I-028

Sparse Decomposition of EPI by Using Greedy Pursuit Algorithm

Qiang YAO Toshiaki Fujii
The Graduate School of Engineering, Nagoya University
Furo-cho, Chikusa-ku, Nagoya, 464-8603

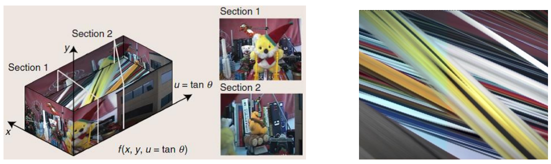
ABSTRACT

EPI (Epipolar Plane Image) is the basic element in representation of 3D image and FTV (Free-view point TV). In order to generate the FTV, a huge amount of data has to be acquired, which has been posing a great challenge for the application of FTV. Numerous methods have been conducted to handle the capturing issue until the concept of compressed sensing was proposed, which has triggered a great revolution in signal acquisition field. Fortunately, the compressed sensing can also be used for capture of Ray space to generate FTV. In addition, for reducing the number of measurements in compressed sensing, it is necessary to analyze the sparsity of signal first, and the special properties of EPI give us a platform to do the special process and analysis to this category of image. In this paper, we focus on the sparse decomposition of EPI, and try to represent EPI by the combination of few atoms in the overcomplete dictionary by greedy pursuit algorithm. The experimental result gives the comparison of the sparse decompositions by using different dictionaries.

Key words Sparse decomposition, Overcomplete dictionary, EPI, Orthogonal Matching Pursuit

1 Introduction

EPI (Epipolar Plane Image) is a significant component in the construction of Ray space [1], which was proposed as early as in 1996 in order to study for the visual 3D communication. The basic idea of Ray space is that the captured views from different positions could be gathered together to form a cubic space which could be viewed as a sealed space with infinite lights going through. EPI can be viewed as the combination of light routes in one intersection of the cubic. Besides, FTV (Free viewpoint TV) [2] can be obtained by vertically cutting the group of EPIs in the Ray space as shown in **Figure 1**. Traditionally, for FTV system generation, camera array is required to capture different views from various angles, thus, a huge amount of data will be generated, which poses a great challenge for data compression and transmission. Recently, the research on the theory and application of compressed sensing [3, 4]has brought great interests in alternatives to traditional capturing conception. Fortunately, compressed sensing theory can be also used in the capture process of FTV system, so that the burden of data could be alleviated greatly in the acquisition process. To be specific, the number of measurements in compressed sensing procedure depends greatly on the sparsity of original signal. Thus, the research and analyze about the sparsity of EPI can provide great assistance in the capture of Ray space and generation of FTV.



(a) Ray Space (b) EPI

Fig 1: The generation of FTV by cutting Ray space

Truly, EPI indeed has several unique features which are not equipped in normal pictures, and one of the best ways to evaluate the sparsity of EPI is the image decomposition[5]. Traditionally, the complete dictionary, also called basis, could provide a good result for signal analysis and decomposition. However, sparser approximation of signal can be obtained by using overcomplete dictionary. In literature, the sparse solution can be roughly classified as Convex Relaxation methods[6], Bayesian Framework[7], and Greedy Pursuit methods.[8, 9]

In this paper, based on the previous research and for simplicity, greedy pursuit methods are employed to decompose EPI sparsely in an overcomplete dictionary. The whole structure of this paper is as follows. In the first section, we give a brief introduction of EPI with its application in FTV generation, and also present the overview of signal sparse decomposition. In the second part, the greedy pursuit methods, including matching pursuit and orthogonal matching pursuit, are mentioned respectively. Meanwhile, the dictionary selection and parameters adjustment are also proposed in this part. Then, the experiment setting and results are presented and discussed in the third section. Finally, several short conclusions and future works are mentioned in last part.

2 Sparsity Decomposition Method

2.1 Greedy Pursuit Algorithm

Greedy pursuit method, including MP (Matching Pursuit) and OMP (Orthogonal Matching Pursuit), can provide a great flexible method in image representation and it can offer a simple way to iteratively decompose the signal along the most significant features. Thus, the signal can be represented as the combination of a group of elements with the corresponding coefficients in the selected dictionary. Besides, the energy of signal can also be concentrated on the smallest number of elements through the method of largest inner product in each iteration.

