SENSING AND COMMUNICATION WITH AND WITHOUT BITS

Michael Gastpar,[1] *Martin Vetterli,*[1,2] *Pier Luigi Dragotti*[3]

[1] Dept. of EECS, University of California, Berkeley, CA 94720
[2] Faculté Informatique et Communications, Ecole Polytechnique Fédérale, CH-1015 Lausanne
[3] Department of Electrical and Electronic Engineering, Imperial College, London SW7-2AZ

ABSTRACT

The successful design of sensor network architectures depends crucially on the structure of the sampling, observation, and communication processes. One of the most fundamental questions concerns the sufficiency of *discrete* approximations in time, space, and amplitude. In the case of space and time, the question can be rephrased as whether there is a *spatio-temporal sampling theorem* for typical data sets in sensor networks. This question has a positive answer in many cases of interest. The issue of discretization of amplitudes is more subtle and can be expressed as the question of whether there is a *(source/channel) separation theorem* for typical sensor networks. We show that this question has a negative answer in general and that the price of separation can be large. To illustrate these issues, we review the underlying theory and discuss specific examples.

1. INTRODUCTION

In this paper, we take a structure-driven end-to-end approach to the sensor network problem, illustrated in Figure 1. Underlying the whole problem is the physics of the process of interest. This structures the data sets, points to sampling schemes, and indicates what types of correlation will be present in the sensor data. After sampling using the sensors, we are faced with the classic dilemma of the communication engineer: "to separate or not to separate". That is, we either go to the digital domain, and apply discretization of the data through quantization and source compression, or we keep data in analog form. The former implicitly assumes a separation into source and channel coding, and can be optimal in certain scenarios, while the latter permits any form of joint source-channel coding. Thus, the main focus and goal is to show how the structure of the distributed sensing and communication problem dictates new processing architectures. The key challenge lies in the discretization of space, time, and amplitude, since most
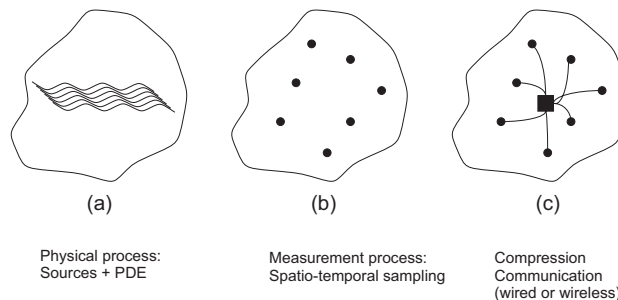
Fig. 1. The end-to-end sensor network problem. (a) A physical environment, possibly driven by sources, generates a distributed data set. (b) A number of sensors acquire, through spatio-temporal sampling, a discrete time-space version of the physical data set. (c) The data set needs to be conveyed to a central location, via distributed compression combined with communication, or via a joint source channel scheme.

of the advanced signal processing systems operate in discrete domain. In the paper, we investigate and illustrate the sufficiency of such discretization, but also the lack thereof.

2. TIME AND SPACE

Sampling is so common that we sometimes forget it is a little miracle, and that it comes with a few strings attached. In the case of sensor networks, the critical issue is certainly the sampling in space, inherent in the discrete nature of the sensors. Also, distributed signals exist in time and space, and are thus inherently multidimensional. Distributed signal acquisition is thus the spatio-temporal sampling of such signals. Of course, the field of array signal processing has dealt with such problems in the past (see e.g. [1]), but with a perspective that is different from the one used in sensor networks. In typical array signal processing, the array is one-dimensional, regular, and the signals are often narrow band. In sensor networks, the array is irregular and two dimensional (random sensor placement on a plane), and the signals can be wideband (e.g. sound, images, etc.). The obvious question is one

of spatial sampling, with the twist that there cannot be any spatial lowpass filtering before sampling. Thus, most sensor network data is aliased with respect to spatial frequency.

