

# Conceptualization and Significance Study of a New Application CS-MIR

Kaichun Chang, Carl Barton, Costas Iliopoulos, Jyh-Shing Jang

► **To cite this version:**

Kaichun Chang, Carl Barton, Costas Iliopoulos, Jyh-Shing Jang. Conceptualization and Significance Study of a New Application CS-MIR. 8th International Conference on Artificial Intelligence Applications and Innovations (AIAI), Sep 2012, Halkidiki, Greece. pp.146-156, 10.1007/978-3-642-33409-2\_16 . hal-01521427

**HAL Id: hal-01521427**

**<https://hal.inria.fr/hal-01521427>**

Submitted on 11 May 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# CONCEPTUALIZATION AND SIGNIFICANCE STUDY OF A NEW APPLIATION CS-MIR

*Kaichun K. Chang<sup>1</sup>, Carl Barton<sup>1</sup>, Costas S. Iliopoulos<sup>1,‡</sup>, Jyh-Shing Roger Jang<sup>2</sup>*

<sup>1</sup>Dept. of Informatics, King's College London, Strand, London WC2R 2LS, England

<sup>2</sup> Dept. of Computer Science, Tsing Hua University, Taiwan

{ken.chang, carl.barton, c.iliopoulos}@kcl.ac.uk; jang@cs.nthu.edu.tw

‡Digital Ecosystems & Business Intelligence Institute, Curtin, Perth, Australia

**Abstract.** Numerous researches on Music Information Retrieval (MIR) have been estimated and linked with sparse representation method, few has paid enough attention on the application of compressive sensing and how it affects the reconstruction of MIR. This paper provides solid theoretical and various empirical evidence on the conceptualization, theoretical development, and implication of Compressive Sensing (CS), which to great extent, contributes to the application of the Music Information Retrieval.

**Keywords:** compressive sampling; music information retrieval; music identification

## 1 Introduction

This paper describes the theoretical development of Compressive Sensing (CS) study suggested in the past literature and identifies the need for investigating deficiencies of current audio fingerprinting research. By reviewing researches of audio fingerprint and music information retrieval (MIR) that is affected by Compressive Sensing, we can see a transparent linkage between Compressive Sensing theory and Music Information Retrieval system. Therefore, we create a new application CS-MIR to explore more sophisticated issues related so as to generate more practical results for future music information researchers, within which the theoretical perspectives and empirical evidences derived from compressive sensing technologies as a new application related to MIR will be embedded in the discussion of this new theory. In this paper, we conclude that some evidence prove that the new CS-MIR application could be the feasible solution for music classification and music recognition technologies.

## **2 Deficiencies Of Existing Work In Audio Fingerprinting**

### **2.1. Feature Extraction**

When the sampling signals of audio items are obtained, the next step is to extract features from the audio item. The goal of feature extraction is to simplify a large amount of data required to a small size for describing accurately. For the audio fingerprinting system, the extraction of features, which provide direct access to differentiating different items, is extremely crucial. As to the fingerprint extraction, some studies have explored various kinds of features that are robust to distortions, while the others mainly make an analysis on establishing a precise statistical fingerprint model and more appropriate distance measurements to improve the robustness of the system. Both approaches can achieve robust recognition.

Based on the outcome of these theoretical investigations, some practical audio fingerprinting systems are established, such as the Foosic Algorithm. It comes from Foosic, which is a free and open content project and employs its own free fingerprinting technology named libFooID. As to its feature extractor, firstly, all test music data converted to 32-bit float format and then down mixed to mono, and subsequently the next 100 seconds of audio data feed into a res-amplifier to obtain 90 seconds of 8000Hz sampled output. The output is then converted into a frequency spectrum by using a Hann-windowed DFT (here a custom Split-Radix FFT routine used) applied to 8192 sample blocks. By the above processing, a 90 seconds audio turns into 87 frames with frequency spectra. Another example is the "Audio Spectrum Flatness" LLD (Low Level Descriptor) which uses a time-to-frequency mapping and some further computation on a block by block basis to generate an MPEG-7 compliant fingerprint. The fingerprint is based on the calculation of the Spectral Flatness Measure (SFM).

