



(19)
Bundesrepublik Deutschland
Deutsches Patent- und Markenamt

(10) **DE 699 25 155 T2** 2006.01.12

(12)

Übersetzung der europäischen Patentschrift

(97) **EP 1 138 026 B1**

(21) Deutsches Aktenzeichen: **699 25 155.9**

(86) PCT-Aktenzeichen: **PCT/IB99/02012**

(96) Europäisches Aktenzeichen: **99 958 432.9**

(87) PCT-Veröffentlichungs-Nr.: **WO 00/33262**

(86) PCT-Anmeldetag: **01.12.1999**

(87) Veröffentlichungstag

der PCT-Anmeldung: **08.06.2000**

(97) Erstveröffentlichung durch das EPA: **04.10.2001**

(97) Veröffentlichungstag

der Patenterteilung beim EPA: **04.05.2005**

(47) Veröffentlichungstag im Patentblatt: **12.01.2006**

(51) Int Cl.⁸: **G07D 5/00** (2006.01)

G07D 7/00 (2006.01)

G06K 9/62 (2006.01)

(30) Unionspriorität:

9826494 02.12.1998 GB

(73) Patentinhaber:

Mars Inc., McLean, Va., US

(74) Vertreter:

Strehl, Schübel-Hopf & Partner, 80538 München

(84) Benannte Vertragsstaaten:

DE, ES, FR, GB, IT

(72) Erfinder:

BAUDAT, Gaston, CH-1202 Genève, CH

(54) Bezeichnung: **Verfahren und Vorrichtung zum Klassifizieren von Gegenständen**

Anmerkung: Innerhalb von neun Monaten nach der Bekanntmachung des Hinweises auf die Erteilung des europäischen Patents kann jedermann beim Europäischen Patentamt gegen das erteilte europäische Patent Einspruch einlegen. Der Einspruch ist schriftlich einzureichen und zu begründen. Er gilt erst als eingelegt, wenn die Einspruchsgebühr entrichtet worden ist (Art. 99 (1) Europäisches Patentübereinkommen).

Die Übersetzung ist gemäß Artikel II § 3 Abs. 1 IntPatÜG 1991 vom Patentinhaber eingereicht worden. Sie wurde vom Deutschen Patent- und Markenamt inhaltlich nicht geprüft.

Beschreibung

[0001] Die Erfindung betrifft ein Verfahren und eine Vorrichtung zum Klassifizieren von Gegenständen. Die Erfindung betrifft insbesondere die Klassifikation von Münzen und Banknoten.

[0002] In Mechanismen wie Verkaufsautomaten, Wechselautomaten und dergleichen eingeworfene Münzen und Banknoten werden einerseits entsprechend ihrem Wert und andererseits dahingehend klassifiziert, ob es sich um Originale oder Kopien und Fälschungen handelt. Es sind verschiedene Verfahren zum Durchführen einer solchen Klassifikation bekannt. Ein Beispiel beschreibt die GB 2 238 152 A, deren Inhalt hier durch Bezugnahme eingeschlossen wird. Zum Beispiel werden an einer eingeworfenen Münze Messungen vorgenommen, die verschiedene Merkmale der Münze betreffen, etwa das Material und die Dicke. Diese Messungen werden dann jeweils mit einem gespeicherten Wertepaar verglichen, wobei jedes Wertepaar einem akzeptierbaren Nennwert einer Münze entspricht. Wenn der gemessene Wert in den jeweiligen Bereich für einen gegebenen Nennwert fällt, wird die eingeführte Münze durch die Klassifikation diesem Nennwert zugeordnet.

[0003] Bei dieser Art der Klassifikation können die gemessenen Werte als die Elemente eines Merkmalsvektors betrachtet werden, wobei dann die akzeptablen Meßwerte für verschiedene Nennwerte Bereichen im Merkmalsraum entsprechen, die als Akzeptanzbereiche bekannt sind. Bei dem obigen Beispiel ist der Merkmalsraum zweidimensional, und die Akzeptanzbereiche sind Rechtecke. Der Merkmalsraum kann jedoch eine beliebige Dimension haben, wobei dann die Akzeptanzbereiche entsprechend komplex werden. Zum Beispiel beschreibt die GB 2 254 949 A elliptische Akzeptanzbereiche in einem dreidimensionalen Merkmalsraum.

[0004] Andere Beispiele für Verfahren und Vorrichtungen zum Klassifizieren von Banknoten und Münzen sind in der EP 0 067 898 A, der EP 0 472 192 A und der EP 0 165 734 A beschrieben. Wieder andere Verfahren für die Klassifikation umfassen die Verwendung eines neuronalen Netzwerks, wie es zum Beispiel in der EP 0 553 402 A und der EP 0 671 040 A beschrieben ist.

[0005] Ein wesentliches Problem bei der Klassifizierung von Münzen ist die Schwierigkeit beim Trennen von verschiedenen Nennwerten. Die Populationsverteilungen der verschiedenen Nennwerte, die von Interesse sind, können derart sein, daß es nicht leicht möglich ist, geeignete Akzeptanzgrenzen festzulegen, mit denen die Nennwerte angemessen getrennt werden können. Ein anderes Problem ist, daß es zum Erreichen einer adäquaten Trennung erforderlich sein kann, Merkmalsvektoren mit einer großen Anzahl von Elementen in Betracht zu ziehen, was es erschwert, die verschiedenen Verteilungen zu verstehen, wodurch es noch schwieriger wird, geeignete Akzeptanzgrenzen zu erhalten. Diese Probleme ähneln den allgemeinen Klassifikationsproblemen in der Datenanalyse, die untersucht wurden und die zu verschiedenen Techniken einschließlich statistischer Methoden geführt haben.

[0006] Ein Beispiel für ein statistisches Verfahren in der Datenanalyse ist die Hauptkomponentenanalyse ("PCA"). Es ist dies ein Verfahren, bei dem die in einem Raum ausgedrückten Daten unter Verwendung einer linearen Transformation in einen neuen Raum übergeführt werden, in dem sich die meisten Variationen in den Daten mit weniger Dimensionen erklären lassen als im ersten Raum. Das PCA-Verfahren umfaßt das Herausfinden der Eigenvektoren und Eigenwerte der Kovarianzmatrix der Variablen. Die Eigenvektoren sind die Achsen im neuen Raum, wobei der Eigenvektor mit dem höchsten Eigenwert die erste "Hauptkomponente" ist und soweit mit abnehmender Größe. Einzelheiten über das PCA-Verfahren können in Lehrbüchern für die multivariate Analyse nachgeschlagen werden, etwa in "Introduction to Multivariate Analysis" von Chatfield und Collins, siehe dort Kapitel 4.

