Large Solvent and Noise Peak Suppression by Combined SVD-Harr Wavelet Transform

Daesung Kim, Dai-Gyoung Kim,† Yongwoo Lee,‡ and Hoshik Won*

Department of Chemistry, Hanyang University, Ansan 425-791, Korea
†Department of Applied Mathematics, Hanyang University, Ansan 425-791, Korea
‡R & D Center, Samsung Engineering Co., Ltd, Yongin 449-844, Korea
Received March 26, 2003

By utilizing singular value decomposition (SVD) and shift averaged Harr wavelet transform (WT) with a set of Daubechies wavelet coefficients (1/2, -1/2), a method that can simultaneously eliminate an unwanted large solvent peak and noise peaks from NMR data has been developed. Noise elimination was accomplished by shift-averaging the time domain NMR data after a large solvent peak was suppressed by SVD. The algorithms took advantage of the WT, giving excellent results for the noise elimination in the Gaussian type NMR spectral lines of NMR data pretreated with SVD, providing superb results in the adjustment of phase and magnitude of the spectrum. SVD and shift averaged Haar wavelet methods were quantitatively evaluated in terms of threshold values and signal to noise (S/N) ratio values.

Key Words : NMR, Singular value decomposition, Harr Wavelet, Peak suppression

Introduction

During time domain NMR data collection, unwanted large solvent signals and noise signals are easily included for many reasons, such as instrumental imperfection and the impurity of sample. NMR instruments, including filed gradient system, digital filtering, and many hardware related pulse sequences, have been developed to overcome the suppression of noisy peaks and solvent peaks over the last few decades. On the other hand, many efforts on NMR signal enhancement have been made regarding signal processing techniques and algorithms to improve the S/N ratio and resolution because of the enormous cost of high field NMR instruments. For example, for quite some time, singular value decomposition (SVD) has been developed to obtain linear prediction (LP) coefficients in NMR signal processing.[1,2] In recent years, noise elimination methods[3] and large solvent peak suppression methods[4,5] were developed using the basic properties of SVD. The SVD matrix A for a given NMR data can be represented by $m \times n$ dimensional elements as in equation (1),

$$A = U\Sigma V^T = \sum_{i=1}^{n} u_i \sigma_i v_i^T \qquad (1)$$

where U and V are the square matrices carrying $m \times m$ and $n \times n$ dimensions, respectively. These matrices basically have an orthonormal property, *i.e.*, $U^TU = V^TV = I_n$, and are related to the phase of the matrix A. The capital letter T refers to the transverse, and $\Sigma = diag\ (\sigma_1, \sigma_2, \sigma_3, \cdots, \sigma_n)$ refers to a diagonal matrix where diagonal elements are given by σ_i. The singular values implying the magnitude of the elements of matrix A satisfy the condition $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$. The phase of original NMR data can be altered by the operation of matrixes U and V. In addition, proper adjustment of the matrix Σ enables the suppression of unwanted signals. Although both large solvent peaks and noise peaks in an NMR spectrum can be eliminated by SVD, the degree of noise reduction is relatively poor compared with the results of solvent peak suppression. So, a shift averaged Harr wavelet transform that is widely used in the noise suppression of graphic image processing was adopted to reduce noise peaks for NMR data containing both a large solvent peak and noise in the present study.

Since Morlet and his coworkers introduced the wavelet transform in 1984,[6,7] the many applications of wavelet transform have received great attention in various areas, such as image compression and noise suppression in signal processing. Attempts have been made in the signal analysis of Gaussian type spectral lines and possible applications to NMR spectroscopy.[8,9] The discrete wavelet transform (DWT) method was successfully used in the extraction of dynamic behavior in NMR signals.[10] For convenience, orthogonal wavelet transform, which can easily return to original domain, has been used to eliminate noise signals in image processing and other signals in processing areas.

By utilizing the Harr WT with a set of Daubechies wavelet coefficients (1/2, -1/2) and shift averaging of NMR data in combination with a SVD solvent peak suppression algorithm, a new algorithm of NMR data processing was developed. Great improvement in the S/N ratio of signals, fast processing time, and high resolution was achieved. For the purpose of noise elimination, three steps of operation, such as wavelet transform, noise elimination in wavelet space, and inverse wavelet transform, were basically required.[11] Therefore, forward and backward (inverse) wavelet transform were considered simultaneously in mathematical computation. The subsequent shift signal averaging methods were adopted in this work in

*To whom correspondence should be addressed. e-mail: hswon@hanyang.ac.kr

order to avoid serious peak distortion and incomplete noise suppression near the real NMR peak.

