Proceeding of the 2015 IEEE
International Conference on Information and Automation
Lijiang, China, August 2015

A Reconstruction Algorithm based on 3D Tree-Structure Bayesian Compressive Sensing for Underwater Videos *

Xianjian Xiao and Yanbin Zhuang
College of Computer & Information Engineering
Changzhou Institute of Technology
Changzhou, Jiangsu Province, China
{xiaoxj & zhuangyb}@czu.cn

Zunzhi Wang and Xuewu Zhang
College of Computer & Information
University of HoHai
Nanjing, Jiangsu Province, China
553641782@qq.com and zhangxw@hhu.edu.cn

***Abstract* - The traditional underwater video coding has relative high demand of Underwater Acoustic Channel and the underwater video scenes are complex and instable. To deal with these problems, this paper presents a reconstruction algorithm based on three dimension (3D) tree-structure Bayesian compressive sensing for underwater videos. On the encoder side, analog color coded aperture compressive temporal imaging system (CACTI) encodes the video signal. On the decoder end, based on the model of Bayesian compressive sensing, by exploiting the 3D tree structure of the wavelet and Discrete Cosine Transformation(DCT) coefficients, a Bayesian compressive sensing inverse transform algorithm is derived to reconstruct color video frames from single-channel compression measurements. The experimental results show that the algorithm is able to reconstruct complex video scenes more accurately.**

Index Terms* - *Compressive sensing, Underwater videos, 3D tree structure, Wavelet, Bayesian.

I. INTRODUCTION

Underwater detection and operations are the main research topics of lakes and oceans. In deepwater detection , the most commonly used method is video image observation based on underwater robots[1]. Video image observation is the most direct, vivid and effective short distance observation methods. The reason is that video frames have rich details and color information which are easy to be manually recognized and understood. Usually, two measures are taken for collected videos, including compressed access or real-time transmission. To save device storage space, and due to the limitation of underwater acoustic channel bandwidth, the videos must be compressed and encoded with high efficiency. Compressed sensing (CS) theory[2] is a kind of new information acquisition theory breaking the traditional Nyquist/Shannon sampling theory, and it realizes the dimension reduction of signal acquisition at the same time of signal compression. The CS theory demonstrates that if a signal is sparse and compressable, it is possible to losslessly restore the original signal from a few random measurements with great probability. Applying CS to underwater video encode and decode can effectively solve the problems of insufficient memory and transmission bandwidth limitation. The system structure of encoding end is relatively simple, and it moves the complex reconstruction process to the decoding end. The reconstruction algorithm has high complexity, and the demand on the quality and time consumption of video reconstruction is high. Therefore, the main focus of this study is on reconstruction algorithm. Prevalent reconstruction algorithm are categorized into deterministic model method and Bayesian statistical probability model method. Methods based on MP[3] and OMP[4] are deterministic reconstruction methods. However, in many real applications it is difficult use deterministic models to describe interested signals. Therefore, compared with deterministic models, Bayesian statistical probability methods are more reliable[5, 6].

Currently, the study of CS algorithms in video surveillance has achieved important progress, such as traffic video surveillance[7], the feature of which is its objective is clear, and the scene is fixed. The collected videos usually have great time redundancy and the strong frame correlation. The distributed video CS[8] proposed Li-Wei Kang et al. can effectively solve this problem. However, the above algorithm cannot be effectively applied to underwater videos. Because underwater videos are mainly dynamically collected by underwater robots, underwater video monitors, instead of collecting for a fixed scene, since the underwater videos have complex scenes and unfixed, uneven illumination.

For the above characteristics, a CS algorithm for underwater videos is proposed in this study. A coded aperture compressive temporal image (CACTI) system[9] is adopted which uses a camera able to collect low frame rate video measurements, so that the CS encoding of video frames is realized. Based on this, a COLOR-CACTI system is proposed in this study. In the encoding end, the COLOR-CACTI encoding is simulated. In the decoding end, the three-dimensional tree structure of wavelet and DCT coefficients on the basis of Bayesian CS (BCS) framework is used to derive a BCS inverse transform algorithm. The wavelet decomposition is improved to reduce visual redundancy, and the color video frames are reconstructed from single-channel measurements.

