

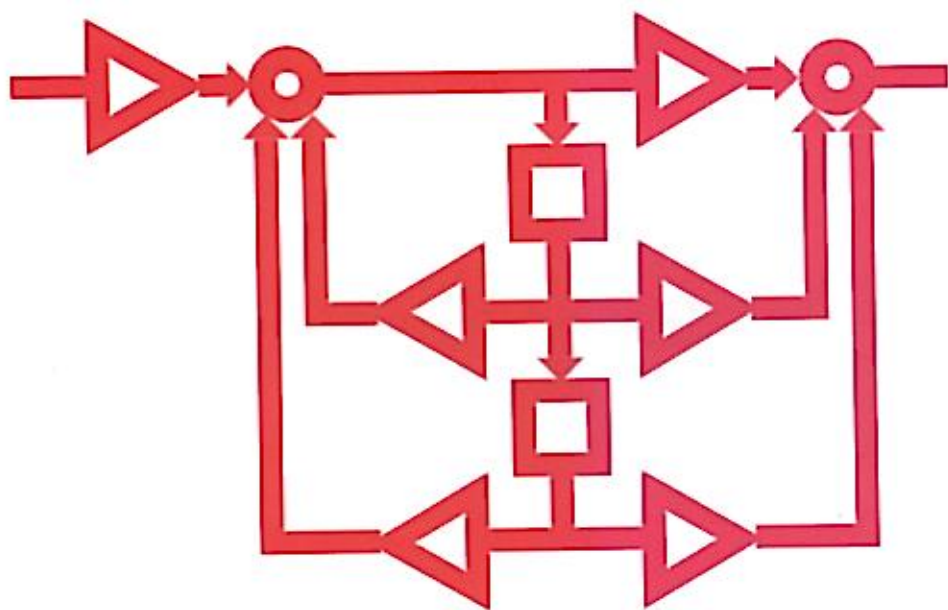
2010년도

제23회 신호처리합동학술대회 논문집

Proceeding of the 2010 Korea Signal Processing Conference

일시 : 2010년 10월 2일 (토)

장소 : 중앙대학교



- 주최 : 사단법인 한국방송공학회 3DTV 연구회
사단법인 대한전자공학회 음향 및 신호처리연구회, 화상처리 및 텔레비전연구회
사단법인 한국통신학회 신호처리연구회, 영상통신연구회
사단법인 한국음향학회 신호처리연구회, 음성처리연구회
- 주관 : 사단법인 한국방송공학회
- 후원 : 중앙대학교 (문화콘텐츠기술연구원, BK21 한국형디지털영상사업단, 서울미래형 콘텐츠컨버전스클러스터, 미래형영상콘텐츠기술연구단, 시각 및 지능시스템연구실), IEEE Korea SP Chapter, IEEE Korea CE Chapter
- 협찬 : (주)네오시드넷, (주)스타넥스, (주)에디텍, 인피니티박스, (주)하이트론시스템즈

1정보 이론 관점에서의 압축 센싱 기반의 다중 센서 시스템 분석

박상준, 이준호, 이흥노²
광주과학기술원 정보통신공학부

Information-Theoretical Analysis for Multiple Sensor System based on Compressive Sensing

Sangjun Park, Junho Lee, Heung-No Lee

School of Information and Communications, Gwangju Institute of Science and Technology
sjpark1@gist.ac.kr, ljh8097@gmail.com and heungno@gist.ac.kr

Abstract: Compressed Sensing(CS)은 활발하게 연구가 진행되고 있는 분야로써, 압축 가능한 신호를 샘플링 하는 시점에서 압축된 샘플을 얻게 한다. 우리는 CS 를 다중 센서 시스템에 적용하고자 한다. 다중 센서 시스템이란 여러 개의 센서가 한정된 지역에서 공동된 현상 등을 관측하는 시스템이다. 이러한 다중 센서 시스템은 동일한 현상을 관측 하기 때문에 센서에서 얻는 정보들 사이에 중복이 필연적으로 발생 한다. 한편 하나의 센서에 대해 [5]는 필요한 샘플의 개수가 신호의 길이가 아닌 신호의 성분들 중에 0 이 아닌 개수에 비례하는 것을 보였다. 우리는 [5]에서 제시된 방법을 이용하여, 다중 센서 시스템에 CS 를 적용하고, 또한 나아가서 정보 이론 관점에서 분석을 하여서, 필요한 샘플의 수가 어떻게 변화하는 지 보이고자 한다.

