

ПРАВИЛО ПРИНЯТИЯ РЕШЕНИЙ ПРИ РАСПОЗНАВАНИИ ОБЪЕКТОВ НА ВИДЕОИЗОБРАЖЕНИИ

Введение

Построение систем распознавания образов на данный момент является одним из наиболее значимых направлений в развитии науки и техники, поскольку они позволяют распознавать ситуации, приводящие к природным и техногенным катастрофам, определять виды и типы движущихся объектов и даже прогнозировать социальные конфликты. Ключевой составляющей такой системы является модуль, реализующий разработанное правило принятия решений, а проектирование новых решающих правил позволяет повысить качество и увеличить скорость распознавания образов, что очень важно для построения телевизионной системы распознавания образов. Правила принятия решений могут быть сгруппированы по категориям и используют для классификации входных объектов дистанционные метрики, статистические методы и нейронные сети.

Анализ методов построения правил принятия решений

Дистанционные метрики позволяют количественно сравнить насколько далеко расположены один от другого два образа. Анализ данных, полученных в процессе сравнения входных образов и образов классов из базы данных, позволяет построить решающее правило на основе Евклидова расстояния [1]

$$d(A, B)_{\text{Евк.}} = \sqrt{\sum_{i=1}^n |A_i - B_i|^2}, \quad (1)$$

где A, B – набор признаков входного образа и образа из базы данных соответственно. Реализация данной метрики проста, но расчет разницы одинаковых видеоизображений уже при малых искажениях может давать большее расстояние, чем между разными изображениями. В таком случае стоит рассмотреть дистанционную метрику Манхэттена, которая используется в сетке тщательно спланированного города [2]. Передвижение может осуществляться по разным маршрутам, но невозможно пересекать квартал по диагонали. Расстояние между координатами двух образов рассчитывается по формуле

$$d(A, B)_{\text{Манх.}} = \sum_{i=1}^n |A_i - B_i|, \quad (2)$$

а дистанционная метрика Минковского – по формуле, являющейся обобщением метрики Манхэттена:

$$d(A, B)_{\text{Минк.}} = \sqrt[k]{\sum_{i=1}^n |A_i - B_i|^k}. \quad (3)$$

Когда параметр k в формуле Минковского стремится к бесконечности, она приобретает вид дистанционной метрики Чебышева

$$d(A, B)_{\text{Чеб.}} = \max_i |A_i - B_i|. \quad (4)$$

При построении классификаторов с использованием статистического подхода, фундаментальным является применение байесовской теории принятия решений. Решающее правило может быть основано на максимуме апостериорной вероятности [3]

$$P(X | R_i) p(R_i) = \max_j \left(P(X | R_j) P(R_j) \right), \quad (5)$$

где X – входной образ; R_i, R_j – образы классов, находящиеся в базе данных системы. Изображение входного объекта относится к образу R_i , если значение апостериорной вероятности является наибольшим относительно остальных образов из базы данных. Когда в коллекции находится большое количество классов, становится возможным выделение основного образа класса, для которого разница параметров относительно входного образа достаточна для принятия решения классификации. В таком случае для расчета байесовской вероятности используются два класса [4]. Если обозначить класс образов из базы данных, принимающий участие в текущей классификации R_i , а R_j – обобщенный класс, сформированный на базе признаков образов остальных классов. Тогда байесовская вероятность рассчитывается по формуле

$$P(R_i | X) = \frac{P(X | R_i) P(R_i)}{P(X | R_i) P(R_i) + P(X | R_j) P(R_j)}. \quad (6)$$

Входной образ соответствует выбранному классу, если полученное значение вероятности больше 1/2. Применение описанных выше методов на практике оказывается затруднительным, поскольку невозможно достоверно предположить значение априорной вероятности. Критерий Неймана - Пирсона, использующий вероятность ложной тревоги ($P_{ЛТ}$) и правильного обнаружения ($P_{ПО}$), является возможным решением возникшей проблемы [5]. Данный критерий стремится минимизировать вероятность ложной тревоги, при которой вероятность правильного обнаружения будет достигать максимального значения. Однако когда классы из коллекции разделяются при помощи гиперплоскостей, то для принятия решения используется минимаксный критерий. основополагающим принципом данного критерия является минимизация вероятности максимума ошибочной классификации. Для этого необходимо рассчитать параметры линейного классификатора, используемого при распознавании [6]

$$f(x) = a^T x + b, \quad (8)$$

где $a \in R^n \setminus \{0\}$, $b \in R$, $x \in R^n$, R^n – Евклидово пространство размерности n . Тогда для каждого из N классов коллекции строится отдельный классификатор, и входной образ x относится к тому классу, для которого значение выражения будет минимальным.

