

網膜神経線維層厚の特徴を用いた緑内障疾患領域の検出

Detection of Affected Segments from Glaucoma Using Features of Nerve Fiber Layer

精密工学専攻 33号 滝田 清
Kiyoshi Takita

1. 序論

近年、高齢化が進んでいるわが国では国民の健康への意識が高まり、いかに健康寿命を延ばすかに関心が向いている。このような背景から健康診断や人間ドックなどを受診する人の数、およびその頻度は増加傾向にある。今後もこのことは続くと予想することができ、大きくなり続けている医師への負担が問題として挙げられる。この問題を軽減する為、Computer-aided diagnosis (CAD)の開発が進められている。CADとして乳癌⁽¹⁾や蜘蛛膜下出血⁽²⁾に対するものが一般的であるが緑内障に対するCADも存在する。緑内障は岐阜県多治見市における疫学調査の結果、40歳以上の20人に1人が患っていると報告されており^(3,4)、2020年までに全世界で緑内障の罹患者数が約8000万人に達すると予測されている⁽⁵⁾極めて身近な病気である。緑内障に対するCADでは乳頭輪郭に対する陥凹部の比(Cup/Disc比)⁽⁶⁾や乳頭出血、網膜神経線維欠損の検出^(7~9)などがあるが、これらの診断補助では緑内障の初期段階での診断を行うことが困難である。そこで眼と緑内障の特徴に注目し、乳頭と黄斑を通る直線を対称線と定義した時の網膜神経線維層厚の対称性⁽¹⁰⁾を評価することで、病状が初期の場合であっても緑内障の可能性を示唆出来ることが示されている。しかし診察の際、最終的な判断には医師の所見は必要不可欠であり、この手助けを行うためには単に病状を判定するだけでなく疾患領域を検出することが重要であると考えられる。先行研究⁽¹¹⁾では網膜神経線維層の厚みと非対称性に注目し疾患領域の検出を行っているが十分な検出性能ではない。そこで本研究では、網膜神経線維層の厚み、非対称性に加えて差分、ばらつきを考慮し、複数の異なった検出アルゴリズムを用いることで疾患領域の検出性能の向上を目指す。

2. 評価値の取得

疾患領域を検出するために、網膜神経線維層の厚み、非対称性、差分、ばらつきを表す四つの評価値を取得する。非対称性を表す評価値を定義するため、Fig.1に示す様に乳頭と黄斑の中心を通る直線を対称線として取得する。網膜神経線維層の厚みはFig.2で示すOptical coherence tomography (OCT)画像から取得できるNerve fiber layer (NFL)の厚みを用いる。スキャン範囲を S 、 S 内のデータ数を $N(S)$ 、ある位置 A の座標を (x, y) 、 A でのNFLの厚さを $thick(A)$ とする。対称性を評価するには、線対称となる位置がスキャン範囲内にある必要がある。対称位置がそれぞれスキャン範囲にあるデータ範囲を S' とし、 A の対称位置を B とする。差分を求める際に注目

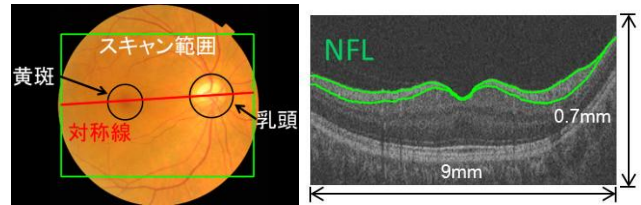


Fig.1 Fundus image

Fig.2 OCT image

点 A に対して A を中心に $U \times V$ の範囲で計算を行う。厚み $thickness$ 、非対称性 $asymmetry$ 、差分 $difference$ という評価値を次式で定義する。

$$thickness(A) = thick(A) \quad (1)$$

$$asymmetry(A) = \frac{thick(A) - thick(B)}{\{thick(A) + thick(B)\}/2} \quad (2)$$

$$difference(A) = \frac{1}{(U \times V) - 1} \sum_{u=1}^U \sum_{v=1}^V \left| thick(x, y) - thick\left(x - \frac{U+1}{2} + u, y - \frac{V+1}{2} + v\right) \right| \quad (3)$$