Keywords: Compressed Sensing, Sparse Signal, Joint Typicality, Multiple Sensor System

I. 서론

다중 센서 시스템이란 여러 개의 센서들이 하나의 공동된 현상을 관측하는 시스템을 일컫는다. 예를 들어서, 어떤 지역의 온도를 기록하기 위하여 많은 센서들을 넓게 퍼뜨려서 해당 지역에 배치하는 것을 생각 할 수 있다. 이러한 다중 센서 시스템의 대표적인 특징은 센서들이 획득한 정보들이 매우 유사할 수 있다는 것이다. 앞에서 사용한 예를 들어 임의의 지역에서 온도를 측정하기 위한 센서들을 5개를 배치 하였다고 가정을 하자. 이 때, 5개의 센서들이 감지한 온도들은 동일한 지역에서 감지했기 때문에 매우 비슷할 것 이다. 즉 센서들의 위치가 지리적으로 가깝고, 동일한 현상을 관측 하기 때문에, 정보들 사이의 유사성이 필연적으로 발생 한다. 하지만, 정보를 취합 라는 쪽에 보내기 위하여, 센서들은 에너지를 소모하면서 통신해야 한다. 이때, 보내는 정보의 양이 많다면, 그 만큼의 전력소모가 필요해진다. 따라서, 센서들이 보내는 정보의 양은 되도록이면 최소가 되도록 압축을 해야 한다. 또한 더 나아가서, 정

보들 사이의 유사성이 있기 때문에, 이러한 유사성을 이용한다면, 좀더 고 효율의 압축을 할 수 있을 것이다.

우리는 이러한 다중 센서 시스템에서, 앞에서 말하고자 하는 효과를 얻기 위하여, 압축 센싱[1]이라는 새로운 샘플링 방식을 응용하고자 한다. 이 방식은 기존의 Shannon-Nyquist rate에 따라 균등하게 샘플을 얻고 난 후에 정보를 보내기 위해, 각각의 센서들이 압축을 하는 방식이 아닌, 샘플을 얻는 시점에서 바로 압축된 샘플을 얻는 방식이다. 또한 개별적으로 압축을 할 때, 정보들 사이의 유사성을 이용하면, 원래의 신호로 되돌아 갈 때, 필요한 샘플의 수를 더 줄일 수 있게 한다. 압축 센싱에 대한 자세한 설명은 섹션 2에서 언급 하겠다.

우리는 압축 센싱을 정보 이론 관점에서 분석 하기 위하여, [5]에서 사용된 Joint Typicality 방식을 이용하고자 한다. [5]에 대해 간략히 언급을 하자면, 하나의 센서에 대해서 압축 센싱을 분석 한 것으로써, 압축 샘플의 개수가 원래 신호의 0이 아닌 성분들의 개수에 선형 비례인 것을 보였다.

이 논문의 나머지는 다음과 같이 조직화 되어 있다. 섹션 2에서는 압축 센싱에 대해서 간략하게 소개 하고, 섹션 3에서는 시스템 모델 및 논문에서 사용될 기호들을 언급 하고, 섹션 4에서는 Joint typicality와 Error Metric에 대해서 말하고자 한다. 섹션 5에서는 실험 및 토론이 있고 섹션 6에서는 결론 및 미래를 언급한다.