[0007] Ein anderes Verfahren für die Datenanalyse zu Klassifikationszwecken ist die lineare Diskriminantenanalyse ("LDA"). Die LDA ist von Nutzen, wenn bekannt ist, daß die Daten in separate Gruppen fallen. Die LDA hilft dabei, die Daten in einen neuen Raum zu transformieren, um den Abstand zwischen dem Mittelpunkt jeder Gruppe von Daten bei der Projektion auf die Achsen im neuen Raum zu maximieren und um die Streuung für jede Gruppe längs der Achsen zu minimieren. Verfahren dafür sind zum Beispiel in "Introduction to Statistical Pattern Recognition" von Fukunaga ("Fukunaga") beschrieben. Zum Beispiel erfolgt die Maximierung durch Herausfinden der linearen Transformation, die den Wert für den Durchstoßpunkt von $C^{-1}V$ maximiert, wobei V die Zwischenklassen-Kovarianzmatrix und C die Kovarianzmatrix aller Proben ist. Wie in Fukunaga erklärt, entspricht dies dem Herausfinden der Eigenvektoren und Eigenwerte von $C^{-1}V$. Die Eigenvektoren sind die Achsen des neuen Raumes. Wie in der Druckschrift beschrieben, hat der neue Raum für N Klassen $N-1$ Dimensionen.

[0008] In vielen Fällen führt jedoch werden die PCA noch die LDA zu einer angemessenen Trennung der Da-

tengruppen. Ein weiteres Verfahren für die Datenanalyse ist die nichtlineare Komponentanalyse (NCA), die auf der PCA beruht. In der NCA werden die Daten mittels einer nichtlinearen Abbildung in einen neuen Raum projiziert und dann in dem neuen Raum die PCA ausgeführt. Die Einzelheiten der NCA sind in dem Artikel "Nonlinear Component Analysis as a Kernel Eigenvalue Problem" von Bernhard Scholkopf, Alexander Smola und Klaus-Robert Müller, Neural Computation 10, 1299–1319 (1998) ("Scholkopf") beschrieben.

[0009] Ein Problem bei der NCA ist, daß die Dimension des nichtlinearen Raumes sehr groß sein kann, so daß die Anzahl der Hauptkomponenten ebenfalls sehr groß wird. Für ein gegebenes Problem ist nicht bekannt, wie viele Hauptkomponenten für eine gute Klassifikation erforderlich sind.

[0010] Die Aspekte der Erfindung sind aus den anhängenden Patentansprüchen 1, 8, 9, 16 und 17 ersichtlich.

[0011] Eine Ausführungsform der Erfindung wird anhand der beiliegenden Zeichnungen beschrieben. Es zeigen:

[0012] [Fig. 1](#) eine Blockdarstellung eines Klassifikationssystems.

[0013] [Fig. 2](#) graphisch die Verteilung von Daten für Münzen; und

[0014] [Fig. 3](#) graphisch eine Projektion der Daten der [Fig. 2](#) auf neue Achsen.

[0015] Die Erfindung wird anhand eines Münzprüfgeräts beschrieben.

[0016] In der [Fig. 1](#) bezeichnet der Kasten 1 ein Meßsystem mit einem Einlaß 2, einem Transportsystem in der Form eines Münzeinlaß- und Münztransportweges (nicht gezeigt) für die Vorlage einer Probe 3 und ein Sensorsystem (nicht gezeigt) zum Messen von physikalischen Größen an der Probe. Das Meßsystem 1 ist über einen Datenbus 5 mit einem Verarbeitungssystem 4 verbunden. Das Verarbeitungssystem 4 ist über einen Datenbus 7 mit einem Klassifizierer 6 verbunden. Das Ausgangssignal des Klassifizierers 6 wird mittels eines Datenausgangsbusses 9 zu einem Verwendungssystem 8 geführt. Das Verwendungssystem 8 ist im vorliegenden Beispiel ein Verkaufsautomat, kann jedoch auch zum Beispiel ein Geldwechselautomat sein.

[0017] Das Meßsystem 1 mißt Merkmale einer eingeworfenen Münze 3. Die gemessenen Merkmale werden zu einem Merkmalsvektor mit n Elementen zusammengesetzt, wobei jedes Element einem gemessenen Merkmal für das Verarbeitungssystem 4 entspricht. Im vorliegenden Beispiel mißt das Sensorsystem Werte für das Material, die Dicke und den Durchmesser einer eingeworfenen Münze mittels bekannter Techniken (siehe zum Beispiel die GB 2 254 949 A). Diese Werte sind die drei Elemente des entsprechenden Merkmalsvektors. Kurz gesagt umfaßt jeder Sensor eine oder mehrere Spulen in einem Schwingkreis. Im Falle des Durchmesser- und Dickensensors bewirkt die Änderung in der Induktanz der einzelnen Spulen, die durch die Nähe einer eingeworfenen Münze verursacht wird, eine Änderung der Schwingungsfrequenz, woraus eine digitale Darstellung der jeweiligen Eigenschaft der Münze abgeleitet werden kann. Im Falle eines Leitfähigkeitssensors bewirkt eine Änderung in der Ladung der Spule, die von der Nähe einer eingeworfenen Münze verursacht wird, daß sich die Spannung an der Spule ändert, woraus eine digitale Ausgangsdarstellung der Leitfähigkeit der Spule abgeleitet werden kann. Auch wenn der Aufbau, die Anordnung und Orientierung jeder Spule und die Frequenz der daran angelegten Spannung so ausgewählt werden, daß die Spule ein Ausgangssignal abgibt, das vor allem von einer bestimmten der Eigenschaften Leitfähigkeit, Durchmesser und Dicke abhängt, ist klar, daß jede Messung in einem gewissen Ausmaß auch von den anderen Eigenschaften der Münze beeinflusst wird.

[0018] Natürlich können viele verschiedene Merkmale von Währungsmitteln gemessen und als Elemente des Merkmalsvektors verwendet werden. Zum Beispiel können im Fall einer Banknote die gemessenen Merkmale die Breite der Banknote, die Länge der Banknote und die Intensität des von der ganzen oder einem Teil der Banknote reflektierten oder transmittierten Lichts umfassen. Das Meßsystem kann beispielsweise so angeordnet sein, daß es eine Banknote entlang N Linien mit optischen Sensoren abtastet. Jede Abtastlinie enthält L individuelle Bereiche, die nacheinander abgetastet werden. In jedem Bereich erfolgt die Messung von M verschiedenen Merkmalen. Das heißt zum Beispiel, daß in jedem Bereich die reflektierte Intensität von roten, grünen und infraroten Lichtstrahlen gemessen wird. Die Gesamtzahl der Messungen für eine Banknote ist daher gleich $L \times M \times N$. Diese Messungen bilden die Komponenten des Merkmalsvektors für die jeweilige Probe, so daß der Merkmalsvektor $L \times M \times N$ Komponenten hat. Alternativ können die Messungen auf eine andere Art verarbeitet werden, um einen Merkmalsvektor zu erhalten, der für die vermessene Probe repräsentativ ist. Zum Beispiel können aus den M Messungen für einen Bereich lokale Merkmalsvektoren für jeden Meßbereich gebildet werden, so daß jeder lokale Merkmalsvektor M Komponenten aufweist. Die lokalen Merkmalsvektoren

können dann über die Fläche der Banknote aufsummiert werden, um einen M-dimensionalen Merkmalsvektor zu erhalten, der für die ganze Probe steht.