Категория методов, спроектированных на основании обобщенных математических моделей нервной системы, использует искусственные нейронные сети. Элементы такой сети, занимающиеся обработкой информации, называются нейронами или узлами и могут быть представлены многослойным персептроном [7]. Процесс обучения нейронной сети необходим для формирования нескольких гиперплоскостей, которые разделяют классы из коллекции. После обучения многослойный персептрон способен определить для каждого набора входных параметров соответствующий набор выходных значений, классифицирующий входной образ. Также существует алгоритм обучения нейронной сети, основанный на использовании радиальных базисных функций [8]. Типичная структура такой сети содержит три слоя: входной, который принимает параметры входного образа, скрытый, который обрабатывает входные параметры, и выходной, формирующий выходные значения образа. Применение нейронных систем при распознавании образов ограничено, поскольку при большом количестве классов в коллекции процесс обучения занимает много времени.

Алгоритм разработанного правила принятия решений

Для принятия решений о принадлежности входного образа, представляющего собой двумерный спектр видеоизображения, к тому или иному классу было разработано и исследовано решающее правило, базирующееся на критерии подобия Дайса [10]. Собственные векторы, получаемые при ортогональном преобразовании двумерных спектров, позволяют рассчитывать коэффициенты подобия. При этом некоторая избыточность набора признаков, используемых для расчета значения коэффициента подобия, позволяет более эффективно классифицировать образы, поскольку использование малого количества признаков приводит к неправильному определению подобия пары образов [11]. Для расчета коэффициента подобия Дайса (D) используется формула

$$D = \frac{b+c}{2a+b+c}, \quad (13)$$

где a – количество совпавших признаков, b – количество признаков первого образа, отсутствующих во втором, c – количество признаков второго образа, отсутствующих в первом. Диапазон принимаемых значений ограничивается в пределах от 0 до 1, а значение, полученное в результате расчетов, определяет количество различных признаков из совокупности элементов наборов сравниваемых образов [12]. После проведенных расчетов набор признаков, поступивший на вход системы, соответствует некоторому классу, если значение подобия больше определенной пороговой величины. Улучшение качества распознавания при использовании коэффициентов подобия Дайса обеспечивается получением больших значений подобия для наиболее соответствующих образов, поскольку в расчете не учитываются мало-значимые элементы набора признаков. Разрабатываемая система распознавания образов использует альтернативную формулу расчета коэффициента подобия

$$D(A, B) = \frac{2 \times \sum_{i=1}^n a_i b_i}{\sum_{i=1}^n a_i^2 + \sum_{i=1}^n b_i^2}. \quad (14)$$

Модуль, реализующий правило принятия решений (рис. 1), обладает доступом к информации о классах из коллекции и когда на вход системы поступает образ для классификации его признаки соответственно сравниваются с признаками образов из коллекции. При определении достаточного уровня подобия идентификатор класса заносится в список совпавших классов входного образа. После проверки всех образов из коллекции выводится результат распознавания.

Экспериментальная оценка правила принятия решений

Прототип системы распознавания образов позволяет качественно оценить сформированное решающее правило. При разработке было выяснено, что качество распознавания, характеризующее вероятностью ошибки распознавания ($P_{ош}$), зависит от количества используемых векторов при определении подобия образов. На рис. 2 отображено, что уровень ошибки распознавания возрастает при использовании большего количества векторов (n). Это связано с тем, что наиболее информативные признаки находятся в первых собственных векторах, которые характеризуют образ, а остальные содержат признаки, значения которых близки к шумовому порогу и их использование при расчете коэффициентов подобия Дайса ухудшает качество распознавания образов системой.

