Improved MRI Mining By Integrating Support Vector Machine Priors In The Bayesian Restoration.

D. A. Karras[1], B.G Mertzios[2], D. Graveron-Demilly[3], D. van Ormondt[4]

[1]Chalkis Institute of Technology, Dept. Automation and Hellenic Open University, Rodu 2, Ano Iliupolis, Athens 16342, Greece, e-mails: dakarras@teihal.gr, dakarras@ieee.org, fax: +30 210 9945231
[2]Thessaloniki Institute of Technology, Thessaloniki, Hellas and Democritus University, Laboratory of Automatic Control Systems, Xanthi, Hellas
[3]Laboratoire de RMN, CNRS, UPRESA 5012, Universite LYON I-CPE, France
[4]Delft University of Technology, Applied Physics Department, P.O Box 5046, 2600 GA Delft, The Netherlands

Abstract. The goal of this paper is to present the development of a new image mining methodology for extracting Magnetic Resonance Images (MRI) from reduced scans in k-space. The proposed approach considers the combined use of Support Vector Machine (SVM) models and Bayesian restoration, in the problem of MR image mining from sparsely sampled k-space, following several different sampling schemes, including spiral and radial. Effective solutions to this problem are indispensable especially when dealing with MRI of dynamic phenomena since then, rapid sampling in k-space is required. The goal in such a case is to make measurement time smaller by reducing scanning trajectories as much as possible. In this way, however, underdetermined equations are introduced and poor image extraction follows. It is suggested here that significant improvements could be achieved, concerning quality of the extracted image, by judiciously applying SVM and Bayesian estimation methods to the k-space data. More specifically, it is demonstrated that SVM neural network techniques could construct efficient priors and introduce them in the procedure of Bayesian restoration. These Priors are independent of specific image properties and probability distributions. They are based on training SVM neural filters to estimate the missing samples of complex k-space and thus, to improve k-space information capacity. Such a neural filter based prior is integrated to the maximum likelihood procedure involved in the Bayesian reconstruction. It is found that the proposed methodology leads to enhanced image extraction results favorably compared to the ones obtained by the traditional Bayesian MRI reconstruction approach as well as by the pure Neural Network (NN) filter based reconstruction approach.

Keywords. MRI Reconstruction, MRI Mining, SVM, MLP, Bayesian Restoration

1. Introduction

A data acquisition process is needed to form the MR images. Such data acquisition occurs in the spatial frequency (**k**-space) domain, where sampling theory determines resolution and field of view, and it results in the formation of the k-space matrix. Strategies for reducing image artifacts are often best developed in this domain. After obtaining such a **k**-space matrix, image reconstruction involves fast multi-dimensional Inverse Fourier transforms, often preceded by data interpolation and re-sampling.

Sampling the **k**-space matrix occurs along suitable trajectories [1,2,3]. Ideally, these trajectories are chosen to completely cover the k-space according to the Nyquist sampling criterion. The measurement time of a single trajectory can be made short. However, prior to initiating a trajectory, return to thermal equilibrium of the nuclear spins needs to be awaited. The latter is governed by an often slow natural relaxation process that is beyond control of the scanner and impedes fast scanning. Therefore, the only way to shorten scan time in MRI when needed, as for instance in functional MRI, is to reduce the overall waiting time by using fewer trajectories, which in turn should individually cover more of k-space through added curvatures. Although, however, such trajectory omissions achieve the primary goal, i.e. more rapid measurements, they entail undersampling and violations of the Nyquist criterion thus, leading to concomitant problems for image reconstruction.

The above mentioned rapid scanning in MRI problem is highly related with two other ones. The first is the selection of the optimal scanning scheme in k-space, that is the problem of finding the shape of sampling trajectories that more fully cover the k-space using fewer trajectories. Mainly three such alternative shapes have been considered in the literature and are used in actual scanners, namely, Cartesian, radial and spiral [1], associated with different reconstruction techniques. More specifically, the Cartesian scheme uses the inverse 2D FFT, while the radial and spiral scanning involve the Projection Reconstruction, the linogram or the SRS-FT approaches [1,2,3].