[0019] Der Merkmalsvektor wird dann in den Klassifizierer **6** eingegeben. Der Klassifizierer **6** bestimmt, ob die Probe zu einer von vorgegebenen Klassen gehört, wozu der Merkmalsvektor und vorgegebene Klassifikationskriterien einschließlich einer Trennfunktion verwendet werden. Wenn die Probe als zu einem akzeptablen Nennwert einer Banknote gehörend identifiziert wird, wird sie akzeptiert und der entsprechende Wert der Banknote gutgeschrieben. Wenn die Probe als zu einer bekannten Gruppe von Fälschungen gehörend identifiziert wird, wird sie zurückgewiesen.

[0020] Bei dem vorliegenden Beispiel kann das System zwei Nennwerte von Münzen und eine bekannte Fälschung feststellen. Die [Fig. 2](#) zeigt eine zweidimensionale Darstellung der Verteilung im Meßraum. Die Kreuze stellen Proben für den ersten Nennwert, die Punkte Fälschungen des ersten Nennwerts und die Kreise Proben für den zweiten Nennwert dar.

[0021] Die Ableitung der Trennfunktion wird im folgenden allgemein beschrieben. Es wird dann das Verfahren zur Klassifikation beschrieben, ebenfalls allgemein, gefolgt von einer Erläuterung der Anwendung des allgemeinen Verfahrens auf das bestimmte Beispiel.

[0022] Kurz gesagt bildet das Verfahren zum Ableiten einer Trennfunktion gemäß der Ausführungsform der Erfindung den Eingangsraum, das heißt den Raum der gemessenen Merkmalsvektoren, mittels einer nichtlinearen Abbildung auf einen höherdimensionalen Raum mit linearen Eigenschaften ab. Im Abbildungsraum werden unter Verwendung von Trainingsdaten und dem Äquivalent einer LDA-Analyse im Abbildungsraum trennende Hyperebenen konstruiert.

[0023] Die Populationsverteilung für die einzelnen Nennwerte wird wie folgt analysiert.

[0024] Zuerst werden für jeden interessierenden Nennwert und jede bekannte Fälschung Proben vermessen und entsprechende Merkmalsvektoren ausgebildet. Die Merkmalsvektoren der einzelnen Proben fallen, wenn sie zum Beispiel in eine n-dimensionale Verteilungsgraphik eingezeichnet werden (wobei n die Anzahl der gemessenen Merkmale ist), grob in Gruppen, die Cluster genannt werden. Die vermessenen Proben werden dann dazu verwendet, eine Trennfunktion wie folgt abzuleiten. Bei dem vorliegenden Beispiel werden 50 Proben für jeden Nennwert und 50 Proben für die Fälschung verwendet.

[0025] Davor erfolgt noch eine allgemeine Erläuterung der verwendeten Notation.

[0026] Der Eingangsraum, das heißt der Raum der Merkmalsvektoren, wird als X definiert.

[0027]

$$X = \bigcup_{i=1}^N X_i,$$

, wobei N die Anzahl der Cluster ist. Die Mächtigkeit des Unterraums X_i wird mit n_i bezeichnet und die Anzahl der Elemente in X mit M. Es ist somit

$$\sum_{i=1}^N n_i = M. \quad x'$$

die Transponierte des Vektors x.

[0028] Im Eingangsraum ist C die Kovarianzmatrix und

$$C = \frac{1}{M} \sum_{j=1}^M x_j x_j' \quad (1)$$

[0029] Bei dem erfindungsgemäßen Verfahren wird eine Kernfunktion k verwendet, die im Abbildungsraum das Skalarprodukt definiert. ϕ sei eine nichtlineare Funktion, die X in den Hilbertraum F abbildet.

$$\phi: X \rightarrow F$$

$$x \rightarrow \phi(x)$$

(2)

mit

$$k(x,y) = \phi(x) \cdot \phi(y) = \phi^T(x) \phi(y)$$

[0030] Wie aus der folgenden Diskussion hervorgeht, ist es nicht notwendig, für ein gegebenes k das ϕ zu konstruieren, obwohl mit Mercer's Theorem gezeigt werden kann, daß ϕ existiert, wenn k der kontinuierliche Kern eines positiven Integraloperators ist (siehe Schölkopf Abschnitt 3 und Anhang C). Auch ist es nicht erforderlich, die Skalarprodukte explizit in F auszuführen, der ein unendlichdimensionaler Raum sein kann.

[0031] In F ist V die Kovarianzmatrix mit

$$V = \frac{1}{M} \sum_{j=1}^M \phi(x_j) \phi'(x_j) \quad (3)$$

[0032] Wir nehmen an, daß die Beobachtungen in F zentriert sind, das heißt daß

$$\sum_{j=1}^M \phi(x_j) = 0$$

ist. Ein Verfahren zum Zentrieren von Daten wird später noch beschrieben.

[0033] B ist die Kovarianzmatrix der Clusterzentren mit

$$B = \frac{1}{M} \sum_{l=1}^N n_l \overline{\phi_l} \overline{\phi_l'} \quad (4)$$

wobei $\overline{\phi_l}$ der Mittelwert des Clusters 1 ist, das heißt

$$\overline{\phi_l} = \frac{1}{n_l} \sum_{k=1}^{n_l} \phi(x_{lk}) \quad (5)$$

wobei x_{lj} das Element j des Clusters 1 ist.

[0034] B stellt die Zwischenclusterträgheit in F dar.

[0035] V kann mittels der Cluster auch ausgedrückt werden als

$$V = \frac{1}{M} \sum_{l=1}^N \sum_{k=1}^{n_l} \phi(x_{lk}) \phi'(x_{lk}) \quad (6)$$

[0036] V steht für die Gesamtträgheit in F .

[0037] Es sei $k_{ij} = k(x_i, x_j)$ und $(k_{ij})_{pq} = (\phi^T(x_{pi}) \phi(x_{qj}))$.