Generally, the signal is named as f, decomposition space is named as space $\mathcal{H}$, and the element in the overcomplete dictionary is called atom, named as x_i. Besides, the x_i is a unitary vector with the norm of 1. Thus, the first decomposition of f can be represented as $f = \langle f, x_0\rangle x_0 + R^0 f$, where, x_0 is the optimal vector which can obtain the largest inner product with signal f , and $R^0 f$ is the residual of signal f. Clearly, if the iteration continues, it can obtain $R^n f = \langle f, x_n\rangle x_n + R^{n+1} f$. Thus, another representation can be obtained after N iterations that $f = \sum_{i=0}^{N-1} \langle f, x_i\rangle x_i + R^N f$. Besides, considering the energy, we have$\|f\|^2 = \sum_{i=0}^{N-1} |\langle f, x_i\rangle|^2 + \|R^N f\|^2$ to guarantee the energy conservation.

MP is one of the simplest methods of realizing signal decomposing. By using this method, the final residual signal will converge to certain predefined threshold ε. However, if the convergent speed is quite slow, it will lose the advantage of sparse representation. Fortunately, OMP (Orthogonal Matching Pursuit) is an optimal approximation with respect to the selected subset of dictionary, because it ensures that the residual of signal is orthogonal to all the previous selected atoms in each iteration. The main difference between MP and OMP is that there is an update of coefficients in order to always keep the least error, rather than choosing the largest inner product as the coefficient. The specific procedure of orthogonal matching pursuit is illustrated in Algorithm-1

Algorithm-1 OMP

Input: $f_{rec} = 0, R^0 f = f, D = \{\emptyset\}, n = 0,$
$Ite_{num} = K$, X= {all the candidate atoms}

Repeat:

1: Compute inner product between candidate atoms and residual signal { $\langle R^n f, x_i\rangle; x_i \in X$};

2: Find the largest inner product with the corresponding atom x_n,
$\|\langle R^n f, x_n\rangle\| \geq \alpha \sup_i \langle R^n f, x_i\rangle, 0 < \alpha < 1$;

3:Record the atom and inner product in the set of D and C respectively;
$\mathrm{D}_{n+1} = D_n \bigcup x_n, C_{n+1} = C_n \bigcup \langle R^n f, x_n\rangle$,

4: Find the optimal solution $\hat{C}_{n+1}$for $argmin\{e = \|f - D\hat{C}_{n+1}\|_{l_2}\}$, update the coefficient set $C_{n+1} \rightarrow \hat{C}_{n+1}$ and residual signal $R^{n+1} f = f - D\hat{C}_{n+1}, f_{rec} = D\hat{C}_{n+1}$;

Until: $n > K$

Output: $\hat{C}, D$

2.2 Overcomplete Dictionary Selection

Actually, it makes sense that signal is decomposed along the significant feature by using greedy pursuit algorithm. Thus, the selection of overcomplete dictionary is of great importance because sparse decomposition will achieve a quite high efficiency if the chosen dictionary can represent basic features of signal. Dozens of interesting dictionaries have been proposed over the last decades [10, 11], and roughly, theses proposed dictionaries could be classified into frequency dictionaries, time-scale dictionaries, and time-frequency dictionaries with Fourier dictionary, wavelet dictionaries, and Gabor dictionaries as typical examples for each type of dictionary respectively. Later, several kinds of packages were proposed, which could merge various dictionaries together to obtain better performance.

Image decompositions among the Gabor dictionary could characterize the local scale, shift, and orientation of the image variations. Besides, for EPI, it has the explicit features of local fluctuation and orientation. Thus, in this paper, we adopt Gabor dictionary with the parameters of scale, translation and rotation in time domain and modulation and phase in frequency domain. The Gabor function can be written as $g(x, y) = e^{-(x^2+y^2)} cos(2\pi f x + \varphi)$. Here, we would like to do some modification of this mother function. Firstly, due to the specific feature of EPI, it only fluctuates in one direction while it is quite smooth in the another direction. Thus, the contribution of y is supposed to be deleted. Besides, for the edge detection, the second derivative operator is always employed to obtain good performance[12]. Thus, we adopt the second derivative of Gabor function to be the final mother function to generate redundant atoms, shown as $g(x, y) = (4x^2 - 2)e^{-x^2} cos(2\pi f x + \varphi)$. Moreover, the operator of Γ is given by $\Gamma g = g(s^{-1} r_\theta(x - u))$ and also the parameters f and φ are tunable. The parameter set can be represented as $P = \{s, u, \theta, f, \varphi\}$, and numerous of atoms can be generated by tuning these parameters. Basically, the pattern of mother function is illustrated in **Figure 2** with the size of $N = 16$ and parameters of