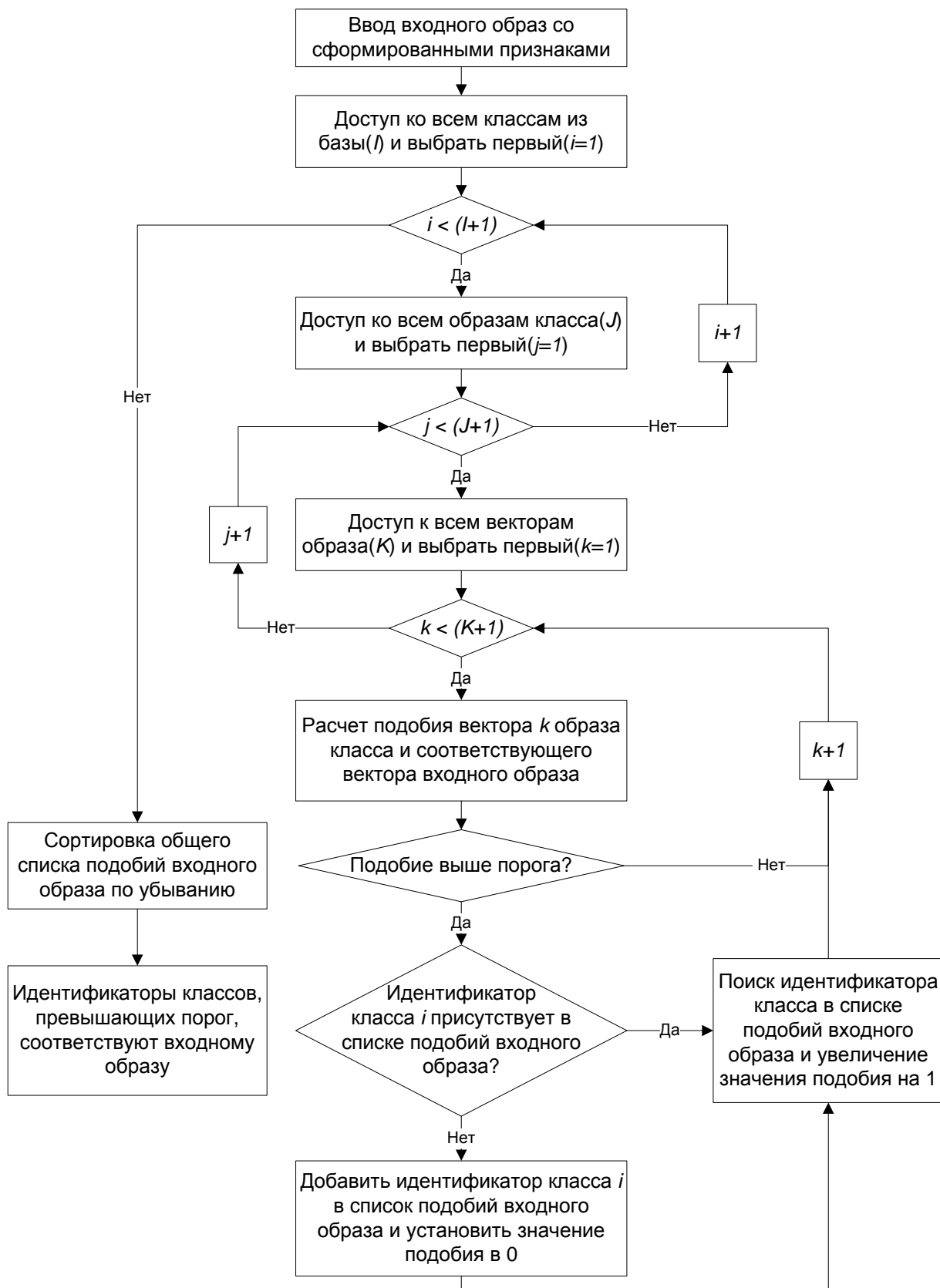


Рис. 1. Правило принятия решений для классификации объектов

Перед предоставлением результатов распознавания классы из списка подобий входного образа ограничиваются порогом идентификации (Thr). Увеличение порога приводит к снижению вероятности ошибки распознавания.

