International Journal of Research in Engineering and Science (IJRES)
ISSN (Online): 2320-9364, ISSN (Print): 2320-9356
www.ijres.org Volume 5 Issue 4 ǁ Apr. 2017 ǁ PP.01-07

Application of Bayes Compressed Sensing in Image rocessing

Liu BingYao[1], Lei JuYang[2], Geng YingBo[3]

[1,2]*(College of Mechanical Engineering, Shanghai University of Engineering Science, China)*

PROJRCT NUMBER: 16KY0111

ABSTRACT: As the demand for information increases, signal bandwidth carrying messages becomes increasingly wider, requirements on acquisition rate and processing rate become increasingly higher and the difficulty in broadband signal processing increases, existing analog-digital converters, transmission bandwidths, software and hardware systems and data storage devices cannot satisfy the needs, acquisition, storage, transmission and processing of signal bring huge pressure. Based on this, a nonparametric hierarchical Bayes learning method of image compression sparse representation was proposed, and a nonparametric hierarchical Bayes mixed factor model under Dirichle process distribution was established specific to the sparsity of geometrical model data of Bayes learning of lower-dimensional signal model, the consistency of subspaces, manifold and analysis of mixed factors, etc. The model can study the low-order Gaussian Mixture Model of high-dimensional image data restricted in low-dimensional subspaces, obtain the number of mixed factors and factors through automatic learning of given data set and use it as prior knowledge of data for the reestablishment of image compressed sensing. Effectiveness of the model was analyzed through simulation experiment.

Keywords: sparsification, nonparametric, hierarchical model, Bayes learning, mixed factor analysis, compressed sensing

I. INTRODUCTION

It is well known, with the rapid development of mobile communication, 2G speech communication has translated into 3G/4G data communication; data information such as signal and image play an increasingly important part in medical science, internet, military communication and daily life, etc. People have higher requirements on the quality of information. [1]Therefore, basic framework of signal processing, higher requirements on acquisition rate and processing rate and bigger difficulties in broadband signal processing impose huge pressure on acquisition, storage, transmission and processing of signal. Existing analog-digital converters, transmission bandwidths, hardware and software systems and data storage devices are heavily challenged by how to process plenty of high-dimensional data contained in image information rapidly and accurately. Thus, how to translate high-dimensional data contained in image information into low-dimensional data and avoid loss of interesting information contained in high-dimensional data, learn a kind of low-dimensional single model and express high-dimensional data using less data characteristics become the core of signal and image processing and the key to process signals and images (such as compressing, denoising and feature extraction, etc.) efficiently and accurately. [2]

II. COMPRESSED SENSING BASED ON BAYES ESTIMATION

2.1 Problem description

Suppose f is $N \times N$ image signal, ψ is the sparse transformation, W is the projection coefficient in the transform domain (it is equal to $N \times N$, it is 0 in most cases), that is, $f = \psi w$. Projection matrix $\Phi' = \{\phi_1^{'}\ \phi_2^{'}\ \phi_3^{'} \cdots \phi_N^{'}\}$ is $M \times N$, where, $M \langle\langle N$, then, the projection process of signal can be expressed as:

$y = \Phi' f + n$ (1) or $y = \Phi w + n$ (2)

Where, $\Phi = \Phi'\Psi$; n is the sum of internal and external noise and is subject to Gaussian distribution. Whereas $M \langle\langle N$, summation is unavailable using underdetermined system of equations (2). Suppose sparsification of w , $\Phi = \Phi'\Psi$ and meets the restricted isometry property (RIP), the second-best solution can be figured out using the solution of L_1 , that is: $\hat{w} = \arg\min\{\|y - \Phi w\|_2^2 + \tau\|w\|_1\}$, (3)