### **2.2. Fingerprint Modelling**

Many works on audio fingerprint have discussed the modelling of audio fingerprints, and in most of the present fingerprint models, fingerprints are modelled as a series of binary numbers [1]. For example, Haitsma et al. modelled the fingerprint database as binary vectors. When testing an unknown audio, the fingerprints of the test audio are compared with the fingerprints of the other items in the database, and then the item which presents a low bit-error rate (BER) (commonly a certain threshold is set beforehand) is chosen. As mentioned above, the sign of the energy difference between the frame and block energies are taken as the fingerprint by Venkatachalam et al. This is the most fundamental approach for the fingerprint modelling.

Recently the researchers are aware of the importance of establishing a statistic model for the extracted features of audio clips or audio clips themselves, because the fingerprints should perform a dimensionality reduction of the original data significantly, provide accurate discrimination for different audio clips, and be invariant to distorted versions. For example, Kulesh et al. [3]. have used Mel-frequency spectral coefficients (MFCC) as the features of fingerprints, and the features are then modelled by a Gaussian mixture models (GMM). According to Cano et al. and Batlle et al. [2], the fingerprints are modelled as a sequence of hidden Markov models (HMM). Firstly, an alphabet of sounds describing an audio song (or item, clip) is extracted, and then these audio units are modelled with hidden Markov model. The unknown audio songs and the set of songs are divided into these audio

units; these audio units are ended up with some symbols for the unlabelled song and a database of sequences representing the original songs. The unlabelled song is recognized using the Viterbi decoding of the unlabelled item fingerprint against the repository of fingerprints. In paper [4], Arunan Ramalingam et al. designed fingerprints by modelling the audio clip as a Gaussian mixture models (GMM). In paper [5], Hui Lin, et al. modeled the fingerprint with precise common component Gaussian mixture models (CCGMMs). By stabilizing these statistical models for audio clips, more accurate recognition of audio items can be obtained.

### 2.3. Search Method for Efficient Matching

Recognition of fingerprints relies on the fingerprint matching technology. Item matching is another important aspect in audio fingerprints, which is based on the distance measurement to decide which item in the database is mostly matching to the item. So the distance measurement for comparing fingerprints is also an important problem to be addressed in the fingerprint matching. In the available audio fingerprinting system, fingerprint matching is often performed using the square of the Euclidean distance for a fast computation and mathematical tractability for analysis. The Euclidean distance between vectors  $f$  and  $g$  is defined as:

$$d = \sqrt{\sum_{i=1}^N (f_i - g_i)^2} \quad (1)$$

Some other distance metric can also be considered in the system such as L1, L2 and KL metric. Once the measurement is determined, we will use the distance measurement to perform an efficient searching. It is very important for the fingerprinting algorithm that the algorithm needs both few bit errors and efficient searching [6]. In the well-known audio fingerprints systems, some searching algorithms are employed.

Philip Database search: a two-phase search algorithm that is based on only performing full fingerprint comparisons at candidate positions pre-selected by a sub-fingerprint search.

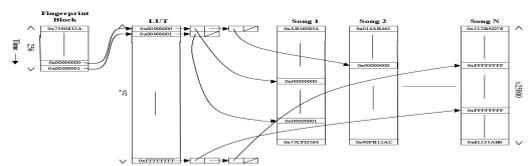


Fig. 2 Database layout

As shown in fig.2, the fingerprint database contains a lookup table (LUT) with all possible 32 bit sub-fingerprints as an entry. A hash table is often used instead of a lookup table in practice. In order to obtain an efficient searching algorithm, the sub-fingerprints that extracted from the database are registered in hash tables. Coarse search in the database is performed firstly, and output  $N$  songs as results. Distances between each song of these  $N$  songs and the query song are top  $N$  smallest. Then the exact search algorithm is used, as shown in fig.3.