The methodology can then be summarized as follows: First, consider the physical process producing the quantity of interest for the sensor network. This leads to a specific spatio-temporal behavior. From there, analyze the possible sampling and aliasing, especially in the spatial dimension. A concrete example is outlined in the next paragraph for the case of distributed audio signals and the plenacoustic function [2]. Other examples of interest where such an analysis can be applied include distributed camera systems, where the plenoptic function [3] plays a key role. This function can be used as an underlying model for distributed image or video acquisition. Interestingly, a sampling theorem for spatial sampling can also be derived in this case [4]. The distribution of temperature, where the heat equation is central, has been considered in [5].

Example 1: **The plenacoustic function.** To make matters specific, we first consider the concrete case of acoustics signals and microphone arrays. The sound field, be it in open space or inside a room, is the solution of a second order partial differential equation called the wave equation. The driving term in the differential equation is given by the various sound sources. The key is thus the kernel of the wave equation, since the source distribution is convolved with the kernel to produce the actual acoustic field. This kernel, also known as the Green function, has a particular form. Its Fourier transform for a particular temporal frequency is essentially bandlimited in spatial frequency. For a concrete example, consider a line in a room, and the spatio-temporal room impulse response $h(x,t)$ with respect to a source. The Fourier transform $H(\omega)$ is essentially supported on a triangle with

$$\phi \leq \omega/c \tag{1}$$

where c is the speed of sound, ϕ and ω are the spatial and temporal frequencies, respectively . Figure 2 shows a sketch of the Fourier transform support, as well as a simulated and a measured spectrum. For details, we refer to [2].

Now we are in a position to address the sampling question. First, it is worth remembering that while the temporal frequency can be limited using lowpass filtering, there is no such possibility over space. That is, spatial sampling cannot be preceded by any spatial filtering. Nonetheless, thanks to the shape of the spectrum, if the maximal temporal frequency is ω_0, then the spatial spectrum is limited to ω_0/c. That is, spatial sampling with a distance between microphones of the order of $d = c/\omega_0$ is adequate to obtain a good representation of the acoustic field. Such a rule of thumb is well known in array signal processing [1]. A precise analysis is given in [2], where the decay of the spectrum and the analysis of the resulting SNR is given. It is to be noted that the discrete spectrum over time and space is not white, and thus residual correlation is present and can be used in distributed compression. One such scheme is analyzed in [6], where it
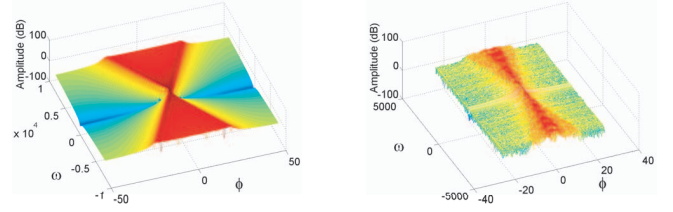
Fig. 2. The Fourier transform of the plenacoustic function, with spatial and temporal frequencies. Simulated (left) and measured (right) plenacoustic function of a room. The triangular shape of the Fourier transform is clearly visible, which leads to a sampling theorem over space when the temporal frequency is limited.

is shown that distributed compression using quincunx sampling achieves the same $D(R)$ as centralized compression. This points to the close interaction of signal structure, spatio-temporal sampling, and distributed compression.

3. BITS

The term *digital* has become so pervasive that it is sometimes assumed to be a given. That is, we go from whatever analog values to some discrete representation. This is what we will assume in the present section. But then, in the next section, we will show that things are not always so simple in general.

In order to discuss digital architectures, we need to define them somewhat more formally. A digital architecture is a two-stage procedure, where each stage is designed independently, the only link between the two stages being *digital rate constraints*. Intuitively, any scheme whose performance can be expressed in terms of a rate-distortion behavior combined with a capacity region will be considered a "digital" architecture. In more detail, this can be described as follows:
1. The *source code* is designed with only the capacity (region) of the channel network available. No further information about the finer structure of the channel can be used.
2. The *channel code* is designed without any knowledge about the source at all. Its goal is to communicate messages in such a way as to avoid errors.

