間引きを用いたバーク法に対する最適重みの導出とその多層薄膜の膜厚測定への適用*

工藤 重樹[†]・酒井 英昭[‡]

Derivation of the Optimum Tapered Weight for the Burg Algorithm with Decimation and Its Application for Measurement of Thickness of a Thin Multi-layer Film*

Shigeki KUDO[†] and Hideaki SAKAI[‡]

The optimum tapered Burg (OTB) algorithm is based on minimization of the average variance of the estimated frequency of a sinusoid. The decimation technique has also been used to resolve clearly sinusoids of closely spaced frequencies. This paper newly derives the optimum tapered weight for the Burg algorithm with decimation and shows that the variance of the estimated frequency becomes even smaller than that of the OTB algorithm. The decimation technique with this optimum tapered weight is applied to measure sub-micron thickness of multi layer film and its effectiveness is shown.

1. はじめに

近年, 大型テレビに代表される情報機器に位相差フィルム, 反射防止膜, カラーフィルタなど多層の光学薄膜が盛んに用いられている. これらのフィルムは, 機能を維持するのにミクロンオーダーからナノオーダーの精度で各膜厚が確保される必要がある. そのため, インラインすなわち製造中に, 多層膜の膜厚を測定する装置の必要性が求められている. その要求に応えるべく白色光を用いた干渉方式に基づく「インライン多層膜の新しい測定手法」[1]が開発されている. そこでは, 最適重み付きバーク法[2]が用いられており, 膜厚に対応する周波数推定の精度が向上している.

また, 従来から, 接近した周波数の分離性能を向上する目的として, 間引き手法[3]が用いられている. 本論文では, 最初に, 従来の最適重み付きバーク法や間引きを用いたバーク法を説明し, それらを多層膜測定に適用した際の問題点について述べる. つぎに, 周波数分解能を向上する目的で, 間引きを用いたバーク法に対して周

波数推定誤差の分散を最小にする最適重みを新たに導出する. そして, 本提案方法により周波数推定の分散がより小さく押さえられることを示す. その後, シミュレーションデータを用いてこのことを検証する. 最後に, 本提案法を多層薄膜の膜厚測定に適用し, 従来[1]より薄い $0.35\mu\text{m}$ の薄膜が安定して測定できることを示す.

2. 従来法

2.1 最適重みつきバーク法[2]

前向き, 後向き予測誤差の二乗和を最小にすることにより, 自己回帰 (以降 AR <Auto Regression> と記す) モデルのパラメータを求め, それで構成される AR モデルを信号に適合させるのがバーク (Burg) 法[2,4]である. そして, 正弦波状の信号から周波数を推定する場合に適用するにあたって, その推定分散を最小にするよう改善された最適重み付きバーク法 (以下 OTB <Optimum Tapered Burg> 法と記す) が開発されている[2]. これは, 時系列データ x_k ($k=0,1,\dots,N$) に対して, つぎのように定式化される. 次数 m の線形予測係数を $(1, -a_{m,1}, \dots, -a_{m,m})$ とする. このとき, 前向き予測誤差 $E_{m,k}$, 後向き予測誤差 $B_{m,k}$ は, つぎのように定義され,

$$E_{m,k} = x_{k+m} - a_{m,1}x_{k+m-1} \cdots - a_{m,m}x_k \quad (1)$$

$$B_{m,k} = x_k - a_{m,1}x_{k+1} \cdots - a_{m,m}x_{k+m} \quad (2)$$

それらの間には, つぎの関係がなりたつ.

* 原稿受付 2008 年 1 月 30 日

[†] 東レエンジニアリング(株) エレクトロニクス事業本部 Electronics Div., Toray Engineering Co., Ltd.; Otsu, Shiga 520-2141, JAPAN

[‡] 京都大学 大学院 情報学研究科 Graduate School of Informatics, Kyoto University; Sakyo, Kyoto 606-8501, JAPAN

Key Words: spectrum analysis, optimum tapered Burg algorithm, decimation, multi-layer film, submicron-thickness.

$$\begin{aligned} E_{m,k} &= E_{m-1,k+1} - a_{m,m} B_{m-1,k} \quad (k=0, \dots, N-m) \\ B_{m,k} &= B_{m-1,k} - a_{m,m} E_{m-1,k+1} \end{aligned}$$

ただし, $E_{0,k} = B_{0,k} = x_k$ である. 一方, 線形予測係数は Levinson-Durbin の再帰式を満足する制約のもとで,

$$a_{m,i} = a_{m-1,i} - a_{m,m} a_{m-1,m-i} \quad (i=1, \dots, m-1)$$

となる. OTB 法は, つぎの重み付き和 S_m が最小になるよう $a_{m,m}$ を求める方法である.

$$S_m = \sum_{k=0}^{N-m} w_{m,k} [|B_{m,k}|^2 + |E_{m,k}|^2] \quad (3)$$

ただし, 重み $w_{m,k}$ はつぎの条件を満たす.

$$\begin{aligned} \sum_{k=0}^{N-m} w_{m,k} &= 1 \quad \forall m \\ w_{m,k} &= w_{m,k-(N-m)/2} \quad (k=0, \dots, N-m) \end{aligned} \quad (4)$$