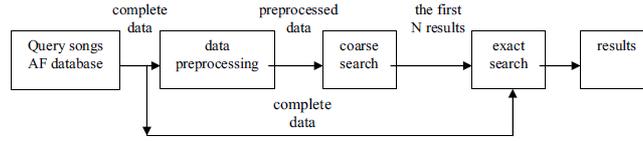


Fig.3. Philip Database search

### 3 Theoretical Development of Compressive Sensing Theory

#### 3.1. Rationalization of Compressive Sensing Model

Just as we have discussed above, the dimension reduction of musical signals are very important in audio fingerprinting and musical classification. Now many dimension reduction approaches have been proposed, such as LCA, which is a non-linear dimensionality reduction method, and NMF (Non-Negative Matrix Factorization) method. Compressive sensing theory provides a random measurement of signals, and proves to be able to not lose the information of the signals under the condition of enough number of measurement and incoherence between the measurement matrix and the transform matrix. Consequently, it is a natural compressive process of signals, which can also be regarded as the process of dimension reduction. Different with LCA and NMF, it is a linear method, which is characteristic of easy realization. In this section, firstly we present a detailed introduction on the compressive sensing theory.

Considering a real valued signal  $x$  (with length  $N$ ), given that it is  $K$ -sparse in the sparse basis matrix  $\Psi$ . Then considering an  $M \times N$  measurement (random or determined) matrix  $\phi$  (here  $M \ll N$ ,  $M$  is far less than  $N$ ), where the rows of  $\Psi$  are incoherent with the columns of  $\Psi$ . In terms of matrix notation, we have  $x = \Psi \alpha$ , in which  $\alpha$  can be approximately using only  $K \ll N$  ( $K$  is far less than  $N$ ) nonzero entries. The compression sensing theory states that such a  $K$ -sparse signal  $x$  can be reconstructed by taking only  $M = O(K \log N)$  number of liner, non-adaptive measurements, which can be described as follows:

$$y = \phi x = \phi \Psi \alpha = A \alpha \quad (2)$$

where  $y$  represents an  $M \times 1$  sample vector,  $A = \phi \Psi$  is an  $M \times N$  matrix. From it we can see that the implementation of compression sensing of a signal can be divided into three steps: first, transform the signal  $x$  to some sparse domain to make it  $K$  sparse under certain dictionary; second, make an observation of the sparse transformation coefficients to obtain the measurements; third, apply an optimization algorithm to retrieve the sparse coefficients, and then recover the original signal  $x$ . From the model we can see that there are three important things in realizing a compressive sensing

of signals, that is, finding a representation space in which the signal is sparse, determining a measurement matrix that is incoherent with the transform matrix, and using an efficient optimization algorithm to recover the sample from the limited number of measurement. In the following we discuss the measurement and optimization.

### 3.2. The Measurements Basis Matrix

The measurement basis matrix  $\phi$  must allow the reconstruction of the length  $N$  signal from  $M < N$  measurement (the vector  $y$ ). From the model we can see that  $M < N$ , so this problem appears to be an ill-conditioned problem. If, however,  $x$  is  $K$ -sparse in a given space and the  $K$  locations of the nonzero coefficients in are known, then the problem can be solved provided that  $M \geq K$ , a necessary and sufficient condition for this poorly defined problem to be well conditioned is that, for any vector  $v$  sharing the same  $K$  nonzero entries as and for some  $\varepsilon > 0$ , the following equation has to be satisfied,

$$1 - \varepsilon \leq \frac{\|Av\|_2}{\|v\|_2} \leq 1 + \varepsilon \quad (3)$$

That is, the matrix  $A$  must preserve the lengths of these particular  $K$ -sparse vectors. Of course, in general the locations of the  $K$  nonzero entries in are not known. However, a sufficient condition for a stable solution for both  $K$ -sparse and compressible signals is that the matrix  $A$  satisfies the above formula for an arbitrary  $3K$ -sparse vector  $v$ . This condition refers to as the restricted isometry property (RIP condition) [7]. Paper [8] said that if the measurement matrix is incoherence with the sparse basis matrix  $\Psi$ , it means that the rows  $\{\phi_i\}$  of  $\phi$  cannot sparsely represent the columns  $\{\psi_i\}$  of  $\Psi$ , so the matrix  $A$  will have the restricted isometry property.