Clearly, any traditional digital communication strategy falls under this category, such as a system where the source is first passed through a vector quantizer, followed by, for example, an entropy coder, and where the resulting bit stream is communicated via an error-correcting code that avoids (block) errors on the noisy channel. On the other hand, a strategy where the channel provides *soft* information, and the source code is designed to work with such soft information, is not considered a digital architecture since the two stages are not truly designed independently of each other. It is clear that such a strategy really constitutes a joint source-channel code. In the next two paragraphs, we discuss digital processing for paradig-

matic digital sensor network scenarios and evaluate the resulting performance. Other digital approaches have been studied for example in [7, 8, 9].

Example 2: **The Camera Sensor Network.** Consider the simplified camera sensor network set-up shown in Figure 3. There are M digital pinhole cameras which are located along a line. We assume that camera locations are known and denote with α the distance between two consecutive cameras. The visual scene which is perpendicular to this line is made of L Lambertian planes. Plane locations are unknown, but the minimum and maximum possible distances of the planes to the line are known and are denoted by z_{min} and z_{max} (z_{max} can be infinity and $z_{min} > 0$). Cameras communicate to a single base station through a classical multi-access Gaussian channel with capacity $C = \frac{1}{2}\log_2(1 + \frac{MP}{\sigma^2})$ where σ^2 is the variance of the noise and P is the power used by each sensor.
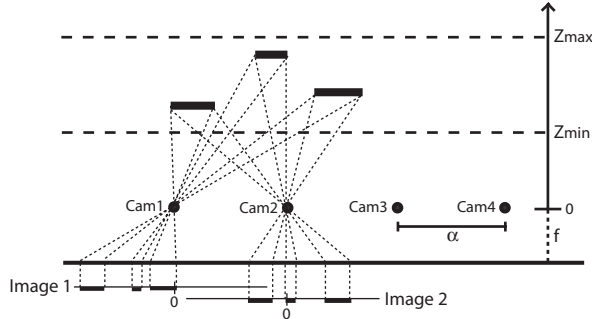

Fig. 3. Our camera sensor network configuration.

Because of the pinhole model, each camera observes a perspective projection of the visual scene. Since the scene is made of Lambertian planes, these projections are piecewise constant functions. The acquisition process at each camera can be modeled as a linear filtering followed by sampling (we assume noiseless measurements for the sake of simplicity). Thus, each camera observes a *blurred* and sampled version of the original piecewise constant projection and it is possible to show that, in many case, exact reconstruction of the original projection from the samples is possible [10]. The reconstruction of the original visual scene is then obtained by back-projecting the reconstructed perspective projections, and is exact when there are at least $M \geq 2L + 1$ cameras observing the visual scene without suffering occlusion. This means that, in this particular context, there exists an exact answer to the sampling problem.

Now assume that no occlusion occurs at any of the M cameras. The perspective projections have been reconstructed and each projection is piecewise constant with L pieces and $2L$ discontinuities. Each projection is therefore specified by $3L$ parameters. The distributed compression is then performed as follows: each sensor quantizes the $3L$ parameters independently and then a Slepian and Wolf (S-W) encoder is used to
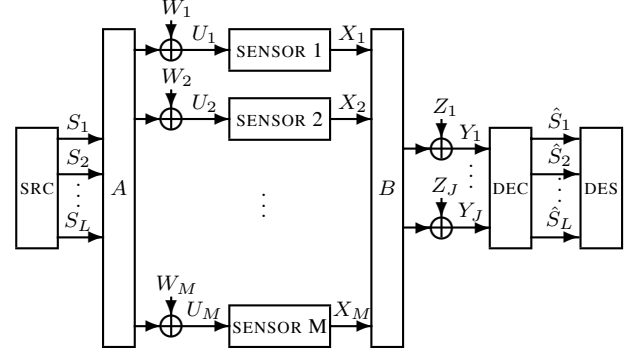

