ANALYSIS OF NOISE REDUCTION IN REDUNDANT EXPANSIONS UNDER DISTRIBUTED PROCESSING REQUIREMENTS

Christopher J. Rozell and Don H. Johnson

Rice University
Department of Electrical and Computer Engineering
Houston, TX 77251-1892

ABSTRACT

We considered signal reconstruction with redundant expansions under distributed processing in noisy environments. Redundant expansions have the ability to reduce noise corrupting the coefficients, but distributed processing schemes will not be able to take full advantage of the redundancy present. We apply frame theory and a generalization called "frames of subspaces" to find conditions when distributed reconstruction suffers no loss in noise reduction ability, and we bound performance loss in more general cases.

1. INTRODUCTION

If a signal is represented as an expansion of redundant vectors with coefficients corrupted by noise, the redundancy can be exploited during reconstruction to reduce the error in the estimate of the original signal. However, making optimal use of the inherent noise reduction properties requires a centralized processing approach where all vectors and coefficients must be known at the same place. There are important and interesting examples of systems that have redundant total representations but are limited to distributed processing environments. One example is a sensor network, where sensor proximity can produce redundant observations but constraints prohibit centralized processing. Sensorineural systems also use redundant representations and process data in distributed ways. While motivated by such practical applications, we limit this paper to a general mathematical analysis of redundant representations.

We show here the potential impact one type of distributed processing has on the noise reduction properties of redundant expansions. We find conditions under which there is no loss in noise reduction ability, and we bound the performance loss in more general cases. We use frame theory [3] to describe this notion of a redundant expansion. To perform our analysis, we employ a recent generalization of frames called "frames of subspaces" [1], that links the properties of a global frame (centralized structures) and local collections of frame elements (distributed structures).

Email: {*crozell,dhj*}@*rice.edu*

This work supported by the National Science Foundation and a fellowship from Texas Instruments.

2. FRAME THEORY

2.1. Basic definitions

A collection of vectors $\{\phi_m\}_{m=1}^{M}$ is called a *frame* [3] for the N-dimensional Hilbert space $\mathcal{H}$ if it yields a stable and complete representation for every signal $s \in \mathcal{H}$. This condition is equivalently expressed as an energy preservation condition resembling the well-known Parseval's theorem,

$$A\|s\|^2 \le \sum_m |\langle s, \phi_m\rangle|^2 \le B\|s\|^2,$$

where $0 < A \le B < \infty$ are constants called *frame bounds*. Unlike a basis, a frame can have elements that are linearly dependent. When the frame vectors are normalized $\|\phi_m\|^2 = 1$ (which we will always assume here), the frame bounds measure the minimum and maximum redundancy. A frame is said to be *tight* when $A = B$, and it is an orthonormal basis if and only if $A = B = 1$. Highly redundant frames will therefore have $A, B \gg 1$.

With linearly dependent elements, there are an infinite number of ways that a signal s can be represented in terms of linear combinations of frame vectors. However, a signal is normally decomposed by means of the *analysis operator*, $F : \mathcal{H} \to l^2$, $Fs = \{\langle s, \phi_m\rangle\} = \{c_m\}$, which calculates the inner product of the vector with each frame element to produce a scalar coefficient. The adjoint of the analysis operator is known as the *synthesis operator*, $F^* : l^2 \to \mathcal{H}$, $F^*\{c_m\} = \sum_m c_m\phi_m$. The composition of F^* and F is known as the *frame operator*, $G : \mathcal{H} \to \mathcal{H}$, $Gs = F^*Fs = \sum_m \langle s, \phi_m\rangle\phi_m$. From the definition of a frame, the analysis operator is bounded for all $\|s\| = 1$: $\sqrt{A} \le \|Fs\| \le \sqrt{B}$. The frame operator is therefore also bounded and invertible, with bounds $A \le \|Gs\| \le B$ and $\frac{1}{B} \le \|G^{-1}s\| \le \frac{1}{A}$ for all $\|s\| = 1$. The frame bounds are $A = \lambda_{\min}$ and $B = \lambda_{\max}$, where $\{\lambda_n\}_{n=1}^{N}$ are the eigenvalues of G.

