Contents

0.1 Time-Frequency Compressive Sensing
(E. Sejdić) 1
0.1.1 Compressive Sensing 1
0.1.2 Compressive Sensing and Time-Frequency Analysis 2
0.1.2.1 Compressive Sensing Based on the Ambiguity Domain 2
0.1.2.2 Compressive Sensing Based on Time-Frequency Dictionary and Matching Pursuit 3
0.1.3 Illustrating Examples 4
0.1.4 Summary and Conclusions 7

0.1 TIME-FREQUENCY COMPRESSIVE SENSING [0]

Traditional sampling approaches rely on the Shannon-Nyquist sampling theorem which mandates that a signal needs to be sampled at least twice the highest frequency present in the signal in order to reconstruct it accurately. This approach often results in a large number of samples, and compression schemes are used shortly after sampling to reduce the number of data points that need to be transmitted and/or stored. This is clearly a redundant approach, as most samples are disregarded soon after acquiring them. In recent years, a new sensing approach called compressive sensing (CS) has been proposed to avoid the aforementioned issues. CS postulates that a signal can be reconstructed using a fewer number of randomly chosen signal samples (e.g., [1], [2], [3]).

0.1.1 Compressive Sensing

The idea behind compressive sensing is to diminish the number of steps during a data acquisition process by combining sampling and compression into a single step [2], [3]. Specifically, compressive sensing approaches acquire signals at sub-Nyquist rates and tend to recover such signals accurately from sparse samples [2]. Compressive sensing approaches are well suited for K-sparse signals, i.e., signals that can be represented by significant K coefficients over an N-dimensional basis, as the compressive sensing approaches will encode a K-sparse, discrete-time signal of dimension N by computing a measurement vector y that consists of $M << N$ linear projections of the vector s [4], [5]:

$$y = \Phi s \tag{0.1.1}$$

where Φ represents an $M \times N$ sensing matrix [3]. To recover s from sparse samples, various contributions suggested to utilize a norm minimization framework:

$$\min ||s||_0 \text{ subject to } \|y - \Phi s\|_2 < \eta \tag{0.1.2}$$

where η is the expected noise of measurements, $||s||_0$ counts the number of nonzero entries of s and $\|\bullet\|_2$ is the Euclidian norm. Eqns. (0.1.1) and (0.1.2) will not always yield an accurate representation of sparse signals, and in some applications, such as medical diagnostic applications, it is desired to achieve a "precise" recovery of such signals. In other words, it is desirable to obtain a very small error, which can be accomplished by compressively sampling signals in a different domain. To denote the change of a domain, we can rewrite eqn. (0.1.1) as [6]:

$$y = \Phi s = \Phi \Psi x \tag{0.1.3}$$

where $s = \Psi x$ represents a sparse representation of a signal in a domain given by Ψ and x represents expansion coefficients.

[0] Author: **Ervin Sejdić**, Department of Electrical and Computer Engineering, University of Pittsburgh, Pittsburgh, PA 15261, USA (esejdic@ieee.org). Reviewers: Igor Djurović and Aydin Akan.

An important aspect of compressive sensing is to understand how the sparsity in the transform domain regulates the number of measurements needed to reconstruct a signal. This is assessed using the so-called coherence measure between the matrices Φ and Ψ:

$$\mu(\Phi, \Psi) = \sqrt{N} \max |\langle \phi_k, \psi_j \rangle| \tag{0.1.4}$$

where N is the signal length, ϕ_k is the k^{th} row of Φ, and ψ_j is the j^{th} row of Ψ. A simple rule is that a lower coherence value indicates a smaller number of random measurements needed to reconstruct a signal.

0.1.2 Compressive Sensing and Time-Frequency Analysis

Nonstationary signals typically do not have a sparse representation either in time or frequency domains, but may have a sparse representation in the time-frequency (t, f) domain (please see the article 14.7). A typical example of a sparse signal in the (t, f) domain is a linear frequency modulated signal. If a (t, f) representation of such signal is obtained using the Wigner-Ville distribution, then the signal is concentrated along a linear line in the (t, f) domain, and most of other points are equal to zero. Hence, it would be advantageous to compressively sample such a signal in the (t, f) domain.