[0038] K sei eine $(M \times M)$ -Matrix, die an den Clusterelementen durch

$$(K_{pq})_{p=1 \dots N, q=1 \dots N}$$

definiert ist, wobei (K_{pq}) die Kovarianzmatrix zwischen dem Cluster _{p} und dem Cluster _{q} ist.

$$K = (K_{pq})_{p=1 \dots N, q=1 \dots N} \text{ where } K_{pq} = (k_{ij})_{i=1 \dots n_p, j=1 \dots n_q} \quad (8)$$

[0039] K_{pq} ist eine $(n_p \times n_q)$ -Matrix, und K ist symmetrisch, so daß $K'_{pq} = K_{pq}$ ist.

[0040] W ist das Matrixzentrum mit

$$W = (W_1)_{1=1 \dots N}$$

(9),

wobei W_1 eine $(n_1 \times n_1)$ -Matrix ist, deren Terme alle gleich $1/n_1$ sind.

[0041] W ist eine diagonale $M \times M$ -Blockmatrix.

[0042] Bei dem Verfahren erfolgt im wesentlichen eine lineare Diskriminantenanalyse im Abbildungsraum F , um die Interclustertr  gheit zu maximieren und die Intraclustertr  gheit zu minimieren. Dies ist der Eigenwertaufl  sung   quivalent, wie Fukunaga zeigt. Es kann dann eine geeignete Trennfunktion abgeleitet werden.

[0043] Das hei  t, das das Verfahren das Finden der Eigenwerte λ und der Eigenvektoren v umfa  t, die

$$\lambda Vv = Bv \quad (10)$$

erf  llen.

[0044] Die Eigenvektoren sind lineare Kombinationen der Elemente von F , so da   es Koeffizienten α_{pq} ($p = 1 \dots N$, $q = 1 \dots n_p$) gibt, derart, da  

$$v = \sum_{p=1}^N \sum_{q=1}^{n_p} \alpha_{pq} \phi(x_{pq}) \quad (11)$$

[0045] Die Eigenvektoren der Gleichung (10) sind die die gleichen wie die Eigenvektoren von

$$\lambda \phi'(x_{ij}) Vv = \phi'(x_{ij}) Bv \quad (12)$$

(siehe Sch  lkopf).

[0046] Unter Verwendung der Definitionen f  r K und W und der Gleichungen (6) und (11) kann die linke Seite von (12) wie folgt ausgedr  ckt werden:

$$\begin{aligned} Vv &= \frac{1}{M} \sum_{l=1}^N \sum_{k=1}^{n_l} \phi(x_{lk}) \phi'(x_{lk}) \sum_{p=1}^N \sum_{q=1}^{n_p} \alpha_{pq} \phi(x_{pq}) \\ &= \frac{1}{M} \sum_{p=1}^N \sum_{q=1}^{n_p} \alpha_{pq} \sum_{l=1}^N \sum_{k=1}^{n_l} \phi(x_{lk}) [\phi'(x_{lk}) \phi(x_{pq})] \end{aligned}$$

mit

$$\begin{aligned} \lambda \phi'(x_{ij}) Vv &= \frac{\lambda}{M} \sum_{p=1}^N \sum_{q=1}^{n_p} \alpha_{pq} \phi'(x_{ij}) \sum_{l=1}^N \sum_{k=1}^{n_l} \phi(x_{lk}) [\phi'(x_{lk}) \phi(x_{pq})] \\ &= \frac{\lambda}{M} \sum_{p=1}^N \sum_{q=1}^{n_p} \alpha_{pq} \sum_{l=1}^N \sum_{k=1}^{n_l} [\phi'(x_{ij}) \phi(x_{lk})] [\phi'(x_{lk}) \phi(x_{pq})] \end{aligned}$$

[0047] Unter Verwendung dieser Formeln f  r alle Cluster i und f  r alle Elemente j erhalten

$$\lambda (\phi'(x_{11}), \dots, \phi'(x_{1n_1}), \dots, \phi'(x_{ij}), \dots, \phi'(x_{N1}), \dots, \phi'(x_{Nn_N})) Vv = \frac{\lambda}{M} K K \alpha$$

wobei

$$\alpha = (\alpha_{pq})_{p=1 \dots N} \\ q = 1 \dots n_p$$

$$= (\alpha_p)_{p=1 \dots N}$$

wobei

$$\alpha_p = (\alpha_{pq})_{q=1 \dots n_p}$$

[0048] Das Verwenden der Gleichungen (4), (5) und (11) für die rechte Seite von (14) ergibt:

$$Bv = \frac{1}{M} \sum_{p=1}^N \sum_{q=1}^{n_p} \alpha_{pq} \phi(x_{pq}) \sum_{l=1}^N n_l \left[\frac{1}{n_l} \sum_{k=1}^{n_l} \phi(x_{lk}) \right] \left[\frac{1}{n_l} \sum_{k=1}^{n_l} \phi(x_{lk}) \right]'$$

$$= \frac{1}{M} \sum_{p=1}^N \sum_{q=1}^{n_p} \alpha_{pq} \sum_{l=1}^N \left[\sum_{k=1}^{n_l} \phi(x_{lk}) \right] \left[\frac{1}{n_l} \right] \left[\sum_{k=1}^{n_l} \phi'(x_{lk}) \phi(x_{pq}) \right]$$

mit $\phi'(x_{ij}) Bv = \frac{1}{M} \sum_{p=1}^N \sum_{q=1}^{n_p} \alpha_{pq} \sum_{l=1}^N \left[\sum_{k=1}^{n_l} \phi'(x_{ij}) \phi(x_{lk}) \right] \left[\frac{1}{n_l} \right] \left[\sum_{k=1}^{n_l} \phi'(x_{lk}) \phi(x_{pq}) \right]$

[0049] Für alle Cluster i und alle Elemente j erhalten wir:

$$(\phi'(x_{i1}), \dots, \phi'(x_{in_i}), \dots, \phi'(x_{ij}), \dots, \phi'(x_{N1}), \dots, \phi'(x_{Nn_s})) Bv = \frac{1}{M} KWK\alpha \quad (14)$$

[0050] Zusammensetzen von (13) und (14) ergibt:

$$\lambda K\alpha = KWK\alpha.$$

[0051] Damit ist

$$\lambda = \frac{\alpha' KWK\alpha}{\alpha' K\alpha} \quad (15)$$

[0052] K kann zerlegt werden in $K = QR$ (Wilkinson, 1971), so daß $K\alpha = QR\alpha$ ist.

[0053] R ist oben dreieckig und Q orthonormal, das heißt $Q'Q = I$.

[0054] Q ist eine $M \times r$ -Matrix und R eine $r \times M$ -Matrix, wobei r der Rang von K ist. Es ist bekannt, daß es für eine allgemeine rechteckige Matrix immer eine QR-Zerlegung gibt.

