# HIGH RESOLUTION DYNAMIC MRI USING MOTION ESTIMATED AND COMPENSATED COMPRESSED SENSING

*Hong Jung, Jong Chul Ye*

Bio-Imaging & Signal Processing Lab., Dept. of Bio and Brain engineering
Korea Adv. Inst. of Science & Technology (KAIST)
373-1 Guseong-dong Yuseong-gu, Daejon 305-701, Republic of Korea

## ABSTRACT

A model based approach called k-t BLAST/SENSE, has drawn significant attentions from MR imaging community due to its improved spatio-temporal resolution. Recently, we showed that k-t BLAST/SENSE corresponds to the special case of a new dynamic MRI algorithm called k-t FOCUSS that is asymptotically optimal from compressed sensing perspective. The k-t FOCUSS exploits the sparsity of x-f support of dynamic scene and converts imaging problem into an $L_1$ minimization problem that can be solved using FOCal Underdetermined System Solver (FOCUSS). In this paper, we extend the idea of k-t FOCUSS and introduce motion estimation and compensation (ME/MC) based prediction step and residual encoding step. The ME/MC based prediction step exploits the temporal redundancies using the motion field estimation and provides much sparser residual signals. The sparse residual signal can then be effectively encoded using much smaller number of k-t samples. Simulation results demonstrate that high resolution dynamic MR images can be accurately obtained even from very limited data samples.

***Index Terms—*** compressed sensing, k-t FOCUSS, k-t BLAST/SENSE, MPEG video, motion estimation/compensation

## 1. INTRODUCTION

In dynamic MRI, the spatio-temporal resolution is the most important quality measure. Basically, MRI acquires data on Fourier domain called k-space. Hence, if some data acquisition steps are skipped for high acceleration, aliasing artifacts often appears due to Nyquist sampling limit. In order to resolve this problem, there have been many investigations.

Recently, a model-based approach k-t BLAST/SENSE has been proposed that outperforms the classical dynamic MR methods [1]. Specifically, this uses a *diagonal* form of the signal covariance matrix obtained from training data and impose it as *a priori* information for the acquisition phase.

Another recent development in dynamic MRI has taken place by introduction of the "compressed sensing (CS)" theroy from signal processing community [2]. According to the compressed sensing theory, perfect reconstruction is possible even from sampling rates dramatically smaller than the Nyquist sampling limit, as long as the non-zero spectral signal is sparse and the samples are obtained with *incoherent* basis [2]. Moreover, even if the signal is not sparse, we can still recover the significant features of the signals if the signals are compressible. The optimal sparse solutions can be then obtained using computationally feasible $L_1$ minimization algorithms rather than resorting to computationally expensive combinatorial optimization algorithms [2]. Hence, the compressed sensing theory has great potential.

Interestingly, even though the aforementioned two methods appear drastically different, a close look at the algorithms reveals striking similarity between them. Specifically, Jung et al [3] showed that the *diagonal* signal covariance matrix in k-t BLAST/SENSE is indeed originated from FOCal Underdetermined System Solver (FOCUSS)[4] originally designed to obtain a *sparse* solution by successively solving quadratic optimization problems. Furthermore, they showed that the k-t BLAST/SENSE corresponds to the first iteration of so called k-t FOCUSS that is asymptotically optimal from compressed sensing perspective. The implementation of k-t FOCUSS is so simple that using only a few additional FOCUSS iterations the remaining residual aliasing artifacts of k-t BLAST/SENSE can be effectively suppressed and high spatio-temporal resolution can be achieved [3]. This suggests that k-t BLAST/SENSE indeed has very close relation with the compressed sensing theory even though it was not revealed in original k-t BLAST/SENSE.

The main contribution of this paper is an extension of k-t FOCUSS to a more general framework with prediction and residual encoding, where the prediction approximately estimates the dynamic images and the residual encoding takes care of the remaining residual signals. As a new prediction method, a motion estimation and compensation scheme is proposed, which estimates the correlation between the frames using motion vector similar to video coding. This method significantly sparsifies the residual signal compared to the temporal average often used in k-t BLAST/SENSE or even in

This work was supported by grant No. 2004-020-12 from the Korea Ministry of Science and Technology (MOST).

k-t FOCUSS. Furthermore, using more sophisticated random sampling pattern and optimized temporal transform, the residual signal can be effectively estimated from very small number of random k-t samples by exploiting the sparsity.

