

Arbeitsgruppe III/2

Helmut Schumacher
Klein & Stekl GmbH
7000 Stuttgart

EIN ALGORITHMUS ZUR REDUKTION DER RECHENZEIT BEI DER STANDARD-MAXIMUM-LIKELIHOOD-KLASSIFIZIERUNG

Zusammenfassung

Es wird ein Verfahren vorgestellt, nach dem die Anzahl der notwendigen Berechnungen der Mahalanobis-Abstände bei der Standard-Maximum-Likelihood-Klassifizierung erheblich reduziert wird. Für jede Klasse werden im Merkmalsraum Hyperellipsoide bestimmt, deren innere Punkte vollständig in diese Klasse fallen.

Das Verfahren arbeitet unabhängig von Table-Lookup-Methoden und eignet sich insbesondere auch für die Installation auf einem Minicomputer.

1. Einleitung

In der Bildverarbeitung gilt die Maximum-Likelihood-Methode als ein Klassifizierungsverfahren, das akzeptable Ergebnisse ermöglicht, aber einen relativ hohen Rechenaufwand erfordert. Soll eine n -kanalige Szene klassifiziert werden, so müssen für alle Bildelemente \underline{x} die Werte der Trennungsfunktionen

$$(1.1) \quad d_i(\underline{x}) = \ln(p(\omega_i)) - \frac{1}{2} \ln(\det(C_i)) - \frac{1}{2} (\underline{x} - \underline{m}_i)^T C_i^{-1} (\underline{x} - \underline{m}_i) \quad (i=1, \dots, m)$$

für alle Musterklassen ω_i berechnet werden. Dabei ist das Klassenzentrum \underline{m}_i der Mittelwertsvektor und C_i die Kovarianzmatrix der Musterklasse ω_i . $p(\omega_i)$ ist die a-priori-Wahrscheinlichkeit für die Zuweisung eines Bildelements zur Musterklasse ω_i . \underline{x} wird der Musterklasse ω_j zugeordnet,

falls

$$d_j(\underline{x}) \geq d_i(\underline{x}) \quad (i=1, \dots, m)$$

und

$$d_j(\underline{x}) \geq s_j,$$

wobei s_j einen Schwellwert darstellt, der die Musterklasse ω_j gegen die Zurückweisungsklasse abgrenzt.

Da die Berechnung der Trennungsfunktionen relativ aufwendig ist, haben wir nach einem Verfahren gesucht, das exakt dieselben Ergebnisse wie das Standard Maximum-Likelihood-Verfahren ergibt, das aber in der Regel sehr viel weniger Berechnungen der Trennungsfunktionen verlangt. Die Grundidee eines solchen Verfahrens wird in Kapitel 2 vorgestellt. In den Kapiteln 3 und 4 wird das vorgeschlagene modifizierte Verfahren ausführlich diskutiert.

2. Kerne einer Musterklasse

Die Trennungsfunktion einer Musterklasse ω_i kann geometrisch so veranschaulicht werden: Alle Bildelemente \underline{x} mit konstantem $d_i(\underline{x})$ liegen auf einem Hyperellipsoid, dessen Mittelpunkt das Klassenzentrum \underline{m}_i ist. Jedem Funktionswert

$$d_i < \ln(p(\omega_i)) - \frac{1}{2} \ln(\det(C_i))$$

ist eindeutig ein Hyperellipsoid aus einer Schar ähnlicher Hyperellipsoide mit Mittelpunkt \underline{m}_i zugeordnet.

Betrachten wir den einfachen Fall mit nur zwei Musterklassen ω_1 und ω_2 ohne Zurückweisungsklasse. Wir suchen Bereiche H_1 und H_2 , in denen mit nur einer Trennfunktionswertberechnung die Entscheidung $\underline{x} \in \omega_1$ bzw. $\underline{x} \in \omega_2$ möglich ist. Dazu sind verschiedene Ansätze möglich. Wir wollen uns auf Bereiche beschränken, die durch Hyperellipsoide in den Scharen der beiden Musterklassen festgelegt sind.

