e_d, the missing bits of the k first bits of x_1 and x_2 are easily obtained as: $[a_2 \quad b_1] = [a_1 \quad b_2] \oplus e_d^k$, where e_d^k corresponds to the k first bits of e_d.

We know that the syndrome of x_1 corresponds to: $s_1 = \mathbf{H}_1[a_1 \quad b_1]^T \oplus \mathbf{H}_2 q_1^T$. Let z_1 be defined as: $z_1 = s_1 \oplus \mathbf{H}_1[a_1 \quad b_1]^T$. We can now retrieve q_1 by computing: $q_1^T = \mathbf{H}_2^{-1} z_1$. Notice that $\mathbf{H}_2^{-1}$ can be obtained using Gaussian Elimination and that, if $\mathcal{C}$ is systematic, $\mathbf{H}_2 = \mathbf{I}$ and $q_1 = z_1^T$. Knowing q_1, we have now completely recovered x_1 and we can easily obtain x_2 as $x_2 = x_1 \oplus e_d$. We can now summarize our coding approach with the following proposition:

Proposition 1: Assume X and Y are two binary sequences of length n, correlated such that their Hamming distance is at most m. Consider an (n, k) linear channel code $\mathcal{C}$ that can correct up to $M \geq m$ errors per n-bit block. The following distributed coding strategy uses a total of $2n-k$ bits to encode the two sequences and is sufficient to allow for a perfect reconstruction of them at the decoder:

- Send the syndromes of X and Y from their respective encoders.
- Send only complementary subsets of their first k bits.

In terms of performance, we can say that the ability of our distributed source coding technique to work close to the Slepian-Wolf bound only depends on the quality of the channel code used. More specifically, if X and Y are uniformly distributed and $p(Y|X)$ is the transition probability, then the closer the channel code $\mathcal{C}$ gets to the capacity of the binary channel $p(Y|X)$, the closer our system gets to the Slepian-Wolf bound. The design of capacity achieving channel codes, however, is beyond the scope of this paper.

Notice that linear codes used in our approach can be either systematic or non-systematic. A similar coding strategy which, however, works only with systematic codes has been recently proposed in [9].

C. Generalization to more than two sources

Our coding strategy proposed in Section III-B can be extended to an M sources scenario (see Figure 4) through the following proposition:

Proposition 2: Assume $x_1, \ldots, x_M$ are M binary sequences of length n correlated such that the Hamming distance between two consecutive sequences is at most m (i.e., $d_H(x_i, x_{i+1}) \leq m$ for $i = 1, \ldots, M-1$). Consider an (n, k) linear channel code $\mathcal{C}$ that can correct perfectly up to $M \geq m$ errors per n-bit block. The following distributed coding strategy uses a total of $n + (M-1)(n-k)$ bits to encode the M sequences and is sufficient to allow for a perfect reconstruction of all of them at the decoder:

- From each encoder, send the syndrome s_i of the corresponding block x_i.
- Send only complementary subsets of their first k bits such that each bit position is sent from only one encoder.

The decoding procedure in that case is similar to the one proposed in Section III-B. Using the M syndromes, we can recover the $M-1$ difference patterns. Then, using these difference patterns, all the missing bits of the a_i's can be recovered for all the sequences. Finally, the original blocks $x_i = [a_i \quad b_i]$ are completed by computing:

$$b_i^T = \mathbf{H}_2^{-1}(s_i \oplus \mathbf{H}_1 a_i^T). \tag{8}$$

This coding strategy is in some cases optimal. For instance, if the M sources $(x_1, x_2, \ldots, x_M)$ are uniformly distributed, the Hamming distance of two consecutive sequences is at most m and the sequences $x_1, x_2, \ldots, x_M$ form a Markov chain;

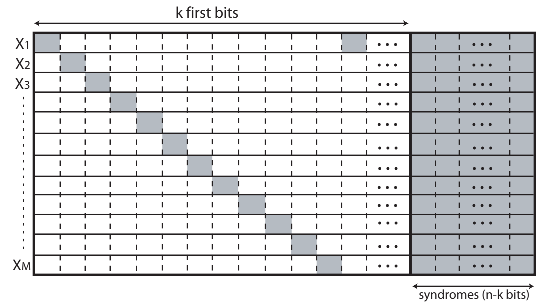

Fig. 4. Our encoding strategy for M correlated binary sources. Each encoder sends the syndrome and a subset of the first k bits of their input block. The subsets are chosen such that each bit position is sent from only one source.

then, if the code $\mathcal{C}$ is such that $M = m$ and Equation (7) is satisfied with equality (i.e., $\mathcal{C}$ is perfect), we can write:

$$\begin{aligned} H(x_1, x_2, \ldots, x_M) &= H(x_1) + \sum_{i=2}^{M} H(x_i|x_{i-1}) && (9) \\ &= n + (M-1)(n-k). && (10) \end{aligned}$$

Our scheme, as indicated in Proposition 2, uses $n + (M-1)(n-k)$ bits and is therefore optimal in this case.

IV. Conclusions

We have proposed a distributed compression technique for camera sensor networks. Our approach estimates the correlation of the different views using some geometrical side information and allows for a flexible allocation of the bit-rates amongst the encoders. An extension of our coding technique to any binary sources has also been proposed. Ongoing research is focusing on the development of efficient distributed compression algorithm for camera sensor networks in natural scenes.

References

[1] D. Slepian and J. Wolf, “Noiseless coding of correlated information sources,” *IEEE Trans. on Inf. Theory*, vol. 19, no. 4, pp. 471–480, July 1973.

[2] A. Wyner and J. Ziv, “The rate-distortion function for source coding with side information at the decoder,” *IEEE Trans. on Inf. Theory*, vol. 22, no. 1, pp. 1–10, January 1976.

[3] S. Pradhan and K. Ramchandran, “Distributed source coding using syndromes (DISCUS): Design and construction,” in *IEEE Int. Conf. on Data Compression*, 1999, pp. 158–167.

[4] J. Garcia-Frias, “Compression of correlated binary sources using turbo codes,” *IEEE Comm. Letters*, vol. 5, no. 10, pp. 417–419, October 2001.

[5] A. Aaron and B. Girod, “Compression with Side Information Using Turbo Codes,” in *IEEE Int. Conf. on Data Compression*, April 2002, pp. 252–261.

[6] A. Liveris, Z. Xiong, and C. Georghiades, “Compression of binary sources with side information at the decoder using LDPC codes,” *IEEE Comm. Letters*, vol. 6, no. 10, pp. 440–442, October 2002.

[7] E. Adelson and J. Bergen, “The plenoptic function and the elements of early vision,” in *Computational Models of Visual Processing*. MIT Press, 1991, pp. 3–20.

[8] T. Richardson and R. Urbanke, “The capacity of low-density parity-check codes under message-passing decoding,” *IEEE Trans. on Inf. Theory*, vol. 47, no. 2, pp. 599–618, February 2001.

[9] V. Stankovic, A. Liveris, Z. Xiong, and C. Georghiades, “Design of Slepian-Wolf codes by channel code partitioning,” in *IEEE Int. Conf. on Data Compression*, March 2004.