Fig. 4. The "Wireless" Sensor Network example.

remove the remaining redundancy. The interesting element here is that the design of the S-W encoders depends on the properties of the physical phenomenon and, since we are assuming that z_{min}, z_{max} and α are known, the practical implementation of the S-W encoders is almost straightforward [11]. It is then possible to show that , if the total bit budget is R, the distortion-rate behaviour at high rates is given by[1]

$$D(R) \sim c_1 2^{-\frac{2R}{2L(2L+1)+L}}$$

and does not depend on the number of sensors involved [11]. Since we can only transmit $R = C$ bits per channel use, we obtain that the distortion at the base station behaves like

$$D(M, P, L) \sim c_1 \left(\frac{1}{1 + MP/\sigma^2}\right)^{\frac{1}{2L(2L+1)+L}} .$$

Example 3: **A "Wireless" Sensor Network.** A second example is illustrated in Figure 4 and has different structural properties. A vector source (with arbitrary distribution) is observed M-fold through a matrix A and in additive noise, independently by M sensors. The M sensors communicate over an additive noise MIMO channel, characterized by the matrix B, to a base station that houses the central estimation officer. The sensors may have (generally limited) cooperation capabilities. For this sensor network, the source coding can be analyzed in extension of the well-known results on the CEO problem [12, 13]. Recent such extensions to non-Gaussian sources appear in [14]. For the simplest case where there is only one source ($L = 1$) with Gaussian distribution and all the noises in Figure 4 also follow a Gaussian law, the rate-distortion function for the source coding problem is well known, and the total available rate is clearly upper bounded by the capacity of the MIMO channel characterized by the matrix B, and if J is held fixed while M increases, this capacity increases at best logarithmically with M, hence as a function of the total sensor power P_{tot}, we find

---

[1]For scaling law relationships, we use the notation $f(x) \sim g(x)$ if $\lim_{x\to\infty} f(x)/g(x) = c$ for a finite non-zero constant c.

$D_{digital} \sim 1/(\log(MP_{tot}))$. The general case of Figure 4 is discussed in more detail in [15].

4. BEYOND BITS

By contrast to the digital architectures discussed in Section 3, there are ways of "coding" that are not based on the representation of all information in terms of discrete messages (such as bit streams). For the purpose of this exposition, we will refer to any such approach as *analog* architecture. Specifically, it should be noted that analog is not taken to imply *linear* processing nor any other constraint of this form. Rather, *analog* should be defined negatively as *non-digital*, and the point of the paper is to show that some sensor network scenarios strictly require non-digital architectures. Such non-digital architectures are, in certain contexts, also referred to as joint source-channel coding.

As we have argued in Section 3, a set of powerful tools has been developed over the past five decades that facilitate the design of algorithms for handling discrete information, including source codes as well as channel codes. No similarly general tools are known for the design of analog architectures. Rather, these techniques are usually designed on a case-by-case basis, and it is often hard to analyze their performance in a precise fashion. Nevertheless, there is a set of paradigmatic exemplary cases available to date that illustrate the need for the development of a more systematic framework for the design of non-digital communication system architectures. The simplest one is illustrated below, and more general cases can be found in [15].

Example 3, revisited. Consider again the linear additive sensor network model of Example 3. For the same simple case ($L = 1$ and the source and all noises are Gaussian), it is easy to show that under a *fixed* total power P_{tot}, to be shared between all the sensors, a simple direct forwarding strategy permits to recover the underlying source to within a mean-squared error distortion of $D_{analog} \approx \sigma_S^2\sigma_W^2/(M\sigma_S^2 + \sigma_W^2)$. In order to compare this to the digital architectures discussed in Section 3, suppose now that a minimum tolerable distortion D_0 and a power budget P_0 is fixed. How many sensors M_{analog} and $M_{digital}$ do the analog and the digital architectures, respectively, require? By comparing the digital and to the analog results, we find that

$$M_{digital} = e^{M_{analog}}. \qquad (2)$$

That is, the digital architecture will require exponentially more sensors than the analog.