### 3.3. The Optimization Based Reconstruction

To recover the signals from the measurement of signals, one should use some optimization algorithms to reconstruct the signal. Reconstruction aims to get the recovered signal from measurements  $y$  making use of the sparse prior or compressibility of signal. Although the sampling process is simply a linear projection, the reconstruction algorithm is highly non-linear. The step of reconstruction is equivalent to find the signal's sparse coefficient vectors  $\alpha$ , which can be written as  $\ell_0$  optimization:

$$\min \|\alpha\|_0 \quad s.t. \quad y = \phi \Psi \alpha, \quad x = \Psi \alpha$$

(4)

where  $\|\alpha\|_0$  represents the number of nonzero elements in vector  $\alpha$ . But unfortunately the above

equation is in general NP-hard. So the basis pursuit (BP) [16] optimization aims to minimize  $\ell_1$ :

$$\min \|\alpha\|_1 \quad s.t. \quad y = \phi \psi \alpha, \quad x = \psi \alpha$$

(5)

where  $\|\alpha\|_1 = \sum_i |\alpha_i|$ .

There are many optimization algorithms to solve the above problem, and some fundamental algorithm includes the greedy pursuit based algorithm, orthogonal greedy pursuit based algorithm, convex optimization algorithm and some variation of them. For two-dimensional images, another popular reconstruction algorithm is through the minimization of total variation (often called as the min-TV method), which offer the reconstructed images with better visual quality at much higher computational cost. Several fast greedy algorithms have also been proposed, such as the matching pursuit algorithm and orthogonal matching pursuit (OMP), the stage-wise orthogonal matching pursuit (StOMP) and iterative thresholding.

In this dissertation, the two compressive sensing based technologies are both based on the above model. Taking the iterative hard thresholding (IHT), a remarkably straightforward and rapid nonlinear signal reconstruction algorithm, as an example, it involves the application of the operators  $\phi$  and  $\phi^T$  once in each iteration, the iteration operation of updating the estimation of  $x$  can be described as:

$$x' = x + \phi^T (y - \phi x) \quad (6)$$

After iteration, the image is first transformed through  $\psi$  to yield  $\alpha$ . Then, the largest  $K$  coefficients of were kept, while the rest were set to zeros [11]. After that, the inverse transform of the linear sparse transformation matrix  $\psi^{-1}$  is applied to yield the reconstructed image. In this optimization algorithm, the initial solution influences the final result to a large scale. We can refer to [10] for other crucial factor about the sensing method and recovery optimization algorithm.

### 3.4. Empirical Results of Implications of Compressive Sensing on MIR

As to the applications of audio fingerprints, several practical requirements have been discussed to realize a successful audio fingerprinting system. Firstly, reliable audio fingerprinting is desirable in automatic music recognition. The audio fingerprinting system should be capable of identifying the distorted or corrupted audio clips in the case of degradations, that is, the system should be very robust. Robust fingerprint extraction method should be capable of dealing with severe degradations such as audio compression and large signal-to-noise ratios should be employed to represent the musical signal. Secondly, it should identify the clips of the item in the database in a few seconds, that is, the system should be time efficient or an audio fingerprinting system should be compact, finally, the system

should be of low computational complexity in forming the fingerprints and matching algorithm for finding the best match item in the database. [10] proposed a general framework for what we call compressive signal processing (CSP), an alternative approach in which signal processing problems are solved directly in the compressive measurement domain without first resorting to a full scale signal reconstruction. They concluded from their empirical results that detection, classification, and estimation enable the extraction of information from the samples, while the filtering enables the removal of irrelevant information and separation of signals into distinct components. While these choices do not exhaust the set of canonical signal processing operations, they believe that they provide a strong initial foundation. In practical applications of compressive sensing, an alternative approach has been made by [11]. as a new spectral compressive sensing (SCS) theory for general frequency-sparse signals, by which their new SCS algorithms significantly outperform the current state-of-the-art CS algorithms while providing provable bounds on the number of measurements required for stable recovery. Compared with this method, other researchers also attempted to introduce a fast and efficient framework for practical compressive sensing, on which the framework is mainly based a novel design of Structurally Random Matrix (SRM). This framework showcases that the number of measurements for exact signal reconstruction is almost minimal. Simulation results with several interesting SRM under various practical settings are also presented to verify the validity of the theory as well as to illustrate the promising potentials of the proposed framework.

