

В.А. Гороховатский, Д.В. Пупченко

Харьковский национальный университет радиоэлектроники, Харьков

КЛАССИФИКАЦИЯ ИЗОБРАЖЕНИЙ ВИЗУАЛЬНЫХ ОБЪЕКТОВ ПО МНОЖЕСТВУ ДЕСКРИПТОРОВ ОСОБЕННЫХ ТОЧЕК НА ОСНОВЕ НЕЙРОННОЙ СЕТИ КОХОНЕНА

Решена задача распознавания визуальных объектов на основе классификации дескрипторов изображений с применением сети Кохонена. Обсуждаются варианты и результаты обработки в плане различного числа подстраиваемых нейронов. Анализируются критерии оценивания уровня правильной классификации. Приводятся результаты программного моделирования сети с разными вариантами обработки и формирования центров классов.

Ключевые слова: компьютерное зрение, особые точки изображения, дескрипторы особых точек, детектор ORB, сеть Кохонена, программное моделирование, число подстраиваемых нейронов, критерий правильной классификации.

Введение

В системах компьютерного зрения представляется актуальным применение методов структурной классификации изображений в ситуации, где число классов (эталонов) равно числу кластеров, созданных на множестве характерных признаков (особенных точек) базы образцов [1-4]. В таком случае применяемая кластеризация непосредственно выполняет группирование множества дескрипторов эталонных изображений в фиксированное множество классов, т.е. решает задачу классификации. На этапе кластеризации осуществляется разбиение множества элементов структурных описаний на классы (подмножества признаков) и построение связанного с ним набора центров классов. Классификация вновь поступивших описаний воплощается путем конкурентного отнесения их элементов к одному из сформированных центров классов, а затем определения класса, набравшего наибольшее число дескрипторов (голосов) [4, 6].

Качество классификации напрямую зависит от результатов обучения системы распознавания на имеющемся множестве признаков структурных описаний базы образцов (обучающая выборка). Представляет интерес изучение возможностей и свойств такого процесса обучения в плане наиболее эффективного использования содержащейся в описаниях эталонной информации, а также исследование углубленных схем обучения, учитывающих степень близости элементов разных классов в пространстве признаков.

В последнее время популярны методы ORB и BRISK, которые обеспечивают детектирование особенных точек (ОТ) изображений [3,4]. ОТ – это числовой вектор, координатно зафиксированный в некоторой точке. BRISK по сравнению с ORB имеет набор альтернатив для формы масок при выявлении

ключевых точек. Основные преимущества детекторов ORB и BRISK заключаются в обеспечении высокого быстродействия за счет существенного упрощения процесса обработки в пространстве дескрипторов бинарного типа.

В связи с рядом прикладных преимуществ актуально применение конкурентных структур при обучении и классификации, а именно – сетей Кохонена, которые реализованы в пакетном и онлайн-режимах [1, 2, 7]. Сети Кохонена могут универсально и с успехом настраиваться на произвольные наборы данных, в том числе полученные в системах визуальной обработки [4]. Представляет интерес изучение адаптационных свойств и параметрических характеристик сетей Кохонена применительно к обучению в задачах классификации изображений по множеству ОТ, где фактически наблюдается неограниченное разнообразие анализируемых данных.

Целью работы есть изучение модификаций сети Кохонена применительно к задаче классификации изображений в пространстве дескрипторов особых точек и оценивание их результативности для конкретных баз изображений.

Задачи исследования – изучение адаптационных возможностей и особенностей применения сети Кохонена для множества дескрипторов, сравнение разных вариантов формирования центров классов, оценивание эффективности работы сети путем программного моделирования.

Применение сети Кохонена для классификации

Пусть $W = \{x \mid x \in R^n\}$, $W \subseteq R^n$ – пространство структурных признаков (дескрипторов) изображений. База описаний образцов (эталонов) $Z \subset W$ задана в виде множества $Z = \{Z^j\}_{j=1}^J$, $s_j = \text{card } Z^j$, $s = \text{card } Z$, $s = \sum_j s_j$. При этом каждый признак

$x_i^j \in Z^j$ асоціюється з еталоном Z^j в плані включення в його склад, ібо база образців задана.

Возьмемо за основу модель класифікації, коли для кожного класу, задаваного еталоном, попередньо синтезується його деякий «центр», а класифікація здійснюється шляхом віднесення до класу з найбільш близьким центром [4]. Приміємо для навчання і отримання центрів класів конкурентну структуру – нейронну мережу Кохонена [1, 2, 5]. Он-лайн-кластеризатор на основі мережі Кохонена і його застосування для множини структурних ознак зображень докладно обговорюються в [2]. Розглянемо процедуру навчання мережі в наступному вигляді.

1. Ініціюємо матрицю $M = \{m_j\}_{j=1}^J$ ваг, рядками якої є формовані вектори нейронів m_j – центрів класів так, що $m_j = x_1^j$, $x_1^j \in Z^j$, де i – номер довільного вектора з класу образців Z^j .