Es sei $R\alpha = \beta$ (16).

[0055] Da die Zeilen von R linear unabhängig sind, gibt für ein bestimmtes β wenigstens eine Lösung α .

[0056] Es ist daher $K\alpha = Q\beta$ und $\alpha'K = \beta'Q'$ (K ist symmetrisch).

[0057] Einsetzen in (15) ergibt

$$\lambda = \frac{\alpha' KWK\alpha}{\alpha' K\alpha} \quad (17)$$

Q ist orthonormal, so daß $\lambda\beta = Q'WQ\beta$ (18).

[0058] Die Gleichung (18) hat die Form einer Standard-Eigenvektorgleichung. Da K singular ist, erlaubt die QR-Zerlegung die Arbeit in einem Unterraum $Q\beta$, was die Lösung erleichtert.

[0059] Die Koeffizienten α können dann mittels der Gleichung (16) aus β abgeleitet werden und anschließend die Eigenvektoren aus der Gleichung (11).

[0060] Die Koeffizienten α werden dadurch normalisiert, daß gefordert wird, daß die entsprechenden Vektoren v in F normalisiert werden. Das heißt:

$v'v = 1$ (19)

oder (aus der Gleichung 11):

$$\begin{aligned}
 V'V &= \sum_{p=1}^N \sum_{q=1}^{n_p} \sum_{l=1}^N \sum_{h=1}^{n_l} \alpha_{pq} \alpha_{lh} \phi'(x_{pq}) \phi(x_{lh}) = 1 \\
 &= \sum_{p=1}^N \sum_{l=1}^N \alpha'_p K_{pl} \alpha_l = 1 \\
 &= \alpha' K \alpha
 \end{aligned}$$

so daß (19) $\Rightarrow \alpha' K \alpha = 1$ (20)

[0061] Die obigen Schritte geben an, wie die Eigenvektoren v der Gleichung (10) gefunden werden können.

[0062] Aus der linearen Diskriminantenanalyse (siehe zum Beispiel Fukunaga) ist bekannt, daß die Anzahl der Eigenvektoren gleich $N-1$ ist, wenn N die Anzahl der Cluster ist. Die Abbildung der Cluster in den von den Eigenvektoren aufgespannten Unterraum wird durch Projektion auf die Eigenvektoren gefunden. Dies erfolgt mit der folgenden Gleichung: Für einen Eigenvektor v und einen Merkmalsvektor x ist

$$\begin{aligned}
 (\phi'(x)v) &= \sum_{p=1}^N \sum_{q=1}^{n_p} \alpha_{pq} \phi'(x_{pq}) \phi(x) \\
 &= \sum_{p=1}^N \sum_{q=1}^{n_p} \alpha_{pq} k(x_{pq}, x)
 \end{aligned} \tag{21}$$

[0063] Daraus ist zu ersehen, daß die Berechnung nicht voraussetzt, daß ϕ bekannt ist oder daß ein Skalarprodukt in F zu berechnen ist.

[0064] In Experimenten wurde gezeigt, daß durch die Verwendung einer geeigneten Kernfunktion die Abbildungen der Cluster im Eigenvektor-Unterraum gut getrennt werden und daß sie insbesondere linear separabel sind, das heißt daß sie durch Linien, Ebenen oder Hyperebenen getrennt werden können.

[0065] Zum Klassifizieren vermessener Artikel kann dann mit bekannten Techniken wie Untersuchen, Mitteln, dem Mahalanobis-Abstand und dem Vergleich mit den k nächsten Nachbarn leicht eine geeignete Trennfunktion abgeleitet werden.

[0066] Wie erwähnt wurde angenommen, daß die Beobachtungen in F zentriert sind. Das Zentrieren wird nun genauer beschrieben. Zuerst wird für eine gegebene Beobachtung x_{ij} : Element j des Clusters i die Abbildung $\phi(x_{ij})$ zentriert gemäß

$$\bar{\phi}(x_{ij}) = \phi(x_{ij}) - \frac{1}{M} \sum_{l=1}^N \sum_{k=1}^{n_l} \phi(x_{lk}) \tag{22}$$

[0067] Wir müssen dann die Kovarianzmatrix K mit zentrierten Punkten definieren:

$$(\bar{k}_{ij})_{pq} = (\bar{\phi}(x_{pi}) \cdot \bar{\phi}(x_{qj}))$$

für einen gegebenen Cluster p und q .

$$(\tilde{k}_{ij})_{pq} = \left[\phi(x_{pi}) - \frac{1}{M} \sum_{l=1}^N \sum_{k=1}^{n_l} \phi(x_{lk}) \right] \left[\phi(x_{qj}) - \frac{1}{M} \sum_{h=1}^N \sum_{m=1}^{n_m} \phi(x_{hm}) \right]$$

$$(\tilde{k}_{ij})_{pq} = (k_{ij})_{pq} - \frac{1}{M} \sum_{l=1}^N \sum_{k=1}^{n_l} (1_{ik})_{pl} (k_{kj})_{lq} - \frac{1}{M} \sum_{h=1}^N \sum_{m=1}^{n_m} (k_{im})_{ph} (1_{mj})_{hq} + \frac{1}{M^2} \sum_{l=1}^N \sum_{k=1}^{n_l} \sum_{h=1}^N \sum_{m=1}^{n_m} (1_{ik})_{pl} (k_{km})_{lk} (1_{mj})_{hq}$$

$$\tilde{K}_{pq} = K_{pq} - \frac{1}{M} \sum_{l=1}^N 1_{pl} K_{lq} - \sum_{h=1}^N K_{ph} 1_{hq} + \frac{1}{M^2} \sum_{l=1}^N \sum_{h=1}^N 1_{pl} K_{lh} 1_{hq}$$

$$\tilde{K} = K - \frac{1}{M} 1_N K - \frac{1}{M} K 1_N + \frac{1}{M^2} 1_N K 1_N$$

[0068] Dabei haben wir die folgende Matrix eingeführt:

[0069] Die

$$l_{pl} = (l_{ik})_{i=1, \dots, n_p; k=1, \dots, n_l}, (n_p \times n_l)$$

-Matrix, deren Elemente alle gleich 1 sind.

[0070] Die

$$l_N = (l_{pl})_{p=1, \dots, N; l=1, \dots, N}, (M \times M)$$

-Matrix, deren Elemente Blockmatrizen sind.

[0071] Für die nicht zentrierten Punkte $\phi(x_{ij})$ können wir damit K aus \tilde{K} ableiten und dann nach den Eigenvektoren von K auflösen. Für einen Merkmalsvektor x ist die Projektion der zentrierten ϕ -Abbildung von x auf die Eigenvektoren v gegeben durch:

$$(\tilde{\phi}'(x)v) = \sum_{p=1}^N \sum_{q=1}^{n_p} \alpha_{pq} \tilde{\phi}'(x_{pq}) \tilde{\phi}(x)$$

[0072] Die obige Beschreibung gibt allgemein das Verfahren der allgemeinen Diskriminantenanalyse wieder. Dieses allgemeine Prinzip wird nun anhand des speziellen Beispiels des Münzprüfers erläutert.