2.2Bayes model

For the perspective of Bayes model, all unknown parameters are construed as satisfying the random distribution of certain prior information. Whereas projection signal y will be affected by internal and external noise, suppose it is subject to Gaussian distribution, whose variance is $\sigma^2 = 1/\beta$ and mean value is Φw , that is, $p(y|w,\beta) = N(y|\Phi w, \beta^{-1})$, (4). According to statistics, the conjugate distribution of inverse variance of Gaussian distribution is Gamma distribution. For the convenience of subsequent calculation, suppose parameter $\sigma^2 = 1/\beta$ is subject to Gamma distribution, that is, $p(\beta|a^\beta, b^\beta) = \Gamma(\beta|a^\beta, b^\beta)$, (5). According to the thought of maximum posterior probability, introduce Lapras prior distribution in order to maximize the sparsification of vector w , that is, $p(w|\lambda) = \frac{\lambda}{2}\exp(-\frac{\lambda}{2}\|w\|_1)$, (6) model analysis, so hierarchical prior thought is introduced. Whereas the conjugate distribution of the mean value of Gaussian distribution is still Gaussian distribution, $p(w|\gamma) = \prod_{i=1}^{N} N(w_i|0, \gamma_i^{-1})$, (7). Where, $\gamma = (\gamma_1, \gamma_2, \cdots \gamma_N)$. In order to maximize the sparsification of vector w , suppose each w_i is subject to Gaussian distribution of different parameters. Gamma distribution of the inverse varianceis stilled used in the second level prior distribution, that is, $p(\gamma_i|\lambda) = \Gamma(\gamma_i|1, \lambda/2) = \frac{\lambda}{2}\exp(-\frac{\lambda\gamma_i}{2})$, $\gamma_i \geq 0, \lambda \geq 0$. Multiply (7) and (8) and integrate γ , then,

$$p(w|\lambda) = \int p(w|\gamma)p(\gamma|\lambda)d\gamma = \prod_i \int p(w_i|\gamma_i)p(\gamma_i|\lambda)d\gamma_i = \frac{\lambda^{N/2}}{2^N}\exp(-\sqrt{\lambda}\sum_i |w_i|)$$ (9)[3]-[4]Finally, suppose prior distribution of hyper-parameter λ . The distribution proposed should enable enough flexibility and a large variation range of λ . Thus, suppose λ is subject to Gamma distribution, whose mean value and variance are 1 and $2/v$ respectively, $p(\lambda|v) = \Gamma(\lambda|v/2, v/2)$. (10) Through combining the prior function of all hierarchies of the aforesaid equations (4), (5), (6), (7), (8) and (10), $p(w,\gamma,\lambda,\beta,y) = p(y|w,\beta)p(\beta)p(w|\gamma)p(\gamma|\lambda)p(\lambda)$. (11)

2.3Nonparametric estimation

Nonparametric Bayes model is a probability model dispense with parameter hypothesis. HDP model Hierarchical Dirichlet Process is a multilayer form based on Dirichlet Process Mixture Model, Dirichlet Process is about distribution of distribution, and that is, sampling of the process is a random process. Whereas countless disperse probabilities can be obtained from the sampled distribution, it cannot be described using limited parameters. Hierarchical Dirichlet Process model is shown in Fig. 1. Stick-breaking process function of HDP model is described as below:

β - $GEM(\gamma)$ (12)

$\pi_k|\beta$ - $DP(\alpha,\beta)$ (13)

ϕ_e - H (14)

$s_k|s_{k-1}$ - $Multinomial(\pi_{s_{k-1}})$ (15)

$y_k|s_k$ - $F(\phi_e)$ (16)

In (12), γ is $Concentration$ parameter of basic distribution H, they constitute a Dirichlet Process G_0 - $DP(\gamma,H)$. In (13), α is $Concentration$ parameter of Dirichlet Process G_j - $DP(\alpha,G_0)$ based on basic distribution G_0, π_k is an independent distribution in relation to $Dirichlet$ Process of $DP(\alpha,\beta)$. s_k is indicative factor, parameters are subject to distribution G_j and the value is ϕ_k by the probability π_{jk}, F is the distribution function of observed data.

III. Realization of Bayes compressed sensing

3.1 Image sparse representation model

According to calculation and harmonic analysis theory, image $u \in R^{N\times1}$ can be expressed as the linear combination of a set of atoms $\{\phi_i\}_{i\in I}$, atoms are taken as column vectors and constitute the dictionary $\Phi \in R^{N\times1}$, and image u can be expressed as $u = \Phi\alpha$ (16). Where, $I = \{1,......,I\}$ $(I \in N)$, atomic parameter index set is a finite set, $I \geq N$. According to sparse representation theory of image, image u has sparse representation in a proper dictionary Φ, that is, coefficient $\alpha = (\alpha_i)_{i\in I}$ should have only a few nonzero elements, and the number of nonzero elements should be far less than dimensions of image, namely $\|\alpha\|_0 \langle\langle N$. Over-complete sparse representation of image is a new image model which can effectively describe internal structure and features of signal and has been widely applied to denoising, defuzzification and repair, etc. of image.[5]

