Supercompact Multiwavelets 을 이용한 CFD 데이터 압축의 Thresholding.

권 도 훈† · 이 도 형**

Thresholding for CFD data compression based on Supercompact Multiwavelets

Dohoon Kwon, Dohyung Lee

Key Words : Supercompact Multiwavelets, Thresholding method(임계처리법), CFD(전산유체역학)

Abstract

CFD data compression method based on supercompact multiwavelets is presented. High data compression can be achieved when taking advantage of the compact nature of multiwavelets. Thresholding technique is also a matter of primary concern in determining pressure ratio. In this paper, we apply thresholding for multiwavelets that considers the coefficient vector as a whole rather than thresholding individual elements. Various thresholding methods are described briefly. CFD data compression suggests that the multivariate thresholding method is suitable for supercompact multiwavelets.

기호설명

H^0, H^1 : Orthonomal wavelets 기저를 사용하여 averaging tranformation 으로 변환시켜주는 가중계수

G^0, G^1 : 가중계수 H^0, H^1 로부터 얻어지는 계수

T^T, T : pre-transformation 과 post-transformation

V : covariance 행렬

λ : 임계값(threshold value)

θ : 참조값(positive reference value)

α_j, r_j : 평균(average) 벡터값과 나머지(residual) 벡터값

$\alpha_0, \alpha_1, \alpha_2$: 이산 근사값(discrete approximation)

† 한양대학교 대학원 기계공학과

* 한양대학교 대학원 기계공학과

E-mail : dohyung@hanyang.ac.kr

TEL : (031)400-5289 FAX : (031)406-5550

1. 서 론

최근 전산유체역학(이하 CFD)은 정확도에 있어서 많은 발전을 거듭하고 있지만, 높은 정확도를 얻기 위해서는 불가피하게 격자(grid)를 많이 생성해야 한다든지, Full Navier-Stokes 방정식을 풀어야 하는 경우가 많이 생기게 되었다. 이러한 경우 해석 data 의 양은 기하급수적으로 증가하게 되고, 결국 계산 시간의 증대 및 data 저장 장치(memory)가 더 많이 필요하게 되는 문제점이 야기된다. 이러한 문제점을 개선하기 위해서 CFD data 압축이 연구되고 있다. 이러한 연구를 위해서는 최근 음성 및 영상, 신호처리 분야에서 많이 쓰이는 wavelets 을 사용하게 된다. 그러나, 영상 및 신호처리에 사용되는 데이터는 거의 모든 곳에서 불연속성을 가지며, 약간의 내재적 매끄러움을 지니는 반면, CFD data 는 거의 모든 곳에서 매끄럽고, 충격파, 와동, 전단층 등의 불연속성을 갖게 되므로 CFD data 의 특성을 고려한 wavelets 을 적용해야 한다. 본 연구에서는 이러한 조건들을 만족시키고 효율적인 CFD data 압축을 수행하기 위해 Supercompact Multiwavelet 을 적용한다. 특히 중요한 것은 data 압축의 실질적인

process 인 임계처리법(thresholding method)이다. 최근 다양한 thresholding method 가 소개되고 있고, 각기 다른 장점을 갖고 있다. Supercompact Multiwavelet 에 적합한 thresholding method 를 알고 적용하는 것은 상당히 중요하며 의미 있는 일이다.

2. Supercompact Multiwavelet

Supercompact Multiwavelet 은 1996 년에 Beam 과 Warming[(8)]에 의해서 개발되었고, 그 후 Dohyung Lee[(1~3)]에 의해 3D 로 확장되었다. Supercompact Multiwavelet 은 다음과 같은 특징을 지닌다. 우선, 높은 차수의 정확도를 유지하면서도 wavelet 계산에 사용되는 support 가 매우 적다는 것이다. 이로 인해 shock 과 같은 불연속면에서의 계산에서도 그 강도의 감소 없이 비교적 정확하게 계산을 수행하게 된다. 또한, 하나의 mother wavelet 이 아닌 여러 종류의 다양한 mother wavelet 을 사용하므로 직교성(orthogonality), 정확성, 대칭성을 갖는다. 그러나, Multiwavelet 을 사용하기 위해서는 기존의 data 영역을 wavelet 영역으로 변환시키는 pre-transformation 과정과 역과정(reverse process)인 post-transformation 과정이 요구되어진다. 따라서, 전체 변환과정은 다음과 같다.