Für alle $\underline{x} \in H_1$ muß gelten:

$$d_1(\underline{x}) \geq d_2(\underline{x});$$

(im Falle von Gleichheit ist die Zuordnung zu ω_1 oder ω_2 willkürlich).

Analog muß für alle $\underline{x} \in H_2$ gelten:

$$d_2(\underline{x}) \geq d_1(\underline{x}).$$

Daraus folgt unmittelbar $H_1 \cap H_2 = \emptyset$ bis auf Bildelemente \underline{x} mit $d_1(\underline{x}) = d_2(\underline{x})$. Geometrisch gesprochen stellen H_1 und H_2 ein Paar von Hyperellipsoiden dar, die sich berühren und denen derselbe Trennfunktionswert für die jeweilige Klasse zugeordnet ist. Alle Bildelemente innerhalb von H_1 werden der Musterklasse ω_1 , alle innerhalb von H_2 ω_2 zugewiesen.

Die Bereiche H_1 bzw. H_2 wollen wir in der Folge als Kerne der Musterklassen bezeichnen. Der Kern einer Musterklasse kann in entarteten Fällen auch leer sein. Hierauf werden wir in Kapitel 3 näher eingehen.

Treten mehr als zwei Musterklassen auf, so können wir zu jedem Paar von Musterklassen (ω_i, ω_j) ein Paar von Kernen, die jeweils auf die andere Musterklasse bezogen sind, bestimmen. Diese Kerne, die sich auf ein Paar von Musterklassen beziehen, wollen wir in der Folge mit K_{ij} bzw. K_{ji} bezeichnen, wobei der erste Index die Musterklasse angibt, deren Mittelpunkt mit dem Kernmittelpunkt übereinstimmt. Für den Bereich H_i , in dem eine Klassenzuteilung $\underline{x} \in \omega_i$ vorgenommen werden kann, sobald $\underline{x} \in H_i$ feststeht, gilt:

$$H_i = \bigcap_{j=1, m (j \neq i)} K_{ij}$$

Die Kerne K_{ij} und damit auch H_i sind durch Angabe von Trennungsfunktionswerten k_{ij} bzw. h_i festgelegt. Kapitel 3 gibt einen Weg zur Bestimmung der Kerne an. Ein Algorithmus, der sowohl H_i als auch K_{ij} verwendet, wird in Kapitel 4 beschrieben.

3. Bestimmung der Kerne

In diesem Kapitel wollen wir die Kerne K_{ij} und K_{ji} für die Musterklassen ω_i und ω_j bestimmen. Die ij Schar ji der Hyperellipsoide für ω_i sei

$$(3.1) \quad E_i: (\underline{x} - \underline{m}_i)^T D_i (\underline{x} - \underline{m}_i) = b_i \geq 0$$

mit $D_i = C_i^{-1}$.

Die Zuordnung zwischen einem Hyperellipsoid aus dieser Schar und dem Trennfunktionswert ist dann (nach (1.1))

$$(3.2) \quad d_i(\underline{x}) = -\frac{1}{2} b_i \ln(p(\omega_i)) - \frac{1}{2} \ln(\det(C_i)).$$

Zunächst wird geprüft, ob die Bedingungen

$$(3.3) \quad d_i(\underline{m}_i) \geq d_j(\underline{m}_i) \quad \text{und}$$

$$(3.4) \quad d_j(\underline{m}_j) \geq d_i(\underline{m}_j)$$

erfüllt sind. In der Regel wird dies der Fall sein.

Es sei jedoch etwa (3.3) nicht erfüllt, also $d_i(\underline{m}_i) < d_j(\underline{m}_i)$. Dies bedeutet, daß der Mittelpunkt \underline{m}_i der Klasse ω_i der Musterklasse ω_j zugewiesen wird. Es folgt jedoch nicht notwendig, daß

$$(3.5) \quad d_j(\underline{x}) > d_i(\underline{x})$$

für alle \underline{x} gilt. Man könnte nun versuchen, zu entscheiden, ob (3.5) gilt. In diesem Falle könnte dann die Musterklasse ω_i gestrichen werden. Wir gehen diesen Weg nicht, sondern setzen als Kernwert $k_{ij} = d_j(\underline{m}_i)$. Der Kernwert k_{ij} wird auf 0 oder einen positiven Wert gesetzt, womit K_{ij} leer ist.

