РАЗПОЗНАВАЩО ПРАВИЛО ЗА НОРМАЛНО РАЗПРЕДЕЛЕНИ ВЕКТОРИ СЛЕД ВТОРИЧНИ ОРТОГОНАЛНИ ПРЕОБРАЗОВАНИЯ

Димитър Тянев

Abstract: *Въз основа на критичен анализ на използуваните в задачите за разпознаване на образи ортогонални преобразования е предложен един нов подход, целящ оптимизирането на ортогоналните признакови пространства. Подходът се състои в извършване на последователни вторични ортогонални преобразования чрез манипулиране на параметрите на собствените базисни системи до постигане на оптималната гледна точка. Последната е определена чрез понятието вторичен скаларен признак. Във връзка с въведеното представяне на класовете е формулирано разпознаващо правило.*

Key words: Orthogonal feature space; Orthogonal transformations; Pattern recognition problems.

1. Проблеми на традиционния подход

Разпознаването на даден образ **x** (в неговия абстрактен смисъл) е процес, в който се установява степента на неговата прилика спрямо известно множество от p възможни за него класове Ω. Тук се имат предвид класове от нормално разпределени в k-мерното пространство вектори, представени чрез своите закони на разпределение:

(1) $$N(\mathbf{C}^{(\Omega)},\mathbf{m}^{(\Omega)}), \quad \Omega = \overline{1,p} \ ,$$

чиито параметри са ковариационните матрици **C** и векторите на математическото очакване **m.**

Приликата или още съответствието се оценява с помощта на различни формални подходи в смисъла на разнообразни математически критерии. Силата на тези критерии обаче е пряко зависима от информативните способности на описващите образа формални признаци. Ето защо винаги са се търсели преходи в такива признакови пространства, които са оптимални в смисъла на информативността или с други думи това са пространства, в които структурните характеристики на даден клас се изявяват по-добре. Едни от най-често прилаганите преобразования на данните от изходното признаково пространство са ортогоналните. Известно е [1], че в съответствие с изказаното разбиране, оптимално ортогонално пространство е пространството на главните компоненти. Проектирането на образите в това пространство се осъществява с помощта на линеен оператор от вида: $\mathbf{y} = \mathbf{A}^T.\mathbf{x}$, където матрицата $\mathbf{A}^T$ е ортонормираният векторен базис на новото пространство. Това преобразование дава възможност всеки отделен клас да се представя от своя собствена признакова система. Едновременно с получаване на векторния базис се решава и задачата за избор на новите признаци. Практически това решение се постига чрез избор само на част от главните компоненти, с което с формира едно подпространство $\mathbf{A}_0$, наричано пространство на Карунен-Лоев. Тогава за разпознаване на даден образ, проектиран в пространството на Карунен-Лоев, е необходимо да бъде получена оценка за това, с каква точност той се изобразява в пространството на всеки отделен клас. Образът принадлежи на онзи клас, в чието пространство се изобразява най-точно. Формалният математически апарат за получаване на такива оценки е известен [1÷6 и др.].

Информативните свойства на пространството на главните компоненти се изразяват пряко чрез собствените стойности $\lambda_i, i = \overline{1,k}$ на определящата го ковариационна матрица **C**. Тъй като тези стойности са реални числа и всяко от тях е свързано с точно определен базисен вектор (признак), това дава възможност лесно да бъде оценен информативният дял на всеки отделен признак.

Непосредственото построяване на линейния оператор, чрез който данните се изобразяват в новото признаково пространство, изисква да се извърши ортогонална декомпозиция на ковариационната матрица на всеки клас поотделно:

(2) $$\mathbf{C}^{(\Omega)} = \mathbf{A}^{(\Omega)}.\mathbf{D}^{(\Omega)}.\left(\mathbf{A}^{(\Omega)}\right)^{T},\ \Omega = \overline{1,p}$$

където p е броят на известните за образите класове. Ако елементите на матрицата

(3) $$\mathbf{D}^{(\Omega)} = \mathrm{diag}\left(\lambda_i^{(\Omega)}\right), i = \overline{1,k}$$

са подредени в низходящ ред, то ортонормираната матрица $\mathbf{A}^{(\Omega)}(k,k)$ може да бъде разделена на два блока:

(4) $$\mathbf{A}^{(\Omega)}(k,k) = \left[\mathbf{A}_0^{(\Omega)}(k,n) \ \vdots \ \mathbf{A}_1^{(\Omega)}(k,k\text{-}n)\right]$$

в съответствие с оценъчното разбиване на матрицата **D** на две части:

(5) $$\mathbf{D}^{(\Omega)} = \begin{bmatrix} \mathbf{D}1^{(\Omega)}(n,n) & \vdots & \mathbf{0} \\ \cdots & \vdots & \cdots \\ \mathbf{0} & \vdots & \mathbf{D}2^{(\Omega)}(k-n,k-n) \end{bmatrix}$$

при което в матрицата $\mathbf{D}2^{(\Omega)}$ попадат оценените като несъществени (k-n) на брой собствени стойности. Числото $n^{(\Omega)}$ е индивидуално за всеки отделен клас. То изразява размерността на собственото пространство, приета като съществена за дадения клас.