Compressed sensing of nonstationary signals can be accomplished using two different approaches. One approach relies on the idea that a signal is sampled in the ambiguity domain [7]. This approach is limited to quadratic (t, f) representations. The second approach is to utilize (t, f) dictionaries in order to find a sparse signal representation.

0.1.2.1 Compressive Sensing Based on the Ambiguity Domain

Let us consider the Wigner-Ville distribution, as one of the commonly used (t, f) distributions. One approach for obtaining these representations of compressively sampled signals is to use the so-called ambiguity domain. Let's consider the Wigner-Ville distribution, $WVD(t, f)$, and its two-dimensional Fourier transform, i.e., the ambiguity function (please see chapter 3) as follows [8]:

$$A_x(\nu, \tau) = \mathcal{F}_{2D}\{WVD(t, f)\} \tag{0.1.5}$$

where $\mathcal{F}_{2D}$ is the two-dimensional Fourier operator (forward and inverse). The advantage of using the ambiguity domain is due to the fact that the cross-terms are usually dislocated from the origin in the ambiguity domain and can be suppressed, or significantly attenuated, by the use of low-pass filtering. This is achieved by applying the kernel $g(\nu, \theta)$:

$$\mathcal{A}_x(\nu, \tau) = A_x(\nu, \tau) g(\nu, \tau) \tag{0.1.6}$$

Improved (t, f) signal power localization can be achieved by using the compressed sensing approach and exploiting sparsity in the (t, f) domain. Namely, we can collect a set of samples from the ambiguity domain and solve the l_1-norm minimization problem to obtain the sparsest (t, f) distribution. By using the compressed sensing approach, the desired (t, f) distribution $\widehat{\rho}_x(t, f)$ can be obtained as:

$$\widehat{\rho}_x(t, f) = \arg \min_{\rho_x(t,f)} \|\rho_x(t, f)\|_1 \tag{0.1.7}$$

$$F_{2D}^{-1}\{\rho_x(t, f)\} - \mathcal{A}_x^M = 0\big|_{(\nu,\tau)\in\Omega} \tag{0.1.8}$$

where $\mathcal{A}_x^M$ denotes the set of samples from the ambiguity domain in the region defined by the mask $(\nu, \tau) \in \Omega$, $\rho_x(t, f)$ denotes the (t, f) distribution, and $\| \cdot \|_1$ denotes the ℓ_1 norm. In the presence of noise, we may use the approximation:

$$\widehat{\rho}_x(t, f) = \arg \min_{\rho_x(t,f)} \|\rho_x(t, f)\|_1 \tag{0.1.9}$$

$$\|F_{2D}^{-1}\{\rho_x(t, f)\} - \mathcal{A}_x^M\|_2 \leq \epsilon\big|_{(\nu,\tau)\in\Omega} \tag{0.1.10}$$

Here, it is important to select a suitable set of ambiguity domain samples which is done by an appropriate ambiguity function masking. The mask can be formed as a small area around the origin of the ambiguity plane, as done for designing high-resolution (t, f) distributions [9]. In other words, the mask is designed to pass any auto-terms, while reducing the cross-terms in the ambiguity domain.

0.1.2.2 Compressive Sensing Based on Time-Frequency Dictionary and Matching Pursuit

The first step in this approach is to construct a (t, f) dictionary. Using the (t, f) dictionary, the next challenge is to obtain a sparse signal representation via eqns. (0.1.2) and (0.1.3). However, this approach can be extremely computationally expensive, which is not suitable for many applications. Therefore, matching pursuit [10] or its recent variants (e.g., [11]) can be utilized in order to avoid computational burdens associated with traditional compressive sensing approaches.

