

Выбор базиса в задаче беспризнакового распознавания личности по фотопортрету

Ермаков А. С.

yermaikov@uic.tula.ru

Тула, Тульский государственный университет

Основным носителем информации о распознаваемых объектах в беспризнаковом подходе к распознаванию образов служит потенциальная функция $K(\omega', \omega'')$, играющая роль скалярного произведения для любых двух объектов распознавания $\omega', \omega'' \in \Omega$.

Потенциальная функция является симметричной действительной функцией, удовлетворяющей условию Мерсера. Последнее условие выражается в неотрицательной определенности матрицы значений потенциальной функции $M = \{m_{ij} = K(\omega_i, \omega_j)\}$ для всякой конечной совокупности объектов $\{\omega_1, \dots, \omega_N\} \subset \Omega$. При этих условиях можно говорить, что потенциальная функция $K(\omega', \omega'')$ погружает множество распознаваемых объектов $\omega \in \Omega$ в линейное пространство.

Потенциальная функция определяет способ измерения несходства объектов между собой, и именно от её выбора зависит, будет ли выполняться гипотеза компактности для решаемой задачи. Именно это соображение является основным при формировании потенциальных функций для решения многих прикладных задач. В то же время, выполнением условий Мерсера для получаемой функции часто пренебрегают. В этом случае двойственная задача оптимизации, к которой сводится задача обучения при использовании метода потенциальных функций, перестает быть выпуклой, что приводит к трудностям получения решения.

Одним из распространенных способов формирования потенциальной функции является скалярное произведение яркостей $\omega_{\mathbf{t}}$ соответствующих пикселей $\mathbf{t} \in T$ изображения. Однако такой способ не учитывает возможности сдвигов и локальных деформаций изображений распознаваемых объектов, в частности, фотопортретов. Поэтому предварительно растры сравниваемых фотопортретов подвергают деформации таким образом, чтобы максимально совместить изображения фотопортретов. Поскольку для каждой пары изображений деформация раstra v_t выбирается независимо, то, даже несмотря на то, что, вычислительно, яркостная потенциальная функция

$$K^B(\omega', \omega'') = \sum_{\mathbf{t} \in T} \omega'_{\mathbf{t} + \mathbf{v}_t} \cdot \omega''_{\mathbf{t} - \mathbf{v}_t}$$

представляет собой скалярное произведение яркостей, это уже не гарантирует выполнения условия Мерсера.

Однако введение базиса в линейном пространстве, в которое погружает объекты потенциальная функция $K^B(\omega', \omega'')$, позволяет обойти это препятствие [1].

Наряду с рассмотренной яркостной функцией $K^B(\omega', \omega'')$, также будем рассматривать деформационную потенциальную функцию $K^T(\omega', \omega'') = \exp(-\alpha|T(\omega', \omega'')|)$, где $|T(\omega', \omega'')|$ — мера деформации растров пары изображений ω' и ω'' , а также комбинированную потенциальную функцию $K(\omega', \omega'') = \delta K^T(\omega', \omega'') + (1 - \delta)K^B(\omega', \omega'')$ где δ — весовой коэффициент, $0 \leq \delta \leq 1$.

Обычно исследователь имеет дело с некоторым множеством $\Omega^0 = \{\omega_1, \dots, \omega_n\}$ реально существующих объектов, причем не все из них предполагаются классифицированными. Это множество названо в [1] базисной совокупностью. Базисная совокупность играет роль конечного базиса в линейном пространстве Ω . Таким образом, каждому элементу линейного пространства Ω соответствует вектор его скалярных произведений с объектами базисной совокупности $\mathbf{x}(\omega) = (K(\omega, \omega_1), \dots, K(\omega, \omega_n))$. Каждый объект ω_i базисной совокупности формирует свой проекционный признак $x_i(\omega) = K(\omega, \omega_i)$.

Очевидно, при обучении в пространстве проекционных признаков, для потенциальной функции $K(\omega', \omega'')$ уже не требуется выполнение условий Мерсера. Недостатком такого подхода является то, что получаемое при этом решающее правило выражается через все элементы базисной совокупности. В частности, в задаче распознавания личности по фотопортрету роль объектов распознавания играют растровые изображения, и процедура вычисления значения потенциальной функции для пары изображений является вычислительно сложной. Это приводит к большой вычислительной сложности алгоритма на этапе распознавания.

Будем считать, что каждый объект ω_i базисной совокупности формирует свою собственную потенциальную функцию $\tilde{K}_i(\omega', \omega'') = K(\omega', \omega_i)K(\omega_i, \omega'')$. В данной работе подмножество наиболее «информативных» потенциальных функций выбирается из всей базисной совокупности с помощью методологии комбинирования потенциальных функций, предложенной в работе [2], которая позволяет оценить информативность каждой отдельной потенциальной функции в виде веса? с которым она входит в общую функцию. Оценки весов вычисляются итерационным алгоритмом. В результате ненулевыми оказываются лишь небольшое число весов при потенциальных функциях тех объектов, которые и будут составлять искомый базис. Проверка предложенного подхода производилась на базе, состоящей из фотопортретов 295 человек (Таблица 1). Было построено 200 решающих правил, по одному на каждого человека. Результаты верификации личности на тестовой выбор-

295 человек по 8 фотографий			
200 клиентов по 8 фотографий		95 самозванцев по 8 фотографий	
$200 \times 6 = 1200$ для обучения	$200 \times 2 = 400$ для проверки	$25 \times 8 = 200$ для обучения	$70 \times 8 = 560$ для проверки

Таблица 1. Структура базы фотопортретов.

	(1)	(2)
Количество элементов в базисе	1400	11
Яростная потенциальная функция, % ошибок	2.3	2.6
Комбинированная потенциальная функция, % ошибок	1.4	1.7

Таблица 2. Результаты распознавания на тестовой выборке: (1) все элементы обучающей совокупности в базисе; (2) после выбора базиса. Во втором случае количество элементов в базисе для каждого решающего правила разное, в среднем 11.

ке, приведенные в Таблице 2, показывают, что удаётся выбрать базис со значительно меньшей размерностью при небольшом ухудшении качества распознавания, что позволяет сильно уменьшить вычислительную сложность решающих правил.

Работа выполнена при поддержке РФФИ, проекты № 05-01-00679, № 06-01-08042, № 06-01-00412, № 06-07-89249.

Литература

- [1] *Mottl V. V., Dvoenko S. D., Seredin O. S., Kulikowski C. A., Muchnik I. B.* Featureless Pattern Recognition in an Imaginary Hilbert Space and Its Application to Protein Fold Classification. // Second Int. Workshop on Mach. Learn. and Data Mining in Patt. Rec., Leipzig, July 2001. — Pp. 322–336.
- [2] *Моттль В. В., Середин О. С., Красоткина О. В., Мучник И. Б.* Комбинирование потенциальных функций в задачах восстановления зависимостей по эмпирическим данным // Докл. РАН. — 2005. — Т. 401, № 5. — С. 607–612.