CFD data ➔ Pre-transformation ➔ Decomposition ➔ Thresholding ➔ Reconstruction ➔ Post-trasnformation ➔ CFD data

2.1 Pre- and Post-transformation

Pre- and Post-transformation 에 사용되어지는 식은

$$\vec{\alpha}_j = T^T \vec{u}_j \ , \ \vec{u}_j = T\vec{\alpha}_j \qquad (1)$$

이다.

2.2 분해(Decomposition)와 복원(Reconstruction)

분해과정에서는 다중해상도법(MRA)가 적용되는데, 각각의 단계마다 적용되는 식(2)~(4)은 다음과 같다.

$$\mathbf{R}^{p-1} = \mathbf{L}\vec{\mathbf{a}}^p , \qquad (2)$$

이때, $\vec{\mathbf{a}}^p$ 는 원본 data 를 나타내고, $\mathbf{L}$ 행렬은 분해행렬이 된다. 평균값(average value)들은 $\mathbf{R}$ 행렬의 $\vec{\alpha}_{ijk}$,즉 Fig.1 에서 000 부분에 저장되고, 나머지 값(residual value)들은 나머지 부분에 저장된다. 복원과정은 분해과정의 역으로 구성된다. 또한, H, G 행렬의 직교성에 의해 사용되는 wavelet 행렬은 transpose 를 취함으로써 구한다. 이 과정에 적용되는 식(5)는 다음과 같다.

$$\mathbf{L} = \begin{bmatrix} \mathbf{H} \\ \mathbf{G} \end{bmatrix}, \mathbf{R} = \begin{bmatrix} \vec{\alpha} \\ \vec{\mathbf{r}} \end{bmatrix} = \begin{bmatrix} \vec{\alpha}_{ijk} \\ \vec{r}_{100,ijk} \\ \vec{r}_{010,ijk} \\ \vec{r}_{110,ijk} \\ \vec{r}_{001,ijk} \\ \vec{r}_{101,ijk} \\ \vec{r}_{011,ijk} \\ \vec{r}_{111,ijk} \end{bmatrix} \qquad (3)$$

$$\vec{\mathbf{a}} = \begin{bmatrix} \vec{\alpha}_{2i-1,2j-1,2k-1} \\ \vec{\alpha}_{2i,2j-1,2k-1} \\ \vec{\alpha}_{2i-1,2j,2k-1} \\ \vec{\alpha}_{2i,2j,2k-1} \\ \vec{\alpha}_{2i-1,2j-1,2k} \\ \vec{\alpha}_{2i,2j-1,2k} \\ \vec{\alpha}_{2i-1,2j,2k} \\ \vec{\alpha}_{2i,2j,2k} \end{bmatrix} = \begin{bmatrix} \vec{\alpha}_{000} \\ \vec{\alpha}_{100} \\ \vec{\alpha}_{010} \\ \vec{\alpha}_{110} \\ \vec{\alpha}_{001} \\ \vec{\alpha}_{101} \\ \vec{\alpha}_{011} \\ \vec{\alpha}_{111} \end{bmatrix} \qquad (4)$$

$$\vec{\mathbf{a}}^p = H^T \vec{\alpha}^{p-1} + G^T \vec{r}^{p-1} \qquad (5)$$

3. Thresholding Method

임계법(Thresholding method)의 개념은 잡음(noise)으로 간주되는 중요하지 않은 wavelet 계수를 버리는 것이다. CFD 데이터 압축은 바로 이 과정에서 일어나게 되므로 일련의 어떠한 과정보다도 중요하다. 보통의 압축 과정에서는 적절한 가정에 의하여 분해(Decomposition)된 데이터들 중 사용자가 정한 값보다 작은 값들을 0 으로 생각하여 없애버리고, 0 으로 치환된 부분은 그 위치만을 기억하는 과정을 통하여 데이터를 압축하게 된다. 최근에 소개된 다양한 임계법은 재각기 다른 특징을 갖고 있으며, 그에 맞는 wavelet basis 를 적용할 때 비로소 압축률의 향상을 기대할 수 있다. 본 논문에서 소개하고 있는 supercompact multiwavelet 에 적합한 thresholding method 를 적용하는 것은 의미있는 일

이다.