また、ばらつきを表す評価値 $thick_SD$ を定義するため、スキャン範囲の行、列ごとの厚みの標準偏差 $row_SD(A)$ 、 $column_SD(A)$ を算出する。 $row_SD(A)$ 、 $column_SD(A)$ を用いて $thick_SD(A)$ を次式で定義する。

$$thick_SD(A) = \frac{column_SD(A)}{row_SD(A)} \quad (4)$$

3. 疾患領域検出

評価値を取得した後、スキャン範囲を分割し、分割した領域ごとで評価値の平均値を算出する。各分割領域で学習用データを用いて「正常」、「緑内障」クラスを作成し、作成したクラスを用いて疾患領域を検出する。

3.1 評価値の平均値算出

疾患領域を検出するためスキャン範囲を $M \times N$ に分割する。分割はFig.3に示す様に乳頭と黄斑の中心を基準とし、乳頭と黄斑を通る直線に対して平行、垂直に行う。分割した後、各領域内で取得した評価値から (m, n) 番目の領域の平均値 $Ave_thick(m, n)$ 、 $Ave_asym(m, n)$ 、 $Ave_diff(m, n)$ 、 $Ave_SD(m, n)$ ($m=1, \dots, M$ 、 $n=1, \dots, N$)を算出する。

3.2 クラスの作成

学習用データを用いて各分割領域で「正常」、「緑内障」の二クラスを作成する。「正常」クラスでは、各分割領域で全てのデータを用いてクラスを作成する。「緑内障」クラスで

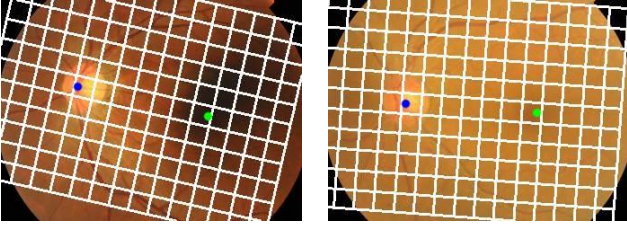


Fig.3 Results of area segmentation

は、各分割領域で疾患領域の割合が適当なしきい値以上の箇所のデータのみを用いてクラスを作成する。疾患領域か否かの判断は眼科医の協力のもと行った。

3.3 判別手法

作成した二クラスを用いて以下に説明する4種類の分類アルゴリズムで疾患領域か否かを判別する。

3.3.1 マハラノビス距離法

各分割領域に対して「正常」, 「緑内障」のクラスを作成し、クラス C における学習用評価値ベクトルを $\mathbf{X}^{(C)}$, $\mathbf{X}^{(C)}$ の次元数を K とする。クラス C における学習用評価値のデータ数を $T^{(C)}$ とし、平均値ベクトル $\mathbf{A}^{(C)}$, 共分散行列 $\mathbf{S}^{(C)}$ を式(5), (6), (7)より求める。

$$\mathbf{X}_t^{(C)} = [x_{t1}, \dots, x_{tK}]^T \quad (t=1, \dots, T^{(C)}) \quad (5)$$

$$\mathbf{A}^{(C)} = \frac{1}{T^{(C)}} \sum_{t=1}^{T^{(C)}} \mathbf{X}_t^{(C)} \quad (6)$$

$$\mathbf{S}^{(C)} = \begin{bmatrix} S_{11}^{(C)} & \dots & S_{1K}^{(C)} \\ \vdots & \ddots & \vdots \\ S_{K1}^{(C)} & \dots & S_{KK}^{(C)} \end{bmatrix} \quad (7)$$

$$S_{ij}^{(C)} = \frac{1}{T^{(C)}} \sum_{t=1}^{T^{(C)}} (x_{ti} - \mathbf{A}_i^{(C)}) (x_{tj} - \mathbf{A}_j^{(C)})$$

各クラスに対するマハラノビス距離 $d_m^{(C)}$ を次式により求める。

$$d_m^{(C)2} = (\mathbf{x} - \mathbf{A}^{(C)})^T (\mathbf{S}^{(C)})^{-1} (\mathbf{x} - \mathbf{A}^{(C)}) \quad (8)$$

各分割領域で「緑内障」クラスに対するマハラノビス距離の値の方が小さい場合、その分割領域を疾患領域とする。

3.3.2 最尤法 (MLE)

