A Study on CFD Data Compression Using Hybrid Supercompact Wavelets

Hyungmin Kang

Graduate student, School of Mechanical and Aerospace Engineering, Seoul National University, Seoul, Korea

Dongho Lee*

Head, Professor, School of Mechanical and Aerospace Engineering, Seoul National University, Seoul, Korea

Dohyung Lee

Professor, Department of Mechanical Engineering, Hanyang University, Ansan, KyungKi-Do, Korea

A hybrid method with supercompact multiwavelets is suggested as an efficient and practical method to compress CFD dataset. Supercompact multiwavelets provide various advantages such as compact support and orthogonality in CFD data compression. The compactness is a crucial condition for approximated representation of CFD data to avoid unnecessary interaction between remotely spaced data across various singularities such as shock and vortices. But the supercompact multiwavelet method has to fit the CFD grid size to a product of integer and power of two, $m \times 2^n$. To resolve this problem, the hybrid method with combination of 3, 2 and 1 dimensional version of wavelets is studied. With the hybrid method, any arbitrary size can be handled without any shrinkage or expansion of the original problem. The presented method allows high data compression ratio for fluid simulation data. Several numerical tests substantiate large data compression ratios for flow field simulation successfully.

Key Words: CFD, Data Compression, Hybrid Method, Supercompact Multiwavelet

Nomenclature -

<i>i</i> , <i>j</i> , <i>k</i> :	Index parameter		
ū :	CFD data		
<i>ā</i> :	Data transformed to wavelet field		
	from CFD data		
<i>r</i> :	Residual or wavelet coefficients		
$\varphi_n(x, y, z)$:	Polynomial shape function		
T :	Post transformation matrix		
H , G :	Decomposition matrices		
Xijk :	Volume		
<i>q</i> :	Vanishing moment		
Ngrid :	Number of grid points		

* Corresponding Author, E-mail : donghlee@snu.ac.kr TEL : +82-2-880-1900; FAX : +82-2-880-1910 School of Mechanical and Aerospace Engineering, Seoul National University, Seoul, Korea. (Manuscript Received April 8, 2001; Revised August 21, 2003)

Max_{data}	: Maximum value in original data			
P_{data}	: Processed data			
O_{data}	: Original data			
L ₂ Error	$\frac{1}{\sqrt{\sum [P_{data} - O_{data}]^2}} / N_{grid}$			
L ₂ Ratio Error :				
	$\sqrt{\sum[(P_{data} - O_{data})/Max_{data}]^2}/N_{grid}$			
Sizing parameter SP :				
	$SP = \frac{N_{grid} - 1}{q - 1}$			

1. Introduction

Recently, CFD dataset size has sharply increased in many practical cases such as full-scale integrated aero-vehicle simulation or turbo machinery analysis. In these cases, many grid points and high fidelity simulations are adopted for better level of accuracy. Even the CFD analysis for basic researches by Lee, S. and Kim, K. (2000) and Shin, S. (2000) needs many grid points and complex governing equations. In Gridpoints published by NAS Division at NASA (2001), the computed CFD results often have as huge data size as giga byte or even tera byte in NASA's 'Information Power Grid (IPG) Project'.

These big data sets, sometimes, create tremendous rendering time as well as technical difficulties in interactive post-processing/visualizing the data. In addition, due to limited network bandwidth constraints, data transmission time between remote computers in distributed systems sharply increases.

To resolve these problems, the supercompact multiwavelet scheme is used for CFD data compression. It is firstly suggested by Beam, R. M. and Warming, R. F. (1996) for 1D case and extended to 3D by Lee, D. (2000, 2001) through the construction of proper 3D approximation and residual operators for the decomposition process. This method can allow that CFD dataset is approximated with small number of supports and higher degree of accuracy. It also provides fundamental advantages such as orthogonality and symmetry.

In the research by Lee, D. (2000), the supercompact multiwavelet scheme has its own limitation that the applied model size should be the product of an integer and power of two. In other word, SP has to be $m \times 2^n (n \in \mathbb{Z})$ along all direction. If the grid size does not meet this requirement, the grid should be either shrunken or expanded for fitting. However, whatever option is chosen, the overall compression ratio drops since some parts of the solution remain untouched in shrinkage case or extra memory is required in expanding case.

In this paper, we propose a hybrid supercompact multiwavelet method. The hybrid method combines the various dimensional versions of multi-resolution (3D, 2D and 1D) in decomposition and reconstruction routines. So it can increase the coverage area directly touched by the various versions and eventually increases the data compression ratio.