(3) 式を最小にする PARCOR 係数 $\hat{a}_{m,m}$ はつぎの式になる.

$$\hat{a}_{m,m} = \frac{2 \sum_{k=0}^{N-m} w_{m,k} B_{m-1,k} E_{m-1,k+1}}{\sum_{k=0}^{N-m} w_{m,k} (|B_{m-1,k}|^2 + |E_{m-1,k+1}|^2)} \quad (5)$$

ここで, 時系列データとして, 余弦波状の信号

$$x_k = \cos(k\theta + \phi) \quad (k=0, \dots, N) \quad (6)$$

の場合を考える. ただし, ϕ はランダムな位相で $[0, 2\pi)$ 上の一様分布に従うと仮定する. この場合, 線形予測係数は 2 次となり, それは $(1, -a_{2,1}, 1)$ で表され, $a_{2,1} = 2a_{1,1}$ なので, $a_{1,1}$ を評価すればよく, それは $a_{1,1} = \cos\theta$ である. そこで, (5) 式により, その推定値 $\hat{a}_{1,1}$ を求める.

$$\hat{a}_{1,1} = \cos(\theta - \Delta\theta) \quad |\Delta\theta| \ll 1$$

と近似して, 角周波数 θ の推定誤差 $\Delta\theta$ を求める. その分散の θ に関する $[0, \pi)$ 上の平均が最小になる重み $w_{1,k}$ が

$$w_{1,k} = \frac{6(k+1)(N-k)}{N(N+1)(N+2)} \quad (k=0, \dots, N-1) \quad (7)$$

と表され, また, このときの推定誤差分散 V は

$$V = \frac{12}{8N(N+1)(N+2)} \quad (8)$$

であることが文献 [2] で示された.

2.2 間引きを用いたバーク法 [3]

従来から, 接近した周波数の分離性能を向上する目的で間引きが用いられている. これは, データ系列を K 個ごとに間引くことで, 周波数を見かけ上 K 倍する方法で, 間引かれたデータにバーク法を適用する. したがって, 接近した周波数の間隔は K 倍になり, 検出がしやすくなる. また, 間引きによるエイリアシングを防ぐた

め, バンドパスフィルタを通して処理する. この結果, データ系列には特定の周波数帯域しか含まれず, SN 比を向上させる機能を有する. 以降この方法を BD (Burg with Decimation) 法と記す.

2.3 従来法の問題点

多層薄膜の膜厚測定において, OTB 法は, 間引きなしでバーク法の重みを最適化した技法であるが, これによる「インライン多層膜の測定手法」[1] を開発し, $0.5\mu\text{m}$ 程度の膜厚以上の膜厚が測定できることを示したが, 更なる薄膜測定の要求には期待できないことがわかった. 実際, 4 章の Fig. 8 に示すようにこの OTB 法では, $0.3\mu\text{m}$ のより薄膜に対して膜厚に対応する周波数の検出ができなかった. また, BD 法は従来のバーク法に単純に間引きの技法を組み入れたものであり, ここの重みは最適なものになっておらず, これを今回の膜厚測定に適用したが, 安定して薄膜に対する周波数が検出できなかった. そこで, 間引き法に対して, 新規に最適重みを導出する必要が生じた. 次章で提案する手法はこの理論的不備を解消するものである.

3. 最適重みつき間引き法

3.1 間引きを用いた最適重みの導出

そこで, 従来の BD 法に対して推定誤差分散を最小にする最適窓の導出を行う. 具体的には, 時系列 x_k に対して K 個ごとに間引きを行い, その時の最適重みを求める. 以下では簡単のため N は K で割り切れ, $N = KL$ であるとする. $k = Ki + j$ とおく. ここで, $i=0, \dots, L-1, j=0, \dots, K-1, K=0, \dots, N-1$ である. このとき, K 個ごとの間引きを行うことから, 長さ L のデータセットが K 組できるが, それぞれに重みをかけ, 全体の重みの総和が 1 となるようにして, $\hat{a}_{m,m}$ をつぎのように定義し, (5) 式と同様に計算する.

$$\begin{aligned} \hat{a}_{m,m} &= 2\text{stn}_m / \text{std}_m \\ \text{stn}_m &= \sum_{j=0}^{K-1} \sum_{i=0}^{L-m} w_{m,Ki+j} B_{m-1,i,j} E_{m-1,i+1,j} \\ \text{std}_m &= \sum_{j=0}^{K-1} \sum_{i=0}^{L-m} w_{m,Ki+j} (|B_{m-1,i,j}|^2 + |E_{m-1,i+1,j}|^2) \\ \sum_{j=0}^{K-1} \sum_{i=0}^{L-m} w_{m,Ki+j} &= 1 \end{aligned}$$

ここで, $E_{m,i,j}, B_{m,i,j}$ は (1), (2) 式で $x_{k+m}, x_{k+m-1}, \dots, x_k$ をそれぞれ $x_{K(i+m)+j}, x_{K(i+m-1)+j}, \dots, x_{Ki+j}$ で置き換えたものと定義する. とくに, $m=1$ のときは,

$$\begin{aligned} \hat{a}_{1,1} &= 2\text{stn}_1 / \text{std}_1 \\ \text{stn}_1 &= \sum_{j=0}^{K-1} \sum_{i=0}^{L-1} w_{1,Ki+j} x_{Ki+j} \\ \text{std}_1 &= \sum_{j=0}^{K-1} \sum_{i=0}^{L-1} w_{1,Ki+j} x_{Ki+j}^2 \end{aligned}$$