5. SOURCE-CHANNEL SEPARATION: SCALING-LAW PERSPECTIVE

It is well-known that the digital communication paradigm (i.e., the source-channel separation theorem) only exceptionally leads to (exactly) optimal overall network designs. In many cases, the digital communication paradigm leads to *scaling-law optimal* designs, but, as illustrated in this paper, not in all cases. The main goal is therefore a characterization of these cases. Several insights can be gained from studying the examples presented in this paper. First of all, with reference to Figure 4, if the number of degrees of freedom of the underlying process, L, is fixed while the number of sensors increases, separation fails. This is to say, in the regime where the sensor density is increased in order to "zoom in" and get a higher resolution, digital communication will be vastly suboptimal. Conversely, if the number of degrees of freedom is increased at the same rate as the sensors, then in a scaling sense, separation holds. In other words, if more sensors are added to cover more area, then digital communication will only incur a bounded penalty.

6. REFERENCES

[1] D. H. Johnson and D. E. Dugeon, *Array Signal Processing: Concepts and Techniques*, Prentice Hall, New Jersey, 1993.

[2] T. Ajdler, L. Sbaiz, and M. Vetterli, "The plenacoustic function and its sampling," *IEEE Transactions on Signal Processing*, February 2005.

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

[4] J. Chai, X. Tong, S. Chan, and H. Shum, "Plenoptic sampling," in *Proc. Conf. on Computer Graphics*, 2000, pp. 307–318.

[5] B. Beferull-Lozano, Robert L. Konsbruck, and Martin Vetterli, "Rate-distortion problem for physics based distributed sensing,," in *IPSN'04*, Berkeley, CA, April 2004.

[6] R.L.Konsbruck, E.Telatar, and M.Vetterli, "On the multiterminal rate-distortion function for acoustic sensing," in *Proc. IEEE Int. Conf. Acoust. Speech, Sig. Proc.*, Toulouse, France, May 2006.

[7] D. Marco, E. J. Duarte-Melo, M. Liu, and D. L. Neuhoff, "On the many-to-one transport capacity of a dense wireless sensor network and compressibility of its data," in *IPSN'03*, Lect. Notes in Comp. Sci., vol. 2634, Springer, New York, NY, April 2003.

[8] D. Marco and D. Neuhoff, "Reliability vs. efficiency in distributed source coding for field-gathering sensor networks," in *IPSN'04* Berkeley, CA, April 2004.

[9] P. Ishwar, A. Kumar, and K. Ramchandran, "Distributed sampling for dense sensor networks: A bit-conservation principle," in *IPSN'03* Lect. Notes in Comp. Sci., vol. 2634, Springer, New York, NY, April 2003.

[10] M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with finite rate of innovation," *IEEE Trans. Signal Processing*, vol. 50(6), pp. 1417–1428, June 2002.

[11] N. Gehrig and P.L. Dragotti, "DIFFERENT-DIstributed and Fully Flexible image EncodeRs for camEra sensor NeTworks," in *Proc. of IEEE Int. Conf. on Image Processing(ICIP)*, Genova (Italy), September 2005.

[12] T. Berger, Z. Zhang, and H. Viswanathan, "The CEO problem," *IEEE Trans. Info. Theory*, vol. IT–42, pp. 887–902, May 1996.

[13] Y. Oohama, "The rate-distortion function for the quadratic Gaussian CEO problem," *IEEE Transactions on Information Theory*, vol. IT–44, no. 3, pp. 1057–1070, May 1998.

[14] K. Eswaran, "Remote source coding and AWGN CEO problems," M.S. thesis, University of California, Berkeley, CA, 2005.

[15] M. Gastpar and M. Vetterli, "Power, spatio-temporal bandwidth, and distortion in large sensor networks," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 4, pp. 745–754, April 2005.