2. Сформуємо елементи навчального множини в вигляді сукупності $Z = \{Z^j\}_{j=1}^J$ всіх s елементів бази еталонів.

3. Вибіримо поточний елемент $z \in Z$, для кожного $j \in [1, 2, \dots, J]$ вивчислимо відстань $q_j = \rho(z, m_j)$ і визначимо клас d нейрона-переможця: $d = \arg \min_j q_j$. Вид метрики визначається властивостями дескрипторів.

4. Вивчислимо зміни ваг для нейронів вихідного шару мережі

$$\Delta m_j = h(j, d, t) \cdot \eta \cdot (z - m_j), \quad (1)$$

де η – задаваний коефіцієнт швидкості навчання; $h(j, d, t)$ – значення функції близькості для нейрона j в момент часу навчання t ; традиційно $h(j, d, t) = \exp[-\rho(j, d) / \sigma(t)]$ визначають в вигляді гауссової функції, а радіус $\sigma(t) = 1 / \exp(t^{-2})$ близькості зменшують з збільшенням параметра t часу навчання $t = 1, \dots, s$.

5. Корректируємо матрицю ваг $M = M + \Delta M$ на кроці t .

6. Продовжуємо навчання (п. 3-5) до завершення списку Z .

7. Перевіряємо виконання критерію зупинки навчання. Це може бути величина помилки кластеризації. При необхідності (невиконання умов зупинки) продовжуємо навчання з п. 2. Це може бути дублювання множини входних даних, вибір даних з множини Z в установленому або в довільному порядку і т.д.

Дослідники відзначають, що нормування даних (приведення норми вектора до одиниці) при застосуванні мережі Кохонена прискорює збіжність

мережі до локального мінімуму помилки [1, 8]. Для мереж з незначительним числом нейронів (2-3) застосування нормування вважається обов'язковим. Якщо обробляються нормовані дані, то нейрони мережі стають нормованими в ході навчання. В даній ситуації ми використовуємо ненормовані дані, прагнучи забезпечити достатнє віддалення нейронів між собою з метою забезпечення результативного розділення класів. В той же час існує досвід застосування для кластеризації і нормованих детекторів ОТ [4]. Значення помилки кластеризації оцінимо критерієм

$$E = \frac{1}{sN} \sum_{j=1}^J \sum_{v=1}^{s(j)} \rho(x_v, m_j), \quad (2)$$

де сумуються відстані елементів до центру всередині кластерів, а потім застосовується усереднення по числу s елементів навчального множини і по розміру N дескриптора ОТ [4].

Іншим ключовим індикатором для оцінювання якості класифікації може бути значення критерію, що підраховує частку елементів навчальної вибірки Z , які потрапили «не в свої» класи. Він визначається як

$$\beta = \sum_{j=1}^J (s_j - a_j) / s, \quad (3)$$

де a_j – кількість ознак з загальної кількості s_j в описанні еталона Z^j , віднесених в процесі класифікації до класу j .

Значення β відображають рівень помилкових рішень при класифікації ознак множини Z . Чим ближче β до нуля, тим вище досягнуто якість класифікації на навчальній вибірці.

Вивчимо в порівняльному плані три базові варіанти побудови класифікатора, що відрізняються кількістю нейронів, які модифікуються в процесі навчання:

- 1) налаштовується тільки нейрон-переможець;
- 2) модифікуються три нейрони, найближчих до відстані до нейрона-переможця;
- 3) модифікується повна мережа нейронів відповідно до відстані до нейрона-переможця.

Оцінимо результативність розпізнавання з застосуванням цих варіантів навчання мережі, що відрізняються глибиною аналізу даних і управлінням мережею.

Обговорювані варіанти навчання класифікатора принципово відрізняються тільки п. 4, де по-різному визначається функція $h(j, d, t)$. В той же час, для досягнення збіжності повинно бути виконано умову $h(j, d, t) \in [0, 1]$. При цьому версія 1 реалізує грубе розпізнавання (налаштовується тільки переможець, принцип «переможець забирає все»). Версія 2 відносить елемент до трьох найбільш близьких центрів (м'якше визначення, «переможець забирає

больше»), а версия 3 осуществляет наиболее тонкую обработку, подобно нечеткому подходу относя анализируемый дескриптор ко всем кластерам одновременно, но с разными весами, пропорциональными расстоянию. Версии 2 и 3 позволяют управлять нейронами – соседями победителя, и фактически реализуют принцип «нейронного газа», в соответствии с которым осуществляют ранжирование элементов сети из окрестности признака обучения [1, 7].

Вычисление центров классов

Результативность работы сети Кохонена в значительной мере определяется начальными условиями ее функционирования, которые определяются значениями начальных центров m_j в п. 1 процедуры. В связи с этим изучим такие варианты выбора центров:

1) произвольно по одному дескриптору из каждого эталона;

2) посредством применения специальной процедуры формирования центров.