である. ここで, $\text{stn}_1, \text{std}_1$ は x_k に (6) 式を代入して,

$$\text{stn}_1 = \sum_{j=0}^{K-1} \sum_{i=0}^{L-1} w_{1,Ki+j} \cos[(Ki+j)\theta + \phi] \cdot \cos[(K(i+1)+j)\theta + \phi]$$

となる。ここで、 $\cos$ の加法定理を用いると、

$$\begin{aligned} \text{stn}_1 &= \frac{1}{2} \sum_{j=0}^{K-1} \sum_{i=0}^{L-1} w_{1,Ki+j} \{ \cos[K\theta] \\ &\quad + \cos[2(Ki+j)\theta + 2\phi + K\theta] \} \\ &= \frac{1}{2} \{ \cos[K\theta] + \sum_{k=0}^{N-1} w_{1,k} \cos[2k\theta + 2\phi + K\theta] \} \end{aligned}$$

となる。ここで、 $Ki+j$ を変数 k に置き換えた。

つぎに、 std_1 を計算する。

$$\begin{aligned} \text{std}_1 &= \sum_{j=0}^{K-1} \sum_{i=0}^{L-1} w_{1,Ki+j} \{ \cos^2[(Ki+j)\theta + \phi] \\ &\quad + \cos^2[(K(i+1)+j)\theta + \phi] \} \end{aligned}$$

ここで、 $\cos$ の倍角の公式を用いると、

$$\begin{aligned} \text{std}_1 &= \sum_{j=0}^{K-1} \sum_{i=0}^{L-1} w_{1,Ki+j} \left\{ \frac{1 + \cos[2(Ki+j)\theta + 2\phi]}{2} \right. \\ &\quad \left. + \frac{1 + \cos[2(Ki+j)\theta + 2\phi + 2K\theta]}{2} \right\} \end{aligned}$$

ここで、 stn_1 の場合と同様に新変数 $k = Ki+j$ に置き換え、 $\cos$ の和を積に変換する公式を用いると、

$$\text{std}_1 = 1 + \cos[K\theta] \cdot \sum_{k=0}^{N-1} w_{1,k} \{ \cos[2k\theta + 2\phi + K\theta] \}$$

となる。 $\text{stn}_1, \text{std}_1$ の式の導出は、付録 1 に記す。文献 [2] と同様にして、 $\text{stn}_1, \text{std}_1$ の第 2 項がそれぞれ十分小さいことが示せるので、結局 $\hat{a}_{1,1}$ は

$$\begin{aligned} \hat{a}_{1,1} &= 2\text{stn}_1 / \text{std}_1 \\ &\simeq \cos[K\theta] + \sin^2[K\theta] \cdot \sum_{k=0}^{N-1} w_{1,k} \{ \cos[2k\theta + 2\phi + K\theta] \} \end{aligned} \quad (9)$$

と近似できる。一方、

$$\hat{a}_{1,1} = \cos[K\theta - \Delta\theta'] \quad |\Delta\theta'| \ll 1$$

と近似して、 $\Delta\theta'$ の分散を求める。ここで、 $\Delta\theta'$ は $K\theta$ の推定誤差である。すなわち、

$$\begin{aligned} E[(\Delta\theta')^2] &= \left(\frac{1}{2} \right) \sin^2[K\theta] \cdot \sum_{k=0}^{N-1} \sum_{k'=0}^{N-1} w_{1,k} w_{1,k'} \cdot \cos[2(k-k')\theta] \end{aligned}$$

ここで、変数として k, k' を用いる。これを、 θ に関して $0 \sim \pi$ まで積分して平均をとると、つぎのようになる。

$$\begin{aligned} \left(\frac{1}{\pi} \right) \int_0^\pi E[(\Delta\theta')^2] d\theta &= \left(\frac{1}{8} \right) \cdot \sum_{k=0}^{N-1} \sum_{k'=0}^{N-1} w_{1,k} w_{1,k'} \\ &\quad \cdot [2\delta(k-k') - \delta(k-k'+K) - \delta(k-k'-K)] \end{aligned} \quad (10)$$

ここに、 $\delta(\cdot)$ は

$$\begin{aligned} \delta(i) &= 1 \quad (i=0) \\ &= 0 \quad (\text{それ以外の } i) \end{aligned}$$

で定義される。ここで、サイズ $N \times N$ の正方行列 A を、対角線上に 2、対角線の両側の $K+1$ 離れた位置に -1 が並び、それ以外は 0 からなる帯行列とする。すなわち、

$$A = \begin{pmatrix} 2 & & & & \\ & 2 & & & \\ -1 & & \ddots & & -1 \\ & \ddots & & \ddots & \\ & & -1 & & 2 \end{pmatrix}$$