# 2. THEORY

## 2.1. Compressed Sensing

Recent theory of compressed sensing tells us that the accurate reconstruction of unknown signal is possible even from sampling rates dramatically smaller than the Nyquist sampling limit as long as the signal is *sparse* in some incoherent basis [2]. Furthermore, the optimal sparse solution can be obtained by solving the $L_1$ minimization [2] rather than complicated combinatorial optimization. Therefore, in order to apply the compressed sensing theory for dynamic MRI, the unknown signal should be sparsified in some basis. We can easily use temporal Fourier transform to make signal sparse for dynamic cardiac MR imaging because heart has periodic motion. Then, the optimal dynamic MR imaging solution from the compressed sensing perspective can be obtained by imposing a sparsity of the solution using $L_1$ norm:

$$\begin{aligned} &\text{minimize} && ||\boldsymbol{\rho}||_1 \\ &\text{subject to} && ||\boldsymbol{v} - \mathbf{F}_y \mathbf{F}_t \boldsymbol{\rho}||_2 \leq \epsilon \end{aligned} \qquad (1)$$

where $\boldsymbol{\rho}$ and $\boldsymbol{v}$ represent sparse x-f image to be reconstructed and k-t measurements, respectively. Here, $\mathbf{F}_y$ and $\mathbf{F}_t$ correspond to Fourier transform along spatial and temporal direction, $\epsilon$ denotes the noise level, $||\cdot||_1$ and $||\cdot||_2$ denote the $L_1$ and $L_2$ norm, respectively.

## 2.2. k-t FOCUSS with Prediction/Residual Encoding

k-t FOCUSS was developed to address the $L_1$ minimization problem (Eq. (1)) using reweighted quadratic optimization technique [3]. This paper generalizes the original formulation in [3] in a more general prediction/residual encoding framework.

More specifically, the unknown signal $\boldsymbol{\rho}$ can be decomposed into the prediction $\boldsymbol{\rho}_0$ and the residual signal $\Delta\boldsymbol{\rho}$:

$$\boldsymbol{\rho} = \boldsymbol{\rho}_0 + \Delta\boldsymbol{\rho} \, . \qquad (2)$$

Our goal is now try to impose the sparsity to the residual signal $\Delta\boldsymbol{\rho}$ rather than the total signal $\boldsymbol{\rho}$. Specifically, we are interested in solving the following compressed sensing problem:

$$\begin{aligned} &\text{minimize} && ||\Delta\boldsymbol{\rho}||_1 \\ &\text{subject to} && ||\boldsymbol{v} - \mathbf{F}\boldsymbol{\rho}_0 - \mathbf{F}\Delta\boldsymbol{\rho}||_2 \leq \epsilon \end{aligned} \qquad (3)$$

where $\mathbf{F} = \mathbf{F}_x \mathbf{F}_t$. The k-t FOCUSS solves Eq. (3) by successively solving reweighted quadratic optimization to find $\Delta\boldsymbol{\rho} = \mathbf{W}\mathbf{q}$.

$$\min ||\mathbf{q}_l||_2, \quad \text{subject to } ||\boldsymbol{v} - \mathbf{F}\boldsymbol{\rho}_0 - \mathbf{F}\mathbf{W}_l\mathbf{q}_l||_2 \leq \epsilon \, . \qquad (4)$$

Then, the update equation at each iteration is given by:

$$\boldsymbol{\rho}_{l+1} = \boldsymbol{\rho}_0 + \boldsymbol{\Theta}_l \mathbf{F}^H \left( \mathbf{F}\boldsymbol{\Theta}_l \mathbf{F}^H + \lambda \mathbf{I} \right)^{-1} (\boldsymbol{v} - \mathbf{F}\boldsymbol{\rho}_0) \, , \quad \text{where } \boldsymbol{\Theta}_l = \mathbf{W}_l \mathbf{W}_l^H \, . \qquad (5)$$

Here, $\mathbf{W}_l$ is the diagonal weighting matrix updated with the solution from the previous step:

$$\mathbf{W}_l = \begin{pmatrix} |\Delta\rho_l(1)|^p & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & |\Delta\rho_l(N)|^p \end{pmatrix}, \; 1/2 \leq p \leq 1 \, . \qquad (6)$$

Note that k-t FOCUSS asymptotically solves $L_1$ minimization problem by setting $p = 0.5$ in Eq. (6) [3]. In k-t FOCUSS, the optimal prediction $\boldsymbol{\rho}_0$ should make the residual signal $\Delta\boldsymbol{\rho}$ as sparse as possible to accelerate the data acquisition. However, to obtain the prediction $\boldsymbol{\rho}_0$, we need additional k-t samples. This illustrates the fundamental trade-off between the prediction and residual encoding, which should be carefully studied for best performance.

