УДК 004.93'12

Н.Ю. Губанов, В.Т. Калайда

Применение методов непараметрической статистики в задачах идентифткации изображения

Рассматривается подход к решению задачи идентификации, основанный на построении опорной гиперплоскости путем размножения формализованных характеристик описания изображения методом непараметрической статистики.

Проблема идентификации личности человека является одной из наиболее актуальных задач (известных как «биометрические системы безопасности») в системах распознавания. Интерес к таким системам вызван бурным развитием информационных технологий, которые находят все большее применение в различных сферах человеческой деятельности: банковская деятельность, шоу-бизнес, в оборонных системах обеспечения доступа к системам вооружения и документации, а также все большую значимость приобретают в связи с нарастающей глобальной угрозой терроризма.

Идентификация человека по чертам лица – одно из самых динамично развивающихся направлений в биометрической индустрии. Привлекательность данного метода основана на том, что он наиболее близок к тому, как люди обычно идентифицируют друг друга.

Развитие мультимедийных технологий, благодаря которым можно увидеть все больше видеокамер, установленных на городских улицах и площадях, аэропортах, вокзалах и других местах скопления людей, определило развитие этого направления как наиболее перспективного.

Среди представленных на рынке программных продуктов следует отметить системы, основанные на методах поэлементного сравнения участков изображения [1], статистическом сравнении (факторный и кластерный анализ [2]) и ряде других. Однако представленные системы обладают рядом существенных недостатков, таких как высокая стоимость и достаточно невысокая достоверность.

В работе предлагается метод идентификации, основанный на принятии решения о принадлежности анализируемого изображения к одному из известных классов по оценке меры близости изображения к опорной гиперплоскости класса.

Под опорной гиперплоскостью класса подразумевается плоскость, наименее уклоняющаяся от всего множества численных характеристик каждого изображения из данного класса.

В качестве формализованных численных характеристик описания изображений используются алгебраические моменты:

$$\mu_{\alpha\beta} = \iint\limits_{\infty} f(x,y)\, x^{\alpha} y^{\beta} dx dy, \quad (\alpha,\beta) = \left(\overline{0,n}\right), \tag{1}$$

где $f(x,y)$ – значение интенсивности изображения в точке (x,y) [3].

Параметры опорной гиперплоскости вычисляют из решения оптимизационной задачи

$$L = \sum_{k=1}^{n} \sum_{j=1}^{m} \left(A_k x_k^j + A_0 \right)^2 \to \min,$$

где x_k – ограниченная (длиной m) последовательность моментов $\mu_{\alpha\beta}$ (1) j-го изображения, принятая для описания данного класса; n – количество изображений в классе.

Для выполнения данного условия необходимо, чтобы все производные целевой функции L, $\frac{\partial L}{\partial A_k} = 0$. В результате получается нормальная система однородных уравнений относительно неизвестных $A_1, A_2, \ldots, A_0$:

$$\begin{cases} \dfrac{\partial L}{\partial A_1} = \sum\limits_{k=1}^{n} \sum\limits_{j=1}^{m} \left(A_k x_k^j + A_0 \right) x_1^j = 0; \\ \dfrac{\partial L}{\partial A_2} = \sum\limits_{k=1}^{n} \sum\limits_{j=1}^{m} \left(A_k x_k^j + A_0 \right) x_2^j = 0; \\ \ldots \\ \dfrac{\partial L}{\partial A_0} = \sum\limits_{k=1}^{n} \sum\limits_{j=1}^{m} \left(A_k x_k^j + A_0 \right) = 0. \end{cases}$$

Полученная нормальная система является однородной и, следовательно, имеет нетривиальное решение только при условии, что ранг матрицы меньше ее размерности [4]. Отличительной особенностью данной нормальной системы является большой разброс коэффициентов ее матрицы. Это обусловлено большим динамическим разбросом моментных характеристик ($10^{-7} - 10^{6}$).

