Perception of Image Characteristics with Compressive Measurements

Jie GUO† , Nonmember, Bin SONG†, a), Member, Fang TIAN†, Haixiao LIU†, and Hao QIN†, Nonmembers

SUMMARY For compressed sensing, to address problems which do not involve reconstruction, a correlation analysis between measurements and the transform coefficients is proposed. It is shown that there is a linear relationship between them, which indicates that we can abstract the inner property of images directly in the measurement domain.

key words: compressed sensing, correlation analysis, measurements, transform coefficients, inner property

1. Introduction

Compressed sensing (CS) enables signal acquisition and reconstruction from incomplete information with high probability provided that the signal is sparsely represented in some basis [1]. In practice, signal recovery is not necessary in many signal processing applications. This has inspired the design of physical systems that directly solve problems in the compressive measurement domain without resorting to a full-scale signal reconstruction. As applied to images or videos, although some methods which only employ random projections to analyze the original image characteristics have been proposed [2], the theoretical analysis and intensive investigation are substantially unexplored. In this work, We find that there is a linear relation of cross-covariance matrices between the measurements and the transform coefficients. What is more, since cross-covariance is a measure of similarity which is able to reflect the smoothness property of an image block, we can directly employ the cross-covariance matrices in the measurement domain to perform image processing tasks, e.g., image segmentation and classification.

2. Linear Relationship between the Cross-Covariance Matrix in Measurement Domain (CMD) and Cross-Covariance Matrix in Transform Domain (CTD)

Signal X is said to be s-sparse with respect to the transform basis Ψ if the number of its non-zero coefficients is s. According to the CS theory, an s-sparse signal can be acquired through the linear random projection

\[ Y = \Phi X \]  

(1)

where \( Y \in \mathbb{R}^{m \times 1} \) is the sampled vector and \( \Phi \) is an \( m \times n \) \((m < n)\) measurement matrix that is incoherent with \( \Psi \). The ratio of \( m/n \) is defined as the measurement rate (MR). The CS theory states that the reconstruction can be formulated as an \( l_1 \) minimization problem by solving: \( \hat{\theta} = \arg \min ||\theta||_1 \) s.t. \( Y = \Phi \Psi \theta = A \theta \), where \( A \) denotes the sensing matrix and \( \theta \) represents the transform domain value of \( X \).

Based on block-based CS [3], after obtaining the measurement vectors \( y_c \) (\( c = 1, 2, \cdots, P \), each has \( m \) samples) block by block for an image, we have the sample matrix \( y \) in the measurement domain

\[ y = [y_1, y_2, \cdots, y_P]^T = [Y_1, Y_2, \cdots, Y_m] \]

(2)

where \( Y_c(w = 1 \cdots m) \) is the column vector of the sample matrix \( y \). Then the cross-variance matrix in the measurement domain \( C_y \) can be expressed as

\[ C_y = \frac{1}{m} [Y_1, Y_2, \cdots, Y_m][Y_1, Y_2, \cdots, Y_m]^T - Y_0 Y_0^T \]

(3)

where

\[ Y_0 = \frac{1}{m} [Y_1 + Y_2 + \cdots + Y_m] \]

\[ = \left[ \frac{1}{m} (A_1 + A_2 + \cdots + A_m) \theta_1 \\ \vdots \\ \frac{1}{m} (A_1 + A_2 + \cdots + A_m) \theta_P \right] \]

(4)

For the reason that the elements of the projection matrix \( A \) follow a Gaussian distribution with mean zero and variance \( 1/m \), by using the central limit theorem we get

\[ \sum_{j=1}^{m} A_{ij} \sim N(0, n) \]

and then deduce that

\[ \frac{1}{m} (A_1 + A_2 + \cdots + A_m) \approx 0 \implies Y_0 Y_0^T = 0 \]

(5)

Note that the value of \( \frac{1}{m} (A_1 + A_2 + \cdots + A_m) \) is closer to 0 with bigger \( m \). For an image or a frame in video sequences, the value of \( m \) is usually dozens or as much as hundreds. By substituting Eq. (5) to Eq. (3), the CMD can be derived as

\[ C_y \approx \frac{1}{m} [Y_1, Y_2, \cdots, Y_m][Y_1, Y_2, \cdots, Y_m]^T \]

\[ = \frac{1}{m} (A_1 \theta_1^T)(A_1 \theta_1^T) = \frac{1}{m} \theta A^T A \theta^T \]