Selection and design of over-complete dictionary is a key problem of sparse representation theory. Presently, there are three major image sparse representation dictionaries: orthogonal system, frame system and over-complete dictionary. Traditionally, image is represented using non-redundant orthogonal dictionaries (orthogonal system) such as Fourier transform, DTC transform and Wavelet Transform. Modern calculation and harmonic analysis show that redundant frame system is conducive to the formation of sparser representation; meanwhile, redundant system can make noise and error more stable. In case two constants $A > 0$ and $B < 0$ exist in atomic sequence $\{\phi_i\}_{i\in I}$ and any $x \in H$ (Hilbert space) satisfies $A\|x\|^2 \leq \sum_{i\in I}|\langle x,\phi_i\rangle|^2 \leq B\|x\|^2$ (17), then, $\{\phi_i\}_{i\in I}$ constitutes the frame of H, A, B are the frame bounds. If $A = B$, $\{\phi_i\}_{i\in I}$ is a tight frame. There are many tight frames. For example, the cascade of M orthogonal basis can constitute a tight frame, and $A = B = M$, and even wavelets of two sets of orthogonal wavelet basis cascade can constitute the frame of $H = L^2(R^2)$. Besides, *Curevlet* can also constitute the tight frame of $L^2(R^2)$, as well as Wave-Atom, Gabor frame, non-subsample Wavelet Transform, etc. Sparse representation theory indicates that sparser representation of image can be realized through further enhancing redundancy of system and forming over-complete dictionary. Generally, we can obtain the dictionary through combining existing orthogonal base or frames, or designing parameterized generation function, transforming parameters, or learning or training algorithm.

3.2 Realization of algorithm

Orthogonal matching pursuit algorithm (OMP) is a typical representative. Based on improvement of OMP, regularization OMP, various reconstructing algorithms such as segmentation OMP, compressed sampling matching pursuit algorithm and subspace pursuit algorithm have been obtained.

With OMP as an example, the process of greedy compressed sensing algorithms wasbriefly introduced in this paper.[6]

Process of OMP algorithm is shown below:

Input: original signal is $X \in R^N$, sparseness is K; measurement matrix is $\Phi \in R^{M \times N}$; measurement vector is $y \in R^M$

Output: reconstructing signal is $\hat{X} \in R^N$.

Initialize all parameters: primary iteration t=1, restoring signal is $\hat{X} = 0$, residual error is $r_0 = y$, index set is $\Lambda_0 = \phi$;

Search index λ_1, find out footer λ corresponding to the maximum value in the inner product of residual error r and measurement matrix φ_i and enable it to satisfy $\lambda_i = \arg \max_{i=1,2,3....N} |< r_{i-1}, \varphi_i >|$;

Renew index set: $\Lambda_i = \Lambda_{i-2} \cup \{\lambda_i\}$, find out the set of reconstructing atoms in the measurement matrix, $\Phi_i = [\Phi_{i-1}, \varphi_{\lambda_i}]$;

According to the Least Squares: $\hat{x}_i = \arg \min_2 \left\| y - \Phi_i \hat{x}_{i-1} \right\|$;

Renew residual error: $r_i = r_{i-1} - \Phi_{\lambda_i} \Phi^+_{\lambda_i} r_{i-1}$, where, $\Phi^+_{\lambda_i}$ is the pseudo-inverse matrix of Φ_{λ_i}, $\Phi^+_{\lambda_i} = (\Phi^T_{\lambda_i} \Phi_{\lambda_i})^{-1} \Phi^T_{\lambda_i}$;

Add an iteration and inspect iteration suspensive condition. If $t > K$, iteration stops; if $t \le K$, repeat step 2;

Output reconstructing signal: signal $\hat{X}$ generated by the position corresponding to X_{λ_τ} is the signal to be reconstructed. [7]

OMP searches the optimal atom in iteration and ensures every residual error renewed is orthogonal to the new atom using Gram-Schmidorthogonalization, which not only ensures the optimal choice of atom and improves the quality of reconstruction, but also ensures the rate of convergence and greatly reduces the processing time of calculation.

OMP adopts the following maximum correlation matching criterion:

$\max\limits_{\theta_j}\left|< r_{i-1},\theta_j >\right|$ Where, r_{i-1} is the residual component of signal, θ_j is the j th atom in matrix Θ . In the following simulation experiment, suppose the original signal is $X \in R^N$, $N = 512$, sparseness is $K = 20$, measurement matrix is $\Phi \in R^{M \times N}$, the times of measurement is $M = 256$, and X is subject to random Gaussian mixed distribution, that is: $x \mid s \sim N(0, R(s))$ Where, $[R(s)]_{n,n} = \sigma_{s_n}^2$, s is a group of ransom variables $s = [s_0, s_1, s_2 \ldots\ldots s_{N-1}]^T$, and s is subject to Bernoulli distribution $Beroulli(p_1)$, p_1 is the probability of $\{s_n\}_{n=0}^{N-1}$ and is equal to 1, $p_1 << 1$, here, $p_1 = K / N = 0.04$.