Wir nehmen in der Folge an, daß (3.3) und (3.4) erfüllt sind.

Zu jedem Hyperellipsoid H_i aus E_i , das nicht \underline{m}_i einschließt, gibt es genau ein H_j aus der Schar E_j , das H_i von außen in einem Punkt \underline{p}_{ij} berührt. Den geometrischen Ort aller dieser Berührungspunkte \underline{p}_{ij} wollen wir in der Folge Berührkurve nennen. Die Berührkurve verbindet \underline{m}_i und \underline{m}_j . In \underline{m}_i ist $d_i(\underline{m}_i) - d_j(\underline{m}_i) > 0$, in \underline{m}_j gilt $d_i(\underline{m}_j) - d_j(\underline{m}_j) < 0$ nach Voraussetzung. Unser Ziel ist es, den Punkt \underline{p} der Berührkurve zu finden, wo $d_i(\underline{p}) - d_j(\underline{p}) = 0$.

Für jeden bekannten Punkt \underline{p} der Berührkurve kann der Wert $d_i(\underline{p}) - d_j(\underline{p})$ bestimmt werden. Die nächste Aufgabe ist daher die Bestimmung der Berührkurve.

3.1 Bestimmung der Berührkurve

\underline{p} sei ein beliebiger Punkt $(\underline{p} + \underline{m}_i, \underline{p} + \underline{m}_j)$. H_i, H_j seien die Hyperellipsoide der beiden Scharen E_i und E_j , auf denen \underline{p} liegt. Die Tangentialhyperebene an H_i in \underline{p} ist

$$(\underline{p} - \underline{m}_i) D_i (\underline{x} - \underline{m}_i)^T = b_i \quad (\geq 0),$$

d.h. $(\underline{p} - \underline{m}_i) D_i$ ist Normalenvektor mit Richtung auf die Tangentialhyperebene zu. Ebenso gilt:

$$(\underline{p} - \underline{m}_j) D_j (\underline{x} - \underline{m}_j)^T = b_j \quad (\geq 0)$$

mit Normalenvektor $(\underline{p} - \underline{m}_j) D_j$.

Im Punkt \underline{p} haben H_i und H_j genau dann eine gemeinsame Tangentialhyperebene, die zwischen H_i und H_j liegt, falls gilt:

$$-\lambda (\underline{p} - \underline{m}_i) D_i = (\underline{p} - \underline{m}_j) D_j \quad (\lambda > 0)$$

$$\text{d.h. } \underline{p}(\lambda) = (\lambda \underline{m}_i D_i + \underline{m}_j D_j) (D_j + \lambda D_i)^{-1}.$$

Damit ist eine Parameterdarstellung für die Berührkurve gewonnen. $\lambda = 0$ ergibt \underline{m}_j , $\lambda \rightarrow \infty$ ergibt \underline{m}_i .

Es ist noch zu zeigen, daß $D_j + \lambda D_i$ regulär ist.

$D_j + \lambda D_i$ ist regulär genau dann, wenn $\text{Det}(D_j + \lambda D_i) \neq 0$ für $\lambda > 0$.

Da D_i als Inverse einer Kovarianzmatrix selbst regulär ist, gibt es eine reguläre Matrix A so, daß

$$A D_i A^T = E.$$

Damit gilt:

$$\text{Det}(D_j + \lambda D_i) \neq 0 \Leftrightarrow \text{Det}(A D_j A^T + \lambda E) \neq 0.$$

$\text{Det}(A D_j A^T - \mu E) = 0$ ist aber die Eigenwertgleichung für die

Matrix $A D_j A^T$, die, wie D_j , einem Hyperellipsoid zugeordnet ist, also nur positive Eigenwerte μ hat. Daraus folgt, daß

$$\text{Det}(A D_j A^T + \lambda E) \neq 0 \quad \text{für } \lambda > 0.$$

3.2 Bestimmung des Berührungspunktes, für den $d_i - d_j = 0$ ist

Anstelle des Parameters λ , der nach oben unbeschränkt ist, führen wir einen Parameter ν ein:

$$\nu = \frac{1}{1 + \lambda}$$

Über eine Intervallschachtelung in ν im Intervall $[0, 1]$ kann damit näherungsweise der Punkt \underline{p} der Berührkurve bestimmt werden, für den $d_i(\underline{p}) - d_j(\underline{p}) = 0$ ist.

