

ВТОРИЧНИ ОРТОГОНАЛНИ ПРЕОБРАЗОВАНИЯ В ЗАДАЧИТЕ ЗА РАЗПОЗНАВАНЕ НА ОБРАЗИ

Димитър Ст. Тянев

Анотация: Въз основа на критичен анализ на използваните в задачите за разпознаване на образи ортогонални преобразования е предложен един нов подход, целящ оптимизирането на ортогоналните признакови пространства на отделните класове. 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. Представен е алгоритъм за формиране на вторичните векторни базиси.

SECONDARY ORTHOGONAL TRANSFORMATIONS IN THE PROBLEMS OF PATTERN RECOGNITION

Димитър Ст. Тянев

Abstract: On the basis of a critically analysis of the orthogonal transformations used in the problems of pattern recognition, a new method of approach is suggested, which is meant for the optimization of the orthogonal feature space of the separate classes. This approach consists in accomplishment of consecutive secondary orthogonal transformations by manipulating the parameters of the proper basis systems to the optimal point of view. The latter is determined by the use of the term secondary scalar feature. In connection with the introduced presentation of classes, a classification rule is formulated and the algorithm for forming of the secondary vector bases is presented.

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

1. Проблеми на традиционния подход

Разпознаването на даден образ (в неговия абстрактен смисъл) е процес, в който се установява степента на прилика спрямо известно множество от възможни за него класове. 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 извяват 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. Преминването на образите в това пространство се осъществява с помощта на линейния оператор

$$(1) \quad \mathbf{y} = \mathbf{A}_0^T \cdot \mathbf{x} \quad ,$$

наричан често преобразование на Карунен-Лоев, където $\mathbf{A}_0^T$ е ортонормираният векторен базис на новото пространство, $\mathbf{x}$ е k-мерен вектор-образ в изходното признаково пространство. Това

преобразоване дава възможност всеки отделен клас да се представя от своя собствена признакова система. Тогава за разпознаване на даден образ 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 тези оценки е известен [1], [2], [3] и др., а споменатите теоретични качества се потвърждават както от наши, така и от чужди изследвания.

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 (1), чрез който данните се изобразяват в новото признаково пространство, изисква да се извърши ортогонална декомпозиция на ковариационната матрица на всеки клас поотделно:

$$(2) \quad \mathbf{C}^{(\Omega)} = \mathbf{A}^{(\Omega)} \cdot \mathbf{D}^{(\Omega)} \cdot (\mathbf{A}^{(\Omega)})^T, \quad \Omega = \overline{1, p},$$

където p е броят на известните за образите класове.

Ако елементите на матрицата

$$(3) \quad \mathbf{D}^{(\Omega)} = \text{diag}(\lambda_i^{(\Omega)}), \quad i = \overline{1, k}$$

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

$$(4) \quad \mathbf{A}^{(\Omega)}(k, k) = \begin{bmatrix} \mathbf{A}_0^{(\Omega)}(k, n) & : & \mathbf{A}_1^{(\Omega)}(k, k - n) \end{bmatrix}$$

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

$$(5) \quad \mathbf{D}^{(\Omega)} = \begin{bmatrix} \mathbf{D}_0^{(\Omega)}(n, n) & : & \mathbf{0} \\ \dots & : & \dots \\ \mathbf{0} & : & \mathbf{D}_1^{(\Omega)}(k - n, k - n) \end{bmatrix},$$

при което в матрицата $\mathbf{D}_1^{(\Omega)}$ 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 [7], [8], [9] 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, случай който изисква специално внимание, тъй като води към неединствено решение за факторизацията (2) [7].

Тези особености, както и някои затруднения в изчислителния процес при реализация на разложението (2) подтиква някои изследователи към използване на други ортогонални преобразования като това на Фурие, 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 ползотворно.