A matching pursuit approach begins with an initial approximation of the signal, $\widehat{x}^{(0)}(m) = 0$, and the residual, $R^{(0)}(m) = x(m)$, where m represent the M time indices that are uniformly or non-uniformly distributed. Then, at each subsequent stage, k, the matching pursuit identifies a dictionary atom with the strongest contribution to the residual and adds it to the current approximation:

$$\widehat{x}^{(k)}(m) = \widehat{x}^{(k-1)}(m) + \alpha_k \phi_k(m) \tag{0.1.11}$$

$$R^{(k)}(m) = x(m) - \widehat{x}^{(k)}(m) \tag{0.1.12}$$

where $\alpha_k = \langle R^{(k-1)}(m), \phi_k(m)\rangle / \|\phi_k(m)\|^2$. The process continues till the norm of the residual $R^{(k)}(m)$ does not exceed required margin of error $\varepsilon > 0$: $||R^{(k)}(m)|| \leq \varepsilon$ [10], or a number of bases, $\mathfrak{n}_{\mathfrak{B}}$, needed for signal approximation should satisfy $\mathfrak{n}_{\mathfrak{B}} \leq \mathcal{K}$.

Regardless of a stopping criterion, a signal approximation is obtained using L bases as

$$x(n) = \sum_{l=1}^{L} \langle x(m), \phi_l(m) \rangle \phi_l(n) + R^{(L)}(n) \tag{0.1.13}$$

where ϕ_l are L bases from the dictionary with the strongest contributions.

The advantage of the proposed decomposition is that other signal operations can be easily carried out. Specifically, we can obtain a (t, f) representation of a signal using its L-bases approximation:

$$\mathcal{TF}\{x(n)\} = \sum_{l=1}^{L} \langle x(m), \phi_l(m) \rangle \mathcal{TF}\{\phi_l(n)\} \tag{0.1.14}$$

where $\mathcal{TF}\{\}$ is a (t, f) operator (e.g., short-time Fourier transform) [8].

Note that these steps require the knowledge of sampling times in order to acquire proper values of members of the (t, f) dictionary. In real-life conditions, we generally do not always have it available and there may exist a need to estimate the sampling time instances. To accomplish this task, the use of annihilating filters has been proposed in several contributions [1], [12], [13]. The annihilating filters approach relies on determining the roots of an autoregressive filter in order to estimate the sampling instances.

0.1.3 Illustrating Examples

For example, let us consider a sinusoidally-modulated signal contaminated with white Gaussian noise. In order to provide faster computations, we use the (t, f) representations of size 60x60 (3600 points). The mask is of size 7x7 (1.4% of the total number of points) in the ambiguity domain. The original (t, f) distribution and ambiguity function are shown in Figure 0.1.1(a) and (b), respectively. The resulting sparse representation is illustrated in Figure 0.1.1(c). Note that the compressed sensing approach provides improved results, reducing the noise influence significantly. The number of non-zero points in the sparse (t, f) representation is approximately between 45 and 50 (estimated from different experiments), which is a quite small percentage of the total number of points in the (t, f) domain.

The second case, depicted in Figure 0.1.2(a), involves a signal consisting of four basis functions from a 25-band dictionary based on modulated discrete prolate spheroidal functions (MDPSS) (e.g., [14], [15]) with the normalized half-bandwidth equal to $W = 0.495$ and $N = 256$. For both uniform and non-uniform sampling, only 42 samples were needed to accurately recover the signal (less than 17% of the total number of samples) and its spectrograms based on regular and irregular sample times as shown in 0.1.2(c) and (d). A greater percentage of samples was required for this case in comparison to the first case as the second case is recovered almost exactly.

Thirdly, we examine the accuracy of the instantaneous frequency (IF) estimator based on the complete (t, f) representations and those representations obtained from

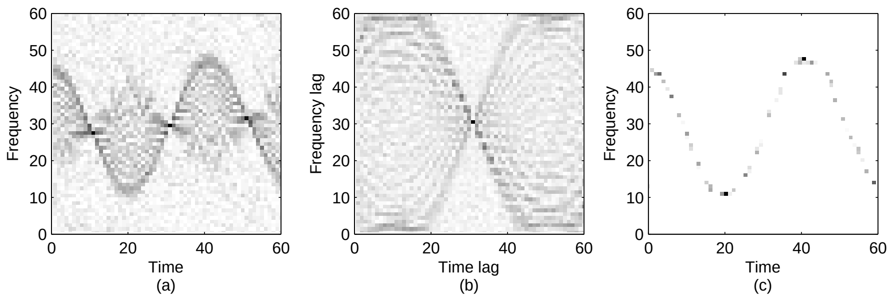

Fig. 0.1.1: (T, f) representations of a sinusoidally-modulated signal using: (a) the Wigner distributions, (b) the ambiguity function, (c) the resulting sparse (t, f) representation.