II. CS THEORY

Underwater detection and operations are the main research topics of lakes and oceans. In deepwater detection, the most commonly used method is video image observation based on underwater robots [1]. Video image observation is the most direct, vivid and effective short distance observation methods. The reason is that video frames have rich details and

* This work is partially supported by National Natural Science Foundation (61273170) , Jiangsu provincial science and technology industry support program (BE2010072)

color information which are easy to be manually recognized and understood. Usually, two measures are taken for collected videos, including compressed access or real-time transmission. To save device storage space, and due to the limitation of underwater acoustic channel bandwidth, the videos must be compressed and encoded with high efficiency. Compressed sensing (CS) theory[2] is a kind of new information acquisition theory breaking the traditional Nyquist/Shannon sampling theory, and it realizes the dimension reduction of signal acquisition at the same time of signal compression. The CS theory demonstrates that if a signal is sparse and compressable, it is possible to losslessly restore the original signal from a few random measurements with great probability. Applying CS to underwater video encode and decode can effectively solve the problems of insufficient memory and transmission bandwidth limitation. The system structure of encoding end is relatively simple, and it moves the complex reconstruction process to the decoding end. The II. CS THEORY

The observation model of CS theory can be described as:

$$y = \Phi f + e = \Phi \Psi^T x + e = A x + e \quad (1)$$

where

- f is the signal to be sampled with a length of N;
- Φ is the observation matrix of $M \times N$ ($M<<N$);
- e is a noise vector;
- x is the expansion coefficients of signal f under sparse transformation matrix Ψ, and only has K non-zero elements ($K<<N$);
- A is the CS information operator.

III. COLOR-CACTI

A. CACTI

CACTI system uses translated binary template to repeatedly modulated video sequence in the measuring time Δt. The speed of video sequence modulation determines the number of high speed video frames reconstructed from encoding exposure measured values. Assuming the modulation speed is n_t times per second, the binary template will be translated by n_t pixels in the measuring time Δt. If the camera collects the compressed measurement at a speed of $1/\Delta t$ frames per second (fps), and n_t video frames should be reconstructed from each compressed measurement, then the encoding speed must be $n_t / \Delta t$ fps. The LCoS modulator[10, 11] can perform exposure encoding at a speed of 3000 fps. However each pixel of encoded images is continuously changing at the time of exposure. Therefore, it consumes large amounts of energy (>3W). However, the CACTI system[9] uses the mechanical modulator, where the periodic translation binary template is used. Therefore, the energy consumption is relatively low ($\sim 0.2W$). The CACTI system has been extended to color video processing[12], which is able to collect R, G, and B color data components. According to appropriate reconstruction algorithms, color video frames can be reconstructed from the monochromatic gray level measurement values.

Fig. 1 shows the encoding/decoding process of COLOR-CACTI. The first row of figure (a) shows the n_t raw RGB video frames; the single-channel color components obtained by mosaic transformation are shown in the second row; the binary templates used to encode video frames are shown in the third row; the modulated frames and a synthetic encoded exposure image are shown in the fourth row. The second row of figure (b) shows the reconstructed data of the color component in the original single-channels. The RGB of the frame estimated by inverse mosaic transform are shown in the first row. Figure (c) shows the color interpolation process by Bayer matrix[13].