II. 압축 센싱 배경

기존의 Shannon-Nyquist 샘플링은, 샘플들을 시간 축에서 균등하게 취하는 방식이다. 이 때 샘플들이 압축이 가능한 샘플들 이더라도, 반드시 균등하게 샘플들을 취하는 방식이 필요하였다. 하지만 압축 센싱에서는 압축이 가능한 신호들을 취할 때, 균등하게 얻는 것이 아닌 압축이 된 샘플을 얻는 것이 핵심이다. 예를 들어서, N 차원의 신호가 있을 때 이 중 K 개의 성분들만 0이 아닌 값 이라고 가정을 하자. 이 때 이러한 신호를 K 최소 신호라 정의할 하고, 이러한 K 최소 신호를 샘플링 할 때, N 개 의 샘플이 아닌 N 보다는 작지만 K 보다 큰 M 개의 샘플을 얻는 것이다. 이러한 M 개의 샘플들은 랜덤 측정 행렬을 이용하여서 손쉽게 얻을 수 있다. 즉 X 를 K 최소 신호라 하고, $F \in \mathbb{R}^{M \times N}$ 를 랜덤 측정 행렬이라고 하고, 우리가 얻고자 하는 M

¹ 이 논문은 2010년도 정부(과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No.2010-00179444)
² 교신저자, FAX: + (82)62-715-2204, Phone: + (82)62-715-2257

차원의 샘플을 \mathbf{Y} 라 하자. 그러면 $\mathbf{Y}=\mathbf{F}\mathbf{X}$ 라는 선형 식을 이용하여서 손쉽게 얻을 수 있다. 이 때 \mathbf{F} 를 생성하는 대표적인 방법으로는 각각의 열 벡터들이 independent 하고 identically distributed (i.i.d) Gaussian $\mathbf{N}(0,1)$ 으로 설계하는 방안이 있다. 이와 같이 랜덤 측정 행렬을 만들게 되면 Restricted Isometry Property (RIP)[2]를 만족시킬 수 있는 확률이 높아진다. 이 RIP는 신호를 복구를 할 때, 복구 성능의 척도이다. 이것에 대해 간략히 언급을 하자면, 모든 K 개 이하의 열 벡터들로 구성된 부분집합이 orthogonal성질을 가지는 것을 말한다.

이제 신호를 복구 하는 쪽에서 생각해보자. 가장 손쉽게 생각해 볼 수 있는 방법 중의 하나는, \mathbf{Y} 를 구성하고 있는 K 개의 열 벡터를 정확하게 찾아서 역 행렬 연산을 통하여 답을 찾는 l_0 norm 방법이 있다. 이 방법을 적용 하기 위해서는 \mathbf{Y} 차원의 크기가 적어도 $2K$ 만큼은 있어야지만 유일하게 \mathbf{X} 신호를 복구 할 수 있다. 하지만 N 개의 열 벡터 중에서 K 개의 열 벡터를 뽑아야 하기 때문에 실용적이지 않다는 문제점이 있다. 하지만 linear programming을 통한 l_1 norm을 이용한 신호를 복구 시에 \mathbf{Y} 의 차원이 $cK \log(N/K)$ 만큼만 있으면 l_1 norm 방법과 똑 같은 신호를 복구해 낸다는 것이 알려져 있다[3]. (일반적으로 c 는 작은 값으로서 3~4정도면 적당하다.) 마지막으로, 또 다른 방법 중 하나는 탐욕 알고리즘을 기초로 한 Orthogonal Matching Pursuit (OMP)[4]복구 방법이 있다.