After summarizing recent results of how random Toeplitz and Circulant matrices can be easily (or even naturally) realized in various applications, [12] have introduced fast algorithms for reconstructing signals from incomplete Toeplitz and circulant measurements; and exhibited computational results showing that Toeplitz and circulant matrices are not only as effective as random matrices for signal encoding, but also permit much faster signal decoding. [13] then conducted a high frequency analysis of (probabilistic) recoverability by the L1-based minimization/regularization principles, in which the absence of noise has shown that the L1-based solution can recover exactly the target of sparsity up to the dimension of the data either with the multiple-input-multiple-output (MIMO) measurement for the Born scattering or with the single-input-multiple-output (SIMO)/ multiple-input-single-output (MISO) measurement for the exact scattering. Meantime, a numerical exploration of compressed sampling theory has come into being, driven by a new greedy pursuit algorithm that computes sparse vectors that are difficult to recover; although it allows us to challenge theoretical identifiability criteria based on poly-topes analysis and on restricted isometry conditions, the theoretical analysis without resorting to Monte-Carlo sampling tends to avoid worst case scenarios. Based on intentional aliasing of the frequency components of the periodic signal while the reconstruction algorithm exploits recent advances in sparse representations and compressive sensing, [14] address the problem of sub-Nyquist sampling of periodic signals and show designs to capture and reconstruct such signals, concluding that for such signals the Nyquist rate constraint can be imposed on strobe-rate rather than the sensor-rate. Although numerous works have emphasized on measuring inequalities of randomized compressive operators, Wakin, et al. took endeavour to derive a concentration of measure bound for block diagonal matrices where the nonzero entries along the main diagonal blocks are sub-Gaussian random variables, concluding that the concentration exponent, in the best case, scales as that for a fully dense matrix and that the energy distribution of the signal plays in distinguishing the best case from the worst. In the paper [15] authors discussed a streaming CS framework and greedy reconstruction algorithm, the Streaming Greedy Pursuit (SGP), to reconstruct signals with

sparse frequency content, in which their experimental results on very long signals demonstrate the good performance of the SGP for validation. To further probe the compressed sensing issue, the paper [16] considered the problem of estimating a sparse signal from a set of quantized, Gaussian noise corrupted measurements by employing two methods and finding that compressed sensing can be carried out when the quantization is very coarse. The recent theory of Compressed Sensing states that a signal, e.g. a sound record or an astronomical image, can be sampled at a rate much smaller than what is commonly prescribed by Shannon-Nyquist. The sampling of a signal can indeed be performed as a function of its intrinsic dimension" rather than according to its cut-off frequency. Probabilistic matching pursuit for compressive sensing has recently caught some attention by certain researchers, who overturned a previous held assumption in compressive sensing research. A novel matching pursuit algorithm has been presented that uses the measurements to probabilistically select a subset of bases that is likely to contain the true bases constituting the signal. The algorithm is successful in recovering the original signal in cases where deterministic matching pursuit algorithms fail. It is also known that exact recovery is possible where the number of nonzero coefficients is up to one less than the number of measurements. Recent results in compressed sensing show that a sparse or compressible signal can be reconstructed from a few incoherent measurements. Since noise is always present in practical data acquisition systems, sensing and reconstruction methods are developed assuming a Gaussian (light-tailed) model for the corrupting noise. However, some researchers have found that when the underlying signal and measurements are corrupted by impulsive noise, commonly employed linear sampling operators fail to recover a close approximation of the signal, in this case, they proposed robust methods for sampling and reconstructing sparse signals in the presence of impulsive noise by employing nonlinear measurement operator based on the weighted myriad estimator and a geometric optimization method. Their simulation results demonstrate that the proposed robust methods significantly outperform commonly used compressed sensing sampling and reconstruction methods in impulsive environments, at the same time, providing comparable performance in less demanding environments.

