상관 벡터를 이용한 압축 센싱 행렬의 성능 분석

홍석범[+0], 박호성[+], 신범규[++], 노종선[+], 정하봉[+++]

서울대학교 전기컴퓨터공학부, 뉴미디어통신공동연구소[+],

삼성전자[++], 홍익대학교 전자전기공학부[+++]

Performance Analysis of Compressed Sensing Matrices Using Correlation Vectors

Seokbeom Hong[+0], Hosung Park[+], Beomkyu Shin[++], Jong-Seon No[+], Habong Chung[+++]

Department of EECS, INMC, Seoul National University[+],

Samsung Electronics, Co., Ltd.[++], School of EE, Hongik University[+++]

fousbyus@ccl.snu.ac.kr, {lovepk98,thechi76,jsno}@snu.ac.kr, habchung@hongik.ac.kr

요 약

본 논문에서는 k-집합 상관 벡터의 l_2-norm 의 표준 편차를 압축 센싱 (compressed sensing) 행렬의 복원 성능에 대한 새로운 측정 기준으로 제시한다. k-집합 상관 벡터는 센싱 행렬의 열 중 k개의 열을 선택하여 생성된 부분행렬 내의 모든 열의 쌍에 대하여 상관값을 구하고 이를 벡터 형태로 구성한 것으로, restricted isometry property (RIP)와 밀접한 관련을 가지고 있다. 또한 본 논문에서는 모의 실험을 통하여 제안하는 측정 기준이 센싱 행렬의 복원 성능에 대한 측정 기준으로써 잘 동작한다는 것을 확인하였다.
.

1. 서론

압축 센싱 시스템에서는, 길이 N의 k-sparse 메시지 벡터 $\mathbf{x}$ $\left(\|\mathbf{x}\|_0 \le k\right)$가 $M \times N$ 센싱 행렬 Φ에 의해 길이 $M\,(M \ll N)$의 관찰 벡터 (measurement vector) $\mathbf{y}$로 압축된다. 적당한 k, M, N에 대하여 우수한 압축 센싱 행렬이 사용되는 경우, 관찰 벡터로부터 기존의 메시지 벡터를 높은 확률로 복원하는 것이 가능하다.

행렬을 구성하고 있는 각 원소가 정규 분포 또는 베르누이 분포를 따르는 랜덤 압축 센싱 행렬은 센싱 행렬로써 우수한 성능을 보인다는 사실은 널리 알려져 있다. 이에 따라 현재까지 압축 센싱 기법에서 사용되는 센싱 행렬은 주로 랜덤하게 생성된 센싱 행렬이 널리 사용되어 왔다. 그러나 최근 들어 우수한 성능과 낮은 복원 복잡도를 가지는 고정 센싱 행렬들이 고안되면서[1]-[3], 우수한 센싱 행렬을 설계하는 문제가 압축 센싱을 적용하는 상황에서 가장 중요한 문제 중 하나로 대두되고 있다. 이에 따라 실제적인 센싱 행렬의 복원 성능을 예측할 수 있는 측정 기준에 대한 필요성이 나타나고 있다.

RIP[4]는 센싱 행렬의 복원 성능에 대한 가장 잘 알려진 측정 기준이다. 만약 모든 열 벡터가 정규화(normalization)된 센싱 행렬 Φ가 가능한 모든 k-sparse 메시지 벡터 $\mathbf{x}$와 상수 δ_k에 대하여 (1) 식을 만족하면, 센싱 행렬 Φ는 δ_k에 대하여 k-RIP 를 만족하는 것으로 간주한다.

$$(1-\delta_k)\|\mathbf{x}\|^2 \le \|\Phi\mathbf{x}\|^2 \le (1+\delta_k)\|\mathbf{x}\|^2 \qquad (1)$$

적절한 δ_k에 대하여 k-RIP 를 만족하는 센싱 행렬을 사용하여 압축 센싱을 수행할 경우, 모든 k-sparse 메시지 벡터에 대한 완벽한 복원이 가능하다[4].

그러나 어떠한 센싱 행렬이 RIP 를 만족하는지의 여부를 확인하는 것은 매우 어려운 작업이다. 또한 현재까지 RIP 외에도 센싱 행렬의 복원 성능에 대한 여러 가지 측정 기준이 개발되었으나, 이러한 측정 기준을 이용하더라도 주어진 센싱 행렬들의 성능을 직접적으로 비교하는 것은 쉽지 않다.