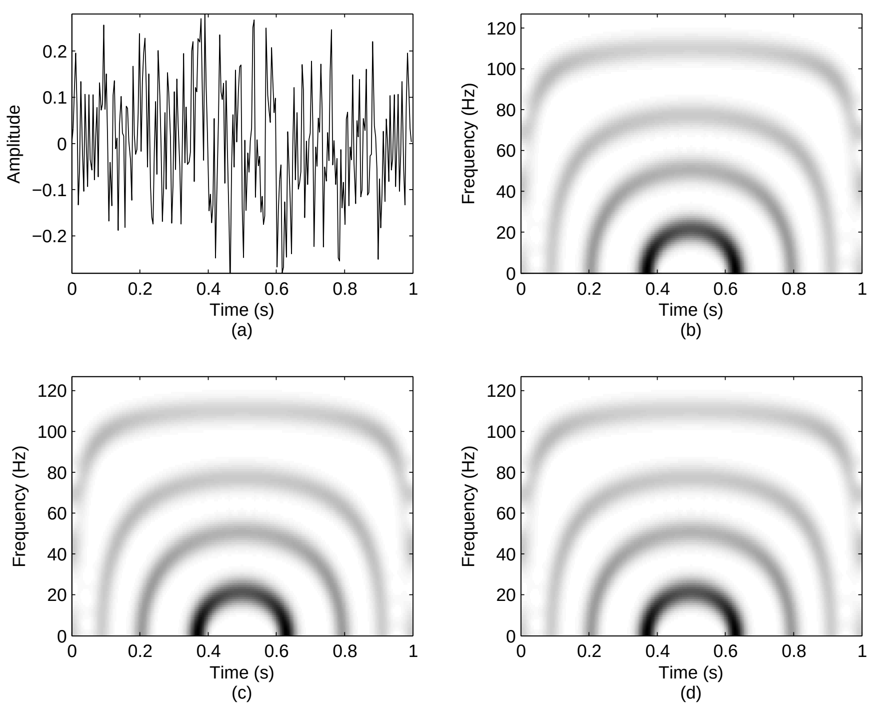

Fig. 0.1.2: The time domain representation of the consider signal is shown in (a). Spectrograms of: (b) the original signal; (c) the signal based on equal distance samples; (d) the signal based on irregular samples.

compressed samples. As a sample signal, we considered a sinusoidally-modulated signal defined as:

$$x(t) = \sin(120\pi t + 2\pi \cos(8\pi t)) \tag{0.1.15}$$

where $0 < t < 1$ and the assumed sampling rate is $T_s = 1/256$ seconds. Figures