[0073] Bei dem Beispiel des Münzprüfers zu Beginn der Beschreibung weist der Merkmalsvektor drei Elemente auf, und es gibt drei Cluster, die jeweils einem der beiden interessierenden Nennwerte und der bekannten Fälschung entsprechen.

[0074] In das Meßsystem 1 werden 50 Proben für jeden Nennwert und 50 Proben der Fälschung eingegeben. Wie erwähnt vermessen die Sensorsysteme die Proben, um für jeden Fall Werte für die Dicke, das Material und den Durchmesser zu erhalten. Aus den an jeder Probe gemessenen Merkmalen werden die entsprechenden Merkmalsvektoren gebildet.

[0075] Aus den 50 Proben-Merkmalsvektoren für jeden Cluster werden für die Verwendung bei der Erzeugung der Trennfunktion 37 zufällig ausgewählt.

[0076] Dann wird eine Kernfunktion gewählt. Die Kernfunktion wird auf der Basis von Versuch und Irrtum danach ausgewählt, welche Funktion die besten Trennergebnisse ergibt. Es gibt eine große Zahl von Kernfunktionen, die Mercer's Theorem erfüllen und die geeignet sein können. Beispiele für die Kernfunktion sind der Polynomkern:

$$k(x, y) = (x \cdot y)^d;$$

der Gauß-Kern:

$$k(x,y) = \exp \frac{(-\|x-y\|^2)}{\sigma^2};$$

der Tangens-Hyperbolikus-Kern:

$$k(x, y) = \tanh((x \cdot y) + \theta);$$

und der Sigmoid-Kern:

$$k(x,y) = \left(\frac{1}{1 + e^{-(x \cdot y) + \theta}} \right).$$

[0077] Bei dem vorliegenden Beispiel wird der Gauß-Kern verwendet mit $\sigma^2 = 0,01$.

[0078] Mit den ausgewählten Proben und der Kernfunktion werden die Matrizen K und W berechnet (Gleichungen (8) und (9)).

[0079] Mit der QR-Zerlegung wird dann K zerlegt.

[0080] Dann werden die Eigenvektoren β und die entsprechenden Eigenvektoren berechnet (Gleichung (18)).

[0081] Anschließend werden die Koeffizienten α berechnet und normalisiert (Gleichungen (16) und (20)).

[0082] Danach werden die Merkmalsvektoren der übrigen 13 Proben für jedes Cluster auf die Eigenvektoren v projiziert (Gleichung 21) und die Ergebnisse zur leichten Untersuchung in eine Graphik eingetragen. Bei dem vorliegenden Beispiel gibt es drei Cluster und damit zwei Eigenvektoren, und die Trennung erfolgt im 2d-Raum. Dies ist in der [Fig. 3](#) gezeigt. Wie zu sehen ist, sind die Cluster gut getrennt. Das heißt, daß jedes Cluster auf einen Punkt projiziert ist, der der Schwerpunkt ist. Die Trennung der Projektion der Cluster mit den Eigenvektoren wird dann analysiert und zur Ableitung einer Trennfunktion verwendet. Bei dem vorliegenden Beispiel läßt sich damit leicht eine lineare Trennfunktion ableiten. Zum Beispiel ist eine geeignete Trennfunktion

für die Eigenvektoren v_1, v_2
und den Eingangsvektor x

[0083] Wenn $[(\phi'(x)v_1) > 0 \text{ und } (\phi'(x)v_2 > 0)]$, dann gehört x zur Gruppe 1 (das heißt ist vom ersten Nennwert);
wenn $[(\phi'(x)v_1 > 0 \text{ und } (\phi'(x)v_2) < 0)]$, dann gehört x zur Gruppe 2 (das heißt ist vom zweiten Nennwert); und
wenn $[(\phi'(x)v_1) < 0 \text{ und } (\phi'(x)v_2) > 0]$, dann gehört x zur Gruppe 3 (das heißt ist eine Fälschung des ersten Nennwerts).

[0084] Die Klassifikation von Münzen mit einem unbekannten Nennwert wird dann wie folgt ausgeführt. Die eingeworfene Münze wird erfaßt, und es erfolgen Messungen hinsichtlich des Materials, der Dicke und des Durchmessers wie bei den Proben. Aus den Meßwerten wird der Merkmalsvektor abgeleitet. Der Merkmalsvektor wird dann (mit der Gleichung 21) auf die berechneten Eigenvektoren projiziert und die Münze wie oben beschrieben gemäß den Projektionswerten und der Trennfunktion klassifiziert.

[0085] Die Analyse der Probenwerte für die anfängliche Datenanalyse und die Ableitung der Trennfunktion kann zum Beispiel mit einem Mikroprozessor erfolgen. Auch kann der Klassifizierer **6** ein Mikroprozessor sein.

[0086] Alternativ kann der Klassifizierer **6** ein neuronales Netzwerk sein, etwa ein neuronales Wahrscheinlichkeitsnetzwerk oder ein Perceptron. Zum Beispiel kann das neuronale Netzwerk N-1 lineare Ausgangsneuronen und M verborgene Neuronen umfassen, wobei jede Kernberechnung ein verborgenes Neuron ist. Die Eingangsgewichtungen sind dann die Werte x_{pq} , und die Koeffizienten α sind die Gewichtungen zwischen den verborgenen Neuronen und der Ausgangsschicht.

[0087] Der Klassifizierer kann auch ein linearer Klassifizierer oder eine Support-Vektormaschine sein.

[0088] Die oben beschriebenen Verfahren der Ausführungsform sind gleichermaßen anwendbar auf eine Banknote oder auf die Klassifikation von anderen Arten von Währungsgegenständen. Es sind auch andere Me-

thoden zum Lösen von (10), zum Beispiel durch Zerlegen von K mit der Eigenvektorzerlegung, möglich.

[0089] Bei der vorliegenden Ausführungsform wird eine nichtlineare Abbildung auf einen höherdimensionalen Raum verwendet. Die Abbildung kann auch auf einen Raum niedrigerer Dimension erfolgen oder auf einen Raum mit der gleichen Dimension wie der Merkmalsvektorraum.