Поставянето на разделителната граница между елементите на матрицата **D** се основава на априорната оценка за допустимата неточност, с която преобразуваните данни ще могат да бъдат използвани. Загубата на точност, която се приема за допустима, зависи от изискванията на обекта и възможностите на инженерното приложение.

Определянето на разделителната граница би могло да бъде улеснено ако елементите на матрицата **D** в първата си част са доминиращи. Така те по естествен начин биха се разделили на значими и незначими. На практика обаче в условията на повечето реални обекти изследвани от нас [10,11,12,13] и др., получаваните собствени стойности са близки помежду си и със слабо затихване, т.е. първичните признакови пространства не притежават силно изразено групиране на признаците, което да спомага по естествен начин разделянето им в смисъла на (5). Подобни ситуации могат да бъдат обяснени единствено с косвения характер на първичните признаци относно физическата природа на реалния обект. В такива случаи изборът кой базисен вектор да бъде пренебрегнат въз основа на относителното сравнение на една собствена стойност с друга е силно затруднен и необективен. Изборът е още по-затруднен когато собствените стойности са кратни, което означава, че собствените вектори на които съответствуват, са равностойни относно критерия за избор. Този случай изисква специално внимание, тъй като води към нееднствено решение за факторизацията (2) [10]. Задачата за намиране на собствено пространство на главните компоненти по естествен начин е свързана със задачата за намаляване на неговия размер, при което тя се решава значително по-лесно, отколкото в изходното признаково пространство, където за оценка на информативната способност на признаците се прилагат разнообразни трудоемки методи [11]. В същото време подборът на признаците според критериите, свързани с вътрешнокласовото и междукласовото разстояние изисква едновременна обработка на образите от всички класове.

Тези особености, както и някои затруднения в изчислителния процес при реализация на разложението (2) подтиква някои изследователи към използуване на други ортогонални преобразования като това на Фурие, на Уолш или на Хаар. Известно е още, че за отделни класове случайни процеси тези преобразования са оптимални в смисъла на преобразованието на Карунен-Лоев. Въпреки това получаването на информативни признаци с помощта на едно ортогонално преобразование в повечето случаи по наше мнение е затруднено или е не особено успешно и полезотворно. Посочените недостатъци при представяне на данните в индивидуалното простаранство на главните компоненти, както и затрудненията свързани с неговото интерпретиране ни навежда на извода, че това пространство все още не е най-подходящото.

Въпреки, че структурните характеристики на класовете изпъкват много по-ясно в пространството на главните компоненти отколкото в изходното признаково пространство ние считаме, че това все още не са пределните възможности на този подход. Основание за това ни мнение е възможността при подходяща интерпретация на дадена ортогонална базисна система и чрез подходящо управление на нейните параметри [7], [8] да бъде формирана друга базисна система, в която структурните характеристики на даден клас да бъдат по-прости и по-ясни. С други думи съществува убеждението, че може да се намери по-добро положение от което да се наблюдава структурата на разпределението. По този начин за всеки клас могат да се формират различни по характер и интерпретация собствени параметрически управляеми базиси. Тук по-долу е изложен един подход за манипулиране на базисните системи.

2. Общ подход към формиране на вторични скаларни признаци

Нека чрез (1) са познати **p** на брой класа чрез техните **k-**мерни разпределения в изходно признаково пространство, което е общо за всички класове. Всяка ковариационна матрица има декомпозицията (2) и може да се използува за построяване на собствен **k-**мерен векторен базис $\mathbf{A}^{(\Omega)}$ на пространството на главните компоненти. Чрез преобразованието:

$$\mathbf{y}_j^{(\Omega)} = \left(\mathbf{A}^{(\Omega)}\right)^T . \mathbf{x}_j^{(\Omega)} \tag{6}$$

всеки j-ти образ от наблюдаваната в изходното признаково пространство статистика се проектира в новото признаково пространство, което е пространство само за дадения клас Ω. По този начин става възможно всеки отделен клас да се представя от двойката

$$\left\{\mathbf{A}^{(\Omega)},\ \mu^{(\Omega)}\right\}, \quad \Omega = \overline{1,p} \ , \tag{7}$$