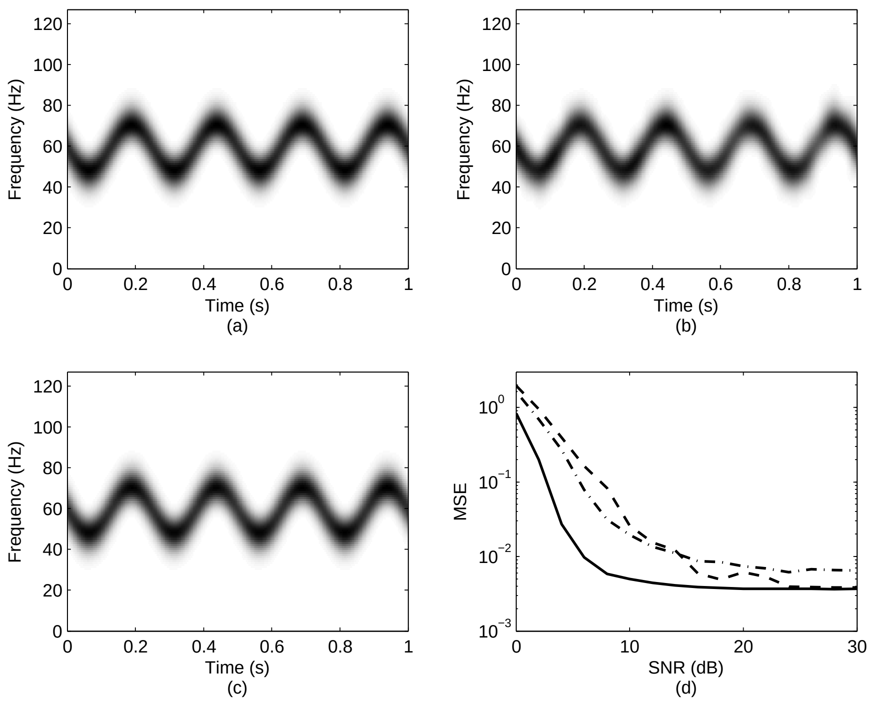

Fig. 0.1.3: Spectrograms of: (a) the original signal; (b) the signal based on equal distance samples; (c) the signal based on irregular samples. The performance of the IF estimator is shown in (d) based on the original signal (solid line), equal-distance compressed samples (dashed line) and nonuniform compressed samples (dash-dotted line).

0.1.3(a)-(c) show sample results when 75% of samples were used for compressive sensing. Next, the signal was contaminated with an additive white Gaussian noise, and its variance was proportional to the considered signal-to-noise ratios (SNR). The accuracy of the estimator was then examined for a number of SNR values as shown in Figure 0.1.3(d), and the presented values were obtained using 500 realizations. As expected, the complete (t, f) representation obtained the lowest values of mean square errors, but (t, f) representations based on compressive samples closely follow the trend of the complete (t, f) representation, especially the (t, f) representation based on uniformly compressed samples. These results clearly demonstrate that even based on compressive samples, we can achieve a reliable estimate of IF. Further studies are needed to understand the effects of the number of acquired samples on mean square errors.

Lastly, the accuracy of compressive sensing schemes based on (t, f) dictionaries is examined using swallowing accelerometry signals, which can be used to infer if a patient has swallowing difficulties. Here, vibrations in three anatomical directions associated with a sample swallow obtained during a videoflouroscopy exam are

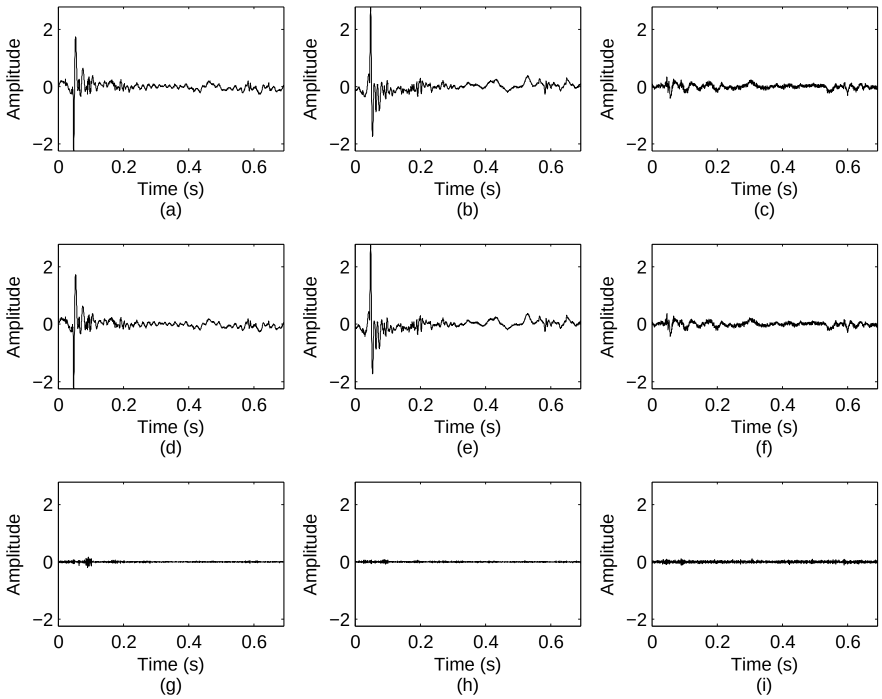

Fig. 0.1.4: Sample swallowing accelerometry signals from a patient with swallowing difficulties: (a) the original signal in the anterior-posterior direction; (b) the original signal in the superior-inferior direction; (c) the original signal in the medio-lateral direction; (d) the recovered signal in the anterior-posterior direction; (e) the recovered signal in the superior-inferior direction; (f) the recovered signal in the medio-lateral direction; (g) the error between the original and the recovered signal in the anterior-posterior direction; (h) the error between the original and the recovered signal in the superior-inferior direction; (i) the error between the original and the recovered signal in the medio-lateral direction.