The second one is associated with image estimation from fewer samples in k-space, that is the problem of omitting as many trajectories as possible without attaining worse reconstruction results. The main result of such scan trajectories omissions is that we have fewer samples in k-space than needed for estimating all pixel intensities in image space. Therefore, there is infinity of MRI images satisfying the sparse k-space data and thus, the image mining problem becomes ill-posed. Additionally, omissions usually cause violation of the Nyquist sampling condition. Despite the fact that solutions are urgently needed, in functional MRI for instance, very few research efforts exist in the literature. The most obvious and simplest such method is the so called “zero-filling the k-space”, that is, all missing points in k-space acquire complex values equal to zero. Subsequently, image mining is achieved as usually, by applying the inverse Fourier transform to the corresponding k-space matrix. Instead of zero-filling the k-space or using linear estimation techniques [2,3] it might be more advantageous to interpolate it by using nonlinear interpolation procedures, like Artificial Neural Networks (ANN). The Bayesian reconstruction approach, developed by two of the authors [1], briefly presented in the next section is another alternative solution. Such a solution could yield good results concerning MR image mining performance [1]. The main contribution, however, of this paper is to develop a novel MR image mining methodology by involving both Bayesian and Neural restoration (based on SVM approximation) techniques and present its competence and advantages over other rival approaches.

2. The Bayesian MRI Restoration Approach.

The Bayesian restoration approach proposed by two of the authors [1], attempts to provide solutions through regularizing the problem by invoking general prior knowledge in the context of Bayesian formalism. The algorithm amounts to minimizing the following objective function [1], by applying the conjugate gradients method,

$$| \underline{\mathbf{S}} - T\,\underline{\mathbf{I}}|^2 / (2\sigma^2) + (3/2) \sum_{x,y} \log \{ \alpha^2 + ({}^{x}\Delta_{xy})^2 + ({}^{y}\Delta_{xy})^2 \} \quad (1)$$

with regards to $\underline{\mathbf{I}}$, which is the unknown image to be reconstructed that fits to the sparse k-space data given in $\underline{\mathbf{S}}$. The first term comes from the likelihood term and the second one from the prior knowledge term of the Bayesian formulation [1]. The parameter σ amounts to the variance encountered in the likelihood term and in the herein conducted simulations is set equal to 1. In the above formula, $T((k_x, k_y),(x,y)) = e^{-2\pi i(xk_x + yk_y)}$ represents the transformation from image to k-space data (through 2-D FFT). The second term symbols arise from the imposed 2D Lorentzian prior knowledge. ${}^{x}\Delta_{xy}$ and ${}^{y}\Delta_{xy}$ are the pixel intensity differences in the x- and y- directions respectively and α is a Lorentz distribution-width parameter. Assuming that P(I) is the prior, imposing prior knowledge conditions for the unknown MRI image, then, the second term of (1) comes as follows.

The starting point is that P(I) could be obviously expanded into $P(I)=P(I_{0,0})\, P(I_{1,0}| I_{0,0})\, P(I_{2,0}| I_{0,0}, I_{1,0})\ldots$ If, now, it is assumed that the intensity $I_{x,y}$ depends only on its left neighbor ($I_{x-1,y}$), then the previous P(I) expansion takes on the form $P(I) = \prod_{(x,y)} P(I_{x,y}| I_{x-1,y})$, provided that the boundaries are ignored. Next, we assume that $P(I_{x,y}| I_{x-1,y})$ is a function only of the difference between the corresponding pixels. This difference is written down as ${}^{x}\Delta_{xy} = I_{x,y} - I_{x-1,y}$. It has been shown that the probability density function of ${}^{x}\Delta_{xy}$ is Lorentzian shaped (see [1,2,3]). These assumptions and calculations lead to computing the prior knowledge in the Bayesian reconstruction as in the second term of (1).

Although this Bayesian restoration approach tackles the problem of handling missing samples in k-space, it exhibits, however, the disadvantage that assumes the existence of special probability distributions, given in closed form descriptions, for representing the unknown ones occurred in MRI, which is an issue under question. In this paper we attempt to remedy this problem by proposing additional priors in the Bayesian formulation in order to capture the probability distribution functions encountered in MRI. These priors are constructed through applying a specifically designed Support Vector Machine (SVM) neural filter for interpolating the sparsely sampled k-space.

3. Design of SVM Neural Network Priors

The method herein suggested for designing efficient Priors for the Bayesian reconstruction formalism, is based on the attempt to extract prior knowledge from the process of filling in the missing complex values in k-space from their neighboring complex values. Thus, instead of assuming a Lorentzian prior knowledge to be extracted from the neighboring pixel intensities in MRI, as a constraint to be applied in the conjugate gradient based Bayesian reconstruction process, the proposed strategy doesn't make any assumption. Instead, it aims at extracting priors without any specific consideration concerning the shape of the distributions involved, by transforming the original reconstruction problem into an approximation one in the complex domain. While linear interpolators have already been used in the literature [2,3], ANN models could offer several advantages when applied as sparsely sampled k-space interpolators. The methodology to extract prior knowledge by applying the ANN filters in MRI reconstruction is described in the following paragraphs.