As the development of the sparse signal recovery research, some researchers [17] have focused on an optimal decentralized algorithm that demonstrates its application in monitoring localized phenomena exploiting energy-constrained large-scale wireless sensor networks. In their proposed algorithm, simulation results just corroborate with previous research outcomes that the sensing performance is globally optimal and attains a high spatial resolution commensurate with the node density of the original network containing both active and inactive sensors. Other researchers [18] studied the number of measurements required to recover a sparse signal in  $C_m$  with  $L$  nonzero coefficients from compressed samples in the presence of noise and proved that  $O(L)$  (an asymptotically linear multiple of  $L$ ) measurements are necessary and sufficient if  $L$  grows linearly as a function of  $M$ . This result improves on the existing literature that is more focused on variants of a specific recovery algorithm based on convex programming and that  $O(L \log(M-L))$  measurements are required in the sub-linear regime ( $L = o(M)$ ). Another group of researchers [16] introduced a model-based Compressive Sensing theory that parallels the conventional theory and provides concrete guidelines on how to create model-based recovery algorithms with provable performance guarantees. This paper highlights a new class of structured compressible signals along with a new sufficient condition for robust structured compressible signal recovery that they dub the restricted amplification property, which is the natural counterpart to the restricted isometry

property of conventional CS.

Recovering or estimating the initial state of a high-dimensional system can require a potentially large number of measurements. Those researchers Error: Reference source not found extrapolated how this controversial issue can be significantly brought down as to certain linear systems when randomized measurement operators are utilised, upon which their work builds recent results from field of Compressive Sensing with a high-dimensional signal containing few nonzero entries that can be efficiently recovered from a small number of random measurements, and they also illustrate their results with simple case study of a diffusion system. Aside from permitting recovery of sparse initial states, their analysis has potential applications in solving inference problems such as detection and classification of more general initial state. There is one aspect pertaining to compressive sensing principles that mainly concerns iterative sparse recovery for inverse and ill-posed problems, researchers in this field have developed their results by providing compressed measurement models for ill-posed problems and recovery accuracy estimates for sparse approximations of the solution of the underlying inverse problem. The main ingredients are formulations that allow the treatment of ill-posed operator equations in the context of compressively sampled data, on which Tikhonov variational and constrained optimization formulations are relied. One important breakthrough to the classical compressive sensing framework lies in the area which the incorporation of joint sparsity measures allow the treatment of infinite dimensional reconstruction spaces, thereby to reassert that theoretical results are furnished with a number of numerical experiments.

Compressive sensing is an emerging field based on the revelation that a small collection of linear projections of a sparse signal contains enough information for stable, sub-Nyquist signal acquisition. When a statistical characterization of the signal is available, Bayesian inference can complement conventional CS methods based on linear programming or greedy algorithms. In recently years, belief propagation (BP) decoding has been utilized to perform approximate Bayesian inference, which implies the CS encoding matrix as a graphical model and that fast computation can be obtained by reducing the size of the graphical model with sparse encoding matrices. And simulation results show that focusing on a two-state mixture Gaussian model will not prevent CS-BP from being easily adapted to other signal models.

## **4 Conclusion**

In conclusion, this paper has explored theoretical foundation and development of compressive sensing model, and has elaborated on how it affects music information retrieval (MIR) system. This paper also reviewed the requirements for compressive sensing by illustrating their natural fit to MIR, and by illustrating and critically analysing four applications of Compressive Sensing in MIR. The paper placed great emphasize on an intuitive understanding of compressive sensing by describing the compressive sensing reconstruction as a process of interference cancellation. There is also a focus on the extrapolation of the driving factors in its applications, ranging from limitations imposed by feature extraction, the characteristics of fingerprint modeling, effective matching concerns in MIR, to practical fingerprint systems and its feasibility. In the end, this paper summed up implications of compressive sensing as a fundamental and crucial theoretical basis and a syndicated branch of compressive sensing for further developing and deepening the further work in the Music Information Retrieval field.