Haar wavelet transform has several advantages in the noise reduction of NMR spectra. First, it is easy to estimate the statistical characteristics of the noise of NMR spectra in the wavelet-domain since Haar wavelet transform is an orthonormal transform. From the computational points of view, Haar wavelet transform provides a simple and efficient transform coding without requiring any boundary treatment. Although the Haar wavelet approximation has the stair-step nature, this feature was significantly reduced through the averaging step in translation-invariant de-noising process.[8]

In noise suppression by wavelet transform, two factors were important for efficiency. The most important factor was a threshold level, and another factor was the number of shifted slices. To find an ideal threshold value automatically, Donoho's universal threshold was introduced by multiplication with weight factor. As shown in equation (2), Donoho and Johnstone suggested the universal threshold (λ) according to statistics.[12]

$$\lambda \cong \sqrt{2\ln N}\hat{\sigma} \qquad (2)$$

Here the $\hat{\sigma}$ is an expected value of standard deviation derived from the equation (3),

$$\hat{\sigma} = \frac{MAD|d_{K-1,j}|_{1 \le j \le 2^{K-1}}}{0.6745} \quad (N = 2^K) \qquad (3)$$

where *MAD* implies a median of given NMR data set.

The SVD method was combined with shift averaged Harr wavelet transform in an attempt to make an automated post processing solution for the best quality of spectrum. Detail results of noise elimination, new algorithms of shifted averaged Harr wavelet, and other results are quantitatively discussed in terms of threshold and signal to noise (S/N) ratio values.

Experiments and Coding

10 mM zinc binding leuteinizing hormone releasing hormone (Zn-LHRH) were dissolved in $DMSO_{-d6}/H_2O$ mixed solvent. The large HOD water peak at off resonance region was used to establish NMR processing algorithms. NOESY experiments were recorded on a Varian UNITY 500 MHz NMR spectrometer with 2048 data points, 256 t1 increments and 300 ms mixing time. The first free induction decay (FID) of NOESY data was selected as a sample NMR signal used in the development of the algorithm.

All of the source coding was accomplished under the IRIX C/C++ compiler of an SGI octane workstation. FID was converted into a HyNMR (in house program) format and double precision text format. The algorithms consist of two major parts, water suppression by SVD and noise filtering using wavelet transform. Both Donoho's universal threshold and uniform threshold functions were included in algorithm development.

Results

Overall processes of algorithm development are summarized in Figure 1. In the first part, SVD was utilized to suppress the large water solvent peak. The time domain FID data were processed to frequency domain by fourier transform, and then phase was corrected in this frequency domain. Toeplitz matrix carrying a size of $m \times (n-m+1)$, where m is an integer rank, was constructed from the frequency domain linear NMR data set having n data points[4] and was then decomposed to obtain a set of eigenvalues, implying the intensity of each peak on the spectrum. After removing the eigenvalues of water peak, Toeplitz matrix was reconstructed by matrix multiplication with two matrices related to the phase of the peaks.

In the second part, shift averaged Harr wavelet transform was utilized to suppress noise peaks. The shifted data matrix was constructed from the Toeplitz matrix obtained at the previous step. After shift averaged Haar wavelet transform, two types of threshold, including Donoho's universal threshold and uniform threshold, were applied to the elimination of noise elements. The noise-filtered matrix can be converted to a noise free pure spectrum by sequential execution of inverse Haar wavelet transform, shift-back operation and signal averaging of all slices. In each step of NMR data processing, S/N ratio values were evaluated for quantitative comparison.

Toeplitz matrix was constructed with a rank of 100 in the present study for the complex 2048 NMR data points. The original NMR spectrum has a large HOD water peak near 712 data points, as shown in Figure 2(A). Four large eigenvalues in front were replaced with zero for the suppression of large water peak. The resulting spectrum of Figure 2(B) obtained by simple SVD treatment gave a S/N ratio of 155. The subsequent suppression of noise by SVD gave a S/N ratio of 180, which exhibits 16% of signal enhancement. As shown in Figure 2 (C) and Figure (D), noise peaks were well eliminated by shift averaged Harr wavelet transform. The uniform threshold method produced more sharp resonance

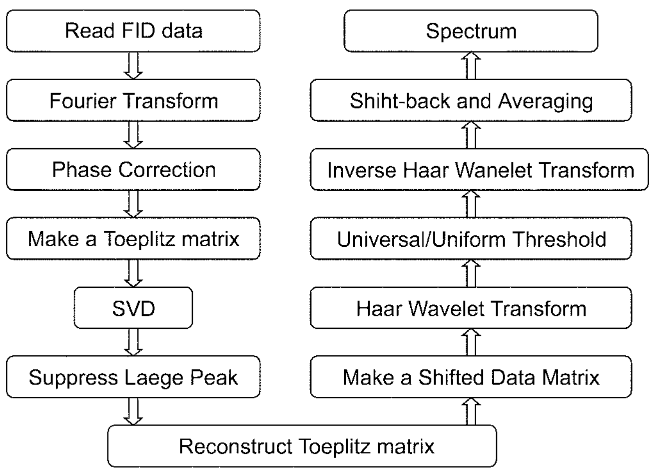

Figure 1. Flow diagram showing overall procedure of SVD and shift averaged Harr wavelet transform.

