

Ю. Б. Б у р к а т о в с к а я, А. В. Д р о б и н с к и й, А. П. С е р ы х
(Томск, ТПУ, ТГУ). **Применение распределений Джонсона в задачах обработки изображений.**

Обычно в задачах обработки изображений в геоинформационных системах либо принимается гипотеза о нормальном распределении значений признаков, характеризующих точки изображения, либо используются методы непараметрической статистики. В работе, представленной данным сообщением, предлагается использовать распределения Джонсона, которые зависят от трех либо четырех параметров и адекватно описывают более широкий класс плотностей, чем нормальное распределение (например, несимметричные плотности и плотности с ограниченным носителем). Таким образом, предлагаемый подход позволит повысить скорость обучения алгоритмов распознавания по сравнению с непараметрическими методами.

Пусть ξ — некоторая случайная величина, ζ — ее преобразование Джонсона, а именно,

$$\zeta = \gamma + \eta\tau(\xi; \varepsilon, \lambda), \quad \eta > 0, \quad -\infty < \gamma < \infty, \quad \lambda > 0, \quad -\infty < \varepsilon < \infty,$$

где $\tau(\cdot)$ — некоторая функция (существуют три различные ее формы), η , γ , λ и ε — параметры распределения, а ζ — стандартная нормальная случайная величина.

Пусть $\xi = [\xi_1, \xi_2, \dots, \xi_m]^T$ — случайный вектор, i -я компонента которого подчиняется распределению Джонсона с функцией $\tau_i(\cdot)$ и параметрами $(\eta_i, \gamma_i, \lambda_i, \varepsilon_i)$. Применяя преобразование Джонсона, получаем случайный вектор $\zeta = [\zeta_1, \zeta_2, \dots, \zeta_m]^T$, компоненты которого имеют стандартное нормальное распределение. Рассмотрим матрицу ковариаций $\Sigma = \|\sigma_{ij}\| = \mathbf{M}\zeta_i\zeta_j$ и предположим, что плотность распределения вектора ζ является нормальной с нулевым вектором математических ожиданий и матрицей ковариаций Σ . Тогда плотность распределения вектора ξ имеет вид

$$f_\xi(x) = \prod_{j=1}^m \tau'_j(x_j; \varepsilon_j, \lambda_j) \frac{1}{\sqrt{(2\pi)^m |D|}} \exp \left\{ -\frac{1}{2}(\tau(x) - \delta)^T D^{-1}(\tau(x) - \delta) \right\},$$

где γ , $\tau(x)$ — $(m \times 1)$ векторы, $\gamma = \|\gamma_i\|$, $\tau(x) = \|\tau_i(x_i; \varepsilon_i, \lambda_i)\|$, η — диагональная $(m \times m)$ -матрица, элементы главной диагонали которой равны η_i , $\delta = -\eta^{-1}\gamma$, $D = \eta^{-1}\Sigma\eta^{-1}$.

Пусть имеется два класса S_1 и S_2 случайных векторов ξ с указанной выше плотностью распределения с параметрами соответственно (δ_1, D_1) и (δ_2, D_2) . Если при этом компоненты ξ_i в обоих классах подчиняются распределению Джонсона с одной и той же функцией $\tau_i(\cdot)$, то расстояние Бхаттачария между этими классами равно

$$d = \frac{1}{8}(\delta_1 - \delta_2)^T [(D_1 + D_2)/2]^{-1}(\delta_1 - \delta_2) + \frac{1}{2} \ln \frac{|(D_1 + D_2)/2|}{\sqrt{|D_1||D_2|}}.$$

На основе полученных результатов был разработан двухэтапный алгоритм кластеризации изображений, который был опробован на модельных и реальных данных. Полученные результаты показывают возможность использования распределений Джонсона для геоинформационных задач.

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований, грант Э 06-07-89190.