본 논문에서 제안하는 압축 센싱 행렬의 복원 성능에 대한 새로운 측정 기준은 k-집합 상관 벡터의 l_2-norm 의 표준 편차이다. 여기서 k-집합 상관 벡터란, 센싱 행렬 내의 k개의 열을 선택하여 생성된 부분행렬 내의 모든 열의 쌍에 대하여 상관값을 구하고 이를 벡터 형태로 구성한 것이다. 본 논문에서는 제안하는 측정 기준과 RIP 상수 간의 연관성을 이론적으로 분석하고, 모의 실험을 통하여 제안하는 기법이 센싱 행렬의 복원 성능에 대한 측정 기준으

로써 잘 동작한다는 것을 확인하였다.

2. 압축 센싱 행렬의 복원 성능에 대한 새로운 측정 기준

센싱 행렬을 위한 새로운 측정 기준을 소개하기에 앞서, 먼저 k-집합 상관 벡터를 정의하겠다. ϕ_j 를 $M \times N$ 압축 센싱 행렬 Φ 의 j 번째 열이라 하고, $\phi_j(i)$ 를 열 벡터 ϕ_j 의 i 번째 항목이라 하자. 센싱 행렬 Φ의 k-집합 상관 벡터 $\mathbf{c}$는 길이 $\binom{k}{2}$의 벡터이다. 열 인덱스의 집합인 $\{1,2,\cdots,N\}$의 크기 k인 부분집합 U와 $g,h \in U, g > h$를 만족하는 g,h에 대하여, 벡터 $\mathbf{c}$ 내의 각 원소들은 $c_{g,h} = \sum_{i=1}^{M} \phi_g(i)\overline{\phi_h(\iota)}$ 와 같은 값을 가진다. 이에 따라 $M \times N$ 센싱 행렬에 대하여 $\binom{N}{k}$개의 서로 다른 k-집합 상관 벡터가 존재하게 된다. 열 인덱스의 집합 U에 대응되는 k-집합 상관 벡터의 l_2-norm 은 $L^{(U)} = \sqrt{\sum_{g,h\in U,g>h} c_{g,h}^2}$ 와 같이 표현된다.

본 논문에서는 k-집합 상관 벡터의 l_2-norm 의 표준편차를 압축 센싱 행렬의 복원 성능에 대한 새로운 측정 기준으로 제안한다. k-집합 상관 벡터의 l_2-norm 은 (1)의 RIP 상수 δ_k와 밀접한 연관성을 갖는다.

$x(i)$를 k-sparse 메시지 벡터 $\mathbf{x}$의 i번째 항목이라 하고, 센싱 행렬의 각 열 ϕ_j 들은 정규화되어 있다고 가정하자. 이 때, 다음의 식이 성립한다.

$$\left\|\Phi\mathbf{x}\right\|^2 = \sum_{i=1}^{M}\left[\left(\sum_{g=1}^{N}\phi_g(i)x(g)\right)\overline{\left(\sum_{h=1}^{N}\phi_h(\iota)x(h)\right)}\right]$$
$$= \left\|\mathbf{x}\right\|^2 + \sum_{i=1}^{M}\sum_{g,h=1,g\neq h}^{N} x(g)\overline{x(h)}\,\phi_g(i)\overline{\phi_h(\iota)}. \quad (2)$$

벡터 $\mathbf{x}$ 내의 원소들 중 0 이 아닌 값을 가지는 것은 최대 k개이다. 이러한 0 이 아닌 원소들의 인덱스를 모아 놓은 집합을 S 라 하고, $g,h \in S$ 에 대해 $x_{g,h} = x(g)x(h)$ 라 정의하자. 이 때, (1), (2)에 의해 다음의 식이 성립한다.

$$\left|\sum_{g,h\in S,g\neq h} x_{g,h}\, c_{g,h}\right| \leq \delta_k \left\|\mathbf{x}\right\|^2 \quad (3)$$

그리고 코시-슈바르츠 (Cauchy-Schwartz) 부등식에 의해 다음과 같은 식을 얻을 수 있다.

$$\left|\sum_{\substack{g,h\in S\\ g\neq h}} x_{g,h}c_{g,h}\right| \leq \sqrt{\left(\sum_{\substack{g,h\in S\\ g\neq h}}\left|x_{g,h}\right|^2\right)\left(\sum_{\substack{g,h\in S\\ g\neq h}}\left|c_{g,h}\right|^2\right)} \quad (4)$$

또한 아래의 식은 (5)의 우변의 2 번째 항에 코시-슈바르츠 부등식을 적용함으로써 어렵지 않게 유도할 수 있다.

$$\sum_{g,h\in S,g\neq h}\left|x_{g,h}\right|^2 = \sum_{g\in S}|x(g)|^2 \sum_{h\in S}\left|\overline{x(h)}\right|^2 - \sum_{g\in S}|x(g)|^4$$
$$= \left\|\mathbf{x}\right\|^4 - \frac{1}{k}\left(\sum_{h\in S}1^2\sum_{g\in S}|x(g)|^4\right) \quad (5)$$
$$\leq \frac{k-1}{k}\left\|\mathbf{x}\right\|^4$$