2. Supercompact Multiwavelet

Wavelet bases determine the efficiency of the wavelet transform. They should be amenable to specific problem constraints. For example, in Daubechies, I.' research (1988), Daubechies wavelet base represents good performance in image compression. However, typical photographic images and fingerprint images are very different from the numerical solutions of the CFD.

The application of wavelets to CFD data should address some constraint issues. CFD data have the discontinuities like shock, vortices and shear layers. The data is also given as a vector quantity, so computation of wavelet should be careful in satisfying the physical laws as conservation laws.

To address these constraints, the supercompact multiwavelet is used in this paper. While the Haar wavelet can exactly represent any piecewise constant function in Haar, A.'s research (1910), the supercompact wavelet can exactly represent any piecewise polynomial functions. Higher level of accuracy is attained by higher order polynomials of supercompact wavelet. And it is based on multiwavelets (family of wavelets). Multiple wavelets conduct decomposition and reconstruction processes using more than single mother wavelet. They could offer fundamental advantages such as orthogonality, symmetry, short support and higher degree of accuracy. Because of these beneficial characteristics, the multiple wavelets are regarded as a good tool in allowing better data compression and feature extraction than a single wavelet.

2.1 Pre- and post-transformations

The use of multiple wavelets has additional steps for transforming given data \vec{u} to vector quantity \vec{a} . Multipoint data is transformed into a vector data by using various orthogonal basis function such as Legendre polynomials. The actual decomposition and reconstruction are performed on this transformed \vec{a} dataset. In Lee, D.'s research (2000), if we choose the order of vanishing moment q, the order of Legendre polynomial *l* is l=q-1. Volume x_{ijk} is 3D physical base domain for multiple wavelet application, composed of substantial points in $x_{i-1} \le x \le x_i$, $y_{j-1} \le y \le y_j$, $z_{k-1} \le z \le z_k$.

$$(x_{ijk})_n = (x_{i-1} + \frac{l_i(n)}{l}, y_{j-1} + \frac{l_j(n)}{l}, z_{k-1} + \frac{l_k(n)}{l})$$
 (1)

where

$$l_i(n) = \operatorname{mod}(n - l_k(n) q^2, q)$$
(2)

$$l_j(n) = \text{quotient}(n - l_k(n)q^2, q) \quad (3)$$

$$l_k(n) = \text{quotient}(n, q^2) \tag{4}$$

And *n* is a sequence integer with an extent like $0 \le n \le q^3 - 1$. Here, **mod** and **quotient** indicate reminder and quotient in integer division, respectively. u_{ijk} is the column vector located at the sequenced positions as in (1), i.e.

$$(\vec{u}_{ijk})_n = u(x_{ijk})_n \text{ for } (x_{ijk})_n \in x_{ijk}$$
 (5)

The transformation from \vec{a}_{ijk} to \vec{u}_{ijk} in matrix notation is expressed as

$$\vec{u}_{ijk} = \mathbf{T} \vec{\sigma}_{ijk} \tag{6}$$

We can choose a polynomial of degree l which adequately represents the data $(\vec{u}_{ijk})_n$ as follows:

$$(\vec{u}_{ijk})_n \approx \sum_{n=0}^{q^3-1} \alpha_n \varphi_n(x, y, z)$$
(7)

The polynomial shape function $\varphi_n(x, y, z)$ is the basis function used in transformation process.

Note that once the basis functions are orthogonal to one another in the same dimension, the orthogonality remains valid to other basis functions in other dimension. Hence the $\varphi_n(x, y, z)$ is also orthogonal to one another and it can be dimensionally spilt as

$$\varphi_n(x, y, z) = \phi_{l_i}(x) \phi_{l_j}(y) \phi_{l_k}(z) \qquad (8)$$

So equation (7) becomes

$$\sum_{n=0}^{q^3-1} \alpha_n \varphi_n(x, y, y)$$

$$= \sum_{l_l}^{l} \alpha_{l_l} \phi_{l_l} \sum_{l_j}^{l} \alpha_{l_j} \phi_{l_j} \sum_{l_k}^{l} \alpha_{l_k} \phi_{l_k}$$
(9)

and the post-transformation is defined.