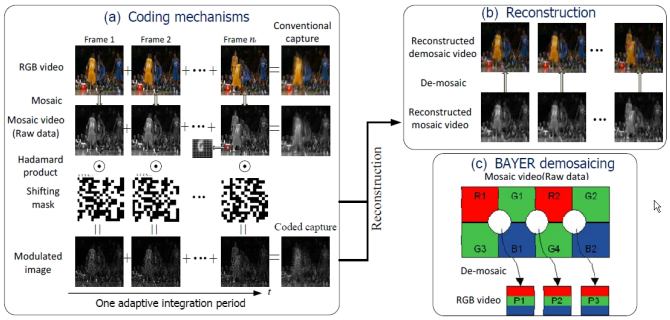

Fig. 1 Encoding/decoding process based on CACTI system.

B. Measurement model

Let $z(x,y,t)$ represent the temporal-spatial analog quantity to be measured. $\phi(x-r(t),y-s(t))$ is the spatial transform of $(r(t),s(t))$ at time t, representing the translated binary template. $h(x,y)$ represents the spatial sampling function of camera with a resolution of $\Delta x \Delta y$. The CACTI system uses translated binary templates to modulate video frame with a period of Δt. Then the modulated video frames are synthesized into a encoded explosure image $(\forall l \geq 1)$:

$$y_{ijl} = \int_0^{\Delta_t} dt \int_0^{n_x \Delta_x} dx \int_0^{n_y \Delta_y} dy \; z(x,y,t+(l-1)\Delta_t) \cdot \phi(x-r(t),y-s(t)) h(x-i\Delta_x, y-j\Delta_y) \quad (2)$$

where

- $i = 1,\ldots,n_x, y = 1,\ldots,n_y$;
- the pixels of camera are $n_x \times n_y$;
- $\{y_{ijl}\}$ is the data set, representing the l-th compressed measurement value, denoted as Y_l afterwards.
- $\phi(x,y)$ is the binary template, representing "pass" and "block" of light (black corresponds to "block")

$z_{ijkl} = z\left(i\Delta_x, j\Delta_y, \frac{k\Delta_t}{n_t} + (l-1)\Delta_t\right)$ represents the sampling of the analog quantity $z(x,y,t)$ of the raw video at (i,j) of the space and time t (n_t discrete time frames, $k = 1,\cdots,n_t$, the l-th time window of compressed measurement values). Define:

$$\phi_{ijk} = \int_0^{n_x\Delta_x}\int_0^{n_y\Delta_y} \phi\left(x - r(\mathrm{t}), \mathrm{y} - \mathrm{s}(\mathrm{t})\right)\Big|_{t=\frac{k\Delta_t}{n_t}} \bullet h\left(x - i\Delta_x, y - \mathrm{j}\Delta_y\right) dxdy \quad (3)$$

Equation (2) can be rewritten as:

$$y_{ijl} = \sum_{k=1}^{n_t} z_{ijkl}\phi_{ijk} + e_{ijl}, \forall i, j \quad (4)$$

$$Y_l = \sum_{k=1}^{n_t} \Phi_k \odot Z_{kl} + E_l \quad (5)$$

where

- e_{ijl} is the noise term;
- $Y_l, \Phi_k, Z_{kl}, E_l \in R^{n_x \times n_y} l$;
- Φ_k represents the *k*-th translation position template;
- Z_{kl} represents the *k*-th video frame data in the CS measurement value l ;

Eq. (4) can be simplified into the following form:

$$y = Hx + e \quad (6)$$

$$H = \left[diag\left(vec\left(\Phi_1\right)\right), \cdots, diag\left(vec\left(\Phi_{n_t}\right)\right)\right] \quad (7)$$

$$x = vec\left(\left[Z_1, \cdots, Z_{n_t} \right]\right) \quad (8)$$

where $y = vec(\mathrm{Y})$ and $vec(\bullet)$ represents the standardization of vectors.

C. Mosaic transform and inversion of color videos

The Bayer matrix is used to record the temporal compressed measurement values of RGB colors. The three processed colors are shown in Fig. 1 (c). An encoded image is composed of four parts, namely, R, B, G1 and G2 (the sizes of which are all 1/4 of the original image). Apply CS to respectively reconstruct each part of the mosaic transformed image, and finally the color video frames are restored by inverse mosaic color interpolation (Fig. 1b). Moreover, if the inter-scale correlation information of wavelet and DCT transform can be obtained, the joint inverse CS transform can be performed for the four parts. The processing results are similar to separately processing R, G, B. That is the main difference between COLOR-CACTI and CACTI[9].