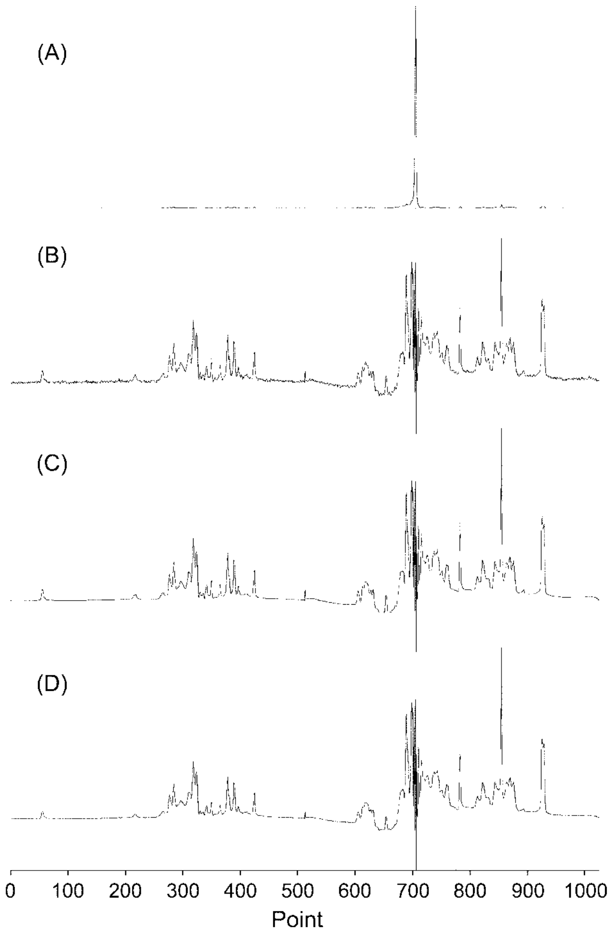

Figure 2. 1D ^{1}H-NMR spectrum obtained from a first FID of NOESY data of Zn-LHRH complex has an off resonance large water peak (A), and spectrum after application of a large solvent peak suppression by using SVD (B), and spectrum after application of noise suppression on the figure-(B) by using uniform threshold (C), and spectrum after treatment of universal threshold method (D).

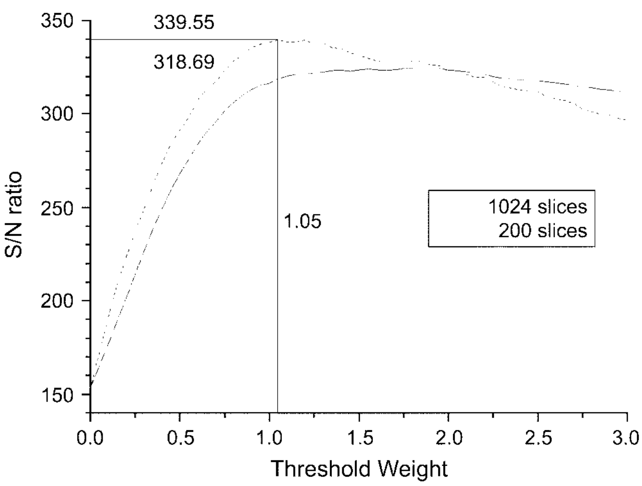

Figure 3. S/N ratio curves as a function of threshold weight value for two different slice values of 200 and 1000. Donoho's universal threshold enables to find the believable optimal threshold level that giving maximum S/N ratio.

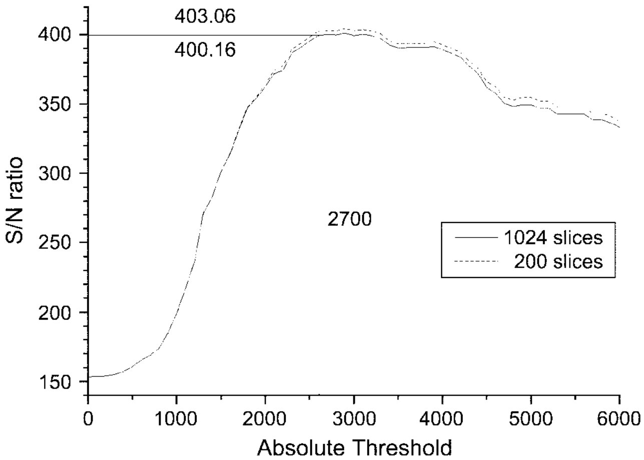

Figure 4. S/N ratio curves as a function of absolute threshold for two different slice values. Uniform threshold gives the nearly overlapped curves between 200 and 1024 shifted slices.

peaks than those of the universal threshold method. S/N ratio values after treatment of the shift averaged Harr wavelet transform increased to 320 for universal threshold and 400 for uniform threshold.