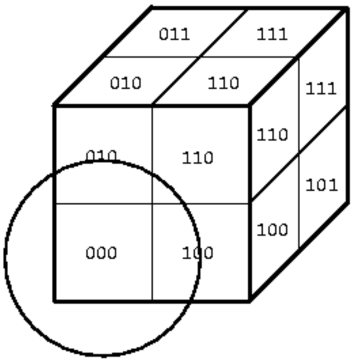


Fig.1 Decomposition 과정 중 3D Subcell 구성.

3.1 표준 임계법(Standard thresholding method)

Donoho[7~8]에 의해 제안된 표준 임계처리방법은 두 가지가 있다. Hard thresholding 과 Soft thresholding 인데 각각 다음과 같이 정의 된다.

$$T^{hard}(x,\lambda) := \begin{cases} x & if\,|x| \ge \lambda \\ 0 & if\,|x| < \lambda \end{cases}, \qquad (1)$$

$$T^{soft}(x,\lambda) := \begin{cases} x - \mathrm{sgn}(x)\lambda & if\,|x| \ge \lambda \\ 0 & if\,|x| < \lambda \end{cases}, \quad (2)$$

임계값(Thresholding value) λ 는 잡음의 표준편차 뿐만 아니라 관측된 표본크기에 따라 결정된다.

Hard thresholding 은 오차의 기대값을 최소화한다는 측면에서 좋은 잡음제거 알고리즘을 제공한다. 그러나, 인위적인 깁스현상이 발생하여 원하는 매끄러움을 갖지 못한다. 그에 비해 Soft thresholding 은 이상적인 축소법(ideal shrinkage method)에 보다 근사한 형태인 연속적인 축소법으로서 매끄러운 조건을 잘 따른다.

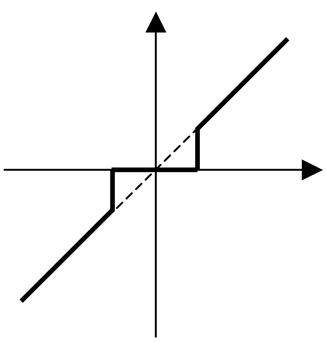


Fig.2 (a) Hard thresholding ; 임계값 $\lambda = 1$ 로 선형함수(점선)에 적용된 축소함수의 결과(굵은 실선)

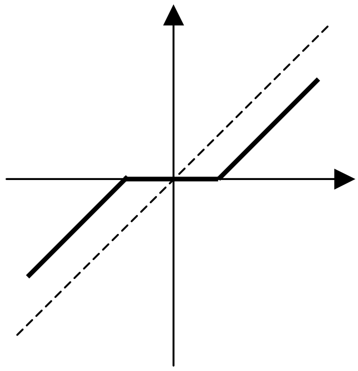


Fig.2 (b) Soft thresholding

3.2 범용 임계값(Universal thresholding method)

Donoho 와 Johnstone[7]이 제안한 범용 임계값 (universal threshold) $\lambda = \lambda^U$ 의 선택은 다음 통계적 결과에 기반을 둔다. $z_i, i = 1,\ldots,N$ 이 $iid\ N(0,1)$ 를 따르는 랜덤변수라 하고,

$$A_N = \left\{ \sqrt{2\ln N} - \frac{\ln\ln N}{\ln N} \le \max_{1\le i\le N} |z_i| \le \sqrt{2\ln N} \right\}$$

라 하면, $\lim_{N\to\infty} \Pr(A_N) = 1$ 이 성립한다. 이 결과에 의하면 분산이 $\sigma^2 = 1$ 이고 표본 크기가 충분히 큰 모든 잡음의 크기가 $\sqrt{2\ln N}$ 보다 작을 확률이 상당히 큼을 알 수 있다. 따라서, 여러 임계처리 방법에 범용적으로 적용할 수 있는 임계값 λ 를 $\lambda = \lambda^U := \sqrt{2\ln N}\hat{\sigma}$ 과 같이 자연스럽게 택할 수 있다. 여기서, $\hat{\sigma}$ 는 잡음의 표준편차 σ 에 대한 추정치 이다. 범용 임계값은 여러 축소 임계처리법에 무리없이 이용될 수 있지만 특히 데이터의 크기가 크거나 σ 가 과추정(over-estimated)되었을 때 λ^U 가 커지므로 열화(degration)또는 일그러짐(distortion)이 발생하는 경향이 있다.