위 식의 결과와 (4)에 의해, 다음의 부등식이 성립한다.

$$\left|\sum_{g,h\in S,g\neq h} x_{g,h}c_{g,h}\right| \leq \sqrt{\frac{k-1}{k}\sum_{g,h\in S,g\neq h}\left|c_{g,h}\right|^2}\cdot\left\|\mathbf{x}\right\|^2 \quad (6)$$

0 이 아닌 원소들의 인덱스가 S에 속하는 모든 메시지 벡터 $\mathbf{x}$에 대하여 (1)을 만족하는 최소의 RIP 상수를 $\delta(S)$라 정의하자. 이 때, $\delta(S)$는 S가 주어졌을 때 모든 $\mathbf{x}$에 대하여 (3)을 만족하는 RIP 상수의 최솟값이 된다. 이 사실과 (6)에 의해, 다음과 같은 부등식을 쉽게 유도할 수 있다.

$$\delta(S) \leq \sqrt{\frac{k-1}{k}\sum_{g,h\in S,g\neq h}\left|c_{g,h}\right|^2} = \sqrt{\frac{2(k-1)}{k}L^{(S)}} \quad (7)$$

즉, $\delta(S)$ 의 상한은 k-집합 상관 벡터의 l_2-norm 에 상수를 곱한 값으로 결정된다. 만약 $|S| = k$인 모든 인덱스 집합 S에 대하여 $L^{(S)} < T$인 경우, 센싱 행렬 Φ 는 $\sqrt{2(k-1)T/k}$ 보다 작거나 같은 값을 가지는 δ_k에 대하여 k-RIP 를 만족한다.

k-집합 상관 벡터의 l_2-norm 의 최댓값을 감소시키는 것이 RIP 를 보장한다는 것은 (7)을 통해 명백히 알 수 있다. 그러나 만약 압축 센싱 기법을 적용함에 있어서 단지 높은 확률의 복원 성공만으로 목표로 하는 경우, 최악의 경우까지 고려하여 완벽한 복원을 보장하는 RIP 는 센싱 행렬 설계 측면에서 필요 이상으로 까다로운 조건이다. 따라서 본 논문에서는 RIP 의 조건을 다소 완화시켜 센싱 행렬의 평균적인 복원 성능을 측정하는 측정 기준을 제안하고자 한다.

(7)의 결과로부터, k-집합 상관 벡터의 l_2-norm L이 큰 값을 가질 확률이 낮은 센싱 행렬의 경우, $\delta(S)$가 큰 값을 가질 확률 역시 낮을 것이다. 즉, 이 경우 센싱 행렬의 k개의 열로 이루어진 부분행렬들은 높은 확률로 RIP 를 만족시키게 된다. 이와 같은 사실로 미루어 볼 때, 큰 값의 L을 가질 확률이 낮은 센싱 행렬일수록 더욱 우수한 복원 성능을 나타낼 것을 예측할 수 있다.

그러나 k-집합 상관 벡터의 l_2-norm 의 분포 자체를 비교하는 것은 어려운 일이기 때문에, 본 논문에서는 이를 대신할 정량적인 측정 기준으로 k-집합 상관 벡터의 l_2-norm L의 표준 편차를 제안한다. 비록 논문의 지면 관계상 그 증명을 수록하지는 못하였으나, 본 논문의 저자들은 센싱 행렬로 주로 사용되는 행렬 내의 원소들의 크기가 일정한 랜덤 센싱 행렬 및 상관 특성이 평균적으로 최적에 가까운 고정 센싱 행렬들에 대하여 L의 평균이 거의 일정하

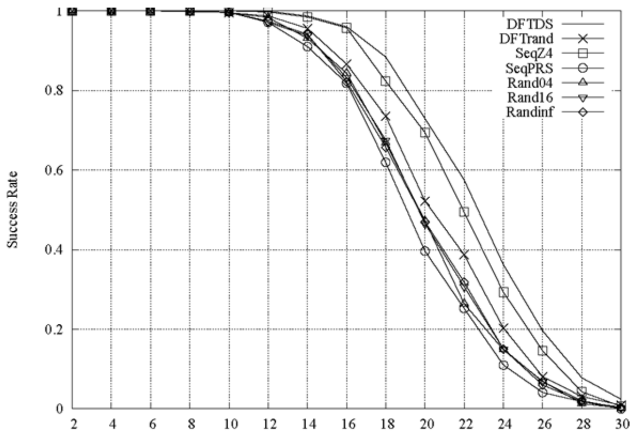

그림 1. OMP 를 이용한 경우의 복원 성공 확률

다는 사실을 증명하였다. 이에 따라 L의 표준 편차를 통해 큰 값의 L을 가질 확률을 추정할 수 있고, 이에 따라 센싱 행렬의 복원 성능을 예측하는 것이 가능하다.