Figure 3 shows the S/N ratio curves as a function of weight factor. The case of 200 shifted slices was more efficient by about 7% than the case of 1024 slices. As shown in Figure 4, when the uniform threshold method was used, the number of shifted slices was not a critical factor because of the fact that the difference of S/N ratio between 200 and 1024 shifted slices was only 0.7%. The S/N ratio value was larger than the universal threshold by about 19% at the most efficient threshold level, 2700.

Conclusion

A new method that can eliminate unwanted large solvent peak and noise from a NMR spectrum by singular value decomposition (SVD) and shift averaged Haar wavelet transform (WT) with a set of Daubechies wavelet coefficients (1/2, -1/2) was described here in detail. By combining the algorithms of simple WT, which can provide superb results in the noise reduction of the Gaussian type NMR spectral lines, into the algorithms of SVD, providing excellent results for adjusting the phase and magnitude of spectrum, we were able to simultaneously eliminate both a large solvent peak and noise.

Although unwanted large solvent peaks and noise peaks can be suppressed by SVD, only 16% of signal enhancement was obtained in noise reduction. However, NMR signals were enhanced up to 157% in terms of S/N ratio by shift averaged Harr wavelet transform. The uniform threshold method produced more sharp resonance peaks than those of the universal threshold method. S/N ratio values after treatment of the shift averaged Harr wavelet transform were 320 for the universal threshold and 400 for uniform threshold, respectively. Since a small number of slices in noise peak suppression causes the wiggle of peaks, the number of slices must be larger than minimum for the correctly shaped peaks. A large number of slices decreases efficiency, and consumes computer resources and computing time. The most reasonable number of slices was about 200. Uniform threshold was more efficient than universal threshold, but the most efficient uniform threshold level could not be directly found. Universal threshold was less efficient than uniform threshold, but has a strong point in that it is applicable to automatic noise suppression.

Combined algorithm of SVD and Haar wavelet transform was optimized with 200 shifted slices and universal threshold providing the best quality of spectrum. It took 370 seconds for both solvent peak and noise suppression by SVD, whereas it took 157 seconds for only the solvent peak suppression. However, it took only 12 seconds for noise elimination by shift averaged Haar wavelet transform under SGI Octane workstation. So, the method used here only consumes 169 seconds in solvent peak and noise suppression for the given NMR data process. Results of computing time and signal enhancement indicate that our new method may be effective in data containing unwanted solvent peaks and noise peaks, and it can be applicable to multidimensional NMR data processes that require good signal enhancement with fast computing time.

References

1. Domoulin, C. L.; Levy, G. C. *Bull. Magn. Reson*. **1984**, *6*, 47.
2. Barkhuijsen, H.; Beer, R. D.; Ormondt, D. V. *J. Magn. Reson*. **1986**, *67*, 371.
3. Fedrigo, M.; Fogolari, F.; Viglino, P.; Esposito, G. *J. Magn. Reson. Series B* **1996**, *113*, 160.
4. Zhu, G.; Smith, D.; Hua, Y. *J. Magn. Reson.* **1997**, *124*, 286.
5. Zhu, G.; Choy, W. Y.; Song, G.; Song, G.; Sanctuary, B. C. *J. Magn. Reson.* **1998**, *132*, 176.
6. Goupillaud, P.; Grossmann, A.; Morlet, J. *Geoexploration* **1984**, *23*, 85.
7. Grossmann, A.; Morlet, J. *SIAM J. Math. Anal.* **1984**, *15*, 723.
8. Delprat, N.; Escudie, B.; Guillemain, P.; Kronland-Martinet, R.; Tchamitchian, P.; Torresani, B. *IEEE Trans. Inform. Theory* **1992**, *38*, 644.
9. Guillemain, P.; Kronland-Martinet, R.; Martens, B. *Wavelets and Applications*; Springer-Verlag: 1991; p 38.
10. Neue, G. *Solid State Nucl. Magn. Reson.* **1996**, *5*, 305.
11. Hansen, P. C.; Sekii, T.; Shibahashi, H. *SIAM J. Sci. Stat. Comput.* **1992**, *13*, 1142.
12. Donoho, D. L.; Johnstone, I. M. *Biometrika* **1994**, *81*, 425.