IV. Bayesian Video Reconstruction by Compressive Sensing

A. Three-dimensional tree structure of Wavelet coefficient

With the appearance of wavelet and the profound study of zero-tree structure of images[14], the three-dimensional tree structure of dynamic video has reached more and more attention[15]. Z. Xiong et al. introduced a tree-based representation method into the blocked DCT of JPEG compression[16]. Inspired by this, we propose to introduce three-dimensional tree structure into video processing, i.e., adopting wavelet transform for the video in spatial domain, and applying DCT in temporal domain, which is shown in Fig. 2:

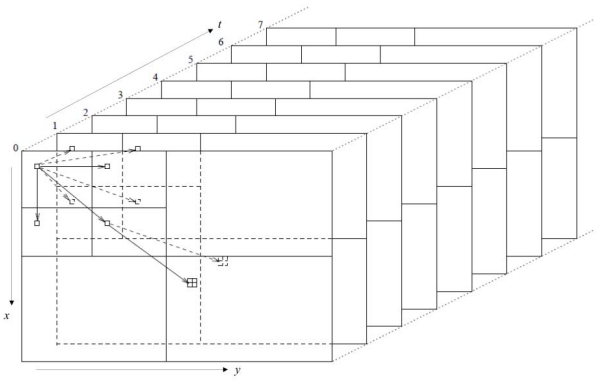

Fig. 2 Three-dimensional tree structure of wavelet coefficients.

Consider a video sequence with n_t frames and spatial pixels being $n_x n_y$. Let (i, j, t) represent the indices of DCT/wavelet coefficients. Assume the decomposition level of the coefficients is L (which is 3 in Fig. 2). The parent-child relationship model of the wavelet coefficients is as follows: (a) a root node (i, j, t) has 7 child nodes, $(i + b_x, j, t)$, $(i, j + b_y, t)$, $(i, j, t + b_t)$, $(i + b_x, j + b_y, t)$, $(i + b_x, j, t + b_t)$, $(i, j + b_y, t + b_t)$, $(i + b_x, j + b_y, t + b_t)$, where (b_x, b_y, b_t) represents the value of coefficients in low-frequency (LL) band; (b) an inner node (i, j, t) has 8 child nodes, $(2i, 2j, 2t)$, $(2i + 1, 2j, 2t)$, $(2i, 2j + 1, 2t)$, $(2i, 2j, 2t + 1)$, $(2i + 1, 2j + 1, 2t)$, $(2i + 1, 2j, 2t + 1)$, $(2i, 2j + 1, 2t + 1)$, $(2i + 1, 2j + 1, 2t + 1)$; (c) leaf node has no child nodes.

Similarly, when tree structure is used for three-dimensional DCT, the size of a block is $\{P_x, P_y, P_t\}$, and $P_x = P_y = P_t = 2^L$. The parent-child relationship model is the same as that of wavelet coefficients.

B. Bayesian statistical model

Based on the full consideration of the relevance of the three-dimensional tree structure of wavelet/DCT coefficients, the parameters of Bayesian statistical model is improved in this study, which are detailedly introduced as follows:

$$\theta = \left(F_t^T \otimes F_y^T \otimes F_x^T\right) x \quad (9)$$

$$\Psi = H\left(F_t \otimes F_y \otimes F_x\right) \quad (10)$$

where

- $F_x \in R^{n_x \times n_x}, F_y \in R^{n_y \times n_y}, F_t \in R^{n_t \times n_t}$ are the orthogonal bases of wavelet/DCT;
- θ represents the coefficients of three-dimensional wavelet/DCT;
- $\otimes$ represents the Kronecker product;
- Ψ is the three-dimensional transform of projection matrix H .