各クラスに対する尤度 $L^{(C)}$ を次式により求める。

$$L^{(C)} = \frac{1}{(2\pi)^{\frac{K}{2}} |\mathbf{S}^{(C)}|^{\frac{1}{2}}} \exp\left(-\frac{1}{2} (\mathbf{x} - \mathbf{A}^{(C)})^T (\mathbf{S}^{(C)})^{-1} (\mathbf{x} - \mathbf{A}^{(C)})\right) \quad (9)$$

各分割領域で「緑内障」クラスに対する尤度の値の方が大きい場合、その分割領域を疾患領域とする。

3.3.3 最短距離法 (NN)

「正常」, 「緑内障」クラスの学習用評価値ベクトルからユークリッド距離が最小となるサンプルを求める。そのサンプルが属しているクラスが「緑内障」クラスであった場合、その分割領域を疾患領域とする。

3.3.4 サポートベクターマシン (SVM)

疾患領域検出は「正常」, 「緑内障」に分類する二クラスの

識別問題とすることが可能なので SVM を適用する。疾患領域検出問題では学習データが線形分離不可能であるため、特徴空間において非線形の識別面が必要となる。そこで本手法では SVM にソフトマージン法とカーネルトリック^(12, 13)を適用する。「正常」と「緑内障」の特徴空間を分離する超平面は以下の条件付き最適化問題の最適解から求める。

$$\text{minimize} \quad L(\mathbf{w}, \zeta) = \frac{1}{2} \|\mathbf{w}\|^2 + \gamma \sum_{a=1}^K (\zeta_a) \quad (10)$$

$$\text{subject to} \quad \zeta_a \geq 0, s_a (\mathbf{w}^T \mathbf{X}_t^{(C)} + b) \geq 1 - \zeta_a \quad (a=1, \dots, K) \quad (11)$$

\mathbf{W} は重みベクトル, γ は識別エラーに対するペナルティ定数, $\zeta_a (\geq 0)$ は不完全な分離を許容するパラメータ, s_a はクラスラベル, b はバイアス項と呼ばれるパラメータである。各分割領域で求められた値を用いて疾患領域か否かの判別を行う。

4. 疾患領域検出性能評価実験

提案した疾患領域検出システムの有用性と疾患領域検出における最適な分類アルゴリズムを検証するため実験を行った。4.1 節では実験条件および提案手法で用いるパラメータの設定について述べる。4.2 節では検出性能を感度 (Recall), 精度 (Precision), 正答率 (Accuracy) の観点から検証する。

4.1 実験条件およびパラメータの設定

評価値 *difference* を取得する際の U , V はそれぞれ 1, 3 とした。

スキャン範囲は 15×11 に分割し、「正常」クラスの作成には正常 13 症例を用いた。「緑内障」クラスの作成には、医師によって疾患領域が特定されている緑内障 14 症例から疾患領域のみを用いた。

「緑内障」クラス作成時の分割領域内での疾患領域の割合に対するしきい値は 0.15 とした。このしきい値は経験的に決定した。

本手法では、カーネルトリックを SVM に適用し学習する際、パラメータの設定が比較的容易である Gauss カーネルを採用した。

$$\mathbf{K}(x_1, x_2) = \exp\left(-\frac{\|x_1 - x_2\|^2}{\sigma^2}\right) \quad (12)$$

ここで、ソフトマージン法で用いる式(10)の γ と、カーネルトリックで用いる式(12)の σ の二つの未知数を決定するために Grid Search を適用した。これは二つのパラメータを γ と σ を対として、それぞれのパラメータを規則的かつ網羅的に変化させ計算を行い、結果をプロットしてその中で最も性能の高い組み合わせを選択する手法である。今回は $\gamma=2^{-20}, 2^{-19}, 2^{-18}, \dots, 2^{10}$, $\sigma=2^{-6}, 2^{-5}, 2^{-4}, \dots, 2^{21}$ と動かして計算を行い、学習データに対する leave-one-out cross-validation の分類性能をプロットしてパラメータを決定した。分類性能を表す指標として d を次式により求めた。

$$d = FPrate^2 + (1 - TPrate)^2 \quad (13)$$

Table 1 Parameters for SVM

| | Thickness | Asymmetry | Difference |
|----------------|-----------|-----------------------|------------|
| $\log(\gamma)$ | -7 | 1 | 3 |
| $\log(\sigma)$ | 16 | 15 | 20 |
| | Thick_SD | Thickness & Asymmetry | All |
| $\log(\gamma)$ | -2 | -9 | -14 |
| $\log(\sigma)$ | 21 | 19 | 19 |

$$TPrate = TP/(TP + FN) \quad (14)$$

$$FPrate = FP/(FP + TN) \quad (15)$$

ここで TP , FP , TN および FN はそれぞれ true positive, false positive, true negative 及び false negative である. 決定したパラメータを Table 1 に示す.