In an orthonormal basis, the frame operator would perform perfect signal reconstruction (i.e., $G = I$). Note that a consequence of this fact is that reconstruction can be performed in a decentralized way: each coefficient only contributes to the vector that generated it. However, because frame elements are linearly dependent, the synthesis operator must be centralized. In fact, because of the inherent redundancy, there are an infinite number inverse operators for F. The most commonly used choice of inverse to perform signal reconstruction is given by the pseudoinverse $\widetilde{F}^{-1} = G^{-1}F^*$, which depends explicitly on *all* of the vectors in the frame. In detail, the reconstruction operator is

$$s = \widetilde{F}^{-1}(Fs) = \sum\nolimits_m \langle s, \phi_m\rangle \widetilde{\phi}_m = \sum\nolimits_m \langle s, \widetilde{\phi}_m\rangle \phi_m,$$

where $\widetilde{\phi}_m = G^{-1}\phi_m$ are known as the *dual frame vectors*. Notice that in the general case, the multiplication by G^{-1} means that each dual vector $\widetilde{\phi}_m$ depends on the whole collection of analysis vectors $\{\phi_m\}_{m=1}^M$. Any change to the set of frame vectors (modification of a vector, or addition/deletion of a vector) affects the whole collection of dual vectors needed for the reconstruction. In a tight frame, the dual vectors are scaled versions of the frame vectors, $\widetilde{\phi}_m = \frac{1}{A}\phi_m$.

2.2. Noise reduction in a frame

It is well known that the pseudoinverse operator $\widetilde{F}^{-1}$ is optimal among all linear operators at reducing noise in the reconstructed signal [3]. The amount of noise power eliminated is related to the amount of redundancy in the elements, which is reflected in the frame bounds [2].

Let $\{\phi_m\}_{m=1}^M$ be a frame with bounds (A, B) for an N-dimensional input space $\mathcal{H}$. Reconstruction coefficients are corrupted by uncorrelated noise $w_c(m)$ having variance σ^2,

$$\hat{s} = \sum_{m=1}^{M} \left[\langle s, \phi_m\rangle + w_c(m)\right] \widetilde{\phi}_m.$$

The total MSE of the reconstructed vector $\hat{s}$ is given by

$$\mathcal{E}\left[\|\hat{s} - s\|^2\right] = \sigma^2 \sum_{m=1}^{M} \|\widetilde{\phi}_m\|^2 = \sigma^2 \sum_{n=1}^{N} \frac{1}{\lambda_n}, \quad (1)$$

where $\{\lambda_n\}_{n=1}^N$ are the eigenvalues of the frame operator G. The MSE per signal dimension is bounded by

$$\frac{\sigma^2}{B} \le \frac{\mathcal{E}\left[\|\hat{s} - s\|^2\right]}{N} \le \frac{\sigma^2}{A}. \quad (2)$$

Notice that in reconstructing with the redundant expansion, the noise is reduced by an amount at least proportional to the minimum redundancy. From the derivation of equation (2) from equation (1), it is clear that the bounds in (2) are not tight bounds. The MSE will only achieve the bounds in (2) in the case of a tight frame when the bounds are equal.

3. "FRAME OF SUBSPACES" THEORY

To begin analyzing a reconstruction from redundant elements in a distributed way, we turn to a new theory of "frames of subspaces" introduced in [1]. Here, signals are decomposed in terms of overlapping subspaces. A family of closed subspaces $\{W_i\}_{i=1}^L$ is a *frame of subspaces* for $\mathcal{H}$ if for every signal $s \in \mathcal{H}$,

$$C\|s\|^2 \le \sum\nolimits_i \|\pi_i(s)\|^2 \le D\|s\|^2, \quad (3)$$

where $0 < C \le D < \infty$ are the frame bounds and $\pi_i(\cdot)$ is the orthogonal projection onto W_i. From (3) it is clear that the frame bounds are themselves bounded by the number of subspaces, $L \ge D$. A frame of subspaces is in many ways analogous to a frame. In frame theory, an input signal is represented by a collection of scalar coefficients that measure the projection of that signal onto each frame vector. In a frame of subspaces, an input signal is represented by a collection of *vector* coefficients that represent the projection (not just the projection energy) onto the each subspace. Formally, for a frame of subspaces, the representation space is defined as $\mathcal{V} = \left\{\{s_i\} | s_i \in W_i \text{ and } s_i \in l^2\right\}$, the space of all collections of finite-energy vectors containing one representative from each subspace.