Because $\varphi_n(x, y, z)$ is orthogonal, the coefficients $(\alpha_n)_{ijk}$ are simply written as

$$(a_{n})_{ijk} = \frac{8}{h_{x}h_{y}h_{z}} \times \int_{x_{i-1}}^{x_{i}} \int_{y_{j-1}}^{y_{j}} \int_{z_{k-1}}^{z_{k}} (u(x, y, z))_{ijk} (\varphi_{n}(x, y, z))_{ijk} dx dy dz$$
(10)

The pre-transformation process is performed as in equation (10).

2.2 Decomposition and reconstruction

In supercompact wavelet case, the number of support points is small. Decomposition and reconstruction of supercompact multiwavelets involve only two vector sequences in obtaining the average and the detail values.

In Lee, D.'s research (2000, 2001), the supercompact wavelets retain the spatial compactness and orthogonality of Haar wavelets. In 3D, decomposition process involves two blocks along each direction on finer grid system in order to compute one average vector and a number of residual vectors on a coarser grid system. One cell on the coarse grid corresponds to 8 sub-cell such as Fig. 1.

One step of the decomposition is performed as the following form :

$$\vec{a}^{p-1} = \mathbf{H} \vec{a}^p \tag{13}$$

$$\vec{r}^{p-1} = \mathbf{G}\vec{a}^{p}$$
 (14)

where

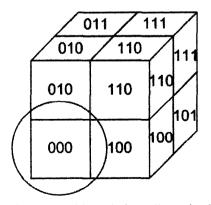


Fig. 1 One step multi-resolution cell merging in 3D (Average values are situated at 000 subcell)

$$\vec{a}^{p} = \begin{bmatrix} \vec{a}_{2i-1,2j-1,2k-1} \\ \vec{a}_{2i,2j-1,2k-1} \\ \vec{a}_{2i,2j,2k-1} \\ \vec{a}_{2i,2j,2k-1} \\ \vec{a}_{2i,2j,2k-1} \\ \vec{a}_{2i,2j-1,2k} \\ \vec{a}_{2i,2j-1,2k} \\ \vec{a}_{2i,2j-1,2k} \\ \vec{a}_{2i,2j,2k-1} \end{bmatrix} = \begin{bmatrix} \vec{a}_{000} \\ \vec{a}_{100} \\ \vec{a}_{010} \\ \vec{a}_{001} \\ \vec{a}_{001} \\ \vec{a}_{011} \\ \vec{a}_{011} \\ \vec{a}_{111} \end{bmatrix}$$
(15)

The 8 \vec{a} in \vec{a}^p are the vectors on the fine grid. They are replaced by one average vector, \vec{a}^{p-1} and 7 residual vectors, \vec{r}^{p-1} on coarse grid. The average vector means interpolation vector and residual vectors represent interpolation errors in decomposition. Due to orthogonality of the Legendre polynomial, the operator **H** is given like this.

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}^{000} \mathbf{H}^{100} \mathbf{H}^{010} \mathbf{H}^{110} \mathbf{H}^{001} \mathbf{H}^{101} \mathbf{H}^{011} \mathbf{H}^{011} \end{bmatrix}$$
(16)

where

To compute the operator matrix \mathbf{B} and \mathbf{D} , these matrices are computed at first using following equations.

$$\mathbf{D} = \begin{bmatrix} -\mathbf{I} & \mathbf{I} & & & \\ -\mathbf{I} & \mathbf{I} & & & \\ -\mathbf{I} & & \mathbf{I} & & \\ -\mathbf{I} & & & \mathbf{I} & & \\ -\mathbf{I} & & & & \mathbf{I} & \\ -\mathbf{I} & & & & & \mathbf{I} \end{bmatrix} - \mathbf{M}^{\mathrm{T}} \mathbf{H}$$
(18)

where
$$\mathbf{M}^{\mathrm{T}} = \begin{bmatrix} \mathbf{H}_{100}^{\mathrm{T}} - \mathbf{H}_{000}^{\mathrm{T}} \\ \mathbf{H}_{010}^{\mathrm{T}} - \mathbf{H}_{000}^{\mathrm{T}} \\ \mathbf{H}_{110}^{\mathrm{T}} - \mathbf{H}_{000}^{\mathrm{T}} \\ \mathbf{H}_{011}^{\mathrm{T}} - \mathbf{H}_{000}^{\mathrm{T}} \\ \mathbf{H}_{011}^{\mathrm{T}} - \mathbf{H}_{000}^{\mathrm{T}} \end{bmatrix}$$

 $\mathbf{B} = (\mathbf{D}\mathbf{D}^{\mathrm{T}})^{-\frac{1}{2}}$ (19)