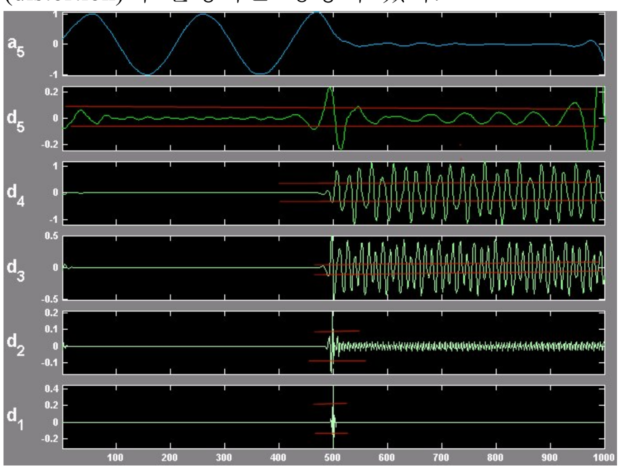


Fig.3 Thresholding 을 적용(red line)

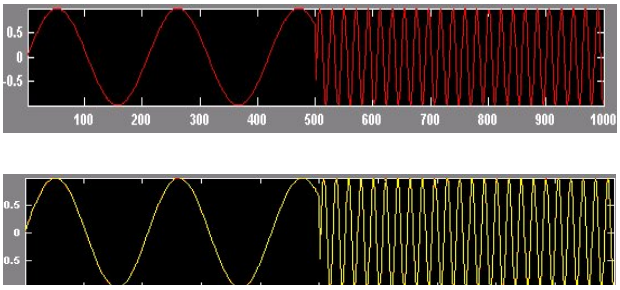


Fig.4 (a) 원본 data, (b) 역변환된 data

만약 압축이나 잡음 제거를 원한다면, Fig.3 의 a5 인 평균값(average)영역의 wavelet 계수와, 나머지 값(residual) 에 해당되는 d1,d2,d3,d4,d5 에 해당되는 각각의 wavelet 계수에 Thresholding 을 적용하여 빨간 선을 벗어나는 지역의 wavelet 계수만을 선택하고 나머지는 버린 후 모두 더해 wavelet 역변환 하면 원본 data(Fig.4 (a))에 비슷한 혹은 완전히 동일한 다양한 시그널과 여러 대역의 압축율을 갖는 시그널(Fig.4 (b))을 완성할 수 있다.

3.3 Multivariate thresholding method

T.Downie 와 B.Silverman[4]에 의해 제안되었으며, 개개의 요소를 thresholding 하기보다는 전체 벡터를 thresholding 한다. 그러므로, 앞에서 언급한 분해과정 후에 생기는 벡터들에 적용할 수 있다. 이 방법은 Multi-wavelet 의 covariance 구조를 이용하는 것으로서 그 과정은 다음과 같다.

$$\theta = R^T V R \qquad (3)$$

이 때, V 는 covariance 행렬을 나타내며, 다음의 식으로 계산된다.

$$V = \sigma^2 T^T T \qquad (4)$$

이 때, T 행렬은 앞에서 언급한 Post-transformation 행렬이 된다. σ 는 잡음의 표준편차이며, CFD 분야에서는 중요하지 않으므로 1 로 놓는다. 이러한 과정에 의하여 양의 값 θ 를 얻은 후에 임계값 λ 를 구하면, 다음과 같은 임계법칙을 적용한다.

$$\hat{R} = R \cdot I \ \ (\theta \geq \lambda) \qquad (5)$$

즉, 식(3)에 의해 계산된 θ 가 사용자가 정한 값 λ 보다 크게 되면 R 행렬 전체를 보존하고 작으면 R 행렬 전체를 0 으로 치환하여 memory 에 저장 시 생략하는 방법을 사용하게 된다. 다른 방법으로는 Soft tresholding 을 사용하여 다음과 같은 식으로 변형한다.

$$\hat{R} = R \cdot \frac{\max(\theta - \lambda, 0)}{\theta} \qquad (6)$$

즉, $\theta - \lambda$ 와 0 중에서 보다 큰 값을 계산하고, difference 값들로 이뤄진 행렬에 곱하여 임계처리를 수행하게 된다. 이 과정에서 0 으로 치환되는 경우에는 memory 상에 값을 기억하지 않는 과정을 통해 압축을 하게 된다.