とする。

つぎに、(10) 式を最小にする重み $w_{1,k}$ を求める。重み $w_{1,k}$ をまとめて $W^t = (w_{1,0}, w_{1,1}, \dots, w_{1,N-1})$ とすると、(10) 式は $(1/8)W^t A W$ となる。ここで、“ t ” は転置の意味で、重みの制約条件 (4) より成分の和は 1 である。ここで、ラグランジュ乗数 λ 、すべての要素が 1 の縦ベクトル $\mathbf{1}$ を導入し、つぎの H を最小にする。

$$H = W^t A W - 2\lambda(W^t \mathbf{1} - 1) \rightarrow \min$$

この問題の解 $W = W_{opt}$ は

$$A W = \lambda \mathbf{1} \quad (11)$$

$$W^t \mathbf{1} = 1 \quad (12)$$

を満足する。そして、このとき (10) 式で、 $k = Ki+j, k' = Ki'+j'$ ($i, i' = 0, \dots, L-1; j, j' = 0, \dots, K-1$) とおき、(11) 式を要素ごとに書き下すと、

$$\begin{aligned} \sum_{i'=0}^{L-1} \sum_{j'=0}^{K-1} [2\delta(Ki+j - Ki'-j') \\ - \delta(Ki+j - Ki'-j'+K) \\ - \delta(Ki+j - Ki'-j'-K)] w_{1,Ki'+j'} = \lambda \end{aligned}$$

となるが、ここから

$$\begin{aligned} \sum_{i'=0}^{L-1} [2\delta(i-i') - \delta(i-i'-1) - \delta(i-i'+1)] w_{1,Ki'+j} \\ = \lambda \end{aligned} \quad (13)$$

となることがわかる。(10) 式で $K=1$ とした式が文献 [2] で取り扱われているが、(13) 式は N を L に置き換えた式にはかならない。つまり、 K 個ごとの間引きを行った場合の最適重みは、データ長が L の場合の (7) 式の重みをおのおの K 個繰り返す形で与えられる。すなわち、

$w_{1,Ki+j}$, λ は

$$w_{1,Ki+j} = \frac{6(i+1)(L-i)}{KL(L+1)(L+2)} \quad (14)$$

$$\lambda = \frac{12}{KL(L+1)(L+2)} \quad (15)$$

となる. その推定誤差分散は (10) 式に (11),(12) 式の関係を用いて,

$$(1/8)W^tAW = (1/8)W^t\lambda\mathbf{1} = \lambda/8$$

なので, K 個ごとの間引きを行った場合, 周波数は $K\theta$ となり, θ の推定分散誤差 V' は,

$$V' = \lambda/(8K^2) \quad (16)$$

である. $L = N/K$ であり, (8) 式とこれを比べると, 両者の推定誤差の分散比は,

$$V'/V = \frac{(N+1)(N+2)}{(N+K)(N+2K)} \leq 1 \quad (17)$$

V : 間引きがない場合の推定分散誤差

V' : 間引きが K の場合の推定分散誤差

N : データ数

K : 間引き数

となり, 間引くことで推定分散誤差は改善される. なお, 間引き K の場合の最適重みを文献 [2] と同じように m 次の場合に, つぎのように拡張する.

$$w_{m,Ki+j} = \frac{6(i+1)(L-m+1-i)}{K(L-m+1)(L-m+2)(L-m+3)} \quad (18)$$

3.2 シミュレーションによる効果の検証

3.1 章の結果を用いて, 最適重みつき間引き法の効果を調べる. モデルとして, つぎの (19) 式で生成した 256 点のデータを $K=4$ で間引き処理し, 最適重みつき, 重みなしの場合の周波数検出能力を比較した. ここで, 以降最適重みつき間引きありの Burg 法を OTBD <Optimum Tapered Burg with Decimation> 法と記す. ツールとして Matlab を用いた.

$$y(k) = \sin\left[\pi k\left(\frac{11}{256}\right)\right] + \sin\left[\pi k\left(\frac{10}{256}\right)\right] + \text{randn}(k) \quad (19)$$

$\text{randn}(k)$: $[0,1]$ の範囲で Gauss 分布に従って

乱数を生成する関数 ($k=1, \dots, 256$)

ここで, ノイズ成分は平均 0, 標準偏差 1 の Gauss 分布に従って, 乱数を出力する関数 randn を用いた.

まず, ノイズがない, すなわち $\text{randn}(k)$ をすべて 0 とした場合を OTBD 法と BD 法とで処理し, 前者を OTBD, 後者を BD としてパワースペクトル (以下 PS と記す) の結果を表した. ここで用いた Burg 法の次数は, つぎに示すノイズなしの場合と同様, 50 とした.

Fig. 1 中のマーク「x」は係数 $2\pi/256$ を除いた周波数 5.0, 5.5 の位置を示している. OTBD 法では, ピーク

Fig. 1 Power spectrum in the OTBD algorithm and the BD algorithm for noiseless data

を与える周波数は正しく検出されているが, BD 法では, 一つのピーク位置が $2\pi \times 0.25/256$ [rad] ずれ, 正しく検出できていない.