т.е. от своя векторен базис и вектора на математическото очакване μ**.** Наредената двойка (7) е в състояние да представя дадения клас, тъй като тя, заедно с получените собствените стойности, отразява неговата статистическа структура и е в пълно съответствие със закона (1). За разпределението на даден клас в това пространство може да се каже, че то е ориентирано в направленията на координатните оси и центрирано относно вектора на математическото очакване μ. В този смисъл разпределението се намира в общо положение спрямо наблюдаващото направление, определено от вектор μ. Преобразованието (6) и представянето (7) могат да бъдат осъществени и в пространството на Карунен-Лоев, ако има подходящи за това условия.

В резултат на изложените в пункт 1 мотиви относно възможността да се намери по-добрата гледна точка, тук се формулира следната задача: да се намери такъв вторичен ортогонален базис **B**, в който векторът на математическото очакване ν да бъде колинеарен на неговия първи базисен вектор. Това означава, че главното направление на равновероятните елипсоиди на разпределението в това ново пространство ще бъде ориентирано в направлението на наблюдателя, т.е. в направлението на вектора на математическото очакване ν**.** При това неговата проекция в това пространство ще има само един коефициент различен от нула.

За решаване на поставената задача, формализираното представяне на даден клас чрез (7), се подлага на параметрична преработка на векторния базис чрез прилагане на обобщени матрични ядра от вида:

$$\begin{bmatrix} e^{i.\varphi_1}.\cos\theta & -e^{i.\varphi_2}.\sin\theta \\ e^{-i.\varphi_2}.\sin\theta & e^{-i.\varphi_1}.\cos\theta \end{bmatrix}$$

където θ, φ_1, φ_2 са реалните параметри, които отразяват желаните свойства на синтезирания базис [7].

Както вече посочихме, тук си поставяме за цел да осъществим такава преработка на векторния базис **A** или $\mathbf{A}_0$, при която всеки клас може да се представи от новата двойка

(8) $$\left\{\mathbf{B}^{(\Omega)},\ s_1^{(\Omega)}\right\},\quad \Omega=\overline{1,p}$$

където $s_1^{(\Omega)} \neq 0$ е скалар. Наредената двойка (8) изразява по същество нашето желание да "наблюдаваме" структурата на дадения клас от най-подходящата гледна точка. Тази двойка ще наричаме вторичен скаларен признак.

Решението на така поставената задача се постига чрез използуване на преобразованието на Хаусхолдер. В [9] е показано, че за всеки k-мерен вектор **y** съществуват такива последователности от матрици на плоски ротации

(9) $$\mathbf{T}_{l_1 q_1},\ \mathbf{T}_{l_2 q_2},\ \ldots,\ \mathbf{T}_{l_r q_r},\quad r \le k-1$$

при което някое тяхно произведение **B** превръща вектор **y** във вектор колинеарен на друг известен вектор **z**, имащ същата евклидова норма.

В този смисъл, съобразявайки се с поставената по-горе задача, без загуба на общност можем да считаме, че търсеното от нас решение е $\mathbf{z}=s_1.\mathbf{e}_1$, където s_1 е скалар и $\mathbf{e}_1=[1,0,\ldots,0]^T$. Тогава линейният оператор

(10) $$\mathbf{B}.\mathbf{y}=\mathbf{z}=s_1.\mathbf{e}_1$$

ще наричаме вторично ортогонално преобразование. Тъй като това преобразование е инвариантно спрямо евклидовата норма, то числото s_1 изразява именно тази норма. С други думи

(11) $$s_1=(\mathbf{y}^T.\mathbf{y})^{1/2}\quad ,$$

при което могат да се използват и двете стойности $s_{11}>0$ и $s_{12}<0$.

Алгоритъмът за формиране на матрицата **B** е изложен в [14].

3. Формулиране на разпознаващо правило

Преобразованието (10), което ще се прилага върху наблюдаваните образи **x**, превръщащи се по схемата $(\mathbf{x}\Rightarrow\mathbf{y}\Rightarrow\mathbf{z})$, изисква формулиране на разпознаващо правило. За целта тук е предложена функция за оценка на точността при изобразяване на образите в новото признаково пространство. Логиката на синтеза на тази функция е следната: тъй като всеки отделен клас ще има свое индивидуално преобразование от вида (10), то принадлежащите към даден клас образи ще се изобразяват в него с доминиращ първи елемент. За всички останали вектори съществени ще се оказват останалите елементи. По силата на тази логика (която е и логиката на синтезираните пространства от тип **B**) степента на точност, с която се преобразува вектор $\mathbf{y}^{(\Omega)}$ във вторичното пространство $\mathbf{B}^{(\Psi)}$ на клас Ψ , $(\Omega\neq\Psi)$ ще оценяваме с величината G, определена както следва:

(12) $$G^{(\Omega)}=\sum_{i=2}^{k}\left(z_i^{(\Omega)}\right)^2,\ \Omega=\overline{1,p}$$

Функцията G използуваме да съставим следното разпознаващо правило:

$$\textbf{\textit{ако}}\quad J_{\Omega\Psi}=G^{(\Omega)}(\mathbf{x})-G^{(\Psi)}(\mathbf{x})<0\quad \textbf{\textit{за всяко}}\quad \Omega\neq\Psi,\quad \Omega,\Psi=\overline{1,p}$$

$$\textbf{\textit{то}}\quad \mathbf{x}\in\Omega$$

4. Заключителни бележки

В края на изложението на формалната страна на разгледаната по-горе задача и нейното приложение ще направим малко по-свободен коментар на получените резултати. По същество беше изложено кратко едно обобщение на наши изследвания, свързани с реалното приложение на различни методи от теорията за разпознаване на образи, както и главните изводи до които сме достигнали, гледайки на тях критично. Естествено е всяко обобщение да породи нови идеи, какъвто е случаят с предлагания тук подход за преход към вторични ортогонални собствени базисни системи. Чрез този подход се дава възможност за избор на възможно най-добрите гледни точки спрямо структурата на всеки клас. Въведеното представяне на класовете чрез вторичния скаларен признак (8) обаче не е единственото - то може да бъде допълнено с втори, дори с трети скалар и да получи вида:

$$\left\{\mathbf{B}^{(\Omega)}, s_1, s_2, s_3\right\}, \qquad (13)$$

което означава, че структурата на класа, получена в пространството **B**, може да бъде наблюдавана в едно двумерно или дори в тримерно подпространство без при това да се губи логиката на синтеза му. Скаларите s_1, s_2, s_3 могат да се избират свободно в алгоритъма за формиране на собствения вторичен базис.

Този подход ни освобождава от задачата (5) за избор на главните компоненти. Може даже да се твърди, че той позволява твърде радикалното й решение, като се има предвид представянето (8) или (13).

Литература

[1]. Ватанабе С., *Разложение Карунена-Лоева и факторный анализ. Теория и приложения*. В книге “Автоматический анализ сложных изображений”, Москва, Издательство “Мир”, 1969.

[2]. Tou J.T., Gonzalez R.C., *Pattern recognition principles*, Addison-Wesley pub. company, London, 1974.

[3]. Дуда Р., Харт П., *Распознавание образов и анализ сцен*, Москва, Издательство “Мир", 1976.

[4]. Патрик Э., *Основы теории распознавания образов*, Москва, Издательство ‘Советское радио", 1980.

[5]. Браверман Э. М., Мучник И. Б., *Структурные методы обработки эмпирических данных*, Москва, Издательство “Наука", 1983.

[6]. Фомин Я. А., Тарловский Г. Р., *Статистическая теория распознавания образов*, Москва, Издательство “Радио и связь", 1986.

[7]. Fukunaga K., *Introduction to statistical pattern recognition*, 2-nd ed., Academic press, New York, 1990.

[8]. Дагман Э. Е., Кухарев Г. А., *Быстрые дискретные ортогональные преобразования*, Академия наук СССР, Издательство “Наука", 1983.

[9]. Солодовников А. И., *Синтез полных систем ортонормированных функций имеющих быстрого преобразования*, В книге “Вопросы теории систем автоматического управления”, Ленинградский университет, 1978.

[10]. Horn R.A., Johnson C.R., - *Matrix analysis*, Cembridge university Press, 1986.

[11]. Тянев Д. С., *Оценка на разделящите свойства на признаците при синтез на разпознаващи алгоритми*, Годишник на ТУ-Варна, 1986.

[12]. Tyanev D.S., Dobrev P.D., - *Two chanel 16-inputs A/D controller with DMA for IBM PC/AT*, 8-th International conference "Systems for automation of engineering and research", X.1994, Varna.

[13]. Недев А., Тенекеджиев К., *Техническа диагностика и разпознаване на образи*, Издателство ТУ – Варна, 1994.

[14]. Тянев Д. С., *Вторични ортогонални преобразования в задачите за разпознаване на образи*, ЮНС’96 “115 години ВВМУ “Н.Й. Вапцаров”, Варна, V.1996.

Този материал е създаден през 1997 год.