4.2 実験

医師によって疾患領域が特定されている 14 症例に対して疾患領域検出を行った. また, 正常 13 症例に対しても疾患領域検出を行った. 性能評価に用いる感度(*Recall*), 精度(*Precision*), 正答率(*Accuracy*)はそれぞれ次式から求めた.

$$Recall = TP/(TP + FN) \quad (16)$$

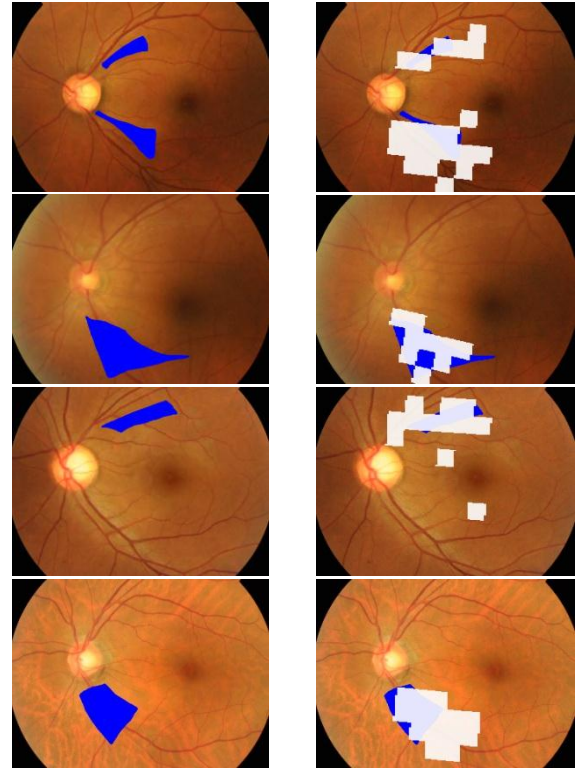
$$Precision = TP/(TP + FP) \quad (17)$$

$$Accuracy = (TP + TN)/(TP + FP + TN + FN) \quad (18)$$

感度は疾患群が陽性になる確率, 精度は陽性と分類したうちの疾患群である確率, 正答率は正しく判断される確率を表している. 疾患領域検出方法ごとの感度, 精度, 正答率の比較を Table 2 に, *Ave_thick* を用いたマハラノビス距離での検出例を Fig.4 に示す. 四つの評価値を組み合わせた際の結果を *All* と記載する. 実験には *leave-one-out cross-validation* を用いた.

本実験では性能を評価するにあたり, 感度, 精度, 正答率と三つの評価指標を用意したが, その中でも感度の結果を最も重要視する. その理由として, 提案手法は医療の現場で用いるものとして想定しており, 疾患群を見逃してしまうことが一番の危険であると考えられるためである.

まずは緑内障の症例に対する結果をみる. 感度では *Ave_thick* を用いたマハラノビス距離法での検出が 71.5% と一番高く, 次いで *All* を用いた SVM での検出が二番目に高く 70.1% という結果を得ることが出来た. 精度では *Ave_asym* を用いた最短距離法での検出が 46.1% と一番高い結果を得ることが出来た. しかしこの時の感度をみると 41.0% と低い結果であった. 正答率では精度の結果が高い検出手法ほど結果が高い傾向が見受けられた. 評価値ごとに結果をみると *Ave_thick* の感度が各手法の中では高い結果であった. これは今回実験に用いた 14 症例全ての病状がある程度進行しており, 網膜神経線維層厚の減少がはっきりしていたためだと考えられる. *Ave_asym* は病状が初期の場合 *Ave_thick* よりも変化に対して敏感に反応する⁽¹⁰⁾ことがわかっている. しかし今回の実験では感度はそれほど高い結果ではなかった. これは *Ave_asym* が乳頭と黄斑を通る直線の上下の網膜神経層厚の厚みの差を表した評価値であり, 緑内障の進行が進み疾患領