Da die Kerne der Musterklassen zu wenige, aber niemals zu viele Bildelemente umfassen dürfen, sollte den Kernkonstanten k_{ij} und k_{ji} nicht eine Näherung, sondern die beste bekannte obere Schranke zugewiesen werden. Ist das ν -Intervall bei Abbruch der Iterationen $[\nu_0, \nu_1]$ und entsprechen diesen Intervallgrenzen die Punkte \underline{p}_0 und \underline{p}_1 der Berührkurve, so sollte $k_{ij} = k_{ji} = \min(d_i(\underline{p}_0), d_j(\underline{p}_1))$ gesetzt werden.

Die Kernkonstanten brauchen nicht übertrieben genau berechnet zu werden, da die Bildelemente \underline{x} durch die Diskretisierung der Grauwerte nur einen Gitterraum aufspannen und die Mächtigkeit der Kerne der Musterklassen im allgemeinen unempfindlich gegenüber geringen Abweichungen der Kernkonstanten ist.

4. Vorgeschlagener Algorithmus

In diesem Kapitel soll beschrieben werden, wie die Kerne der Musterklassen, deren Bestimmung im vorigen Kapitel erläutert wurde, dazu verwendet werden können, um die Zahl der notwendigen Berechnungen von Trennungsfunktionen erheblich zu reduzieren. Wir nehmen an dieser Stelle an, daß folgende Werte bekannt sind:

- alle Konstanten zur Berechnung der Trennungsfunktionen $d_i(\underline{x})$;
- die Schwellwerte s_i ($i=1, \dots, m$) zur Abgrenzung gegenüber der Zurückweisungsklasse ξ ; die zugeordneten Hyperellipsoide seien mit S_i bezeichnet:

$$S_i = \left\{ \underline{x} \mid d_i(\underline{x}) \geq s_i \right\}$$

- die Kernkonstanten h_i mit $H_i = \{ \underline{x} \mid d_i(\underline{x}) \geq h_i \}$ ($i=1, \dots, m$).
- die Kernkonstanten k_{ij} mit

$$K_{ij} = \{ \underline{x} \mid d_i(\underline{x}) \geq k_{ij} \} \quad (i, j=1, \dots, m, i \neq j).$$

Reihenfolge der Trennfunktionswertberechnungen

Im Gegensatz zur Standardmethode brauchen nunmehr in der Regel nicht alle $d_i(\underline{x})$ ($i=1, \dots, m$) für ein Bildelement \underline{x} berechnet werden. Wird eine Musterklasse j gefunden mit $\underline{x} \in H_j$, so kann eine Klassenzuteilung sofort vorgenommen werden, ohne weitere Trennfunktionswerte zu berechnen. Daher muß versucht werden, zunächst die Trennfunktionen für die Klassen zu berechnen, zu denen eine hohe Zuweisungswahrscheinlichkeit besteht. Unter diesem Gesichtspunkt gibt es verschiedene Regeln, nach denen eine geeignete Prüfreihefolge festgelegt werden kann. Die folgenden Regeln (1) und (2) wurden bei unserer Implementierung berücksichtigt. Regel (3) könnte neben andern Regeln eventuell ähnlich erfolgreich sein.