Finally, G matrix is computed as G=BD. H becomes 1×8 block matrix and G does 7×8 block matrix. And H and G are orthogonal due to the orthogonality of Legendre polynomials. So the reconstruction matrices are computed by transposing the decomposition matrices H and G as follows:

$$\vec{a}^{p} = H^{T} \vec{a}^{p-1} + G^{T} \vec{r}^{p-1}$$
(20)

2.3 Thresholding

In the supercompact multiwavelet decomposition, the decomposed data in same volume x_{ijk} is strongly correlated. By this feature, it can be possible to approximate the original data with a few decomposed values. The other values can cut off if they are smaller than a certain threshold value. Data compression is performed in this way.

To get accurate approximation of original data and compress it with large ratio, appropriate thresholding method must be applied. And in the case of multiwavelets, a rather complicated treatment is needed since a vector instead of a scalar must be handled. We follow a thresholding method suggested by Downie, T. R. and Silverman, B. W.(1998), used in noise reduction in signal processing which is based on the covariance structure of multiwavelets.

3. Hybrid Method

Data representation with higher order accurate interpolation requires modification in the vicinity of domain boundaries. In decomposition procedure, two blocks along each direction are involved in computation. So the sizing parameter SP should be power of 2 along all direction. If this requirement is not satisfied, the grid points must be shrunken or expended. But whatever option is chosen, the overall compression ratio drops since some parts of the solution remain untouched in shrinkage case or extra memory is required for the extended problem in expanding case.

One of practical remedies for this size restriction is to divide the problem domain into small blocks where independent version of multi-resolution can be applied to each block. The individual subblocks have flexibility in choosing application sizes and multi-resolution dimensions as 3D, 2D or 1D. The application size varies with the order of accuracy and number of multi-resolution levels.

Depending on the grid size restriction, one divides the original problem's domain into 3D, 2D, and 1D subsets as Fig. 2. At this time, each

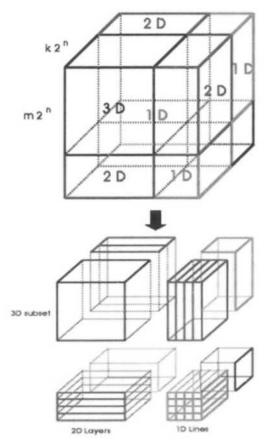


Fig. 2 Hybrid method procedure

dimensional subset has the grid size that makes the SP become power of 2 along each direction. 3D multi-resolution is performed in one piece of 3D subset using full 3D version of supercompact wavelets. 2D multi-resolutions are conducted layer by layer using 2D supercomapct wavelets, and 1D versions are applied line by line using 1D wavelets. The suggested technique is hybrid method that combines 3D, 2D and 1D version of multi-resolutions.

Because of this flexible implementation, most of the problem domains are covered in the multiresolution process, i.e. decomposition or reconstruction. The number of untouched grid points is very small. If the coverage of 3D piece is 12.5% of whole domain, the grid points used in decomposition or reconstruction process are 87.5% of whole domain. Therefore, if the grid size of the problem is worst, hybrid method can increase the compression ratio comparing with grid shrinkage case or grid expanding case.

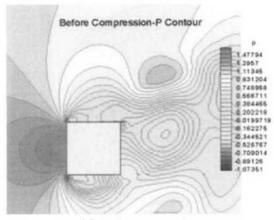
However, the hybrid method could introduce other problems. 3D, 2D and 1D versions of multiresolution go through different processes in decomposition or reconstruction procedure. While the 3D version considers the property variation in all three directions, the other multi-resolutions (2D and 1D) do not have communication between the other layers or lines. Thereby, derivative information is missing across the layers or lines. The shock discontinuity that lies across the sweeping lines does not go through any discontinuity information treatment. However, the accuracy concern turns out to be not much of problem as numerical test (data compression for the solution with shock) shows that the consistent solution is maintained even across the shock where the subdivision interface is located.

4. Numerical Tests and Discussion

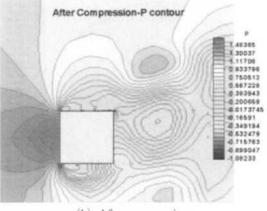
In this research, the hybrid method based on supercompact multiwavelet scheme is implemented for practical data compression. The numerical tests are performed to the datasets with various levels of grid density and fidelity levels of PDE solutions.