Patentansprüche

1. Verfahren zum Herleiten einer Klassifikationsfunktion, um Währungsgegenstände (3) in mindestens zwei Klassen zu klassifizieren, wobei bekannte Probestücke jeder Klasse gemessen und aus den gemessenen Probestücken Merkmalsvektoren hergeleitet werden, eine Abbildungsfunktion ausgewählt wird, die einer nicht-linearen Abbildung des Merkmalsvektorraums in einen zweiten Raum entspricht, Merkmalsvektoren auf Abbildungsvektoren abgebildet werden, wobei die Abbildungsfunktion ein Skalarprodukt im zweiten Raum als Funktion zweier Elemente des Merkmalsvektorraums ausdrückt und N-1 Achsen repräsentierende Koeffizienten im zweiten Raum herleitet, wobei N die Anzahl der Klassen darstellt, wobei die Koeffizienten durch Optimierung der Trennung der Projektionen der Gruppen von Abbildungsvektoren für jede Klasse auf die Achsen hergeleitet werden, Werte erhalten werden, die die Projektionen der Abbildungsvektoren für die gemessenen Probestücke auf die Achsen repräsentieren, und diese Werte verwendet werden, um eine Trennfunktion zum Trennen der Klassen entsprechend einer Trennfunktion in dem (N-1)-dimensionalen Raum herzuleiten.

2. Verfahren nach Anspruch 1, wobei der zweite Raum höherdimensional ist als der erste Raum.

3. Verfahren nach Anspruch 1 oder 2, wobei eine Matrix V, die die Kovarianz-Matrix im zweiten Raum ist, und eine Matrix B, die die Kovarianz-Matrix der Klassenzentren im zweiten Raum ist, und die Lösungen der Gleichung $\lambda Vv = Bv$ und die Koeffizienten aus den Lösungen v hergeleitet werden.

4. Verfahren nach einem der vorstehenden Ansprüche, wobei die Abbildungsfunktion $k(x,y)$ ist und $k(x,y) = (x \cdot y)^d$.

5. Verfahren nach einem der vorstehenden Ansprüche, wobei die Abbildungsfunktion $k(x,y)$ ist und

$$k(x,y) = \exp \frac{(-\|x - y\|^2)}{\sigma^2}.$$

6. Verfahren nach einem der vorstehenden Ansprüche, wobei die Abbildungsfunktion $k(x,y)$ ist und $k(x,y) = \tanh((x \cdot y) + \theta)$.

7. Verfahren nach einem der vorstehenden Ansprüche, wobei die Abbildungsfunktion $k(x,y)$ ist und

$$k(x,y) = \left(\frac{1}{1 + e^{-(x \cdot y) + \theta}} \right).$$

8. Verfahren zum Klassifizieren eines Währungsgegenstandes (3), wobei Merkmale des Gegenstandes gemessen werden, ein Merkmalsvektor aus den gemessenen Werten erzeugt wird und der Gegenstand (3) unter Verwendung einer Klassifikationsfunktion klassifiziert wird, die durch ein Verfahren nach einem der Ansprüche 1 bis 7 hergeleitet wird.

9. Vorrichtung zum Klassifizieren von Währungsgegenständen, mit einer Meßeinrichtung (1) zum Messen von Merkmalen eines Währungsgegenstandes (3), einer Einrichtung (4) zum Erzeugen eines Merkmalsvektors aus den gemessenen Werten, und mit einer Klassifikationseinrichtung (6) zum Klassifizieren des Gegenstandes unter Verwendung von Daten, die eine nichtlineare Abbildungsfunktion entsprechend einer Abbildung des Merkmalsvektorraums in einen zweiten Raum repräsentieren, die Merkmalsvektoren in Abbildungsvektoren abbildet, wobei die Abbildungsfunktion ein Skalarprodukt im zweiten Raum als eine Funktion von zwei Elementen des Merkmalsvektorraums ausdrückt, und die Klassifikationseinrichtung N-1 Achsen repräsentierende Koeffizienten verwendet, wobei N die Anzahl von Klassen darstellt, die durch die Vorrichtung klassifiziert werden können, und mit einer Trennfunktion in dem (N-1)-dimensionalen Raum entsprechenden Funktion zum Trennen der Projektionen der Abbildungsvektoren in diesem (N-1)-dimensionalen Raum.

10. Vorrichtung nach Anspruch 9, wobei die Klassifikationseinrichtung (6) eine Einrichtung zum Herleiten von Werten umfaßt, die die Projektion der Abbildung des Merkmalsvektors des gemessenen Gegenstandes auf die oder jede Achse darstellen.

11. Vorrichtung nach Anspruch 9 oder 10, wobei die Klassifikationseinrichtung **(6)** ein neuronales Netz umfaßt.
12. Vorrichtung nach einem der Ansprüche 9 bis 11, mit einem Münzeinwurf **(2)**, wobei die Meßeinrichtung eine Sensoreinrichtung zum Abtasten einer Münze umfaßt.
13. Vorrichtung nach Anspruch 12, wobei die Sensoreinrichtung zum Abtasten des Materials und/oder der Dicke und/oder des Durchmessers einer Münze ausgelegt ist.
14. Vorrichtung nach einem der Ansprüche 9 bis 11, mit einer Banknoteneingabe **(2)**, wobei die Meßeinrichtung eine Sensoreinrichtung zum Abtasten einer Banknote umfaßt.
15. Vorrichtung nach Anspruch 15, wobei die Sensoreinrichtung zum Erfassen der Intensität von Licht ausgelegt ist, das von der Banknote reflektiert und/oder transmittiert wird.
16. Münzprüfer mit einer Vorrichtung nach einem der Ansprüche 9 bis 13.
17. Banknotenprüfer mit einer Vorrichtung nach einem der Ansprüche 9 bis 11, 14 oder 15.