つぎに, ノイズを付加したデータに対して処理を行った. この場合, 各正弦波の SN 比は -3dB で, ノイズを含むため, 帯域制限に Daubechies wavelet [6] (以下 DB と記す) の次数 16, レベル 6 に分解し, $K=4$ なので, 本来 $A_2 (= A_6 + D_6 + D_5 + D_4 + D_3)$ を用いるべきであるが, モデルから帯域が既知であるので, その一部の $D_5 + D_4 + D_3$, すなわち帯域 $[\pi/32, \pi/4]$ を用いて周波数解析を行った. データセットとしては, 20 組用意した. ここに D_i ($i=1, \dots, 6$) は, DB で分解した $[\pi/2^i, \pi/2^{i-1}]$ の帯域を含む詳細信号で, A_i は $[0, \pi/2^i]$ ($i=1, \dots, 6$) の帯域を含む近似信号である [6]. 今後の説明に単位 [bin] を導入しておく. FFT 処理のデータ数を $2Q$ 個とすると, 周波数間隔は $2\pi/2Q$ [rad] となり, i [bin] は, $i \times 2\pi/2Q$ [rad] に対応する.

Fig. 2 Relation between wavelet level and bin

また, 最適重みは (18) 式から計算したものをを用い, OTBD 法と BD 法の場合の PS を算出して, その結果を評価した. まず, OTBD 法の場合の PS を Fig. 3 に示す.

つぎに, BD 法の場合の PS を示す.

グラフ中の「x」は Fig. 1 と同じ意味である. Fig. 3, Fig. 4 より, OTBD 法の方がピーク周波数の位置のばらつきが小さいように見える. 実際, 上の係数を除いて, $[4.75, 5.75]$ の領域でまず一つ目のピークを与える周波数を求め, その左側にもう一つ目のピークを与える周波数を求め, これを推定周波数としたとき, 二つの周波数が正しい位置に検出された数は, OTBD 法の場合は 10 個で, BD 法の場合は 2 個であった. これより, 各正弦波

Fig. 3 Power spectrum in the OTBD algorithm for noisy data

Fig. 4 Power spectrum in the BD algorithm for noisy data

の SN 比が -3dB でも OTBD 法は接近した周波数の検出に安定しているといえる。

4. 多層膜膜厚測定への適用

3章でシミュレーションにより、接近した周波数の検出に対する OTBD 法の安定性が示された。この方法を薄膜の多層膜膜厚測定に適用する。装置の詳細は文献 [1] に示しているのので、ここでは概要を記す。白色光 (446.5nm から 954.0nm の波長域) を多層膜試料にコア径 $100\mu\text{m}$ の 2 分岐ファイバを用いて垂直に照射し、その反射光を分光器に導く。分光器から得られる光電変換後のデータは各膜の界面で反射した光が干渉したもので、それは余弦波状の信号の線形結合で近似される [1]。また、膜厚と余弦波の周波数は比例関係にあることを文献 [1] で示した。ここで、測定に用いたサンプルは S01, S02, S03 の 3 種で、膜の構成と膜厚はどれも同じで、つぎの通りである。

基盤の上に色のついた薄膜 (LY1 と記す) $0.35\mu\text{m}$ 、その上に透明な膜 (LY2 と記す) $25\mu\text{m}$ からなる。また、光の照射の概略を Fig. 5 に示す。実際の白色光は垂直照射であるが、わかりやすくするため、斜めからの照射と

して、図示している。測定装置の概略図は付録 1 に示す。周波数解析後の PS では、膜と空気、金属の界面の屈折率差が大きいので、LY2+LY1 の膜厚に対する一番高い第 1 ピークが現れ、そのすぐ左に LY2 に対応する第 2 ピークが現れるはずである。LY1 はサブミクロンの膜厚で、屈折率は不明なので、単位面積当たりの重量から膜厚に変換した。光電変換されたデータは 1024 点で、それを DB の次数 16, レベル 6 に分解し、膜厚に対応する周波数成分を含むレベルのデータを 1024 点に再構成して $K=4$ の間引きを行い、256 点のデータセット 4 組に対して次数 200 の OTB 法、すなわち OTBD 法で周波数解析した。参考に、S01 の 1024 点の干渉データと wavelet 分解した後の再構成の 1 例を Fig. 6, Fig. 7 に示す。

Fig. 5 Basic model

Fig. 6 Reflective spectrum of S01

グラフの X 軸は波長の逆数の波数、Y 軸は光源のスペクトルで補正した反射強度を表している。ここで、wavelet 変換を用いたのは、測定対象に色目があり、干渉以外の周波数成分除くためである。Fig. 6 中、S01 の高周波成分は干渉信号で、その中央の線は wavelet 分解後の近似信号 A_6 である。 A_6 が反射スペクトルの中央を通っていることより、変換が妥当であると判断できる。信号から A_6 を引くことで、干渉成分のみが抽出できる。なお、1024 点のデータの前後に 200 点のデータを対称に追加して wavelet 分解をしている。これは、データの端での分解を改善するためである。つぎに、wavelet 分解後、1024 点に再構成した結果を Fig. 7 に示す。

周波数解析については、DB で分解後の詳細信号 $D_2 + D_3$ を用いて 1024 点に再構成し、その後 OTB 法で処理した。求めた予測誤差フィルタ係数を 0 拡張して、8192 個として PS を求めた。ここに、横軸の単位を [bin] とした。前出の Fig. 2 に示すように、膜厚に対応する周波数を含んだ領域は D_3 であるが、 D_2 を周波数の漏れを考慮して処理データに含め、 $D_2 + D_3$ を解析した。間引