4.1 Application to 2 dimensional dataset

The test dataset is the simulation of vortex shedding around a square cylinder with ground proximity by Hwang, J., Lee, B., Park, Y. and Lee, D. (2001). This case is a solution of unsteady incompressible Navier-Stokes' Equation and the grid size is 112×177 . The order of vanishing moment q is 3. In this case, the ideal grid size is 65×129 i.e. $2^5 \times 2^6$ in SP number or 129×257 i.e. $2^6 \times 2^7$ in SP number. The problem's grid size should be shrunken or expanded to become ideal size. In the shrinkage case, the grid points that are not used are about 11,000. In expanding case, about 13,000 more grids are needed in computa-



(a) Before compression



(b) After compression

Fig. 3 Pressure contours of vortex shedding dataset (2D, 112×177); Compression ratio is 14.5: 1; Third order scheme is used; 5 level of multiresolution

tion. So data compression ratio might decrease. The hybrid method can release the size restriction. Using the method, untouched grid points are about 2,200, therefore high compression ratio can be obtained. In this case, the compression ratio is about 14.5:1.

Fig. 3 shows the contours of the pressure field before and after data compression. Comparing the two figures, the pressure contours represent very similar features. But memory required for compressed dataset is only 5.5 percent of the original data. The L₂ Error norm is 5.921. However, in this case the flow variables with large order of values exist in original dataset, yielding big L₂ Error norm and the L₂ Ratio Error is about 2.655×10^{-5} , so the error is very small.

4.2 Application to 3 dimensional dataset

The first dataset is the Rott's solution (1958) of the analytical, self-similar, free vortex propagation. Rott's vortex is an exact solution of the Navier-Stokes' equation, which allows a steady, 3D axisymmetric vortex. This vortex propagates in the streamwise direction with linearly increasing free stream velocity. And the gird size is $11 \times 51 \times 51$.

Using third order scheme, the ideal size of grid set is $9 \times 49 \times 49$ in shrinkage case and $17 \times 65 \times 65$ in expanding case. If the shrinkage case, the untouched grid points are about 7,000. It is about 24.5% of original grid set. In expanding case, using supercompact multiwavelets only, the additional grid points are about 43,000. But with hybrid method, the untouched grid points are only 8. Therefore, high compression ratio can be obtained. In this case, the compression ratio increases to 34.8:1.

Figure 4 shows the solution on a normal plane (51×51) to the streamwise direction. It shows the energy contours of before and after data compression. In this application, the L₂ Error is about 7.736×10^{-6} and the L₂ Ratio Error is 5.564×10^{-6} . The error is very small and it can be possible to obtain very similar data comparing with original data with high compression ratio.

The second case is the transonic flow problem in 3D wing. At this time, the wing is not a publicly known wing, instead designed in our laboratory. The solution is obtained from Navier-Stokes' equation computation. The Mach number is about 0.8. The grid size is $33 \times 33 \times 129$. When third order supercompact multiwavelet being used in this case, the grid size becomes ideal. So, whether hybrid method is used or not, the compression ratio doesn't change.

Fig. 5 shows the density contours of original and reconstructed dataset. In these figures, the shock in original dataset is clearly captured in the reconstructed dataset. And the compression ratio is 14.7 : 1. The L₂ Error is about 1.272×10^{-5} and the L₂ Ratio Error is 1.571×10^{-5} . So the error is quite small.

Fig. 6 shows the density change across the shock along the line that the distance between the line and airfoil surface is about 0.2 times of chord length and it goes to chord direction. It can be shown that the shock strength and density change are very similar between original and reconstructed data set. So it can be found that original data features do not change after data compression.

The total compression ratio and error of each dataset are shown in Table 1.

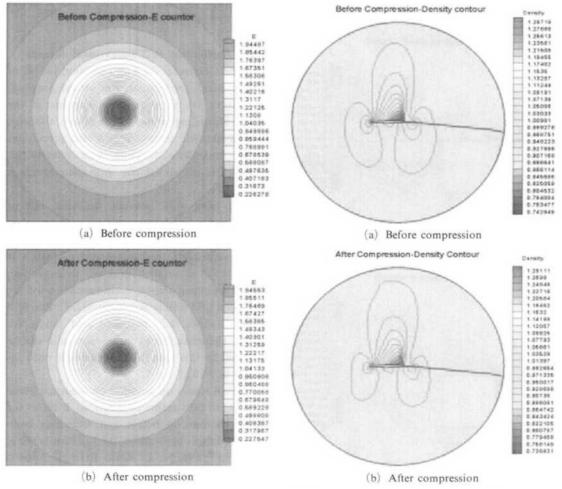
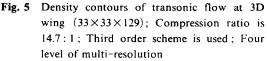