considered. A MDPSS-based dictionary was used with $W = 0.25$ and $N = 128$. In this sample case presented shown in Figure 0.1.4, we considered the error when 33% of samples were available. These sample results that the considered biomedical signals can be recovered almost exactly with a small number of samples. This is due to the fact that bases in the (t, f) dictionary are suitable to capture these non-stationarities accurately.

0.1.4 Summary and Conclusions

In this article, we briefly summarized the techniques of compressive sensing based on the (t, f) analysis. As expected, such compressive sensing approach provided accurate results in the recovery of non-stationary signals. In particular, we considered compressive sensing approaches based on the ambiguity function and (t, f) dictionaries. While the current article considered a (t, f) dictionary based on modulated discrete prolate spheroidal sequences, other bases such as wavelets or modulated Gaussian functions are expected to produce accurate results as well.

References

[1] M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with finite rate of innovation," *IEEE Transactions on Signal Processing*, vol. 50, pp. 1417–1428, June 2002.

[2] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, pp. 1289–1306, Apr. 2006.

[3] W. Dai and O. Milenković, "Subspace pursuit for compressive sensing signal reconstruction," *IEEE Transaction on Information Theory*, vol. 55, pp. 2230–2249, May 2009.

[4] R. G. Baraniuk, "Compressive sensing," *IEEE Signal Processing Magazine*, vol. 24, pp. 118–121, July 2007.

[5] E. J. Candes and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, pp. 21–30, Mar. 2008.

[6] S. Stanković, I. Orović, and E. Sejdić, *Multimedia Signals and Systems.* New York, NY: Springer US, 2012.

[7] P. Flandrin and P. Borgnat, "Time-frequency energy distributions meet compressed sensing," *IEEE Transactions on Signal Processing*, vol. 58, pp. 2974–2982, June 2010.

[8] E. Sejdić, I. Djurović, and J. Jiang, "Time-frequency feature representation using energy concentration: An overview of recent advances," *Digital Signal Processing*, vol. 19, pp. 153–183, Jan. 2009.

[9] B. Boashash, G. Azemi, and J. M. O'Toole, "Time-frequency processing of nonstationary signals: Advanced TFD design to aid diagnosis with highlights from medical applications," *IEEE Signal Processing Magazine*, vol. 30, pp. 108–119, Nov. 2013.

[10] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Transactions on Signal Processing*, vol. 41, pp. 3397–3415, Dec. 1993.

[11] D. Needell and J. A. Tropp, "Cosamp: Iterative signal recovery from incomplete and inaccurate samples," *Applied and Computational Harmonic Analysis*, vol. 26, pp. 301–321, May 2009.

[12] S. Senay, L. F. Chaparro, and L. Durak, "Reconstruction of nonuniformly sampled time-limited signals using prolate spheroidal wave functions," *Signal Processing*, vol. 89, pp. 2585–2595, Dec. 2009.

[13] T. Blu, P.-L. Dragotti, M. Vetterli, P. Marziliano, and L. Coulot, "Sparse sampling of signal innovations," *IEEE Signal Processing Magazine*, vol. 25, pp. 31–40, Mar. 2008.

[14] E. Sejdić, M. Luccini, S. Primak, K. Baddour, and T. Willink, "Channel estimation using DPSS based frames," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2008)*, (Las Vegas, Nevada, USA), pp. 2849–2852, Mar./Apr. 31–4, 2008.

[15] E. Sejdić, A. Can, L. F. Chaparro, C. M. Steele, and T. Chau, "Compressive sampling of swallowing accelerometry signals using time-frequency dictionaries based on modulated discrete prolate spheroidal sequences," *EURASIP Journal on Advances in Signal Processing*, vol. 2012, pp. 101–1–14, May 2012.