(a) Affected Segment (b) Mahalanobis using Ave_thick

Fig.4 Results of glaucoma region detection

Table 2 Experimental results [%]

| Disease condition | | 緑内障 | | | 正常 |
|-------------------|---------------------------------|------|------|------|------|
| 手法 | 評価値 | 感度 | 精度 | 正答率 | 正答率 |
| Mahalanobis | <i>Ave_thick</i> | 71.5 | 39.0 | 91.3 | 97.0 |
| | <i>Ave_asym</i> | 58.3 | 37.7 | 91.4 | 93.7 |
| | <i>Ave_dif</i> | 43.1 | 24.8 | 88.3 | 93.7 |
| | <i>Ave_SD</i> | 59.7 | 21.0 | 83.5 | 85.1 |
| | <i>Ave_thick & Ave_asym</i> | 58.3 | 36.2 | 91.0 | 97.4 |
| | <i>All</i> | 47.2 | 41.7 | 92.6 | 99.2 |
| MLE | <i>Ave_thick</i> | 68.1 | 38.1 | 91.1 | 96.6 |
| | <i>Ave_asym</i> | 52.1 | 40.1 | 92.2 | 93.6 |
| | <i>Ave_dif</i> | 57.6 | 25.3 | 86.8 | 89.7 |
| | <i>Ave_SD</i> | 34.7 | 19.7 | 87.1 | 93.6 |
| | <i>Ave_thick & Ave_asym</i> | 54.2 | 37.7 | 91.6 | 97.5 |
| | <i>All</i> | 46.5 | 42.1 | 92.7 | 99.3 |
| Nearest Neighbor | <i>Ave_thick</i> | 66.0 | 37.0 | 90.9 | 97.4 |
| | <i>Ave_asym</i> | 41.0 | 46.1 | 93.3 | 96.2 |
| | <i>Ave_dif</i> | 29.9 | 26.2 | 90.4 | 95.2 |
| | <i>Ave_SD</i> | 25.7 | 22.2 | 89.7 | 95.3 |
| | <i>Ave_thick & Ave_asym</i> | 52.1 | 43.1 | 92.7 | 97.6 |
| | <i>All</i> | 51.4 | 43.3 | 92.8 | 97.8 |
| SVM | <i>Ave_thick</i> | 68.1 | 40.8 | 91.9 | 99.5 |
| | <i>Ave_asym</i> | 36.8 | 42.7 | 93.0 | 99.5 |
| | <i>Ave_dif</i> | 36.8 | 29.3 | 90.5 | 99.5 |
| | <i>Ave_SD</i> | 28.5 | 26.6 | 90.6 | 99.8 |
| | <i>Ave_thick & Ave_asym</i> | 64.6 | 30.6 | 88.7 | 99.8 |
| | <i>All</i> | 70.1 | 31.1 | 88.4 | 99.8 |

域が眼全体に広がると評価値の変化の差が小さくなるためだと考えられる。そのため、*Ave_asym* を組み合わせ *Ave_thick* & *Ave_asym* や *All* の感度の結果が *Ave_thick* に比べて低くなったと考えられる。初期の症例で実験を行った場合、*Ave_asym* の感度の方が高くなると予想することが出来る。一方精度では *Ave_asym* を用いた場合が各手法で高い結果であった。これは対称性をみることで緑内障の症状の特徴をとらえることが出来ていたためだと考えられる。*Ave_thick* の精度が低い理由としては、層の厚みには個人差が存在し、正常箇所であっても他と比較した時に厚みが薄ければ疾患領域とされてしまうからだと考えられる。

次に正常の結果をみる。正答率が 99%以上の結果を複数の検出手法で得ることが出来た。SVM での検出手法では、評価値によらず 99.5%以上の高い結果であった。中でも *Ave_thick*, *All* を用いた時は緑内障の感度が高いにも関わらずこのような高い結果を得られたことから、正常群に対して正しく陰性であると判断出来ていることがわかる。

以上をまとめると、本手法を用いて疾患領域を検出することが可能であることを確認した。一方で、検出手法ごとでは明確な優劣をつけられる程の差は出なかったが、その中でも *Ave_thick* を用いたマハラノビス距離法や SVM での検出は比較的高い結果を得られた。疾患領域検出の感度、精度ともに実用化するには十分な結果を得ることは出来なかったが、領域検出の根幹となるシステムを構築することが出来たと考えられる。