Besides, along with revealing the fact that the concepts and approaches, points discussed potentially allow entirely new applications of Music Information Retrieval, we can conclude that CS-MIR is still in its infancy. Many crucial issues remain unsettled. These include: optimizing sampling trajectories, developing improved sparse transforms that are incoherent to the sampling operator, studying reconstruction quality in terms of clinical significance, and improving the speed of reconstruction algorithms. The important point is that complex tasks can be addressed easily in the sparse domain for large datasets and future work will explore other information retrieval tasks such as similarity search. Music informatics researchers may recognise the majority of the signal representations and machine learning algorithms applied; however the source material has important differences from musical signals (e.g. its temporal structure) which necessitate differences in approach.

## 5 References

1. J. Haitsma and T. Kalker, "A highly robust audio fingerprinting system," in Proc. 3rd Int. Conf. Music Information Retrieval, Oct. 2002, pp. 107–115.
2. P. Cano, E. Batlle, T. Kalker, and J. Haitsma, "A review of algorithms for audio fingerprinting," in Proc. Int. Workshop Multimedia Signal Processing, 2002.
3. V. Kulesh, I. Sethi, and V. Petrushin, "Indexing and retrieval of music via gaussian mixture models," in Proc. 3rd Int. Workshop on Content Based Multimedia Indexing, Rennes, France, Sep. 2003, pp. 201–205.
4. Arunan Ramalingam, and Sridhar (Sri) Krishnan, Gaussian Mixture Modeling of Short-Time Fourier Transform Features for Audio Fingerprinting, IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, VOL. 1, NO. 4, DECEMBER 2006:457-463.
5. Hui Lin, Zhijian Ou, Member, IEEE, and Xi Xiao, Generalized Time-Series Active Search With Kullback-Leibler Distance for Audio Fingerprinting, IEEE SIGNAL PROCESSING LETTERS, VOL. 13, NO. 8, AUGUST 2006 465-468.
6. J. Haitsma and T. Kalker, "Speed-change resistant audio fingerprinting using auto-correlation," in Proc. Int. Conf. Acoustics, Speech, and Signal Processing, Apr. 2003, vol. 4, pp. 728–731.
7. E. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," IEEE Trans. Information Theory, vol. 52, pp. 489–509, Feb. 2006.
8. J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," IEEE Trans. Information Theory, 53(12) pp.4655-4666, 2007.
9. L. Gan, "Block compressed sensing of natural images," Proc. Int. Conf. on Digital Signal Processing (DSP), Cardiff, UK, 2007.
10. M.A. Davenport et al., "Signal Processing with Compressive Measurements," IEEE Journal of Selected Topics in Signal Processing, ISSN: 1932-4533, pp.1, February 2010.
11. M.F. Duarte, and R.G. Baraniuk, "Spectral Compressive Sensing," Rice University Technical Report TREE-1005, Feb. 2010.
12. W. Yin et al., "Practical Compressive Sensing with Toeplitz and Circulant Matrices," Tech. rep., CAAM, Rice University (2010).
13. A.C. Fannjiang, "Compressive Inverse Scattering I. High Frequency SIMO/MISO and MIMO Measurements", Inverse Problem, vol 26, no. 3, 2010.
14. A. Veeraraghavan et al, "Coded Stroboscopic Photography: Compressive Sensing of High-speed Periodic Events," IEEE Transactions of Pattern Analysis and Machine Intelligence, (PAMI) 33(4), 671–686 (2011).

15. P.T. Boufounos, and M.S. Asif, "Compressive Sampling for Streaming Signals with Sparse Frequency Content," CISS, [10], March 2010.
16. A. Zymnis, S. Boyd, and E. Candes, "Compressed Sensing with Quantized Measurements," IEEE Signal Processing Letters, vol. 17, no. 2, Feb. 2010.
17. Q. Ling and Z. Tian, "Decentralized Sparse Signal Recovery for Compressive Sleeping Wireless Sensor Network," IEEE Transactions on Signal Processing, vol. 58, no. 7, July 2010.
18. M. Akcakaya, and V. Tarokh, "Shannon Theoretic Limits on Noisy Compressive Sampling," IEEE Transactions on Information Theory, vol. 56, issue 1, pp.492-504, Jan. 2010.