IV. Analysis of simulation result

Matlab2015a, core i7 processor was adopted and lean figure (256*256) was used for simulation experiment in this paper.

FIG. 1 Image contrast

Compared with the first figure, the last figure has a lot of spotted noise, and the middle two figures have a better visual effect. Based on removal of nonparametric Bayes Gaussian noise, the filtering effect is better.

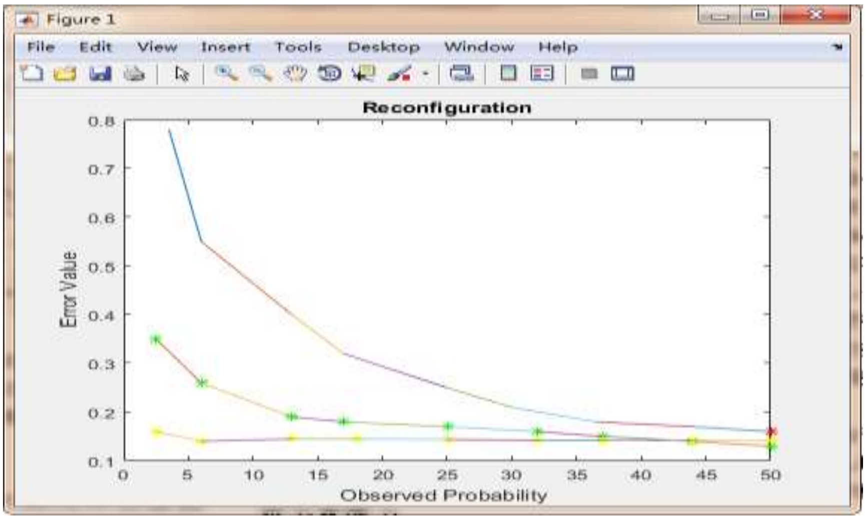

Fig. 2 bcs/ts- bcs/mfacs

Whereas TS- BCS method was used only for structural information of wavelet coefficient and excluded data manifold, and BCS method just estimated hypothesis information and structural information of data was not adopted, accurate results of reconstruction were not obtained. Compressed sensing method based on compressed sensing Bayes model obtained prior assumption probability distribution of reconstruction problem using structural information of data, so accurate reconstruction was obtained. Particularly, the result obtained under a few observed values and based on MFA model was superior to other methods.

V. CONCLUSION

Image processing using sparse nonparametric Bayes compressed sensing reconstructing algorithm overcame the advantage of great Bayes computational burden, improved the calculation efficiency and fulfilled certain practical value. Different from other algorithms, it adopted nonparametric non-supervision self-learning method and automatically added iteration in the sample data training process. Without human interference and through comparison with other algorithms, the superiority of performance was proved and its significant practical value was further verified; follow-up studies can be further optimized specific to the reduction of algorithm complexity.

REFERENCES

[1]. M. Zhou, C. Haojun, P. John, ＆Lu,S. Guillermo, L. Carin, Non. Parametric BayesianDictionary Learning for Sparse Image Representations, in Proe. Neural and InformationProcessing Systems, 2009.

[2]. S. S. Chen, D. L. Donoho, M. A. Saunders, Atomic decomposition by basis pursuit, SIAMJournal on ScientificComputing, V01. 20, No. 1, PP. 33_6 1, 1 998.

[3]. R. Tibshirani, J. IL Statist. Soc. B, Regression shrinkage and selection via the lasso, V01. 58, No. 1, PP. 267—288, 1996.

[4]. D. Manuel, D. Christian, T. Jean-Yves, Classification of chirp signals using hierarchicalbayesian learning and MCMC methods, IEEE Transactions on Signal Processing,V01. 5, No. 2, PP. 377·388,

2002.

[5]. H．Lima，Sparse Bayesian Hierarchical Modeling of high-dimensional cluster problems，Journal ofMultivariate Analysis，V01．101，No．7，PP．1728—1737，2010.

[6]. M．Wang,Y．Shuyuan，W．Yanyan，W．Jing,High Resolution Radar Imaging Based onCompressed Sensing and Fast Bayesian Matching Pursuit,International Workshop onM2RSM，PP．1-5，2011.

[7]. H．Jung,Y．Jong Chul，A sparse Bayesian learning for highly accelerated dynamic MⅪ，IEEE International Symposium on Biomedical Imaging,PP．253-256，2010.