5. 結論と今後の展望

本研究では網膜神経線維層の厚み、非対称性、差分、ばらつきを評価値として取得し、異なる分類アルゴリズムを用いて緑内障疾患領域の検出手法を提案した。提案手法はスキャン範囲内に存在する乳頭と黄斑の位置をもとに領域を分割し、分割領域ごとで取得した評価値の平均値を算出する。各分割領域で平均値を用いて「正常」、「緑内障」クラスの作成を行い、マハラノビス距離法、最尤法、最短距離法、サポートベクターマシンといった異なる分類アルゴリズムにより疾患領域の検出を行った。また、提案手法の有用性を検証するため、予め医師によって疾患領域と特定されている個所と検出結果の比較検討を行った。検出結果から複数の検出手法での明確な優劣の差は現れなかった。正常に対する結果では高い正答率を得ることが出来た。一方、緑内障に対する感度、精度は実際の医療の現場に用いられる程の結果を得ることは出来なかった。しかし、ある程度の領域検出を行えたことから根幹となるシステムの構築を行うことが出来た。

今後の展望として、検出性能の向上が挙げられる。そのため、クラスの作成に用いる症例数を「正常」、「緑内障」とともに増やすことが必要であると考えられる。また、現在は一つ一つの検出手法を単独で用いているが、複数の手法を組み合わせることで検出結果の向上を目指す。

参考文献

- (1) 加野亜紀子, “乳がんの画像診断を支援するコンピュータ自動解析システムの開発”, 医用画像情報学会誌, 21(1), 79-83, 2004.
- (2) 小椋潤, 内山良一, 山内将史, 他, “ベクトル集中度フィルタを用いたMRA画像における脳動脈瘤の検出法”, 医用画像情報学会雑誌, 24(2), pp.84-89, 2007.
- (3) A. Iwase, Y. Suzuki, M. Araie, et al, “The prevalence of primary open-angle glaucoma in Japanese: The Tajimi study,” *Ophthalmology*, 111(9), pp.1641-1648, 2004.
- (4) T. Yamamoto, A. Iwase, M. Araie, et al, “The Tajimi study report 2 prevalence of primary angle closure and secondary glaucoma in a Japanese population,” *Ophthalmology*, 112(10), pp.1661-1669, 2005.
- (5) H. Quigley, and A. Broman, “The number of people with glaucoma worldwide in 2010 and 2020,” *Br J Ophthalmology*, 90(3), pp.262-267, 2006.
- (6) 畑中裕司, 福田圭祐, 納土淳, 他, “眼底画像上の神経乳頭部解析による緑内障リスク推定,” 信学技報(MI), 109(127), pp.25-28, 2009.
- (7) 小林義治, “OCTの見方,” 日本視能訓練士協会誌, 39, pp.21-27, 2010.
- (8) Y. Hayashi, T. Nakagawa, Y. Hatanaka, et al, “Detection of retinal nerve fiber layer defects in retinal fundus images using Gabor filtering,” in Proc. SPIE Medical Imaging 2007: Computer-aided Diagnosis, San Diego, 6514, pp.65142Z-1-65142Z-8, 2007.
- (9) R. Sihota, P. Sony, V. Gupta, et al, “Diagnostic capability of optical coherence tomography in evaluating the degree of glaucomatous retinal nerve fiber damage,” *Invest Ophthalmol Vis Sci*, 47, pp.2006-2010, 2006.
- (10) 滝田清, 西優紀美, 寺林賢司, 梅田和昇, 富所敦男, “OCT画像から得られる網膜の層情報を用いた緑内障診断補助,” 画像センシングシンポジウム 2012, IS1-11, 2012.
- (11) 滝田清, 寺林賢司, 梅田和昇, 富所敦男, “網膜神経線維層厚の非対称性を用いた緑内障疾患領域の検出,” 精密工学会秋季大会 2012, G18, 2012.
- (12) N. Cristianini and J. S. Taylor, “Introduction to Support Vector Machines and Other Kernel-based Learning Methods,” Cambridge University Press, 2000.
- (13) 前田英作, 村瀬洋, “カーネル非線形部分空間法によるパターン認識,” 電子情報通信学会論文誌, J82-D-II(4), pp.600-612, 1999.