Different from literature [6, 17, 18] where the projection matrix is directly applied to wavelet/DCT coefficients in the model, the projection matrix Φ is obtained by capturing the translated binary template by hardware in the encoding strategy of COLOR-CACTI. Then row transform of H is performed in wavelet/DCT domain to obtain Ψ.

Moreover, the following noise model is used for CS:

$$y \mid \theta, \alpha_0 \sim \mathrm{N}\left(\Psi\theta, \alpha_0^{-1}\mathrm{I}\right) \tag{11}$$

$$\alpha_0 \sim Gamma\left(a_0, b_0\right) \tag{12}$$

where

- a_0, b_0 are both hyper-parameters;
- α_0^{-1} represents the noise variance;
- I is an identity matrix;
- $\mathrm{N}\left(\Psi\theta, \alpha_0^{-1}\mathrm{I}\right)$ represents a Gaussian distribution with $\Psi\theta$ mean value and $\alpha_0^{-1}\mathrm{I}$ variance.

To eastablish the three-dimensional sparse coefficient model of wavelet/DCT, θ must satisfy the *spike-and-slab*[19] distribution:

$$\theta = z \odot w \tag{13}$$

where $w \in R^{n_x n_y n_t}$ is a non-sparse coefficient vector and z is a binary vector representing the two states of hidden Markov tree (HMT) [20]. When the coefficient is in low state, is is set to 0, therefore increasing the sparsity.

To establish the relationship model between the scales and the tree structures of wavelet/DCT, the binary vector z is adopted in this study, which satisfies a Bernoulli distribution with a probability parameter π. Meanwhile, w is modeled to satisfy a Gaussian distribution with 0 mean value and α_ℓ^{-1} variance. The complete Bayesian model is:

$$y \mid \theta, \alpha_0 \sim \mathrm{N}\left(\Psi\theta, \alpha_0^{-1}\mathrm{I}\right) \tag{14}$$

$$\theta = z \odot w \tag{15}$$

$$w_{i,\ell} \sim N\left(0, \alpha_\ell^{-1}\right) \tag{16}$$

$$z_{i,\ell} \sim Bernoulli\left(\pi_{i,\ell}\right) \tag{17}$$

$$\pi_{i,\ell} = \begin{cases} \pi^{\ell}, if\ \ell = 0,1 \\ \pi^{p0}, if\ 2 \le \ell \le L, z_{p\alpha(i,\ell)} = 0 \\ \pi^{p1}, if\ 2 \le \ell \le L, z_{p\alpha(i,\ell)} = 1 \end{cases} \tag{18}$$

where

- $\pi_{i,\ell} i, \ell$ represents the probability that the i-th wavelet coefficient is significant at level ℓ;
- π^{ℓ} represents the significance probability of the parent coefficient in status ℓ.

$$\alpha_0 \sim Gamma\left(a_0, b_0\right) \tag{19}$$

$$\alpha_\ell = \prod_{j=0}^{\ell} \tau_j \tag{20}$$

$$\tau_\ell \sim Gamma\left(c_0, d_0\right), \ell = 0, \cdots, L \tag{21}$$

$$\pi^{\ell} \sim Beta\left(e^{\ell}, f^{\ell}\right), \ell = 0,1 \tag{22}$$

$$\pi^{p0} \sim Beta\left(e^{p0}, f^{p0}\right), 2 \le \ell \le L \tag{23}$$

$$\pi^{p1} \sim Beta\left(e^{p1}, f^{p1}\right), 2 \le \ell \le L \tag{24}$$

In the experiments of this study, relative quantities are expressed as follows:

$$e^0 = 1, f^0 = 0 \tag{25}$$

$$e^1 = 0.9N_\ell, f^1 = 0.1N_\ell \tag{26}$$

$$e^{p0} = \tfrac{1}{N} N_\ell, f^{p0} = \tfrac{N-1}{N} N_\ell \tag{27}$$

$$e^{p1} = 0.5N_\ell, f^{p1} = 0.5N_\ell \tag{28}$$

where N_ℓ represents the coefficients at level ℓ and N stands for the length of θ.