Fig. 7 Decomposed data by Daubechies wavelet of S01

Fig. 8 Power spectrum of S01/S02/S03 without decimation in the OTB algorithm

きのない OTB 法を Fig. 8 に示す。Burg 法の次数は 800 である。全体膜厚に対応する第 1 ピークは Fig. 8 中、楕円で囲った部分に現れているが、その左側には、擬似のピークも含めて多数のピークが見られる。S01, S02, S03 の全体膜厚の左に存在するピークで、 $0.35\mu\text{m}$ ($=13[\text{bin}]$) 程度に対応) 程度薄くなった位置のピークを LY2 の膜厚を表すものとした。その結果を Table 1 に示す。なお、膜厚と [bin] との関係は、付録 2 に示す。

Table 1 Thickness of LY1 without decimation in the OTB algorithm [bin]

	LY1+LY2	LY2	LY1
S01	787	769	18
S02	804	784	20
S03	779	764	15

ここで選択したピークは、必ずしも 2 番目に高いものではなかった。その意味で、LY2 の膜厚は安定して得られていない。そのため、LY1 の膜厚変動が大きい。ちなみに 2 番目に高いピークを LY2 として、LY1 を算出すると、S01, S02, S03 に対しては 62, 102, 52[bin] となる。これらは、 $0.35\mu\text{m}$ には対応せず、この場合、膜厚が検出できていないと判断できる。

つぎに、本論文での提案手法である OTBD 法を用いた結果を Fig. 9 に、グラフの横軸を着目する 600~1000[bin] の領域に限定して示す。

Fig. 9 Power spectrum of S01/S02/S03 in the OTBD algorithm

一般に、 K 個ごとの間引きの場合は、対象領域幅を $1/K$ にバンドパスしておく必要がある [3]。ここでは、 $K=4$ の間引きを行い、長さ 256 点のデータセットを四つ作った。この場合は、 D_2 を含むとエイリアシングの影響が生じるが、Fig. 8 に示すように D_2 の高周波域には高いピークが含まれておらず、 $D_2 + D_3$ を用いた。Fig. 9 中、*_burg $\times 10$ は、Burg 法の次数 200 の BD 法による PS の各成分を 10 倍したもので、*_all は Burg 法の次数 200 の OTBD 法による PS である。ここでは、間引きが 4 であるので、周波数を 4 で割っている。Burg 法の次数は、100, 150, 200 と試行して 200 と決めた。* は S01, S02, S03 である。Fig. 9 より、低周波数領域のピークから $0.35\mu\text{m}$ に対応するピークを解釈することは難しい。そこで、全体膜厚 (LY1+LY2) に対応するピーク位置の左側に LY2 に対するピークが現れると解釈し、 $0.35\mu\text{m}$ の膜厚に対応する膜厚は、二つのピーク位置の差から求めた。この場合、LY2 のピーク高さは全体膜厚のピーク高さに対して、2 番目に高いものであった。そのため、*_all の第 1 ピークは正しいと判断した。*_all, *_burg $\times 10$ からの結果をおおの Table 2, Table 3 に

Table 2 Thickness of LY1 in the OTBD algorithm [bin]

	LY1+LY2	LY2	LY1
S01	810.75	796.00	14.75
S02	775.75	761.00	14.75
S03	821.75	807.25	14.50

Table 3 Thickness of LY1 in the BD algorithm [bin]

	LY1+LY2	LY2	LY1
S01	810.78	795.75	14.50
S02	792.25	780.75	11.50
S03	836.75	823.75	13.00

表す。なお、*Burg $\times 10$ の全体膜厚に対するピーク位置は、*allの位置と若干異なっている。また、*allの全体膜厚と似た位置にも低いピークは存在している。Table 2, Table 3より、膜厚 $0.35\mu\text{m}$ は、本論文の提案する OTBD 法で、安定して検出されている。しかし、従来の BD 法では、Table 3 に示すように、S02, S03 のデータに対して LY1 の変動幅は大きい。Fig. 9 中、LY1 が正しく検出された部分を楕円で囲み、処理の名前を「正解」と記載した。

さて、本来、膜厚に対応するピークは 1 本の直線であるが、ノイズにより歪んでいると考えられる。そこで、*all algorithm により得られた PS に対して、検出されたピーク位置を中心に 5 点による放物線近似 [7] $y = a_{20} + a_{21}X + a_{22}X^2$ を行い、この近似曲線の最大値を与える座標 $X = -a_{21}/2a_{22}$ を真のピーク位置と見なして、Table 4 に示した。係数の求め方については付録 3 に示す。

Table 4 Thickness of LY1 by approximation of a parabolic curve with 5 points [bin]

	LY1+LY2	LY2	LY1
S01	810.76	795.95	14.81
S02	775.74	761.01	14.73
S03	821.76	807.19	14.57

放物線近似することで、LY1 の変動幅が若干小さくなり改善された。なお、LY2 に関しても、変動幅は改善されている。

5. 結論

間引きを行ったデータセットに対して、周波数の推定誤差分散を最小にする最適重みを導出した。その最適重みつき間引きあり Burg 法を OTBD 法として、それを用いることで、推定分散誤差が間引きなしの場合に対してより小さくなることを示した。それを、256 点で各正弦波の SN 比が -3dB で、周波数差が $2\pi \times 0.5/256[\text{rad}]$ であるモデル信号を 20 組用いて OTBD 法と BD 法（従来の間引きあり Burg 法）とで処理し、OTBD 法の方が PS のピークを与える周波数の位置のばらつきが小さく、20 組中、正解の周波数の数が多いことを例示した。