(6)

Recall the work in [4], let\( A = \frac{1}{\sqrt{m}} \begin{bmatrix} a_1, a_2, \cdots, a_m \end{bmatrix}^T \), we have

\[ A^T A = \frac{1}{m} \begin{bmatrix} a_1, a_2, \cdots, a_m \end{bmatrix} \begin{bmatrix} a_1, a_2, \cdots, a_m \end{bmatrix}^T \]
\[
\frac{1}{m} \sum_{i=1}^{m} a_i a_i^T \approx E a a^* = 1
\]  

Then, by substituting Eq. (7) into Eq. (6), CMD can be rewritten as

\[
C_y \approx \frac{1}{m} \theta \theta^T A \theta^T \approx \frac{1}{m} \theta \theta^T
\]

Similarly, we can also get the following cross-covariance matrix \(C_\theta\) in the transform domain as (discrete cosine transform has been chosen, namely DCT, as the transform basis \(\Psi\))

\[
C_\theta \approx \frac{1}{n} \theta \theta^T
\]

Combining Eq. (9) and Eq. (8), we have

\[
C_y \approx \frac{n}{m} C_\theta
\]

By now, the linear relation between the CMD and the CTD has been proved. If we use the cross-covariance matrices, such as image type classification with the \(k\)-means clustering method, the computation complexity can be reduced from \(O(P n k t)\) to \(O(P m k t)\) or even less (note that \(m < n\)), where the \(k\) represents the number of output clusters and \(t\) is the iteration times.

3. Simulation Results

Images, including cameraman BMP (256×256) and the first frame of foreman CIF (352×288), are employed to evaluate the proposed correlation analysis. Each image is split into non-overlapping 16×16 blocks. To construct the cross-covariance matrix, we split each 16×16 block into four 8×8 sub-blocks. All sub-blocks are projected independently using the random Gaussian matrix. A \(C_y \in \mathbb{R}^{4 \times 4}\) is obtained after sampling for each 16×16 block. Meanwhile, since the DCT basis is orthogonal, its inverse matrix equals the transposed matrix. We calculated the transform coefficients as \(\theta_c = \Psi^T x_i\), wherein \(x_i\) is the original pixel domain vector of a sub-block. Then the corresponding \(C_\theta \in \mathbb{R}^{4 \times 4}\) for each 16×16 block is obtained. The linear relation between the CMD and the CTD at different measurement rates is shown below. For simplicity, we just illustrate the simulation results when MR equals 0.2, 0.8 respectively, which are shown in Fig. 1 and Fig. 2. The horizontal axis is the standard deviation values of the CTD \(R_f\) of different blocks. While for the vertical axis: the asterisk point represents the standard deviation values of the CMD \(R_y\) and the circle point denotes \((n/m)R_f\). It can be concluded from the figures that \(R_y \approx (n/m)R_f\).

4. Conclusion

A linear relationship of cross-covariance matrices between measurements and transform coefficients is proposed, which provides us a novel way to analyze the original image characteristics directly in the measurement domain. We can employ the measurement-domain cross-covariance matrices to perform image processing tasks, such as texture classification, feature abstraction, with considerably low computation complexity.

Acknowledgements

This work has been supported by the National Natural Science Foundation of China (61271173, 61372068), the Research Fund for the Doctoral Program of Higher Education of China (20130203110005), the Fundamental Research Funds for the Central Universities (K5051301033), the 111 Project (B08038), and also supported by ISN State Key Laboratory.

References

[1] E. Candès and J. Romberg, “Sparsity and incoherence in compressive sampling,” Inverse Prob., vol.23, no.3, pp.969–985, 2007.
[2] L. Liu and P.W. Fieguth, “Texture classification from random features,” IEEE Trans. Pattern Anal. Mach. Intell., vol.34, no.3, pp.574–586, 2012.
[3] L. Gan, “Block compressed sensing of natural images,” Proc. International Conference on Digital Signal Processing, pp.403–406, 2007.
[4] E.J. Candès and Y. Plan, “A probabilistic and RIPless theory of compressed sensing,” IEEE Trans. Inf. Theory, vol.57, no.11, pp.7235–7254, 2011.