C. Corollary

By improving the variational Bayesian method, the model parameters in literature [18] are improved. Finally, the mathematical expectations of multiple samples are used as the reconstructed wavelet coefficients. The improved τ_ℓ is as follows ($\langle \bullet \rangle$ represents expectation):

$$\left\langle \tau_\ell \right\rangle = \frac{c_0 + 0.5\sum_{j=0}^{\ell} N_\ell}{d_0 + 0.5\sum_{j=0}^{\ell}\sum_{i=1}^{N_j}\left\langle w_{j,i}^2 \right\rangle} \tag{29}$$

D. Algorithm steps

1) Use the COLOR-CACTI system to collect each color component of video sequences, in order to capture video frames at a low frame rate.
2) Perform three-dimensional tree structure transform for 12 frames of the video to obtain the compressed measurement values and parameters i, ℓ;
3) Combine the parameters with the Bayesian model to obtain the improved BCS parameter τ_ℓ. Meanwhile, update the BCS reconstruction algorithm;
4) Use the updated BCS reconstruction algoritm to perform joint reconstruction for the compressed measurement values in (2), and obtain the reconstructed video at a high frame rate for each channel;
5) Perform color interpolation for the reconstructed video frames of each channel to obtain color video.

V. Experimental results and analysis

A. Underwater video simulation

In this study, MATLAB is used to carry out experimental simulation for an underwater video showing a turtle preying. The encoding end simulates the COLOR-CACTI system and encodes the video. It performs temporal compression for the video, which is shown in the left part of Fig. 3. Fig. 3 shows the compressed measurement values of the video at a low frame rate, and the high frame rate video sequence reconstructed from the measurement values. The left part shows the four sequential compressed measurement values, while the right part shows the 22 frames of the video reconstructed from each compressed measurement value.

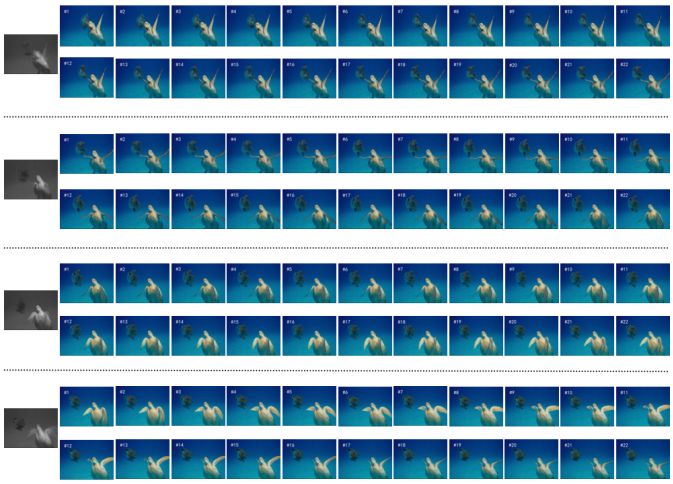

Fig. 3 Reconstruction simulation results of underwater video

It is clear from the figure that the color video frames can be restored from the temporal compressed measurement values of several frames with high quality. The visual effects are also improved to some extent. The comparison of specified parameters is as follows.

B. *Algorithm comparison*

The hyper-parameters a_0, b_0, c_0, d_0 [6, 18] are set to $a_0 = b_0 = c_0 = d_0 = 10^{-6}$. The proposed BCS inverse transform algorithm is compared with the following algorithms: 1) Generalized alternating projection algorithm (GAP) [21]; 2) Two-step iterative algorithm shrinkage/thresholding (TwIST) [22]; 3) K-SVD [23] combined with Orthogonal Matching Pursuit (OMP) inverse transform algorithm; 4) Gaussian mixture model (GMM) inverse transfrom algorithm [25]; 5) Linear Bregman algorithm [26]. The experimental data are shown in Fig. 4.