4. 실제 적용 사례

4.1 Vortex Propagation Solution

이 데이터는 Rott[11]의 vortex propagation 에 관한 data 이다. Rott 의 solution 은 Navier-Stokes 방정식의 정해로서 정상상태에서의 3D 축대칭 vortex 를 나타낸다. Grid 크기는 $11 \times 51 \times 51$ 이고 원본 data 의 크기는 2,626Kb 이다. Fig.5 는 원본 data 의 vortex 전파 양상과 재구성된 data 의 전파양상을 비교한 것이다. L2RationError 는 7.841×10^{-6} 이며, 압축된 data 의 크기는 75.4Kb 이다. 압축률은 1/34.8 이다.

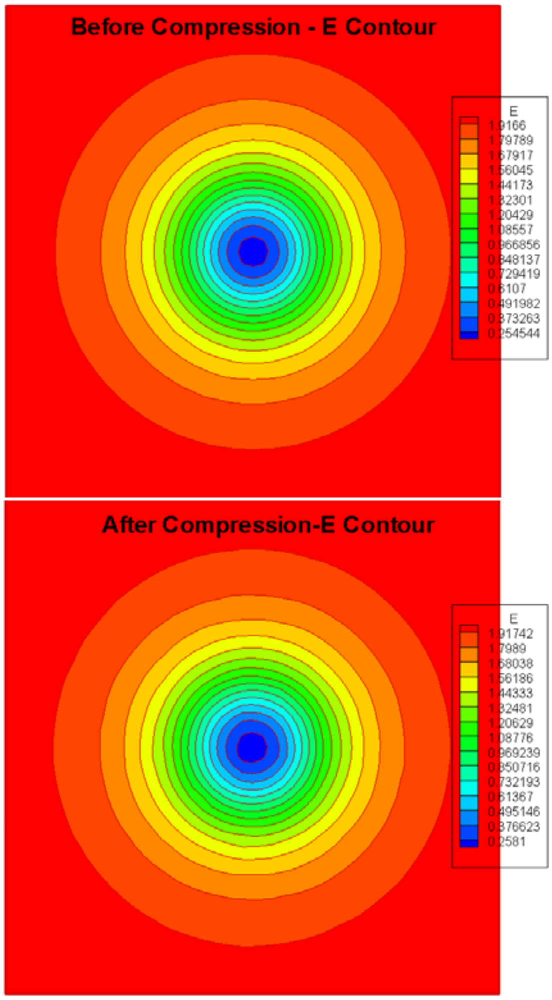
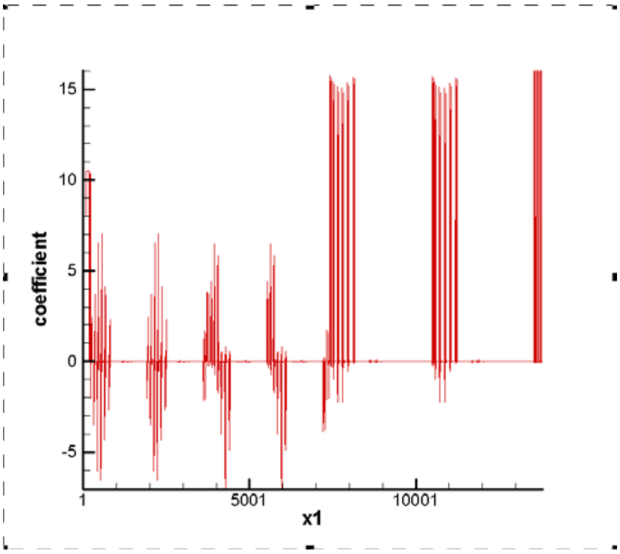
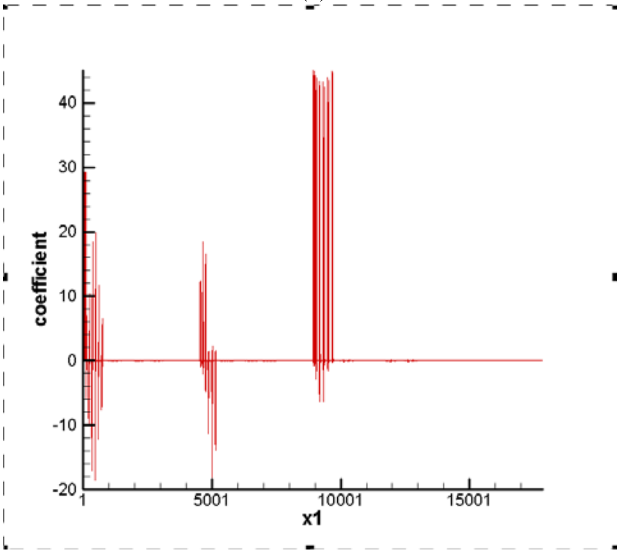