Посочените недостатъци при представяне на данните в индивидуалното пространство 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 да бъде формирана друга базисна система [4], [5], в която структурните характеристики на даден клас да бъдат по-прости и по-ясни, с което да позволяват получаването на по-рационални решения за отделните задачи в теорията на разпознаване на образи. По този начин за всеки клас могат да се формират 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 са познати p на брой класа чрез техните обучаващи извадки $\mathbf{X}^{(\Omega)}(m^{(\Omega)}, k)$, $\Omega = \overline{1, p}$, всяка от които има обем от $m^{(\Omega)}$ на брой вектор-образа, наблюдавани в k -мерното изходно признаково пространство, което е общо за всички класове.

Всяка обучаваща извадка може да се използва за построяване на собствен k -мерен векторен базис $\mathbf{A}^{(\Omega)}$ на пространството на главните компоненти. Чрез преобразованието на Карунен-Лоев:

$$(6) \quad \mathbf{y}_j^{(\Omega)} = \left(\mathbf{A}^{(\Omega)} \right)^T \cdot \mathbf{x}_j^{(\Omega)}, \quad j = \overline{1, m^{(\Omega)}}$$

всички образи от обучаващата извадка се проектират в новото признаково пространство, което е пространство само за дадения клас. По този начин става възможно всеки отделен клас да се представя от двойката

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

т.е. от своя векторен базис и вектора на неговото координатно начало $\boldsymbol{\mu}$. Наредената двойка (7) е в състояние да представя дадения клас, тъй като тя, заедно с получените собствените стойности, отразява неговата статистическа структура. За да бъдем обаче коректни трябва да посочим, че този признаков модел е най-адекватен в случаите когато разпределението на образите, принадлежащи на даден клас, е нормално. Така при тази постановка може да се приеме, че двойката (7) наистина представя дадения клас, тъй като изходната обучаваща извадка е вече консумирана.

Споменатият вектор на координатното начало може да бъде определен по различни начини като се имат предвид статистическите свойства на обучаващата извадка $\mathbf{Y}^{(\Omega)}$.

Формализираното по този начин представяне на даден клас 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 $\theta, \varphi_1, \varphi_2$ са реалните параметри, които отразяват желаните свойства на синтезирания базис [4].

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

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

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.

В резултат на изложените по-горе съждения относно формализацията на поставената цел, тук се формулира следната задача: да се намери такъв вторичен ортогонален базис $\mathbf{B}$, в който векторът на координатното начало μ да бъде колинеарен на първия базисен вектор, при което той ще има само един коефициент различен от нула.

Решението на така поставената задача може да бъде постигнато чрез използване на преобразованията на Хаусхолдер [6], където е показано, че за всеки k -мерен вектор $\mathbf{y}$ съществуват такива последователности от матрици на плоски ротации

$$(9) \quad \mathbf{T}_{l_1 q_1}, \mathbf{T}_{l_2 q_2}, \dots, \mathbf{T}_{l_N q_N}, \quad N \leq k-1,$$

при което някое тяхно произведение $\mathbf{B}$ превръща вектор $\mathbf{y}$ във вектор колинеарен на друг известен вектор $\mathbf{z}$, 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 $\mathbf{e}_1 = [1, 0, \dots, 0]^T$. Тогава линейният оператор

$$(10) \quad \mathbf{B}^T \cdot \mathbf{y} = \mathbf{z} = s_1 \mathbf{e}_1$$

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

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

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

Остава открит въпросът за формиране на матрицата $\mathbf{B}$, чийто вид в смисъла на (9) не е единствен. Проблемът обаче е чисто алгоритмичен и неговото решение е изложено по-нататък.