Such trade-off between prediction/residual encoding has an important parallel in video compression such as MPEG (Motion Picture Expert Group) [5], which exploits the temporal redundancies to compress the video. A simplified coding structure of MPEG without bidirectional (**B**) picture is constituted in Figure 1 (a). There are two types of frames. First, the intra pictures (**I**) are obtained using discrete cosine transform (DCT) by exploiting the spatial correlation as done in still image compression [5]. Then, the predicted pictures (**P**) are coded using past reference frames including intra pictures and previously predicted pictures. The **P** frame can compress the data significantly by exploiting the temporal redundancies using motion estimation and compensation (ME/MC). The remaining residual signals that cannot be estimated from ME/MC are then encoded using DCT. Now, we can observe the fundamental trade-off between prediction and residual coding. As the reference **I** frame becomes more accurate and the motion estimation accuracy increases using very dense motion vector fields, the residual signal can be significantly sparsified. But, such accurate encoding of **I** and motion fields will introduce additional coding bits. On the contrary, if smaller number of bits are used for prediction, the residual signal becomes large, requiring significant bits for residual coding. Therefore, the optimal bit allocation for **I** frame coding, ME/MC and residual coding has been a big issue in video compression [5].

This paper presents a new prediction/residual encoding scheme based on ME/MC to optimize data allocation between prediction and residual encoding and improve the performance of our k-t FOCUSS [3].

### 2.3. Implementation of Prediction/Residual Encoding

Recall that the $(l+1)$-th update of our k-t FOCUSS framework can be summarized as follows:

$$\boldsymbol{\rho}_{l+1} = \underbrace{\boldsymbol{\rho}_0}_{\text{prediction}} + \underbrace{\boldsymbol{\Theta}_l \mathbf{F}^H \left(\mathbf{F}\boldsymbol{\Theta}_l \mathbf{F}^H + \lambda \mathbf{I}\right)^{-1} (\boldsymbol{\upsilon} - \mathbf{F}\boldsymbol{\rho}_0)}_{\text{residual encoding}}, \quad (7)$$

where $\boldsymbol{\Theta}_l = \mathbf{W}_l \mathbf{W}_l^H$. This section explains how to optimize the prediction and residual encoding to improve the reconstruction quality of the k-t FOCUSS.

#### *2.3.1. Motion Estimation / Motion Compensation*

In section 2.2, the trade-off between prediction encoding and residual encoding in video coding was briefly mentioned. In video coding, the number of motion vectors for ME/MC should be limited since these information should be also transmitted using additoinal bits. However, in dynamic MRI we are free to choose many number of motion vectors for ME/MC method as long as it can predict $\boldsymbol{\rho}_0$ more accurately.

In order to exploit ME/MC, at least one reference frame is required. The concept of reference frame is not new in dynamic MRI. For example, Liang et al already proposed dynamic imaging method called RIGR (Reduced-encoding Imaging by Generalized-series Reconstruction) that exploits the two fully encoded reference frames [6].

In our framework, the reference frame/frames can do the same role with $\mathbf{I}$ frames in MPEG coding. However, in order to calculate motion vectors more accurately, dynamic frames that have high correlation with the reference frames are necessary. However, for highly accelerated MR acquisition, direct Fourier inversion of the down-sampled k-t sample does not provide high resolution dynamic frames, which makes the motion estimation inaccurate. Interestingly, this problem can be overcome by recursively applying our k-t FOCUSS algorithm. More specifically, the original k-t FOCUSS results that uses temporal average as a prediction can be used first to provide intermediate quality reconstruction. Then, motion estimation is done using the fully sampled references frames and the k-t FOCUSS reconstructed frames. Using the estimated motion vectors, we apply another step of k-t FOCUSS algorithm with the motion compensated prediction $\boldsymbol{\rho}_0$. Such recursive application of k-t FOCUSS improves the reconstruction quality significantly.