$s = 4, u = N/2, \theta = -\pi/4, f = 0.2, \varphi = 0.$

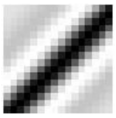


Fig 2: An atom pattern of 2D Gabor dictionary

3 Experimental Results

In this section, firstly, we set the experiment environment, giving out the whole parameter scope of dictionary. Afterwards, we compare the results of 1D decomposition and 2D decomposition. In addition, we also compare the 2D decomposition results of 2D Gabor dictionary and 2D DCT with the same block size. The specific scope of each parameter in 2D Gabor dictionary is shown in **Table-1**,and the complexity of 2D dictionary is $KNlog_2N$.

Table-1 Parameter Set of 2D dictionary

$s = 2^j, 0 \leq j \leq log_2N$
$u = [0, N-1]$
$\theta = -\frac{6\pi}{12}, -\frac{5\pi}{12}, -\frac{4\pi}{12}, -\frac{3\pi}{12}, -\frac{2\pi}{12}, -\frac{\pi}{12}, 0$
$f = 0, 0.1, 0.2, ..., 1.0$
$\varphi = 0, \pi$

Next, the mother function of 1D decomposition is given out, $g_\gamma(x) = Ke^{-x^2}$ where $s = 2^j, 0 < j < log_2N$, and $0 < u < N-1$, and the complexity of 1D dictionary is $Nlog_2N$. The test EPIs are shown is **Figure 3**, with the resolutions of 640x80 and 640x40 respectively. The sparsity of EPI is defined as $S = \frac{T-C}{T}$, where T and C are total number of coefficient and the number of non-zero coefficient, respectively. Thus, the iteration number is set to be 5, 10, 20, and 40 with the block size of 20x20 .

(a) fuzzy EPI

(b) kuma EPI

Fig 3: The Test EPI

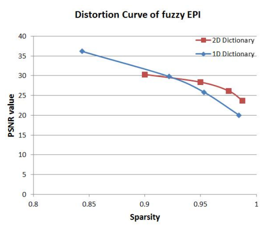
Firstly, we take one block from EPI as an example to illustrate the sparse decomposition. The size of the block is set to be 16 with the parameter set in **Table-1**, and we choose 5 iterations for simplicity to show the approximation. The whole procedure is show in **Figure 4**. The original block is approximated by the combination of these five atoms with the corresponding coefficients.

(a) Original block

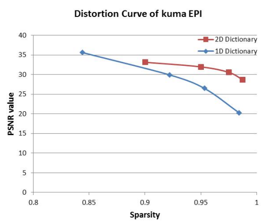
(b) Atom combination

Fig 4: The illustration of 2D decomposition of one block

The comparisons between 1D and 2D dictionary are shown in **Figure 5**. It is clear that the recovery performance by 2D dictionary is almost better than the one by 1D dictionary in the same sparsity. Especially, as the sparsity is quite high, 2D dictionary has outstanding performance.The reasons are basically as follows. For one thing, 2D atoms can present the local feature of signal better, and for another thing, the size of 2D dictionary is larger than 1D dictionary. However, as the decrease of sparsity, the 1D dictionary could achieve better performance, and the probable reason is that 1D dictionary does not consider about the local structure and it can get finer approximation as the increase of atoms.



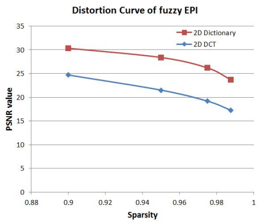
(a) fuzzy EPI



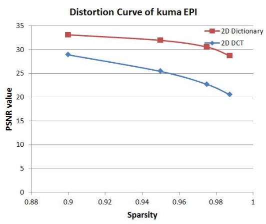
(b) kuma EPI

Fig 5: The comparison between 1D and 2D dictionary

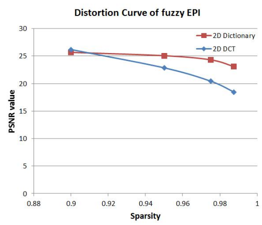
The comparisons between 2D Gabor dictionary and 2D DCT with different block sizes are shown in **Figure 6**. There are several points which are worthy of mentioning. Firstly, as the increase of block size, the performance of 2D DCT is increasing and the main reason is that 2D DCT is a general basis for 2D signal decomposition. Thus, the generality becomes more obvious as the block size increases. Conversely, the performance of 2D Gabor dictionary drops as the expansion of block size and the main reason is that the generated atoms can explicitly represent the features of EPI in a relatively small area. If the size is too large, the local feature will not be obvious, and the performance will decrease. Besides, as the sparsity increases, the 2D Gabor dictionary outperforms 2D DCT mainly due to the large scope of dictionary and the manipulation of parameters.