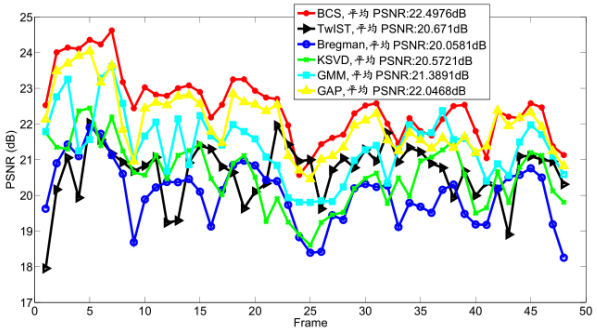

Fig. 4 Comparison of simulation data of algorithms.

It is clear from the data that the proposed BCS inverse transform algorithm achieves relatively satisfactory signal to noise ratios, which are 2dB higher than those of TwIST, K-SVD and Bregman and there is small advantage over GMM and GAP.In this study, the single-channel video measurement values collected by CS cameras at low frame rates are used based on COLOR-CACTI system. The improved Bayesian inverse transform algorithm is adopted to reconstruct high frame rate videos. The Bayesian model can effectively deal with video signals of complex scenes using prior probability models. Meanwhile, the three-dimensional tree structure of wavelet/DCT coefficients is used in this study, and a kind of new BCS inverse transform algorithm is proposed. The simulation and data comparison demonstrate that the proposed algorithm is able to achieve relatively good reconstruction results for underwater videos with relatively complex scenes, and improve the visual redundancy to some extent, having a good application perspective.

VI. Conclusions

In this study, the single-channel video measurement values collected by CS cameras at low frame rates are used based on COLOR-CACTI system. The improved Bayesian inverse transform algorithm is adopted to reconstruct high frame rate videos. The Bayesian model can effectively deal with video signals of complex scenes using prior probability models. Meanwhile, the three-dimensional tree structure of wavelet/DCT coefficients is used in this study, and a kind of new BCS inverse transform algorithm is proposed. The simulation and data comparison demonstrate that the proposed algorithm is able to achieve relatively good reconstruction results for underwater videos with relatively complex scenes, and improve the visual redundancy to some extent, having a good application perspective.

References

[1] Z. Feng, "A review of the development of autonomous underwater vehicles (AUVs) in Western countries," *Torpedo Technology,* vol. 13, no. 1, pp. 5-9, 2005.

[2] David L Donoho, "Compressed sensing," *Information Theory, IEEE Transactions on*, vol. 52, no. 4, pp. 1289-1306, 2006.

[3] S.G. Mallat, and Z. F Zhang, "Matching pursuits with time-frequency dictionaries," *Signal Processing, IEEE Transactions on,* vol. 41, no. 12, pp. 3397-3415, 1993.

[4] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad. "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," In Proc. Conference Record of The Twenty-Seventh Asilomar Conference on, IEEE, 1993, pp. 40-44.

[5] J. Wu, F. Liu, and L. C. Jiao, *et al,* "Compressive sensing SAR image reconstruction based on Bayesian framework and evolutionary computation," *Image Processing, IEEE Transactions on*, vol. No. 7, pp. 1904-1911, 2011.

[6] L. H. He, and L. Carin, "Exploiting structure in wavelet-based Bayesian compressive sensing," *Signal Processing, IEEE Transactions on,* vol. 57, no. 9, pp. 3488-3497, 2009.

[7] X. Wang, Research and Application of the Video Surveillanee Technology on Intelligent Transportation Systems, Northwest University, 2004.

[8] L. W. Kang, and C. S. Lu, "Distributed compressive video sensing," In Proc. ICASSP, IEEE, 2009, pp. 1169-1172.