Step1. We compile a set of R representative N X N MRI images with k-space matrices completely known, which comprise the training set of the SVM approximators. Subsequently, we scan these matrices following the specific sampling schemes mentioned above and then, by randomly omitting trajectories the sparse k-spaces are produced, in order to simulate the real MR data acquisition process.

Step2. The original k-space matrix as well as its corresponding sparse k-space matrix associated with one N X N MRI training image, is raster scanned by a (2M+1) X (2M+1) sliding window containing the associated complex k-space values. The estimation of the complex number in the center of this window from the rest of the complex numbers comprising it is the goal of the proposed approximation procedure. Each position of this sliding window is, therefore, associated with a desired output pattern comprised of the complex number in the original k-space corresponding to the window position, and an input pattern comprised of the complex numbers in k-space corresponding to the rest (2M+1) X (2M+1) -1 window points.

Step3. Each such pattern is then, normalized according to the following procedure. First, the absolute values of the complex numbers in the input pattern are calculated and then, their average absolute value $|z_{aver}|$ is used to normalize all the complex numbers belonging both in the input and the desired output patterns. That is, if z_1 is such a number then this normalization procedure transforms it into the $z_1/|z_{aver}|$. In the case of test patterns we apply the same procedure. That is, the average absolute value $|z_{aver}|$ for the complex numbers z_i of the test input pattern is first calculated. Then, the normalized complex values $z_i/|z_{aver}|$ feed the SVM approximation filter to predict the sliding window central normalized complex number z^{norm}_{centre}. The corresponding unnormalized complex number is simply $z^{norm}_{centre} * |z_{aver}|$.

Step4. The next step is the production of training patterns for the SVM approximators and their training procedure. To this end, by randomly selecting sliding windows from the associated k-spaces of the R training images and producing the corresponding input and desired output training pairs of patterns, as previously defined, we construct the set of training patterns. The assumption underlying such an approach of training SVM approximators is that there are regularities in every k-space sliding window, the same for any MRI image, to be captured by the SVMs without any prior assumption for the probability distributions. SVM training is defined by applying the following procedure.

SVMs is a Neural Network (NN) methodology, introduced by Vapnik in 1992 [5]. They have recently started to be involved in many different classification tasks with success. Few research efforts, however, have used them in nonlinear regression tasks as the MR image mining problem we herein present. One of the goals of the herein study was to evaluate the SVM for Nonlinear Regression approach in such tasks in comparison with other ANN techniques and linear estimation methodologies. The results herein obtained justify that the SVM approach could widely and successfully be involved in function approximation/regression tasks. The task of nonlinear regression using a Support Vector Machine could be defined as follows.

Let $f(\mathbf{X})$ be a multidimensional scalar valued function to be approximated, like the real/imaginary part of the sliding window central normalized complex number z^{norm}_{centre} above defined in step 3. Then, a suitable regression model to be considered is:

$$D = f(\mathbf{X}) + n$$

where, $\mathbf{X}$ is the input vector comprised of (2M+1) X (2M+1) -1 real/imaginary parts of the complex k-space normalized values associated with the window whose central normalized complex value is z^{norm}_{centre}, n is a random variable representing the noise and D denoting a random variable representing the outcome of the regression process. Given, also, the training sample set $\{(\mathbf{X}_i, D_i)\}$ $(i=1,..,N)$ then, the SVM training can be formulated as the optimization problem next outlined:

Find the Lagrange Multipliers $\{\lambda_i\}$ $(i=1, ..,N)$ and $\{\lambda'_i\}$ $(i=1, ..,N)$ that maximize the objective function,

$$Q(\lambda_i , \lambda'_i) = \Sigma_{\iota=1..N} D_i (\lambda_i - \lambda'_i) - e\, \Sigma_{\iota=1..N} (\lambda_i + \lambda'_i) - \tfrac{1}{2}\, \Sigma_{\iota=1..N} \Sigma_{j=1..N} (\lambda_i - \lambda'_i) (\lambda_j - \lambda'_j) K(\mathbf{X}_i, \mathbf{X}_j)$$

subject to the constraints:

$\Sigma_{i=1..N}$ $(\lambda_i - \lambda'_i) = 0$ and $0 <= \lambda_i <= C$, $0 <= \lambda'_i <= C$ for $i=1..N$, where C a user defined constant

In the above definition, $K(\mathbf{X}_i, \mathbf{X}_j)$ are the kernel functions. In the problem at hand we have employed the radial basis kernel

$$K(\mathbf{X}, \mathbf{X}_j) = exp(-1/2\sigma^2 \,||\, \mathbf{X} - \mathbf{X}_j||^2)$$

Taking into account the previous definitions we can then, fully determine the approximating function as

$$F(\mathbf{X}) = \Sigma_{i=1..N} (\lambda_i - \lambda'_i) K(\mathbf{X}, \mathbf{X}_i)$$

which estimates the real and the imaginary part of the complex number z^{norm}_{centre}. Namely, $F_{real}(\mathbf{X}_{real}) = \Sigma_{i=1..N} (\lambda_{i\_real} - \lambda'_{i\_real}) K_{real}(\mathbf{X}_{real}, \mathbf{X}_{i\_real})$ and $F_{imaginary}(\mathbf{X}_{imaginary}) = \Sigma_{i=1..N} (\lambda_{i\_imaginary} - \lambda'_{i\_imaginary}) K_{imaginary}(\mathbf{X}_{imaginary}, \mathbf{X}_{i\_imaginary})$ are the two corresponding SVMs. The former is applied to approximate the real part of z^{norm}_{centre} while the latter for approximating its imaginary part.

Step5. After training phase completion, the SVM filter has been designed and can be applied to any similar test MRI image as follows. To this end, the (2M+1) X (2M+1) sliding window raster scans the sparse k-space matrix associated with this test image, starting from the center. Its central point position moves along the perimeter of rectangles covering completely the k-space, having as center of gravity the center of the k-space array and having distance from their two adjacent ones of 1 pixel. It can move clockwise or counterclockwise or in both directions. For every position of the sliding window, the corresponding input pattern of (2M+1) X (2M+1) – 1 complex numbers is derived following the above described normalization procedure. Subsequently, this normalized pattern feeds the SVM approximator. The wanted complex number corresponding to the sliding window center, is found as $z_{centre} = z_{SVM}^{out} * |z_{aver}|$, where z_{SVM}^{out} is the SVM output and $|z_{aver}|$ the average absolute value of the complex numbers comprising the un-normalized input pattern. For each rectangle covering the k-space, the previously defined filling in process takes place so that it completely covers its perimeter, only once, in both clockwise and counterclockwise directions. The final missing complex values are estimated as the average of their clockwise and counter-clockwise obtained counterparts. The outcome of the SVM filter application is the reconstructed test image, herein named SVM_Img (equation (2) below). Its difference from the image $I^{(t)}$ obtained during the previous step of conjugate gradient optimization in the Bayesian reconstruction formula (1), provides the neural prior to be added for the current optimization step.

4. Incorporation of SVM Neural Prior Knowledge into the Bayesian Formalism

Following the 5 steps above, we can formulate the incorporation of SVM priors to the Bayesian restoration process as follows.

• Design the SVM Neural filter as previously defined

• Consider the Bayesian reconstruction formula (1). The image to be optimized is I given the k-space S. The initial image in the process of conjugate gradient optimization is the zero-filled image. At each step t of the process a different $I^{(t)}$ (the image at the t step, that is, the design variables of the problem) is the result. Based on figure 1 below, by applying the SVM filter on the original sparse k-space, but with the missing points initially filled by the FFT of $I^{(t)}$ (in order to derive the $I^{(t)}$ k-space)- and afterwards refined by the SVM predictions, we could obtain the difference $I^{(t)}$ -SVM_Img$^{(t)}$ as the Neural Prior.

• Therefore, the Neural Network (NN) Prior form, based on SVM approximation is:

$$\sum_{x,y} | \mathrm{SVM\_Im}g^{(t)}(x,y) - I^{(t)}(x,y) | \qquad (2)$$

where, SVM_Img$^{(t)}$(x,y) is the SVM estimated pixel intensity in image space (SVM reconstructed image: Inverse FFT of SVM completed k-space) at step t and $I^{(t)}$(x,y) is the image obtained at step t of the conjugate gradient optimization process in the Bayesian reconstruction

• The proposed Prior in the Bayesian reconstruction is given as

Final Prior = Lorentzian Bayesian Prior + a* SVM_Prior (3)