Analogous to frame theory, a frame of subspaces has an analysis operator, $T : \mathcal{H} \to \mathcal{V}$, $Ts = \{\pi_i(s)\} = \{s_i\}$. The adjoint of the analysis operator is the synthesis operator, $T^* : \mathcal{V} \to \mathcal{H}$, $T^*\{s_i\} = \sum_i s_i$. The composition of the analysis and synthesis operators form the frame operator, $U : \mathcal{H} \to \mathcal{H}$, $Us = T^*Ts = \sum_i \pi_i(s)$. The frame operator is bounded and invertible, with bounds $C \le \|Us\| \le D$ and $\frac{1}{D} \le \|U^{-1}s\| \le \frac{1}{C}$ for all $\|s\|$ =1. As in section 2.1, the unique pseudoinverse for T is given by $\widetilde{T}^{-1} = U^{-1}T^*$.

Frames of subspaces are an interesting mathematical structure on their own, but one of the fundamental results of [1] shows that they also provide a link between the properties of a global frame and local collections of those frame elements. If each subspace W_i has an associated family of vectors that locally form a frame for W_i with bounds (A_i, B_i), then the total collection of vectors globally form a frame for $\mathcal{H}$. It can be shown that the frame bounds (A, B) for this global frame are bounded by $CA_{\min} \le A$ and $B \le DB_{\max}$, with equality when the same frame bounds apply to each local frame.

Because of the link between frames for a local subspace and the global frame of elements taken together, we can use frames of subspaces to represent distributed processing. In this model, the frame vectors local to one subspace are used to reconstruct the orthogonal projection of the signal into that subspace while maximizing the noise reduction properties of the local frame. The collection of reconstructed signals within each subspace are then used to reconstruct the original input signal while maximizing the noise reduction properties of the frame of subspaces. This subspace-based reconstruction distributes the processing by only requiring knowledge of the frame vectors within a subspace to do a local reconstruction. Such a scheme adds a level of robustness to the system because only vectors in one local subspace are affected if a frame vector is added or removed.

3.1. Subspace noise reduction

In section 2.2 we investigated the noise reduction properties inherent in the redundancy between frame vectors. To evaluate a distributed reconstruction scheme where redundant subspace representations are combined, we must determine the noise reduction properties inherent in a frame of subspaces. The redundancy between the subspaces means that (just as in section 2.2) we can use the pseudoinverse to perform reconstruction and reduce the presence of additive noise in the decomposition.

Let $\{W_i\}_{i=1}^L$ be a frame of subspaces for the space $\mathcal{H}$ with bounds (C, D). The "coefficients" in this decomposition are the collection of vectors $\{s_i\} \in \mathcal{V}$, where each vector is the projection of the input signal s onto a subspace, $s_i = \pi_i(s)$. Consider the case when the vector coefficients are corrupted independently with a noise vector,

$$\hat{s} = U^{-1} \sum_i (s_i + w_i),$$

where the vector $w_i \in W_i$ has covariance matrix Γ_i. The linearity of $\widetilde{T}^{-1}$ implies that $(\hat{s} - s) = U^{-1} \sum_i w_i$. Define $\widetilde{w} = \sum_i w_i$, so that $\widetilde{w}$ has covariance matrix $\widetilde{\Gamma} = \sum_i \Gamma_i$. Because multiplication affects the covariance in a quadratic way, the covariance of $U^{-1}\widetilde{w}$ is given by $U^{-1}\widetilde{\Gamma}U^{-1}$. The total MSE of the reconstructed signal therefore equals the trace of the covariance,

$$\mathcal{E}\left[\|\hat{s} - s\|^2\right] = \mathrm{Tr}\left[U^{-1}\widetilde{\Gamma}U^{-1}\right]. \tag{4}$$

As mentioned earlier, U^{-1} is a bounded operator with bounds $\left(\frac{1}{D}, \frac{1}{C}\right)$. Simple calculations using this fact give bounds on the MSE per signal dimension,

$$\frac{\mathrm{Tr}\left[\widetilde{\Gamma}\right]}{ND^2} \leq \frac{\mathcal{E}\left[\|\hat{s} - s\|^2\right]}{N} \leq \frac{\mathrm{Tr}\left[\widetilde{\Gamma}\right]}{NC^2}. \tag{5}$$

Notice that the natural reconstruction for a frame of subspaces also reduces the noise in the reconstructed signal by an amount that depends on the minimum redundancy. As before, the bounds in (5) are not tight bounds. The MSE will only achieve the extreme bounds in (5) when the frame of subspaces is tight ($C = D$).