Es folgen 2 Blatt Zeichnungen

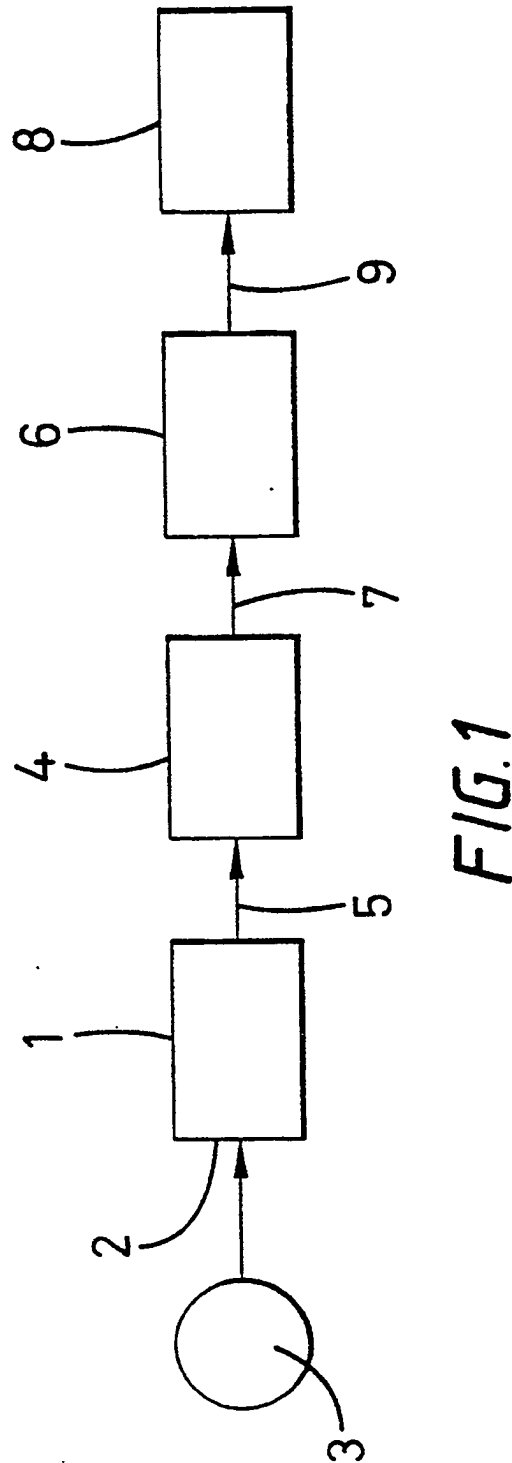


FIG. 2

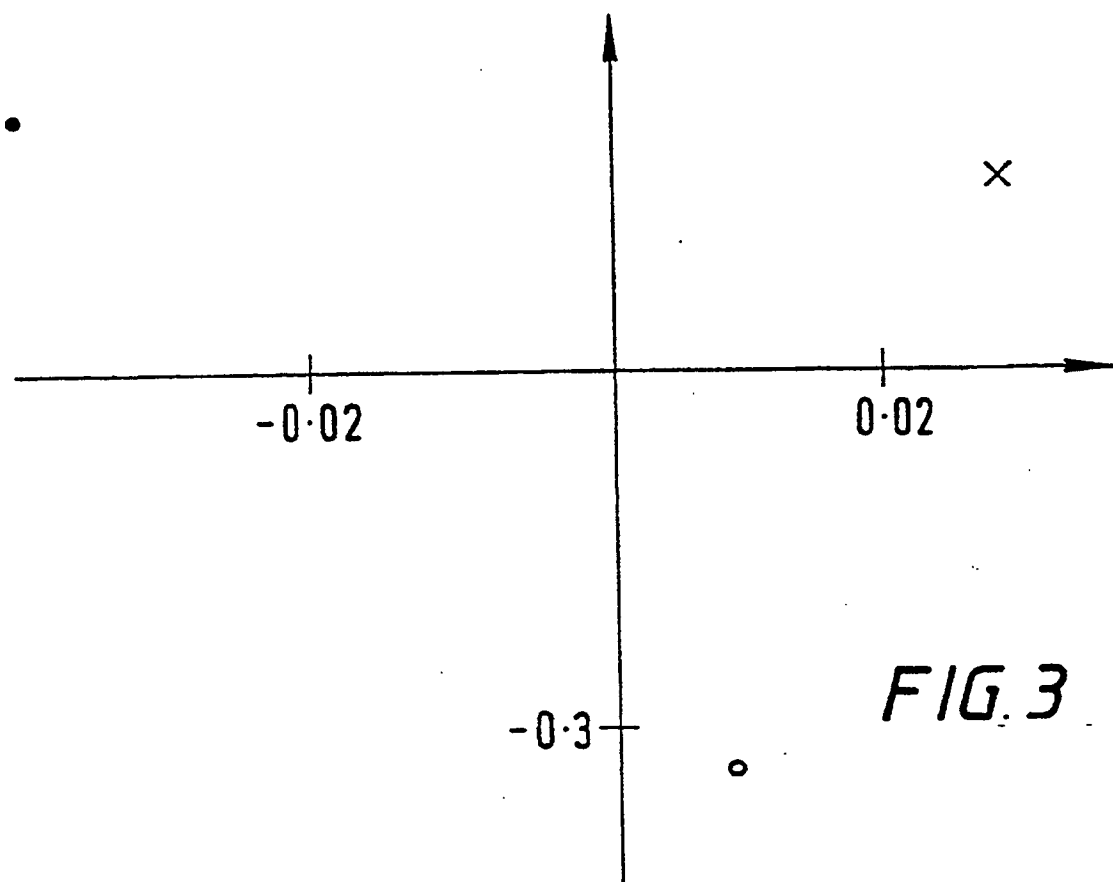
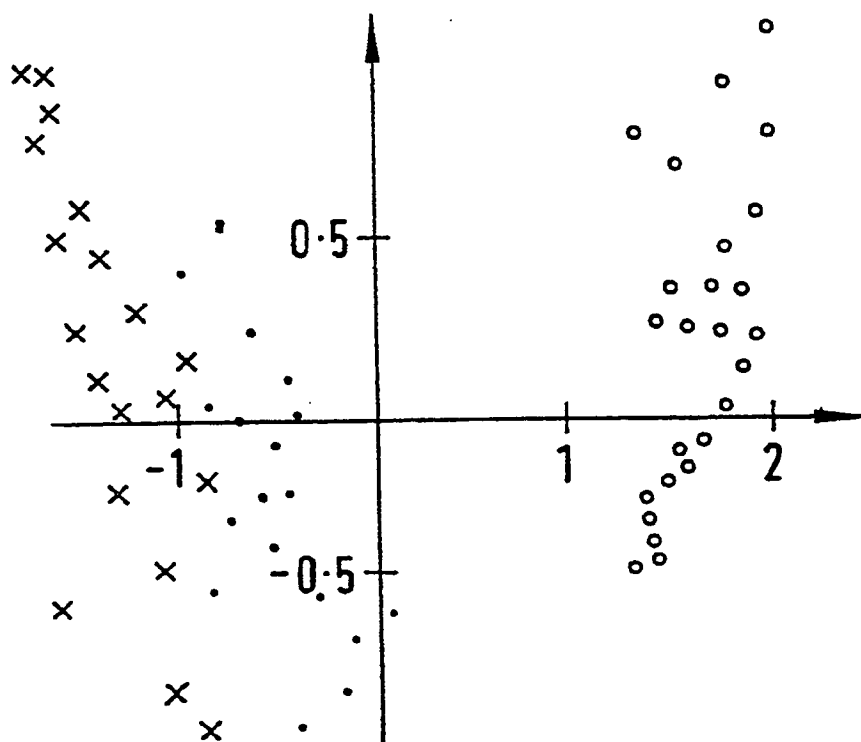


FIG. 3