Для решения системы линейных уравнений наиболее целесообразно использовать метод, основанный на методе разложения матрицы в виде $\mathbf{B} = \mathbf{USV^T}$ (сингулярное разложение), поскольку последнее обладает высокой точностью и устойчивостью вычислительного процесса [5], где $\mathbf{B}$ – действительная матрица размером $m \times n$, матрица $\mathbf{U}$ сформирована из n ортонормированных собственных векторов, соответствующих n наибольшим собственным значениям матрицы $\mathbf{BB^T}$, а матрица $\mathbf{V}$ – из ортонормированных собственных векторов матрицы $\mathbf{B^TB}$. Диагональные элементы матрицы $\mathbf{S}$ – неотрицательные значения квадратных корней из собственных значений матрицы $\mathbf{BB^T}$. Для оценки принадлежности анализируемого изображения к классу будем использовать следующую величину: $\delta = \dfrac{\sum_{k=1}^{N} A_k x'_k + A_0}{\sqrt{\sum_{k=1}^{N} {A_k}^2}}$ – расстояние до гиперплоскости, где x'_k – формализованные характеристики анализируемого изображения.

Принятие решения о принадлежности к классу изображений осуществляется сравниванием внутриклассовых расстояний и расстояния до анализируемого изображения.

Из-за небольшого количества изображений, участвующих в вычислении коэффициентов опорной гиперплоскости, оценка параметров гиперплоскости ненадежна.

Для достоверной идентификации человека необходимо использовать методы, позволяющие повысить точность оценки. Это достигается путем размножения измерений на основе непараметрической оценки функции плотности распределения моментных характеристик класса изображений и генерации дополнительных характеристик по оцененному распределению:

$$\bar{f}(x) = \frac{1}{n\prod_{v=1}^{m} c_v} \prod_{v=1}^{m} \sum_{i=1}^{n} \Phi\left(\frac{x_v - x_v^i}{c_v}\right),$$

где $\bar{f}(x)$ – совместная функция плотности вероятности; Φ – ядерная функция; m – количество моментных характеристик; n – количество изображений; c_ν – коэффициент размытости ядерной функции, который характеризует область её определения (расплывчатость ядра) [6].

В работе используется ядерная функция Епанечникова, вид которой представлен на рис. 1:

$$\Phi(u) = \begin{cases} \dfrac{3}{4\sqrt{5}} - \dfrac{3\,u^2}{20\sqrt{5}} & \forall \quad |u| < \sqrt{5}; \\ 0 & \forall \quad |u| \geq \sqrt{5}. \end{cases}$$

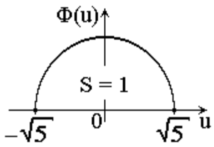

Рис. 1. Вид ядерной функции Епанечникова

Оценка совместной функции плотности распределения проводилась путем вычисления оптимального значения параметра размытости c в функции плотности распределения, обеспечивающего минимальную внутриклассовую дисперсию σ, методом деформируемого многогранника (метод Нелдера–Мида) [7].

Для минимизации функции n переменных $f(x)$ в n-мерном пространстве строится многогранник, содержащий $(n+1)$ вершину. Очевидно, что каждая вершина соответствует некоторому вектору x (в нашем случае вектор параметров размытости c).

Вычисляются значения целевой функции $f(x)$ в каждой из вершин многогранника, определяется максимальное из этих значений и соответствующая ему вершина $x[h]$.

Через эту вершину и центр тяжести остальных вершин проводится проецирующая прямая, на которой находится точка $x[q]$ с меньшим значением целевой функции, чем в вершине $x[h]$ (рис. 2). Затем исключается вершина $x[h]$.

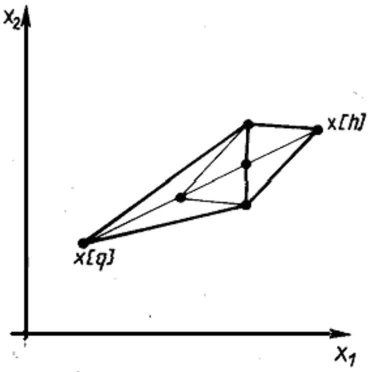

Рис. 2. Геометрическая интерпретация метода деформируемого многогранника

Из оставшихся вершин и точки $x[q]$ строится новый многогранник, с которым повторяется описанная процедура. В процессе выполнения данных операций многогранник изменяет свои размеры, что и обусловило название метода.

Целевой функцией, значение которой вычисляется в каждой из вершин многогранника, является дисперсия

$$\begin{cases} \sigma = +\sqrt{\dfrac{1}{n-1}\sum\limits_{i=1}^{n}(\delta_i - \bar{\delta})^2}, \text{ если } c_i > 0\text{ , } i = \left(\overline{1,n}\right); \\ \sigma = \sigma_{\max}, \text{ если } c_i \le 0, \end{cases}$$

где n – количество изображений; δ_i – отклонение i-го изображения от гиперплоскости класса; $\sigma_{\max}$ – штрафная функция.