つぎに、この OTBD 法を多層膜膜厚測定に用いた。この方法を用いることで、サブミクロンのより薄膜の膜厚が安定して測定できた。その結果、薄膜の最小測定値が $0.5\mu\text{m}[1]$ から $0.35\mu\text{m}$ に下がった。このような SN 比の信号であれば、 $0.35\mu\text{m}$ が 14[bin] 台であることから、概略その $1/4$ の $0.1\mu\text{m}$ では 3[bin] 台となり、この差なら検出でき、したがって、 $0.1\mu\text{m}$ 厚は測定できると思われる。

さて、今後は更なる薄膜の測定を行う予定であるが、その場合は、干渉縞の山の数が 1 以下になる。この状態ではモデルによるフィッティングを行い、その結果から膜

厚を算出することが考えられる。たとえば、得られたスペクトルを $A + B\cos(\omega k + \eta)$ ($k = 1, 2, 3, \dots, 1024$) とモデル化し、パラメータ A, B, ω, η を高速に精度よく求める方法の検討が考えられる。このとき、膜厚は、

$$\omega = 2\pi\nu/N$$

$$(N = 1024, \nu: \text{膜厚に比例する周波数})$$

の ν から求まる。また、使用波長を紫外域に短くすることも重要である。

参考文献

- [1] 工藤：インライン多層膜の新しい測定手法；システム制御情報学会論文誌，Vol. 18, No. 1, pp. 28–38 (2005)
- [2] M. Kaveh and G. A. Lippert: An optimum tapered Burg algorithm for linear prediction and spectral analysis; *IEEE Trans. Acoust., Speech, Signal Process.*, Vol. ASSP-31, No. 2, pp. 438–444 (April 1983)
- [3] M. P. Quirk and B. Lin: Improving resolution for autoregressive spectral estimation by decimation; *IEEE Trans. Acoust., Speech, Signal Process.*, Vol. ASSP-31, No. 3, pp. 630–637 (June 1983)
- [4] 足立，酒井，高橋，飯國：システム工学（機械系大学講義シリーズ 31），第 2 章，コロナ社（1996）
- [5] 工藤，酒井：特願 2005-354267「多層薄膜の膜厚測定方法および膜厚測定装置」（2005）
- [6] 榊原：ウェーブレットビギナーズガイド，東京電機大学（1995）
- [7] A. Savitzky and M. J. E. Golay: Smoothing and differentiation of data by simplified least squares procedures; *Analytical Chemistry*, Vol. 36, No. 8, pp. 1627–1639 (July 1964)

付録 1

ここで、 $\text{stn}_1, \text{std}_1$ の導出について補足説明する。まず、 stn_1 から式の導出を行う。

$$\text{stn}_1 = \sum_{j=0}^{K-1} \sum_{i=0}^{L-1} w_{1,Ki+j} \cdot \cos[(Ki+j)\theta + \phi] \cdot \cos[(K(i+1)+j)\theta + \phi]$$

$\cos$ の加法定理を用いて、

$$\text{stn}_1 = \left(\frac{1}{2}\right) \sum_{j=0}^{K-1} \sum_{i=0}^{L-1} w_{1,Ki+j} \{ \cos[K\theta] + \cos[2(Ki+j)\theta + 2\phi + K\theta] \}$$

ここで、

$$k = Ki + j, \quad 0 \leq k \leq N-1$$

と変数 i, j を変数 k に置き換える。これを用いると、 stn_1 はつぎのようになる。

$$\begin{aligned}
\text{stn}_1 &= \left(\frac{1}{2}\right) \sum_{k=0}^{N-1} w_{1,k} \{\cos[K\theta] + \cos[2k\theta + 2\phi + K\theta]\} \\
&= \left(\frac{1}{2}\right) (\cos[K\theta] + \sum_{k=0}^{N-1} w_{1,k} \cos[2k\theta + 2\phi + K\theta])
\end{aligned} \quad (20)$$

つぎに, std_1 の導出について説明する.

$$\begin{aligned}
\text{std}_1 &= \sum_{j=0}^{K-1} \sum_{i=0}^{L-1} w_{1,Ki+j} \{\cos^2[(Ki+j)\theta + \phi] \\
&\quad + \cos^2[(K(i+1)+j)\theta + \phi]\}
\end{aligned}$$

$\cos$ の倍角の公式を用いて,

$$\begin{aligned}
\text{std}_1 &= \sum_{j=0}^{K-1} \sum_{i=0}^{L-1} w_{1,Ki+j} \left\{ \frac{1 + \cos[2(Ki+j)\theta + 2\phi]}{2} \right. \\
&\quad \left. + \frac{1 + \cos[2(Ki+j)\theta + 2\phi + 2K\theta]}{2} \right\}
\end{aligned}$$

ここで, 上と同様に, $k = Ki + j$, $0 \leq k \leq N-1$ と置き換え, $\cos$ の和を積に変換する公式を用いる.