Fig. 4 Energy contours of vortex propagation solution (11×51×51); Compression ratio is 34.8:1; Third order scheme is used; 5 level of multi-resolution



	Compression ratio	L ₂ Ratio Error
Square cylinder problem	14.5 : 1	2.655×10 ⁻⁵
Rott's vortex propagation	34.8 : 1	5.564×10 ⁻⁶
Flow problem in 3D wing	14.7 : 1	1.571×10 ⁻⁵

 Table 1
 Results of each dataset

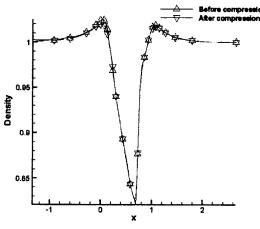


Fig. 6 Density comparison across the shook

5. Concluding Remarks

Hybrid method with supercompact multiwavelet scheme is generalized to three dimensions with the use of multi-scaling wavelet functions. Supercompact multiwavelet can be considered as a proper choice for CFD simulation data in terms of reducing the support points and representing features of dataset. And to resolve the size restriction problem and to expand the multi-resolution coverage area, hybrid multi-resolution is represented. It combines various dimensions of supercompact wavelet schemes such as 3D, 2D and 1D. It can improve the compression ratio significantly by resolving the size restriction problem and expanding the multiresolution coverage area.

This wavelet scheme allows high compression ratio with reasonable error bound. It is also shown that the inaccuracy from the introduction of the hybrid method is negligible, and important features in the original dataset are well preserved in the reconstructed dataset. This method is successfully applied to practical CFD dataset with discontinuities such as shock, shear layers, etc.

Due to these advantages, hybrid method with supercompact multiwavelet scheme can be accepted as a good management tool for visualizing features of huge CFD dataset with handy and swift treatment and also a valuable tool for rapid data transfer in distributed systems.

Acknowledgment

This work was supported by NASA Ames Research Center while the author^{**} visited ARC under RIACS/ARC summer visiting scholar program. This study has been supported by R&D Project for High Speed Railway system as well as Brain Korea 21 Project. Also, it was supported by Center of Innovative Design Optimization Technology (ERC of Korea Science and Engineering Foudation).

References

Beam, R. M. and Warming, R. F., 1996, "Multiresolution Analysis and Supercompact Multiwavelets," NASA, TM 110405.

Daubechies, I., 1988, "Orthonormal Basis of Compactly Supported Wavelets," Comm Pure Appl Math, Vol. 41, pp. 909~996.

Downie, T. R. and Silverman, B. W., 1998, "The Discrete Multiple Wavelet Transform and Thresholding," *IEEE Transaction on Signal Pro*cessing, Vol. 46 Issue: 9, pp. 255~2561.

"Gridpoints", NASA NAS System Division, Spring, 2001.

Haar, A., 1910, "Zur Theorie der Orthogonalen Func-Tionen-Systeme," Math. Annan., Vol. 69, pp. 331~371.

Hwang, J., Lee, B., Park, Y. and Lee, D., 2001, "Numerical Investigations on the Flowfiled around a Square Cylinder at Ground Proximity," AIAA, 2001-0600, 39th Aerospace Sciences Meeting & Exhibit, Reno January 12-15.

Lee, D., 2000, "Supercompact Multi-Wavelets for three Dimensional Flow Field Simulation," AIAA Paper 2000-0396, 38th AIAA Aerosapce Sciences Meeting & Exhibit, Reno, NV, January 10-13.

Lee, D., 2001, "Supercompact Multi-Wavelets for Flow Field Simulation," Computers & Fluids, 30(7-8): 783-805.

Lee, S. and Kim, K., 2000, "Design Optimization of Axial Flow Compressor Blades with Three — Dimensional Navier — Stokes Solver," *KSME Int.* Vol. 14, No. 9, pp. 1005~1012.

Rott, N., 1958, "On the Viscous Core of a Line Vortex," J Appl Math Phys, 9b(5/6): 543-53.

Shin, D., 2000, "Direct Numerical Simulation of 3-Dimensional Axial Turbulent Boundary Layers with Spanwise Curvature," KSME Int. Vol. 14, No. 4, pp. 441~447.