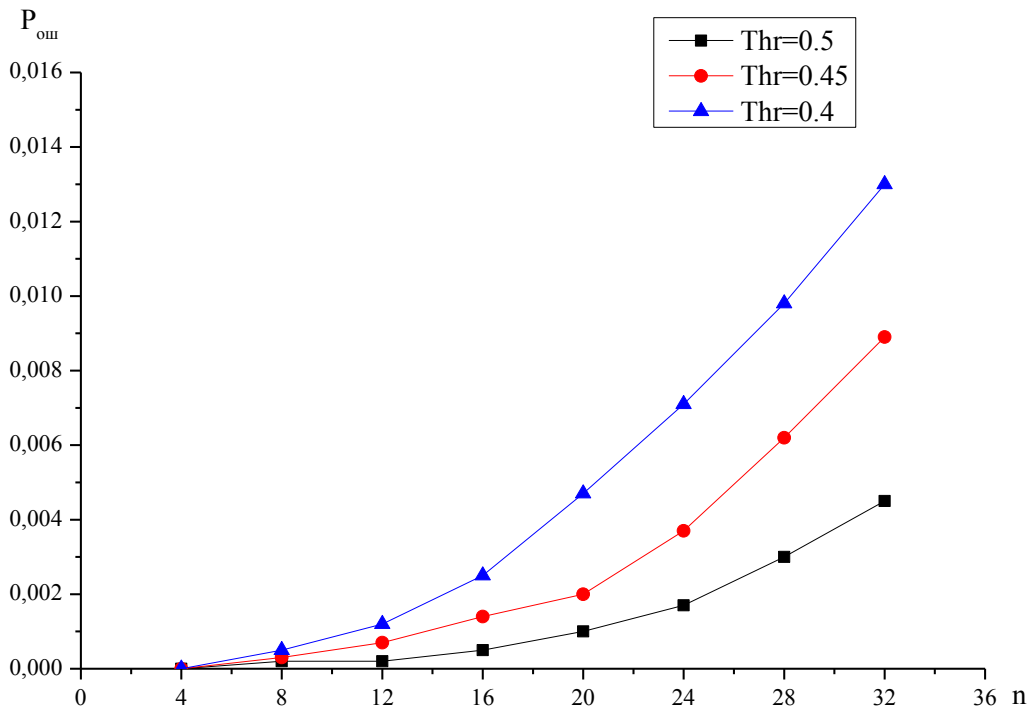


Рис. 2. Зависимость вероятности ошибки распознавания от количества собственных векторов

Влияние порога на качество распознавания, которое характеризуется вероятностью правильного распознавания (P_{np}), отображено на рис. 3.