- (R1) Bei Fernerkundungsbildern ist in der Regel die Wahrscheinlichkeit, daß benachbarte Punkte zur selben Musterklasse gehören, hoch. Gehen wir von einer zeilenweisen Klassifizierung einer Szene aus, so empfiehlt es sich, als erstes die Klassenzugehörigkeit zur Klasse des linken Nachbarbildelementes zu prüfen. Für das erste Bildelement in der Zeile oder im Falle, daß das linke Nachbarbildelement der Zurückweisungsklasse zugehörig ist, versagt diese Regel.
- (R2) Die Prüfreihefolge wird nach der Häufigkeit der Zuweisung von Bildelementen zu den Musterklassen festgelegt. Es kann etwa nach jeder Bildzeile die Reihenfolge der Häufigkeit der Klassenzuteilungen entweder über alle bisher klassifizierten Bildelemente oder nur der Bildelemente der letzten Zeile bestimmt werden. Offensichtlich versagt diese Regel für die Bildelemente der ersten Zeile.
- (R3) Die Indices der Musterklassen werden für jedes Bildelement sortiert anhand eines wesentlich einfacher als die Trennfunktionen zu berechnenden Abstandes, z.B. der Euklidischen Norm. Dies erscheint nur gerechtfertigt bei großer Spektralkanalanzahl.

In der Folge sei I eine geordnete Teilmenge einer Permutation der Indices $1, \dots, m$ der Musterklassen. $I = (i_1, i_2, \dots, i_l)$ ($l \leq m$) gebe die Reihenfolge an, in der für die Klassenzuteilung noch relevante Trennfunktionen berechnet werden sollen. Nach dem Testen der Klassenzugehörigkeit zur Musterklasse i_1 wird i_1 aus I gestrichen. Eventuell können weitere Indices i_1 aufgrund der Kerne aus I gestrichen werden. Zu Beginn der Klassifizierung eines Bildelementes enthält I eine Permutation der Indices, die z.B. nach den Regeln (R1) und (R2) erhalten werde.

Testen der Zugehörigkeit zur Musterklasse ω_i

Es sei i der erste Index in I . Es wird $d_i(x)$ berechnet. Damit kann $x \in H_i$ und $x \in S_i$ entschieden werden. Dann wird nach folgender i Entscheidungstabelle vorgegangen:

Fall	(1)	(2)	(3)	(4)
$x \in H_i$	N	N	J	J
$x \in S_i$	N	J	N	J
(A) $x \in \omega_i$				x
(B) $x \in \rho$			x	
(C) beste bisherige Klasse merken		x		
(D) I reduzieren	x	x		

Bemerkung zu (C):

Es wird der Index l abgespeichert, für den gilt:

$$d_l(\underline{x}) \geq d_j(\underline{x}),$$

wobei j alle bisher geprüften Indices durchläuft, für die

$$\underline{x} \in S_j \text{ gilt.}$$

Bemerkung zu (D):

Index i wird aus I gestrichen. Weiter werden alle Indices j gestrichen, für die gilt:

$$\underline{x} \in K_{ij}, \text{ d.h. } d_i(\underline{x}) \geq k_{ij}.$$

\underline{x} ist klassifiziert, sobald Fall (3) oder (4) eintritt oder \bar{I} leer ist. Im letzteren Fall wird \underline{x} der Musterklasse ω_ℓ zugewiesen, falls ein solcher Index ℓ existiert, sonst zu ϑ .

5. Schlußwort

Die beschriebene Methode wurde bei der Deutschen Forschungs- und Versuchsanstalt für Luft- und Raumfahrt in Oberpfaffenhofen als Bestandteil des DIBIAS-Systems ("Digitales Bild-Auswertungs-System") installiert.

Die bei Anwendung dieser Methode zu erreichende Ersparnis an Rechenzeit ist sehr stark abhängig von der jeweiligen Klassifizierungsaufgabe. Die Erfahrung hat jedoch gezeigt, daß man in der Regel davon ausgehen kann, daß sich die erforderliche Rechenzeit um wenigstens die Hälfte reduzieren läßt. Die Zeit für die Bestimmung der Kernkonstanten ist gegenüber der Zeit für die Klassifizierung vernachlässigbar.

An dieser Stelle möchte ich mich besonders bei Herrn Dr. Peter Haberäcker von der DFVLR für die Zusammenarbeit bei der Entwicklung dieser Methode bedanken. Für das Aufzeigen des Weges, die Berührkurve zu bestimmen, sowie für die Beweisidee möchte ich besonderen Dank sagen an Frau Dr. Roswitha Blind vom Mathematischen Institut B der Universität Stuttgart.