Fig.5 Vortex Propagation Energy contour 비교

4.1.1 Vortex propagation data 의 Thresholding



(a)



(b)

Fig.6 (a) Soft thresholding method, (b) Multivariate thresholding method.

Fig.6 는 각기 다른 thresholding method 를 적용했을 때의 data 를 나타낸다. Multivariate thresholding method 를 적용했을 때가 비교적 noise 도 제거되고 압축률의 향상을 기대할 수 있다.

4.2 천음속(Transonic Flow) 영역에서의 3D wing 압력분포

천음속 영역에서의 3D wing CFD data 를 Super-compact Multiwavelet 에 적용하였다. Mach 수는 0.8 로서, 천음속 영역이고, grid 크기는 $33\times33\times129$ 이다. 또한 밀도 contour 를 비교함으로써 기법의 정확성을 확인하였다. 이 때 사용된 original data 의 크기는 약 10,700Kb 이며 Fig.7 에서 보듯이 날개 뒷부분에서 발생하는 shock 을 압축 후 복원한 data 에서도 비교적 정확하게 잡아냄을 확인할 수 있다. 또한 L2RationError 는 1.333×10^{-5} 으로서 무시할 수 있을 정도로 작음을 알 수 있다. 또한 supercompact multiwavelet 을 이용한 압축기법은 original data 를 약 729Kb 정도로 압축한다. 따라서 아주작은 error 와 함께 1/14.7 이라는 매우 높은 압축률로 original data 를 압축함을 알 수 있다.

4.2.1 3D Wing data 의 Thresholding.

Multivariate thresholding method 에 의해 3D wing 압력분포 data 의 thresholding 을 수행했다. Fig.8 은 0 값에 thresholded data 들이 치환되어져 있다. 즉, 압축률이 크다는 것을 알 수 있다. Multivariate thresholding method 는 Multiple wavelet 에 적합함을 알 수 있다.