3.2. Noise reduction under distributed processing

We now have the tools available to consider the noise reduction capability of a redundant expansion under distributed processing requirements. Let $\{W_i\}_{i=1}^L$ be a frame of subspaces for the space $\mathcal{H}$, with frame bounds (C, D). Let each subspace be spanned by a collection of M_i vectors that locally form a frame for W_i with frame bounds (A_i, B_i). When taken together, these vectors form a frame for $\mathcal{H}$ with bounds (A, B). A signal s is represented by all of the vectors in the global frame, and coefficients are corrupted by additive noise with subspace-dependent variance σ_i^2.

For distributed reconstruction, local frame vectors are first used to reconstruct the projection of the signal onto each subspace, $\hat{s}_i$. From equation (2), we have a bound on the total MSE when reconstructing each $s_i = \pi_i(s)$, $\frac{\sigma_i^2 N}{B_i} \leq \mathcal{E}\left[\|\hat{s}_i - s_i\|^2\right] \leq \frac{\sigma_i^2 N}{A_i}$. Distributed reconstruction $\hat{s}_d$ is completed by using each $\hat{s}_i$ as a corrupted coefficient in the frame of subspaces. The previous bound tells us that we can view the subspace projections s_i as being corrupted independently by an additive noise vector $w_i \in W_i$, with covariance matrix Γ_i that has bounded total variance $\frac{\sigma_i^2 N}{B_i} \leq \mathrm{Tr}\left[\Gamma_i\right] \leq \frac{\sigma_i^2 N}{A_i}$. If we define $\widetilde{\Gamma} = \sum_i \Gamma_i$, applying equation (5) bounds the MSE per signal dimension,

$$\begin{aligned} \frac{L\sigma_{\min}^2}{B_{\max}D^2} \leq \frac{1}{D^2}\sum_i \frac{\sigma_i^2}{B_i} \leq \frac{\mathcal{E}\left[\|\hat{s}_d - s\|^2\right]}{N} \\ \cdots \leq \frac{1}{C^2}\sum_i \frac{\sigma_i^2}{A_i} \leq \frac{L\sigma_{\max}^2}{A_{\min}C^2}. \end{aligned} \tag{6}$$

Frame reconstruction with our distributed processing constraint has the power to reduce noise in the reconstructed signal by an amount that depends on the minimum redundancy of both the frame of subspaces, and the individual local frames that span the subspaces. From the derivation of this bound, it is clear that these bounds are also not tight and are only achieved in the special case when the frame of subspaces is tight (C =D) and the local frames are all tight with the same bounds (A_i =$A_{\min}$ =B_i =$B_{\max}$ for all i).

4. COMPARISON OF CENTRALIZED AND DISTRIBUTED PROCESSING

4.1. Comparison of bounds

Section 2.2 mentions that the pseudoinverse operator for the global frame is the unique linear operator that achieves maximum noise reduction. Consequently, we know that no other linear reconstruction scheme (including the distributed processing of section 3.2) can perform better than centralized processing. In fact, the distributed reconstruction operator is optimal *only* over linear operators that factor (using local frame and synthesis operators G_i and F_i^*) into

$$\widetilde{T}^{-1} = \left[\begin{array}{c} U^{-1} \ldots U^{-1} \end{array}\right] \left[\begin{array}{c} G_1^{-1}F_1^* \\ \vdots \\ G_L^{-1}F_L^* \end{array}\right].$$

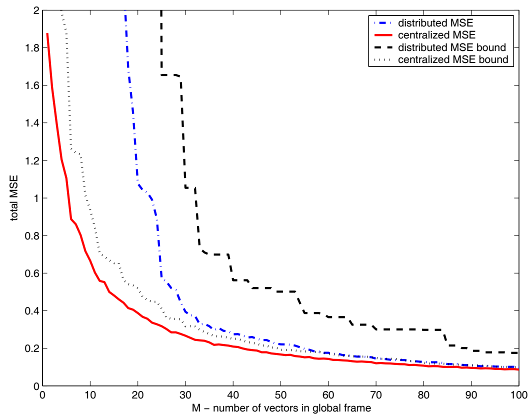

Fig. 1. An example of the noise reduction properties for in centralized and distributed reconstruction. Vectors are randomly placed in $\mathbb{R}^3$ and separated into five subspaces.