3. 모의 실험 결과

이 장에서는 모의 실험을 통해 본 논문에서 제안하는 측정 기준이 센싱 행렬의 복원 성능에 대한 측정 기준으로써 잘 동작한다는 것을 확인할 것이다. 성능 비교를 위해 사용된 센싱 행렬은 다음과 같다;

1. 이산 푸리에 변환 행렬에서 랜덤하게 또는 차집합 (difference set)을 이용하여 행을 선택[5]해서 생성한 65×4161 행렬,
2. 원소들의 크기가 일정한 다양한 알파벳 크기의 63×4096 랜덤 행렬,
3. 멱 잉여류 시퀀스 (power residue sequence)[6]로부터 생성된 61×3844 행렬, $\mathbb{Z}_4$ 상에서 정의된 4 진 family A 시퀀스[7]로부터 생성된 63×4096 행렬.

그림 1 에는 위의 행렬들을 압축 센싱 행렬로 사용했을 때의 복원 성공 확률을 그래프로 나타내었다. 메시지 벡터는 k-sparse 이고, 벡터 내의 0 이 아닌 원소들은 크기가 (0,100)에서 균일하게 분포되고 위상이 $[0,2\pi)$에서 균일하게 분포된 복소수라고 가정하였다. 모의 실험은 각각 1000 개씩의 프레임에 대해 진행되었고, 복원 알고리듬으로는 orthogonal matching pursuit (OMP)[8]이 사용되었다.

표 1 은 그림 1 에 나타난 센싱 행렬들의 k-집합 상관 벡터의 l_2-norm 의 표준 편차를 나타낸다. 이는 각각 10^7개씩의 상관 벡터에 대하여 계산되었다. 표 1 의 결과를 그림 1 과 비교해 보면, l_2-norm 의 표준 편차가 더 작은 값을 갖는 센싱 행렬일수록 더 우수한 복원 성능을 가진다는 것을 알 수 있다. 이를 통해 본 논문에서 제안하는 기법이 센싱 행렬의 복원 성능에 대한 측정 기준으로써 잘 동작한다는 것을 확인할 수 있다.

Type	k=3	k=4	k=8	k=12	k=20
DFTDS	0	0	0	0	0
DFTrand	5.990E-2	6.045E-2	6.078E-2	6.079E-2	6.072E-2
SeqZ4	1.416E-2	1.379E-2	1.357E-2	1.353E-2	1.348E-2
SeqPRS	6.810E-2	6.692E-2	6.552E-2	6.527E-2	6.506E-2
Rand04	6.115E-2	6.178E-2	6.226E-2	6.235E-2	6.242E-2
Rand16	6.123E-2	6.181E-2	6.226E-2	6.235E-2	6.241E-2
Randinf	6.122E-2	6.180E-2	6.226E-2	6.233E-2	6.240E-2

표 1. k-집합 상관 벡터의 l_2-norm 의 표준 편차

4. 감사의 글

본 연구는 2012 년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No. 2012-0000186).

5. 참고 문헌

[1] R. Calderbank, S. Howard, and S. Jafarpour, “Construction of a large class of deterministic sensing matrices that satisfy a statistical isometry property,” *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 2, pp. 358–374, Apr. 2010.

[2] W. U. Bajwa, R. Calderbank, and S. Jafarpour, “Why Gabor frames? Two fundamental measures of coherence and their role in model selection,” *J. Commun. Netw.*, vol. 12, no. 4, pp. 289–307, Aug. 2010.

[3] L. Gan, C. Ling, T. T. Do, and T. D. Tran, Analysis of the Statistical Restricted Isometry Property for Deterministic Sensing Matrices Using Stein’s Method 2009 [Online]. Available: http://www.dsp.ece.rice.edu/files/cs/Gan_StatRIP.pdf

[4] E. J. Candes and T. Tao, “Decoding by linear programming,” *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.

[5] G. Caire, T. Y. Al-Naffouri, and A. K. Narayanan, “Impulse noise cancellation in OFDM: An application of compressed sensing,” in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2008, pp. 1293–1297.

[6] Y. K. Han and K. Yang, “New *M*-ary power residue sequence families with low correlation,” in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2007, pp. 2616–2620.

[7] S. Boztas, A. R. Hammons, and P. V. Kumar, “4-phase sequences with near-optimum correlation properties,” *IEEE Trans. Inf. Theory*, vol. 38, no. 3, pp. 1101–1113, May 1992.

[8] Y. Pati, R. Rezaifar, and P. Krishnaprasad, “Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition,” in *Proc. Asilomar Conf. Signals, Syst., Comput.*, Nov. 1993, vol. 1, pp. 40–44.