$$\begin{aligned}
\text{std}_1 &= 1 + \left(\frac{1}{2}\right) \sum_{j=0}^{K-1} \sum_{i=0}^{L-1} w_{1,Ki+j} \{\cos[2(Ki+j)\theta + 2\phi] \\
&\quad + \cos[2(Ki+j)\theta + 2\phi + 2K\theta]\} \\
&= 1 + \left(\frac{1}{2}\right) \sum_{k=0}^{N-1} w_{1,k} \{\cos[2k\theta + 2\phi] \\
&\quad + \cos[2k\theta + 2\phi + 2K\theta]\} \\
&= 1 + \left(\frac{1}{2}\right) \sum_{k=0}^{N-1} w_{1,k} \{2\cos[2k\theta + 2\phi + K\theta] \cdot \cos[K\theta]\} \\
&= 1 + \cos[K\theta] \cdot \sum_{k=0}^{N-1} w_{1,k} \cos[2k\theta + 2\phi + K\theta] \quad (21)
\end{aligned}$$

ここで, (20),(21) 式を用いて $\hat{a}_{1,1}$ を計算すれば, (9) 式が得られる.

$$\begin{aligned}
\hat{a}_{1,1} &= \frac{2\left(\frac{1}{2}\right)(\cos[K\theta] + \sum_{k=0}^{N-1} w_{1,k} \cos[2k\theta + 2\phi + K\theta])}{1 + \cos[K\theta] \cdot \sum_{k=0}^{N-1} w_{1,k} \cos[2k\theta + 2\phi + K\theta]} \\
&= \cos[K\theta] + \frac{\sin^2[K\theta] \cdot \sum_{k=0}^{N-1} w_{1,k} \cos[2k\theta + 2\phi + K\theta]}{1 + \cos[K\theta] \cdot \sum_{k=0}^{N-1} w_{1,k} \cos[2k\theta + 2\phi + K\theta]}
\end{aligned}$$

ここで, 文献 [2] と同様にして, 分母の第 2 項が小さいと示せるので,

$$\hat{a}_{1,1} \simeq \cos[K\theta] + \sin^2[K\theta] \cdot \sum_{k=0}^{N-1} w_{1,k} \cos[2k\theta + 2\phi + K\theta]$$

となる.

付録 2

装置の概要を Fig. A1 に記す. 詳細は文献 [1] に示すとおりである.

Fig. A1 General drawing of the instrument

分光器の測定領域は, 測定開始波長 (λ_1) として 446.54[nm], 測定終了波長 (λ_2) として 953.95[nm] である. このとき, 単層膜の膜厚 d は, つぎの式で表される.

$$2nd = m / (1/\lambda_1 - 1/\lambda_2) \quad (22)$$

ここで, n は屈折率, m は λ_1 と λ_2 との間に含まれる干渉信号の山の数で, 周波数に対応する. また, 波長の逆数は波数で, (22) 式は波数 $\sigma_1 (=1/\lambda_1)$ と波数 $\sigma_2 (=1/\lambda_2)$ との間に含まれる山の数と膜厚が比例していると解釈できる. この周波数は, 波数の差のみに依存する. このため, 論文の中では, 波長を波数に変換して議論している.

本論文では, σ_1 と σ_2 間の 1024 個のデータに対して, PS を求める際に OTBD 法で求めた予測フィルタの後に 0 を付け加え, 8192 点に拡張して FFT 処理している. この場合, PS のピーク位置は, 8182 点の場合の M に対応し, (22) 式の m とは, $M = 8m$ の関係にある. LY1 の屈折率として 2.0 を用い, 膜厚 d を $0.35\mu\text{m}$ とすると, $8m$ は

$$M = 8m = 8 \times 2nd \times (1/\lambda_1 - 1/\lambda_2) = 13.5[\text{bin}]$$

となる. また, LY2 の屈折率 n は 1.7 近辺であるので, $25\mu\text{m}$ では, おおよそ 820[bin] である. Table 2,3 より LY2 は若干薄めである.

付録 3

文献 [7] の方法に基づいた放物線近似 $y = a_{20} + a_{21}X + a_{22}X^2$ を行う. PS の最大値を与える点を中心に前後の 5 点に対して, それに対する値を y_i ($i = 1, 2, \dots, 5$) とし, 最大値を y_3 とした場合, 放物線の係数 a_{20}, a_{21}, a_{22} はつぎの式で与えられる.

$$a_{20} = [-3(y_1 + y_5) + 12(y_2 + y_4) + 17y_3]/35$$

$$a_{21} = [2(y_5 - y_1) + (y_4 - y_2)]/10$$

$$a_{22} = [2(y_1 + y_5) - (y_2 + y_4) - 2y_3]/14$$

このとき、近似放物線の最大値を与える Xm 座標は、

$$Xm = -a_{21}/(2a_{22}) + 3$$

となる．一例として、S01 のピーク近傍の様子を Fig. A2 に示す．なお、LY1+LY2 と LY2 の膜厚に対応する x 座標は原データの y 値が max になる点であるが、グラフでは、 $x=3$ として表示している．ここで、近似曲線での max を与える位置は、記号「|」で示した．

Fig. A2 Detailed power spectrum near the region of the peak for S01 data and the smoothed curve