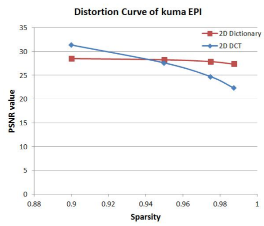
(a) fuzzy EPI(20x20)



(b) kuma EPI(20x20)



(c) fuzzy EPI(40x40)



(d) kuma EPI(40x40)

Fig 6: The comparison between 2D Gabor dictionary and 2D DCT

4 Conclusion

In this paper, the modified Gabor function is adopted to generate the overcomplete dictionary, which is later employed for the sparse decomposition of EPI. Meanwhile, by using greedy pursuit algorithm, the signal could be sparsely decomposed into quite few atoms which represented the main features of the signal. The experimental results show that 2D dictionary performs better than 1D dictionary and also better than the complete basis function, 2D DCT, at high sparsity condition. Meanwhile, we also explain the performance alternate as the change of block size. Next, we will reduce the size of dictionary while preserve the good performance. Besides, the feasibility of EPI sparsity decomposition paves a great way for the implementation of compressed sensing to EPI sparse acquisition for the Ray space reconstruction, which will be also our future work.

References

[1] T. Fujii, T. Kimoto, and M. Tanimoto, “ Ray space coding for 3D visual communication, ”Proc. Picture Coding Symposium (PCS1996), vol.2, pp.447-451, 1996

[2] Masayuki Tanimoto, Mehrdad Panahpour Tehrani, Toshiaki Fujii, and Tomohiro Yendo. “ Free-Viewpoint TV. ”IEEE Signal Processing Magazine, Vol.28, No.1, pp.67-76, 2011.

[3] David L. Donoho. “ Compressed Sensing. ”IEEE Transactions on Information Theory, Vol.52, No.4, pp.1289-1306, 2006.

[4] E. Candes and M. Wakin, “ An Introduction to Compressive Sampling. ”IEEE Signal Processing Magazine, Vol.25, No.2, pp.21-30, 2008.

[5] Toshiaki Fujii, Tadahiko Kimoto, Masayuki Tanimoto, ”Compression and Interpolation of Ray-Space Data Based on Nonorthogonal Expansion of Local EPI Pattern.“ Picture Coding Symposium of Japanese (PCSJ2000), pp.29-30, 2000(In Japanese).

[6] Scott Shaobing Chen, David L. Donoho, Michael, and Saunders.“ Atom decomposition by basis pursuit. ”SIAM Journal on Scientific Computing, Vol.20, pp.33-61, 1998.

[7] Shihao Ji, Ya Xue, and L. Carin,“ Bayesian compressive sensing. ”IEEE Transactions on Signal Processing, Vol.56, No.6, pp. 2346-2356, 2008.

[8] S.G. Mallat and Zhifeng zhang, “ Matching Pursuits with time-frequency dictionaries. ”IEEE Transactions on Signal Processing, Vol.41, No.12, pp. 3397-3415, 1993.

[9] Pati, Y.C. Rezaiifar, R. Krishnaprasad, P.S. “ Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition. ”The Twenty-Seventh Asilomar Conference on Signals, Systems and Computers, Vol.1, pp.40-44, 1993.

[10] M.Yaghoobi, L.Daudet, and M.E.Davies, “ Parametric dictionary design for sparse coding. ”IEEE Transactions on Signal Processing, Vol.57, No.12, pp.4800-4810, 2009.

[11] Ivana Tosic and Pascal Frossard, “ Dictionary Learning. ”IEEE Signal Processing Magazine, Vol.28, No.2, pp.27-38

[12] Pierre Vandergheynst and Pascal Frossard, “ Efficient Image Representation by Anisotropic Refinement in Matching Pursuit. ”2001 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2001. Proceedings. (ICASSP ’01).Vol.3, pp. 1757-1760, 2001.