3. Формулиране на разпознаващо правило

Преобразованието (10), което ще се прилага върху наблюдаваните образи $\mathbf{x}$, 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 $\mathbf{B}$) 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:

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

Функцията G използваме да съставим следното разпознаващо правило:

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

4. Алгоритъм за получаване на вторичен векторен базис

Тъй като последователността (9) ще превръща в нула всички елементи на вектор $\mathbf{z}$ -(10) с изключение на първия, то ще определим това да става в обратен ред, т.е.

$$i = k, k-1, k-2, \dots, 2; \quad i = \overline{k,2}.$$

Ще използваме транспонирано матрично ядро в следния общ вид:

$$(13) \quad \mathbf{T}_{lq} = \begin{bmatrix} c & -s \\ s & c \end{bmatrix}; \quad l = i-1, \quad q = i,$$

което за да бъде ортогонално и инвариантно към евклидовата норма на въртящия се вектор $\mathbf{y}$, трябва да определя елементите си по следния начин:

$$(14) \quad c_i = \frac{-y_{i-1}}{\beta_i}; \quad s_i = \frac{y_i}{\beta_i}$$

където $\beta_i = \sqrt{y_{i-1}^2 + y_i^2}$

Възможно е изчисленията (14), както и изложените по-долу да се проведат чрез направляващия единичен вектор $\mathbf{e}_y$; $\mathbf{y} = \|\mathbf{y}\| \cdot \mathbf{e}_y$, $\|\mathbf{e}_y\| = 1$.

И така синтезираният алгоритъм представлява циклическо повторение за всяко $i = \overline{k,2}$ на следните изчисления:

1. Изчисление на параметрите β_i, c_i и s_i според (14) от съответните елементи на вектор $\mathbf{z}_i$. В началния момент (за $i=k$) $\mathbf{z}_i = \mathbf{y}$.

2. Формиране на матрицата на плоското въртене $T_{i-1,i}$ както следва:

$$T_{i-1,i} = \begin{bmatrix} 1 & \dots & \dots & \dots & \dots & 0 \\ \vdots & 1 & & & & \vdots \\ \vdots & & c_i & -s_i & & \vdots \\ \vdots & & s_i & c_i & & \vdots \\ \vdots & & & & 1 & \vdots \\ 0 & \dots & \dots & \dots & \dots & 1 \end{bmatrix} \begin{matrix} \rightarrow i-1 \\ \rightarrow i \end{matrix}$$

3. Извършване на въртенето $z_{i-1} = T_{i-1,i} \cdot z_i$, при което i -тият елемент на вектор z_{i-1} получава нулева стойност.

4. Натрупване на матричното произведение: $B_{i-1} = B_i \cdot T_{i-1,i}$. В началния момент (за $i=k$) е в сила равенството $B_i = I$.

5. След завършване на итерациите (при $i=2$) матрицата B_2 съдържа произведението на последователността (9), т.е. $B_2 = B$ е търсеният векторен базис.

5. Заключение 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:

$$(15) \quad \left\{ B^{(\Omega)}, s_1, s_2, s_3 \right\},$$

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 (15).

Литература

- [1]. Ватанабе С., - *Разложение Карунена-Лоева и факторный анализ. Теория и приложения.* - В книге "Автоматический анализ сложных изображений". Москва, Издательство "Мир", 1969.
- [2]. Fukunaga K., - *Introduction to statistical pattern recognition*, 2-nd ed., Academic press, New York, 1990.
- [3]. Tou J.T., Gonzalez R.C., - *Pattern recognition principles*, Addison-Wesley pub. company, London, 1974.
- [4]. Дагман Э.Е., Кухарев Г.А., - *Быстрые дискретные ортогональные преобразования*, Академия наук СССР, Издательство "Наука", 1983.
- [5]. Солодовников А.И., - *Синтез полных систем ортонормированных функций имеющих быстрого преобразования*, В книге "Вопросы теории систем автоматического управления", Ленинградский университет, 1978.
- [6]. Horn R.A., Johnson C.R., - *Matrix analysis*, Cembidge university Press, 1986.
- [7]. Тянев Д.С., - *Алгоритми за разпознаване на образи и използването им в техническата диагностика*, Дисертация, Варна, 1990.
- [8]. Тянев Д.С., - *Оценка на разделящите свойства на признаците при синтез на разпознаващи алгоритми*, Годишник на ТУ Варна, 1986.
- [9]. 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.