Структурная схема алгоритма приведена на рис. 3.

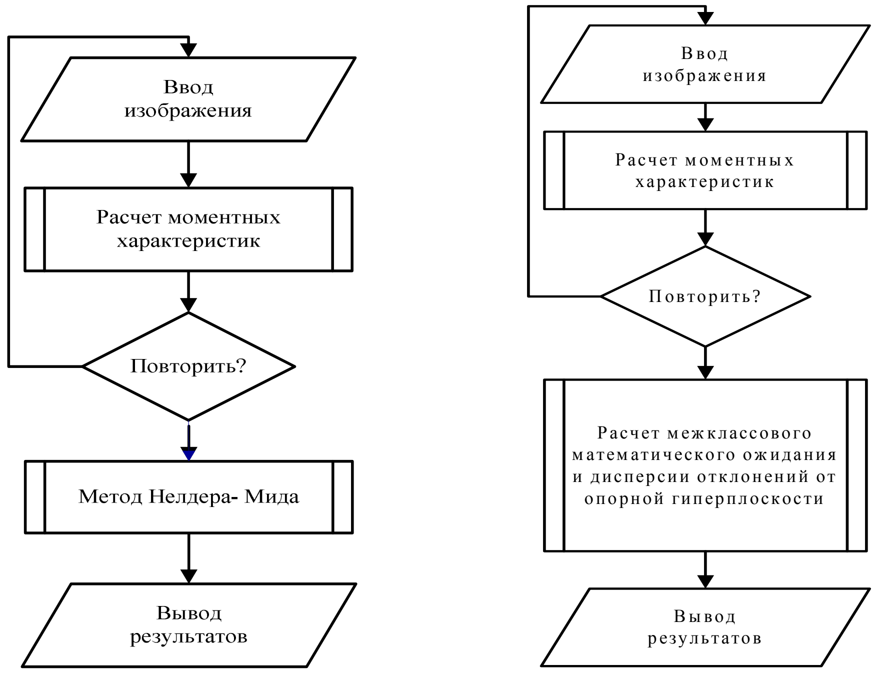

Рис. 3. Структурная схема алгоритма

Результаты тестовых расчетов по 237 изображениям 23 классов показали, что использование размножения параметров с оптимальными значениями коэффициентов размытия ядерных функций позволило уменьшить внутриклассовую дисперсию изображений в 10^6 раз (от первоначального $10^{-5} - 10^{-6}$ до $10^{-11} - 10^{-12}$).

Работа выполнена при поддержке РФФИ, проект № 06–08–00751.

Литература

1. Samal A. and Iyengar P. A. Automatic recognition and analysis of human faces and facial expressions: a survey// Pattern Recognition. – 1992. – Vol. 25, № 1. – P. 65–77.
2. Загоруйко Н.Г. Методы распознавание и их применение. – М.: Сов. радио, 1972.
3. Калайда В.Т. Некоторые проблемы представления и поиска в электронной форме филиграней (водяных знаков) на бумаге книжных памятников / А. И. Елизаров, В. А. Есипова, В. Т. Калайда [и др.] // Охрана и реставрация культурного наследия Сибири. – Томск: Изд. дом «Курсив», 2002. – С. 110–119.
4. Курош А. Г. Курс высшей алгебры. – М.: Наука, 1968. – 431 с.
5. Уилкинсон Р. Справочник алгоритмов на языке АЛГОЛ. – М.: Машиностроение, 1976. – 491 с.
6. Лапко А.В. Непараметрические системы классификации / А.В. Лапко, В.А. Лапко, М.И. Соколов, С.В. Ченцов. – Новосибирск: Наука, 2000. – 240 с.
7. Химмельблау Д. Прикладное нелинейное программирование. – М.: Мир, 1975. – 534 с.

Губанов Николай Юрьевич
Томский государственный университет, студент 5-го курса, гр. 720
Эл. адрес: kolya990@mail.ru.

Калайда Владимир Тимофеевич
ГОУ ВПО «Томский государственный университет систем управления и радиоэлектроники»
Д.т.н., профессор кафедры АСУ

N.U. Gubanov, V.T. Kalayda

Non-parametric statistic method application on picture identification problem

Approach on picture identification problem basis production support hyper-plane by copy formalizing characteristic description of picture of non-parametric statistic method.