III. 시스템 모델

1. 다중 K 회소 신호와 Support Set

다중 센서 시스템에서, S 개의 센서가 있다고 가정하자. 이 때 s 번째 센서가 취득한 K 회소 신호를 \mathbf{X}_s 라고 하자. 그러면 이 때 s 번째 센서의 Support Set은 다음과 같이 정의가 된다

$$I(\mathbf{X}_s) = \{i | \mathbf{X}_s(i) \neq 0\}. \quad (1)$$

즉 s 번째의 Support Set은 \mathbf{X}_s 의 값 중에 0이 아닌 값에 해당하는 인덱스를 원소로 가지고 있는 집합이다. 이제 우리는 모든 S 개의 센서들이 동일한 Support Set(I)을 가진다고 가정 하자. 즉 우리의 시스템 모델은 어느 한 센서의 신호에서 i 번째 위치가 0이었다면 모든 센서들이 취득한 신호에서 i 번째 값은 0인 것이다. 하지만 Support Set이 같다고 0이 아닌 값마저 똑 같은 것은 보장 하지 않는다. 즉 $\mathbf{X}_s(2)$ 가 5라고 해서 반드시 $\mathbf{X}_t(2)$ 가 5일 필요는 없는 것이다.

2. 노이즈 관측 모델

노이즈 관측 모델은 다음과 같다.

$$\mathbf{R}_s = \mathbf{Y}_s + \mathbf{N}_s, \quad (2)$$

여기서, \mathbf{Y}_s 는 압축 샘플로서 $\mathbf{F}_s \mathbf{X}_s$ 와 똑같은 값이다. $\mathbf{Y}_s = \mathbf{F}_s \mathbf{X}_s$. 또한 $\mathbf{N}_s(i)$ 는 i.i.d이고 평균이 0이고 분산이 σ^2 인 Gaussian이다. \mathbf{F}_s 는 각각의 원소들이 평균이 0이고, 분산이 1이고 i.i.d인 Gaussian이고, 모든 센서에 대해서, \mathbf{F}_s 의 원소들과 $\mathbf{N}_s(i)$ 는 상호 독립이다. 각각의

센서들이 압축센싱을 통해 압축된 샘플들(\mathbf{Y}_s)을 얻을 때, 더해지는 노이즈들(\mathbf{N}_s)은 상호 독립이라고 가정 한다.

3. 기호

Support Set(I)와 비슷한 Utility Set(J)을 정의한다. J 의 크기는 K 이고 I 의 교차되는 원소의 개수는 최대 $K-1$ 까지만 허용이 된다. 즉 $0 \leq |J \cap I| \leq K-1$. 그 다음으로 $\mathbf{F}_i(i)$ 는 \mathbf{F}_s 행렬의 i 번째 열 벡터를 뜻하고, $\mathbf{F}_{J,i}$ 는 부분 행렬로서 \mathbf{F}_i 의 열 벡터들 중에서 Support Set의 원소들에 대응하는 것들이 모인 행렬이다. $\mathbf{F}_{J,i}$ 는 $\mathbf{F}_{J,i}$ 와 비슷하게 \mathbf{F}_i 의 부분 행렬로서, Utility Set의 원소들에 대응하는 열 벡터들로 구성된 행렬이다. 마지막으로 Orthogonal Projection 행렬과, Projection 행렬을 정의 한다. \mathbf{A} 행렬이 주어졌을 때, Orthogonal Projection 행렬은 $\mathbf{O}_A = \mathbf{A}(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$ 로 정의가 되고, Projection 행렬은 $\mathbf{Q}_A = \mathbf{I} - \mathbf{O}_A$ 로 정의 한다.

IV. Joint Typicality, Error Metric 및 Theorem

Theorem 및 (4)와 (5)에 대한 증명 및 자세한 내용은 [8]에 수록 되어 있다.

1. Joint Typicality와 정의

각각의 센서들로부터 온 압축된 신호들인 \mathbf{R}_s 들과 Utility Set J 가 δ -Jointly Typical이 되려면 모든 센서들에 대해서, $\mathbf{F}_{J,i}$ 의 모든 열 벡터들과 노이즈들이 상호 독립적이고

$$E\left\{ \left| \frac{1}{S} \sum_{s=1}^S \frac{1}{M} \left(\left| \mathbf{Q}_{J,i} \mathbf{R}_s \right|^2 - (M-K)\sigma^2 \right) \right| \right\} < \delta \quad (3)$$

사건 (3)이 발생하는 확률이 1에 가까워야 한다. 여기서, 사건 (3)의 확률과 관계된 랜덤 변수는 E 안에 있는 변수들이다.