• That is, the optimization process $I^{(t)}$ is attempted to be guided by the SVM_Img$^{(t)}$ produced by the SVM. Therefore, the proposed algorithm amounts to minimizing the following objective function, by applying again the conjugate gradients method,

$$| \underline{\mathbf{S}} - T\ \underline{\mathbf{I}} |^2 / (2\sigma^2) + (3/2) \sum_{x,y} \log \{ \alpha^2 + ({}^{x}\Delta_{xy})^2 + ({}^{y}\Delta_{xy})^2 \} + a^* \sum_{x,y} | \mathrm{SVM\_Im}g^{(t)}(x,y) - I^{(t)}(x,y) | \qquad (4)$$

where, a=3/2 is the value used for the parameter a above in our experiments.

Including SVM neural knowledge into the Bayesian prior

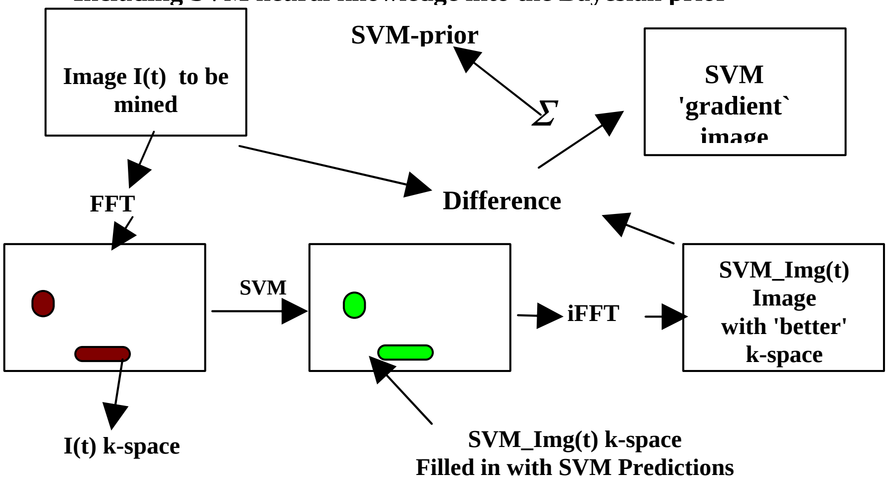

Fig. 1. The difference between the image to be mined I(t) and the SVM mined image SVM_Img(t) constitutes the neural prior.

5. Evaluation Study and Conclusions

An extensive experimental study has been conducted in order to evaluate the above defined novel Bayesian reconstruction methodology. All the methods involved have been applied to a large MRI image database downloaded from the Internet, namely, the Whole Brain Atlas http://www.med.harvard.edu/ AANLIB/home.html (copyright © 1995-1999 Keith A. Johnson and J. Alex Becker). We have used 10 images randomly selected out of this collection for training the SVM approximation filters, and 10 images, again randomly selected for testing the proposed and the rival reconstruction methodologies. All images are 256 by 256. Their k-space matrices have been produced applying the 2D FFT to them. Radial trajectories have been used to scan the resulted 256 X 256 complex k-space arrays. 4 X 256 = 1024 radial trajectories are needed to completely cover such k-spaces. In order to apply the reconstruction techniques involved in this study, each k-space has been sparsely sampled using 128 only radial trajectories. Regarding the sliding window raster scanning the k-space, a 5 X 5 window was the best selection.

Concerning the SVM filter architecture, the 48-17-2 (number of inputs-number of support vectors-number of outputs) one was found after the SVM design stage (step 4, section 3). Actually, as explained in step 4 of section 3 this SVM approximation filter is comprised of two different SVMs (this explains the number two of outputs). The first one is associated with approximating the real part of z^{norm}_{centre} while the second one with approximating its imaginary part. This SVM approximation filter has been trained using 3600 training patterns. The compared reconstruction techniques involved in this study are: the proposed novel Bayesian mining approach, the traditional Bayesian reconstruction technique as well as the SVM filtering approximation technique. In addition, a Multilayer Perceptron (MLP) neural interpolator of 48-12-2 (number of inputs-hidden nodes-number of outputs) architecture (found to be the best one) has been involved in the comparisons. Moreover, the simplest "interpolation" approach, namely filling in the missing samples in k-space with zeroes and then, reconstructing the image, has been invoked. All these methods have been implemented using the MATLAB programming platform.