[9] P. Llull, X. J. Liao, and X. Yuan, *et al,* "Coded aperture compressive temporal imaging," *Optics express*, vol. 21, no. 9, pp. 10526-10545, 2013.

[10] Y. Hitomi, J. W. Gu, and M. Gupta, *et al,* "Video from a single coded exposure photograph using a learned over-complete dictionary," In Proc. ICCV, IEEE International Conference on, 2011, pp. 287-294.

[11] D. Reddy, A. Veeraraghavan, R. Chellappa, "P2C2: Programmable pixel compressive camera for high speed imaging," In Proc. CVPR, IEEE Conference on, 2011, pp. :329-336.

[12] X. Yuan, P. Llull, X. J. Liao *et al*. "Low-Cost Compressive Sensing for Color Video and Depth," In Proc. CVPR, 2014 IEEE Conference on, 2014, pp. 3318-3325.

[13] F. Liu, and P. Du. Research on interpolation algorithm based on Bayer color filter arrays. Chengdu: University of Electronic Science and Technology, 2006.

[14] J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *Signal Processing, IEEE Transactions on,* vol. 41, no. 12, pp. 3445-3462, 1993.

[15] Y. W. Chen, W. A. Pearlman, "Three-dimensional subband coding of video using the zero-tree method," Visual Communications and Image Processing'96, International Society for Optics and Photonics, 1996:1302-1312.

[16] Z. X. Xiong, O. G. Guleryuz, and M. T. Orchard, "A DCT-based embedded image coder," *Signal Processing Letters, IEEE,* vol. 3, no. 11, pp. 289-290, 1996.

[17] R. G. Baraniuk, V. Cevher, and M. F. Duarte, *et al,* "Model-based compressive sensing," *Information Theory, IEEE Transactions on*, vol. 56, no. 4, pp. 1982-2001, 2010.

[18] L. H. He, and H. J. Chen, L. Carin Lawrence, "Tree-structured compressive sensing with variational Bayesian analysis". *Signal Processing Letters, IEEE,* vol. 17, no. 3, pp. 233-236, 2010.

[19] H. Ishwaran, and J. S. Rao, "Spike and slab variable selection: frequentist and Bayesian strategies," *Annals of Statistics*, pp. 730-773, 2005.

[20] M. Crouse, R. D. Nowak, and R. G. Baraniuk, "Wavelet-based statistical signal processing using hidden Markov models", *Signal Processing, IEEE Transactions on*, vol. 46, no. 4, pp. 886-902, 1998.

[21] X. Yuan, J. B. Yang, and P. Llull, et al. "Adaptive temporal compressive sensing for video," Image Processing (ICIP), 2013 20th IEEE International Conference on. IEEE, 2013, pp. 14-18.

[22] J. M. Bioucas-Dias, and M. A. T. Figueiredo, "A new TwIST: two-step iterative shrinkage/thresholding algorithms for image restoration," *Image Processing, IEEE Transactions on*, vol. 16, no. 12, pp. 2992-3004, 2007.

[23] M. Aharon, M. Elad, and A. Bruckstein, "-svd: An algorithm for designing overcomplete dictionaries for sparse representation" *Signal Processing, IEEE Transactions on,* vol. 54, no. 11, pp. 4311-4322, 2006.

[24] J. A. Tropp, "Greed is good: Algorithmic results for sparse approximation," *Information Theory, IEEE Transactions on,* vol. 50, no. 10. pp. 2231-2242, 2004.

[25] K. Xu, K. Qin, and X, Liu, *et al,* "Cloud Transformation Method Based on Gaussian Mixed Model and Its Application to Image Segmentation," *Geomatics and Information Science of Wuhan University*, vol. 38. No. 10, pp. 1163-1166, 2013.

[26] H. Zhang, and L. Cheng, "A- Linearized Bregman Iteration Algorithm," *Mathematica numerica signal,* vol. 6, no. 1, pp. 97-104, 2010.