In order to obtain the motion vectors, Mean Absolute Difference (MAD) between the specified blocks of the reference frames and dynamic frames is calculated as shown in Figure 1 (b). When the search area is determined, the motion vectors for each blocks on individual dynamic frame are calculated by minimizing the MAD as shown in Figure 1 (b). These processes to obtain the motion vectors are called ME. Then, the dynamic frames are newly estimated during MC using the estimated motion vectors. The MC process is done on image domain using only the reference frames and motion vectors. The dynamic images are reconstructed by simply relocating the specified blocks of the reference frames according to the estimated motion vectors. Suppose the motion vector for the $(x, y, t)$ coordinate be $[i, j]$, then MC is following.

$$\boldsymbol{\sigma}_0(x, y, t) = \boldsymbol{\sigma}_{ref}(x+i, y+j)\,, \quad (8)$$

where $\boldsymbol{\sigma}_0$ and $\boldsymbol{\sigma}_{ref}$ indicate the MC reconstruction and reference frames, respectively. Especially, when two reference frames are available, MC can be performed by the linear interpolation of both matched blocks on both reference frames according to the time distance as shown in Figure 1 (c). Suppose that two reference frames are measured on time 0 and $T$, and the motion vectors for a certain pixel positioned on $(x, y, t)$ with respect to each reference frame are $(i_1, j_1)$ and $(i_2, j_2)$, respectively. Then, MC is following.

$$\begin{aligned} & \boldsymbol{\sigma}_0(x,y,t) = \tfrac{t\boldsymbol{\sigma}_{ref1}(x_1,y_1)+(T-t)\boldsymbol{\sigma}_{ref2}(x_2,y_2)}{T} \\ \text{where} \quad & x_1 = x+i_1, y_1 = y+j_1, \\ & x_2 = x+i_2, y_2 = y+j_2, \end{aligned} \quad (9)$$

where $\boldsymbol{\sigma}_{ref1}$ and $\boldsymbol{\sigma}_{ref2}$ represent the measured reference frames on time 0 and $T$.

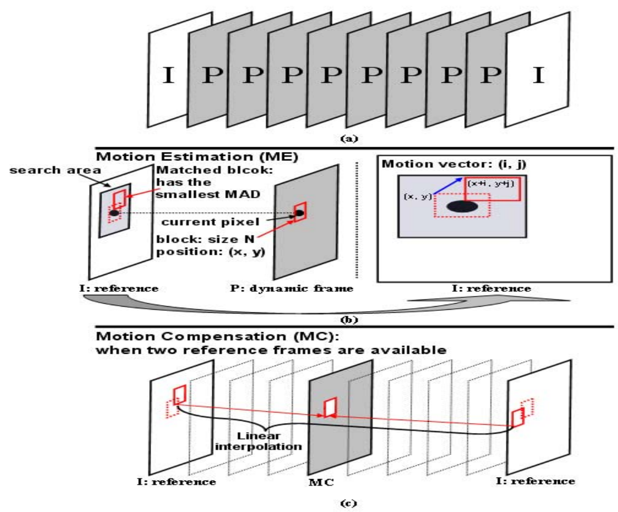


**Fig. 1**. (a) MPEG coding, (b) Motion Estimation (ME), and (c) Motion Compensation (MC)

There are many advantages using ME/MC to obtain the prediction term $\boldsymbol{\rho}_0$. First, the finer structure of dynamic images can be obtained without blurring because MC just relocates the corresponding pixels of the fully sampled reference frames. Second, k-t samples are not fully allocated for prediction step because the estimation is performed on image domain rather than frequency domain. Therefore, the most of the k-t samples can be reused for the following residual encoding step.

#### *2.3.2. Residual Encoding*

The main focus of residual encoding step is to efficiently estimate the residual signal using small number of k-t samples.

This goal can be achieved based on compressed sensing. In order to employ compressed sensing safely, random sampling pattern is used on k-t space resulting in highly incoherent basis as shown in Figure 2.

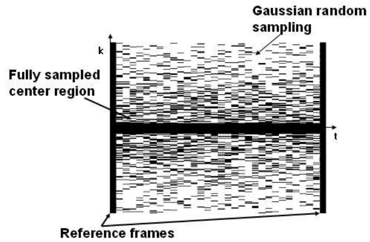


**Fig. 2**. Random sampling pattern

Since the prediction step provides sparse residual, k-t FOCUSS can effectively exploit the sparsity of the residual signal. However, there are still many rooms to improve the residual encoding. In general, the temporal transform $\mathbf{F}_t$ in Eq. (1) is not necessarily Fourier transform. Karhunen-Loeve Transform (KLT) along temporal direction can be also used as an another option of sparsifying transform. As a spatial transform, wavelet transform can be applied along spatial direction as well to make signal much sparser.