Concerning the measures involved to quantitatively compare reconstruction performance, we have employed the usually used Sum of Squared Errors (SSE) between the original MRI image pixel intensities and the corresponding pixel intensities of the reconstructed image. Additionally, another quantitative measure has been used, which expresses performance differences in terms of the RMS error in dB [4]:

```
lambda=(image_recon(:)'*image_orig(:))/(image_recon(:)'*image_recon(:));resid
u=image_orig-lambda*image_recon;
dB=10*log10((image_orig(:)'*image_orig(:))/(residu(:)'*
 residu(:)));
```

The quantitative results obtained by the different reconstruction methods involved are outlined in table 1 (average SSE and RMS errors for the 10 test MRI images). Concerning reconstruction performance qualitative results, a sample is shown in figure 2. Both quantitative and qualitative results clearly demonstrate the superiority of the proposed Bayesian image mining methodology embedding SVM filtering based prior knowledge, in terms of MRI image restoration performance over the other rival methodologies (simple Bayesian restoration, SVM / MLP MRI mining filter and zero-filled reconstructions). Future trends of our research efforts include implementation of the 3-D Bayesian reconstruction with Neural Network priors for f-MRI as well as applications in MRI image segmentation for tumor detection.

MRI Mining Method	**SSE (average in the 10 test MRI images)**	**dB (average in the 10 test MRI images)**
Proposed Bayesian MR Image mining with SVM Prior	**2.63 E3**	**17.52**
Proposed Bayesian MR Image mining with MLP Prior	**2.85 E3**	**16.67**
Traditional Bayesian restoration	**3.40 E3**	**15.92**
SVM restoration	**3.27 E3**	**16.02**
MLP restoration	**3.30 E3**	**15.98**
Zero-filling restoration	**3.71 E3**	**15.26**

Table 1. The quantitative results with regards to reconstruction performance of the various methodologies

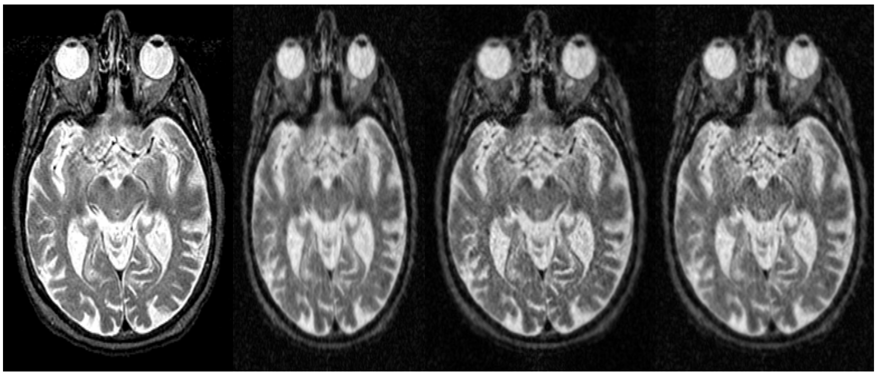

Fig. 2. From left to right: The proposed Bayesian MR Image mining involving SVM priors, the sparsely sampled k-space (nr=128)–zerofilled image reconstruction, the MLP filtering and the traditional Bayesian reconstruction results. The Test Image illustrates a brain slice with Alzheimer’s disease (http://www.med.harvard.edu/ AANLIB/cases/case3/mr1-tc1/020.html).

References

[1]. G.H.L.A. Stijnman, D. Graveron-Demilly, F.T.A.W. Wajer and D. van Ormondt: MR Image Estimation from Sparsely Sampled Radial Scans "Proc. ProRISC/IEEE Benelux workshop on Circuits, Systems and Signal Processing, Mierlo (The Netherlands), 1997" , 603-611

[2]. M.R. Smith, et al.: Application of Autoregressive Modelling in Magnetic Resonance Imaging to Remove Noise and Truncation Artifacts, Magn. Reson. Imaging, 4, 257 (1986).

[3]. P. Barone and G. Sebastiani: A New Method of Magnetic Resonance Image Reconstruction with Short Acquisition Time and Truncation Artifact Reduction, IEEE Trans. Med. Imaging, 11, 250 (1992).

[4]. I. Dologlou, et al.: Spiral MRI Scan-Time Reduction through Non-Uniform Angular Distribution of Interleaves and Multichannel SVD Interpolation, in "Proc. ISMRM, 4th Meeting, N.Y, 1996", p. 1489.

[5]. Haykin S., “Neural Networks: A comprehensive foundation”, Second Edition, Prentice Hall, 1999.