It is interesting to consider how much of a penalty in noise reduction a distributed processing scheme incurs compared to centralized processing. From section 2.2, considering the collection of local frames together yields a global frame for $\mathcal{H}$ with lower and upper bounds $A \geq CA_{\min}$ and $B \leq DB_{\max}$. Using a simple extension of equation (2) (with unequal noise power) tells us that a centralized reconstruction of s using the global frame directly, $\hat{s}_c$, would yield an upper bound on the MSE per signal dimension of $\frac{\mathcal{E}\left[\|\hat{s}_c - s\|^2\right]}{N} \leq \frac{\sigma^2_{\max}}{A} \leq \frac{\sigma^2_{\max}}{CA_{\min}}$. Comparing this bound to the MSE upper bound in (6) when using the distributed scheme, we see that (because $L \geq C$) the upper bound using the centralized approach is better than the distributed reconstruction by a factor of $\frac{L}{C}$. While these are not tight bounds on the error, they hint at the potential for the distributed reconstruction to perform worse than the centralized scheme and give a bound on the performance reduction.

It is also interesting to consider conditions under which the noise reduction ability is the same for distributed and centralized processing. Consider the case when all of the local frames have the same number of vectors $\left(M_i = \frac{M}{L}\right)$ and are tight frames with the same frame bounds, $A_i = A_{\min} = B_i = B_{\max}$ for all i. In this case, the global frame has frame bounds $A = CA_i$ and $B = DB_i$. If the frame of subspaces is also tight ($C = D$), the global frame will additionally be tight with bounds $A = B = \frac{M}{N}$. If the noise has equal power in each subspace ($\sigma_i^2 = \sigma^2$), then we can directly calculate the MSE per signal dimension under both processing schemes and find that they are equal, $\frac{\mathcal{E}\left[\|\hat{s}_c - s\|^2\right]}{N} = \frac{\mathcal{E}\left[\|\hat{s}_d - s\|^2\right]}{N} = \frac{\sigma^2 N}{M}$. Though the conditions proposed here for equal noise reduction may seem restrictive, a result from [2] regarding random frames indicates that frames will become tight asymptotically as more random vectors are added. Therefore, systems where random vectors are randomly assigned to a local subspace will asymptotically meet the conditions for achieving the optimal (centralized) noise reduction.

4.2. Example

Figure 1 compares the noise reduction properties of a centralized and distributed reconstruction in $\mathcal{H} = \mathbb{R}^3$. Frame vectors were generated at random (uniformly distributed in $\mathbb{R}^3$) and assigned to one of five possible subspaces. Centralized and distributed reconstructions using noisy coefficients ($\sigma_i^2 = 1$) are performed and the average reconstruction error is plotted. The centralized scheme always outperforms the decentralized scheme, though the performance tends to become similar as more vectors are added and the frames become tighter.

5. CONCLUSIONS

In redundant expansions where distributed reconstruction is required, assessing the noise reduction capability is an important consideration. It would be intuitive to think that distributed reconstruction would always incur a penalty over the optimal centralized case. However, in this paper we have shown at least one condition (there may be more) where the distributed system can be constructed to suffer no penalty compared to centralized reconstruction. Additionally, we provide bounds on the performance penalty in the more general case when distributed reconstruction is suboptimal.

The notion of a frame of subspaces has played an important role in determining the performance difference between our two cases, and this mathematical formalism may prove useful in future studies of distributed systems. However, our analysis is very abstract and it is not immediately clear how to apply these results to practical situations. While there are obvious connections between frames and filterbanks, we are currently working to understand more about how this analysis can apply to problems in distributed sensing, feature extraction and sensory neuroscience.

6. REFERENCES

[1] P. Casazza and G. Kutyniok. Frames of subspaces. In *Wavelets, Frames and Operator Theory*, volume 345, pages 87–113. American Mathematical Society, 2004.

[2] V. Goyal, J. Kovačević, and J. Kelner. Quantized frame expansions with erasures. *Applied and Computational Harmonic Analysis*, 10:203–233, 2001.

[3] S. Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, New York, NY, second edition, 1999.