## 3. SIMULATION RESULTS

For comparative study, we implemented the new k-t FOCUSS using ME/MC prediction, k-t BLAST [1] and k-t FOCUSS with temporal average [3]. We have acquired 25 frames of full k-space data from a cardiac cine of a patient using a 1.5 T Philips scanner at Yonsei University Medical Center. The field of view (FOV) was $345.00 \times 270.00mm^2$, and the matrix size for scanning was $256 \times 220$, which corresponds to 256 samples in frequency encoding and 220 phase encoding steps. Then, we have compared the reconstruction results for each method at the acceleration factor of 11. Figs. 3(a) and (b) shows the results for k-t BLAST and k-t FOCUSS with ME/MC. k-t FOCUSS with ME/MC (Fig. 3(b)) significantly outperforms the k-t BLAST results in Fig. 3(a). The aliasing artifacts shown in Figs. 3(a)(indicated with white arrows) are greatly reduced in (b). The difference images also demonstrates that the estimation error has been significantly reduced. Furthermore, the calculated mean square error (MSE) in (c) clearly shows the reduced error in k-t FOCUSS with ME/MC prediction, quantitatively. Additionally, we plotted MSE for the result of k-t FOCUSS with temporal average in (c). k-t FOCUSS with ME/MC clearly has the smallest MSE over all frames.

## 4. CONCLUSION

This paper described a generalization of k-t FOCUSS that is optimal from compressed sensing perspective. The k-t FOCUSS reconstruction consists of two terms; one from prediction, and the other from residual encoding. We proposed a new prediction scheme based on ME/MC similar to video. This approach tries to decorrelate the temporal correlation between the frames using motion estimation in image domain.

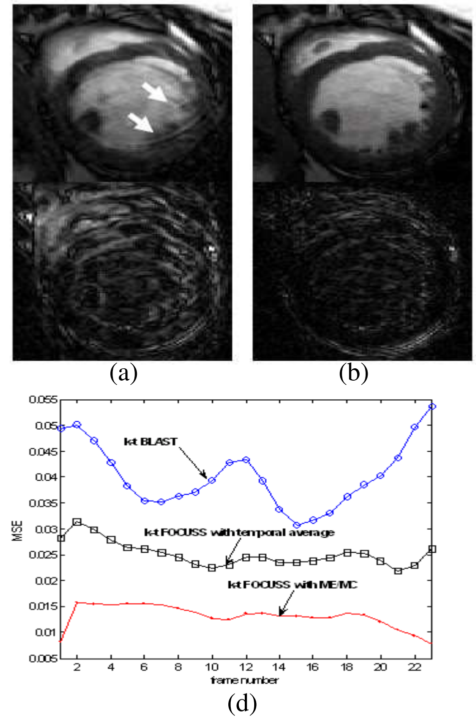


**Fig. 3**. (a) k-t BLAST and (b) k-t FOCUSS with ME/MC. (c) plots MSE for k-t BLAST, original k-t FOCUSS, and k-t FOCUSS with ME/MC.

Experimental results confirmed that high spatio-temporal resolution can be achieved using k-t FOCUSS with ME/MC prediction even from severely undersampled k-t measurements.

## 5. REFERENCES

[1] Jeffrey Tsao, P. Boesiger, and K. P. Pruessmann, "k-t BLAST and k-t SENSE: Dynamic MRI with high frame rate exploiting spatiotemporal correlations," *Magn. Reson. Med*, vol. 50, no. 5, pp. 1031–1042, October 2003.

[2] D. L. Donoho, "Compressed sensing," *IEEE Trans. on Information Theory*, vol. 52, no. 4, pp. 1289–1306, April 2006.

[3] H. Jung, J. C. Ye, and E. Y. Kim, "Improved k-t BLAST and k-t SENSE using FOCUSS," *Phys. Med. Biol.*, vol. 52, no. 11, pp. 3201–3226, 2007.

[4] I. F. Gorodnitsky and B. D. Rao, "Sparse signal reconstruction from limited data using FOCUSS: re-weighted minimum norm algorithm," *IEEE Trans. on Signal Processing*, vol. 45, no. 3, pp. 600–616, March 1997.

[5] D. L. Gall, "MPEG: a video compression standard for multimedia applications," *Communications of the ACM*, vol. 34, no. 4, pp. 46–58, April 1991.

[6] Z. -P. Liang and P. C. Lauterbur, "An efficient method for dynamic magnetic resonance imaging," *IEEE Trans. on Med. Imag*, vol. 13, no. 4, pp. 677–686, 1994.