2. 올바른 Support Set I 와 δ -Jointly Typical Case

양수인 δ 에 대해서, 모든 센서들에 대해서, $\mathbf{F}_{J,i}$ 의 모든 열 벡터들이 선형 독립이면, 다음 아래의 식이 성립이 된다.

$$P\left\{ E\left\{ \left| \mathbf{I}(\mathbf{R}_s, \mathbf{F}_{J,i}) \right|^2 \right\} \leq 2 \exp\left(-\frac{S(M\delta)^2}{2(M-K)+M\delta} \right) \right\} \quad (4)$$

여기서, $\delta_s = \delta/\sigma^2$ 이다.

3. 올바르지 않는 Support Set J 와 δ -Jointly Typical Case

먼저, 올바르지 않는 Support Set은 Utility Set을 이용해서 표현 할 수 있다. 이에 대한 것은 시스템 모델에서 설명 하였다. 이제 Utility Set (J)을 이용하면, 양수인 δ 에 대해서, 모든 센서들에 대해서, $\mathbf{F}_{J,i}$ 의 모든 열 벡터들이 선형 독립이면, 다음 아래의 식이 성립 된다.

$$P(E(J, \mathbf{R}, \mathbf{F})) \leq \exp\left\{-\frac{S(M-K)(\sum_{i=1}^S \sigma_{i,j}^2 - \sigma_i^2 - \delta_i)^2}{2\sum_{i=1}^S \sigma_{i,j}^2}\right\} \quad (5)$$

여기서, $\delta_i = \delta M / (M - K)$, $\sigma_{i,j}^2 = \sum_{m=1}^M X_m(i)^2 + \sigma^2$ 이다.

4. Error Metric 정의

압축된 신호들로부터 원래의 신호들로 되돌아 갔을 때, 오류의 척도로서 우리는 다음 아래와 같은 Metric 을 사용 한다.

$$I\left\{\sum_{i=1}^s \frac{|y_i - \hat{X}_i(i) \neq 0|}{|I|} > S(1-\alpha)\right\} \quad (6)$$

여기서, \hat{X}_i 는 s 번째 센서의 복구된 신호이고, $\alpha \in (0,1)$ 그리고, $I(\cdot)$ 는 Indicator 함수이다.

이 Error Metric은 우리는 Proposed Error Metric (PEM)이라고 정의한다. PEM의 특징은 모든 신호들이 올바르게 복구가 됐을 때 1이고 그 외의 경우는 0이다.

5. Main Theorem

Theorem: 양수인 δ 와 위에서 언급된 랜덤 측정 행렬 \mathbf{F} 와 PEM이 주어졌을 때, 센서의 수가 많아지면 각 센서당 획득해야 할 샘플인 M 이 $K+1$ 로 수렴한다.

이 이론과 비슷한 것을 [6],[7]에서도 볼 수 있다. 간단히 소개를 하자면, [6],[7]에서는 우리의 동일하게 Support Set이 같다는 가정을 이용했지만, 노이즈 관측을 고려를 하지 않았다. 또한 [6]에서 나온 결과와는 적은 수의 센서에서는 잘 동작 하지 않는 문제점이 있고, [7]에서는 알고리즘을 제시를 했을 뿐이지 분석이 되어있지 않다.

V. 실험 및 토론

우리는 위에서 얻은 식 (4)와 (5)를 이용하여서, 실험을 수행 하였다. 실험 환경은 다음과 같다. 센서의 수를 1, 2, 4, 8로 증가 하였고, 신호들의 길이를 50으로, K 와 값도 5로 고정 하였다. 노이즈 에너지와 δ 의 값은 0.01로 지정 하였다. 이런 상황에서 우리는 각 센서당 얼마나 큰 M 이 필요한지 보고자 한다.