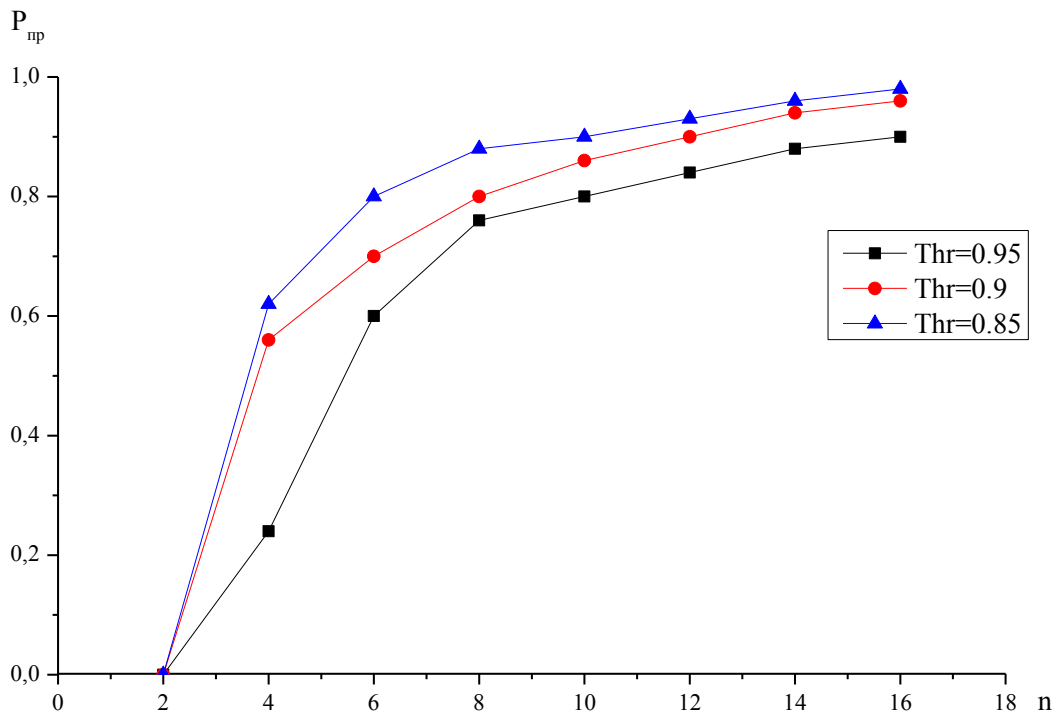


Рис. 3. Зависимость вероятности правильного обнаружения от количества собственных векторов

Выводы

Количество собственных векторов, используемых при распознавания образов, ограничено конфигурацией прототипа системы распознавания. Следовательно, количество векторов входного образа, подобных к образу класса из базы данных, возможно нормировать и при обучении системы сформировать порог идентификации. Требования к системе распознавания описывают стратегию определения оптимального значения порога. Затем данная величина будет использоваться для принятия решения при классификации входных образов, а на рис. 2 и 3 продемонстрированы возможности по настройке системы для обеспечения необходимого качества распознавания образов.

Список литературы: 1. *Sung-Huuk C.* Comprehensive Survey on Distance / Similarity Measures between Probability Density Functions // International Journal of Mathematical Models and Methods in Applied Science, Volume 1, Issue 4, 2007, pp. 300 – 307. 2. *Reynolds E. B.* Taxicab Geometry // The Pi Mu Epsilon Journal, Volume 7, Number 2, 1980, pp. 77 – 88. 3. *Liu C., Wechsler H.* A unified Bayesian framework for face recognition // Image Processing, IEEE International Conference, Vvolume 1, 1998, pp. 151 – 155. 4. *James E. A. K., Annadurai S.* An Efficient Bayesian Approach to Face Recognition based on Wavelet Transform // International Journal of Computer Applications, Volume 15, Number 8, 2011, pp. 22 – 26. 5. *Van Trees H. L.* Detection, Estimation, and Modulation Theory, Part I: Detection, Estimation, and Linear Modulation Theory // John Wiley & Sons, 2001, 710 p. 6. *Kitahara T., Mizuno S., Nakata K.* An Extension of a Minimax Approach to Multiple Classification // Journal of the Operations Research Society of Japan, Volume 50, Number 2, 2007, pp. 123 – 136. 7. *Ososkov G., Stadnik A.* Face Recognition By a New Type of Neural Networks // Proceedings of World Scientific and Engineering Society, 2001, pp. 304 – 308. 8. *Bhowmik M. K., Bhattacharjee D., Nasipuri M., Basu D. K., Kundu M.* Classification of Fused Images using Radial Basis Function Neural Network for Human Face Recognition // Nature & Biologically Inspired Computing, IEEE, 2009, pp. 19-24. 9. *Howell A. J., Buxton H.* Face Recognition using Radial Basis Function Neural Networks // Proceedings of the British Machine Vision Conference, 1996, pp. 455–464. 10. *Lin Z., Lyu M. R., King I.* MatchSim: A Novel Similarity Measure Based on Maximum Neighborhood Matching // Knowledge and Information Systems, Volume 32, Issue 1, 2012, pp. 141 – 166. 11. *Van Deemter K., Gatt A.* Beyond DICE: Measuring The Quality of a Referring Expression // Proceeding of the Workshop on Production of Referring Expressions: bridging Computational and Psycholinguistic Approaches (preCogSci-09), 2009. 12. *Anderson M. J., Ellingsen K. E., McArdle B. H.* Multivariate Dispersion as a Measure of Beta Diversity // Ecology Letters, number 6, 2006, pp. 683 – 693.

Харьковский национальный
университет радиоэлектроники

Поступила в редколлегию 10.03.2014