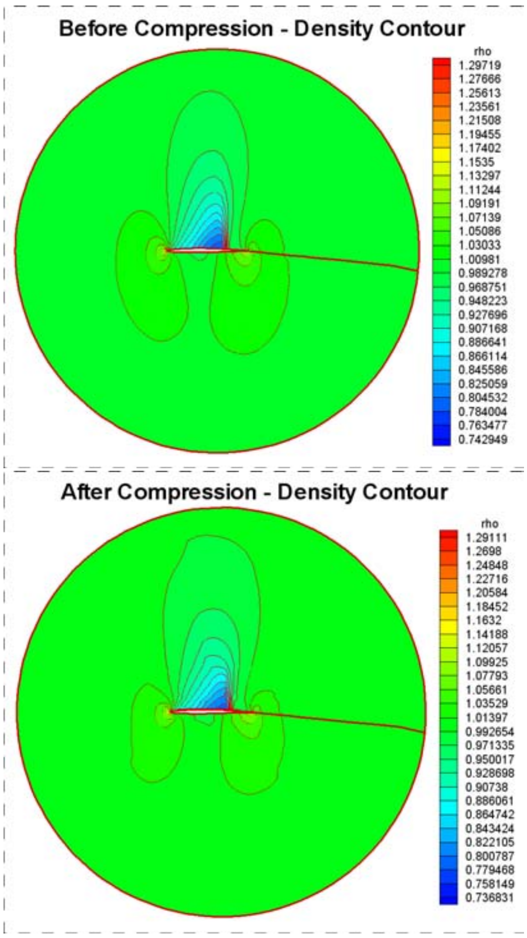


Fig.7 3D wing 밀도 contour 비교

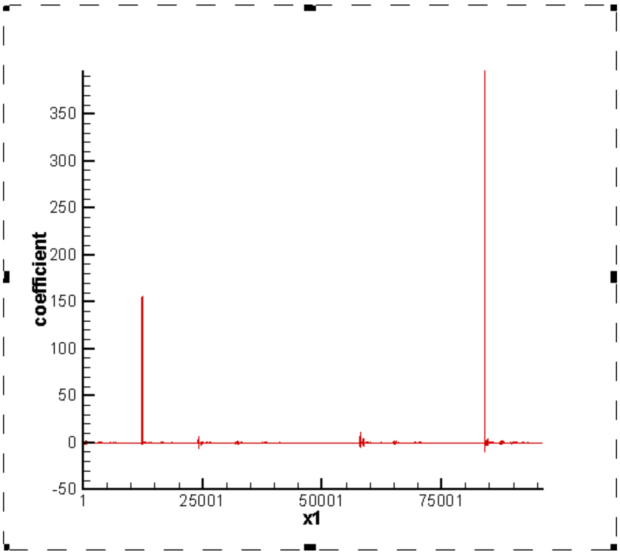


Fig.8 Multivariate thresholding method.

5. 결 론

Supercompact Multiwavelet 에 적합한 Thresholding methode 는 multivariate thresholding 이며, 개개의 요소를 thresholding 하기 보다는 coefficient 벡터를 thresholding 하는 것이 효과적임을 알 수 있다. 원본 data 와 압축된 후 복원된 data 와의 차이가 거의 없으며, shock 의 위치, 파형의 전파 양상 등 중요한 특징을 큰 오차 없이 잘 나타냄을 알 수 있다. 따라서, CFD data 의 경우에는 Supercompact Multiwavelet 기법의 적용과 Multivariate thresholding 의 적용으로 저장공간과 수렴시간을 최소화 할 수 있다.

후 기

본 연구는 한국과학재단 지역대학 우수 과학자 지원 연구(R05-2001-000-01150-0)로 수행되었음.

참고문헌

(1) Dohyung Lee, 2003, "MultiDimensional Super - compact Wavelets For Fluid Dynamics ," Numerical Heat Transfer, Part B, vol.43, pp. 307~329.
(2) Dohyung Lee, Richard M. Beam and Robert F. Warming, 2001, "Supercompact multiwavelets for flow field simulation," Computers & Fluids, vol 30, pp. 783~805.
(3) Dohyung Lee, 2000, "Supercompact Multiwavelets for Three Dimensional Flow Field Simulation," 38th AIAA Aerospace Meeting & Exhibit.
(4) T.R. Downie and B.W. Silverman, 1998, "The Discrete Multiple Wavelet Transform and Thresholding Methods," IEEE Transactions on signal processing, vol. 46, pp. 2558~2561.
(5) Brani Vidakovic and Peter Mueller, 1991, "Wavelets for Kids : A Tutorial Introduction," Duke University.
(6) Wim Sweldens and Peter Schroder, 1995, "Building your own wavelets at home," ACM SIGGRAPH Course Note#13.
(7) David L. Donoho and Iain M. Johnstone, 1993, "Ideal Spatial Adaptation via Wavelet Shrinkage," Stanford University.
(8) D.L.Donoho, 1995, "Denoising by soft thresholding," IEEE Tans. Inform. Theory, vol.41, pp.613~627.
(9) Richard M. Beam and Robert F. Warming, 1996, "Multiresolution Analysis and Supercompact Multiwavelets,", SIAM Annual Meeting.
(10) Michel Misiti, Yves Misiti, Georges Oppenheim and Jean-Michel Poggi, 1996, "Wavelet TOOLBOX," The Math Works, Inc.
(11) Rott N., 1958, "On the viscous core of a line vortex," J Appl Math Phys, vol 9b(5/6), pp.543~553.