그림 1을 보면 우리는 센서의 수가 증가함에 따라 복구 하기 위한 각 센서당 보내야 하는 샘플의 수가 줄어 드는 것을 확인을 할 수 있다. 이 결과는 우리가 얻어 들었던 Theorem과 일치함을 볼 수 있다.

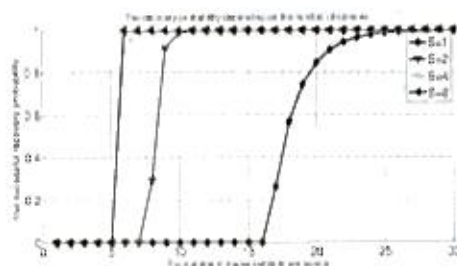


그림 1. $N=50, K=5, \alpha=0.01, \delta=0.01, \sigma_i^2=0.01$. 센서의

수가 증가함에 따라 올바르게 복구하기 위한 필요한 압축 샘플의 수가 급격히 줄어 드는 것을 볼 수 있다.

VI. 결론 및 미래

본 논문을 통해, 우리는 압축 센싱 이론이 다중 센서 시스템에서 발생하는 정보들의 중복을 제거 하기 위한 매력적인 방법이 될 수 있다는 것을 보였다. 또한 센서의 수가 증가함에 따라서, 각 센서당 보내야 하는 압축 샘플의 수가 $K+1$ 개만 있으면 된다는 것을 정보 이론 관점에서 보였다. 또한 우리가 제안한 방법은 모든 가능한 경우들을 조사할 하여 얻은 결과이다. 따라서 이 방법은 우리가 정의한 Metric에서의 최적의 해를 찾는 방법이라고 감히 말 할 수 없다. 더 나아가서, 우리는 이러한 분석결과를 만족 시킬 수 있는 실용적인 알고리즘을 설계할 하고, 분석 해야 할 것이다. 마지막으로, 우리는 좀 더 복잡한 모델에서 분석 수행해야 할 것이다.

참고문헌

- [1] D. Donoho, "Compressed sensing," *IEEE Trans. On Information Theory*, vol. 52, pp. 1289-1306, 2006.
- [2] E. Candes, J. Romberg and T. Tao "Robust Uncertainty Principles: Exact Signal Reconstruction from Highly Incomplete Frequency Information," *IEEE Trans. Inform. Theory*, vol. 52, no. 2, pp. 489-509, Feb. 2006
- [3] E. Candes and T. Tao, "Decoding by linear programming," *IEEE Trans. Inform. Theory*, vol. 51, no. 12, pp. 4203-4215, Dec. 2005.
- [4] J.A Tropp and A. C. Gilbert, "Signal recovery from partial information via Orthogonal Matching Pursuit," *IEEE Trans. Inform. Theory*, vol. 53, no. 12, pp. 4655-4666 Dec. 2007
- [5] Mehmet Akçakaya and Vahid Tarokh, "Shannon-Theoretic Limits on Noisy Compressive Sampling", *IEEE Trans. Inform. Theory*, VOL. 56, NO. 1, Jan. 2010
- [6] S. Sarvotham, M. B. Wakin, D. Baron, M. F. Duarte, and R. G. Baraniuk, "Analysis of the DCS one-stage greedy algorithm for common sparse supports", Technical Report TREE-0503: Rice University, Department of Electrical and Computer Engineering, Oct. 2005
- [7] D. Baron, M. F. Duarte, M. B. Wakin, S. Sarvotham and R. G. Baraniuk, "Distributed compressive sensing," Jan. 2009.
- [8] Sangjun Park, Junho Lee, Younghak Shin, Heung-No Lee, "Per-Sensor Measurements Required for Noisy Distributed Compressive Sensor System", will be submitted to *IEEE Trans. Inform. Theory*.