Учитывая бинарное представление дескрипторов, для каждого Z^i определим вектор центра класса на основании логического правила

$$m_i(b) = \begin{cases} 1, & \sum_{d=1}^{s(i)} x_d(b) \geq s(i)/2, \\ 0, & \sum_{d=1}^{s(i)} x_d(b) < s(i)/2, \end{cases} \quad (4)$$

$$x_d \in Z^i, \quad b = 1, \dots, 256,$$

где $x_d(b)$ – бит с номером b для дескриптора с номером d в описании эталона.

Центр (4) отражает свойства эталона (класса). Значение каждого из битов m_i определяется значениями соответствующих битов дескрипторов, принадлежащих эталону с номером i .

Анализируем элементы $x \in Z$ содержания структурных описаний базы эталонов (обучающая выборка) путем отнесения их к соответствующему классу v с использованием конкурентного способа

$$v = \arg \min_{i=1, \dots, J} \rho(x, m_i). \quad (5)$$

Отметим, что в (5), как и в п. 3 процедуры обучения Кохонена, необходимо определить метрику для вычисления отличия двух дескрипторов. В методе Кохонена естественным есть применение евклидовой метрики или ее квадрата

$$\rho(x, m_i) = \sum_{b=1}^{256} (x(b) - m_i(b))^2, \quad (6)$$

ибо здесь обрабатываются вещественные данные (256 – размер детектора ORB). В то же время при работе с битовыми данными, например в (5), более эффективно с точки зрения объема вычислений использовать метрику Хемминга

$$\rho(x, m_i) = \sum_{b=1}^{256} |x(b) - m_i(b)|, \quad (7)$$

что определяет количество расходящихся битов для двух двоичных последовательностей одинаковой длины. В результате обработки для каждого образца в соответствии с процедурой Кохонена или с соотношением (5) получаем кластерное представление $h[Z^i] = (h_1^i, \dots, h_J^i)$ эталона, где h_a^i – целые числа. Оно соответствует распределению элементов множества Z^i по классам эталонов. Если рассмотреть набор векторов $h[Z^i]$ как строки матрицы размером $J \times J$, то ошибка (3) вычисляется через сумму диагонали этой матрицы, деленной на общее число признаков s базы.

Результаты экспериментальных исследований

В среде C++ нами разработана программная модель для классификации изображений на основании структурных описаний, сформированных детектором ORB. Вычисления, анализ и обработка детекторов ОТ реализованы в виде программного средства на основе модуля features2d кросс-платформенной библиотеки OpenCV [9].

Дескрипторы имеют вид бинарного вектора размером 256, кодируются в Open CV в виде типа uchar (unsigned char – тип данных в C++, используются для хранения символов, объем 8 бит, значение 0 ... 255), а не в битах. Данные хранятся в матрице, где количество строк равно числу выявленных дескрипторов, а число столбцов равно 32 (256 бит дескриптора трансформируются в 32 uchar). В качестве базы эталонов применен набор из 5 изображений денежных знаков (рис. 1).

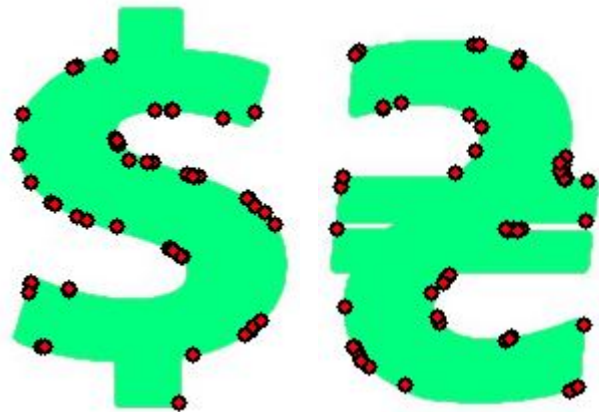


Рис. 1. Примеры изображений с координатами выделенных ОТ

В целях формирования качественных структурных описаний, устойчивых к фоновым и случайным помехам, на основе использования стандартных функций библиотеки OpenCV проведена специальная обработка анализируемых изображений, которую образно можно назвать «ретушированием». В ходе такой обработки изображение подвергалось фильтрации и бинаризации: изображение переводило-

лось в градации оттенков серого цвета, выбран по- порю яркости по критерию бинаризации Отсу, прове- дена бинаризация [10]. Примененная методика на- правлена на выделение контура объекта и снижает влияние помех, пересечений областей полутонов, случайных областей с резким изменением интен- сивности и т.д. Количество выявленных ОТ колеба- лось в пределах 47 ... 50, общий объем обучающей выборки, составило 244 ОТ.

Проведем обработку (4) для множеств дескрип- торов каждого из эталонов и для полученных цен- тров с использованием метрики Хемминга выпол- ним классификацию (5). В результате получена мат- рица распределения ОТ по классам в виде табл. 1.

Таблица 1
Распределение ОТ по классам с центрами (4)

Эталон	Классы				
	1	2	3	4	5
Z^1	14	17	3	4	14
Z^2	9	28	7	1	8
Z^3	10	18	7	7	11
Z^4	16	15	4	12	7
Z^5	7	14	2	1	29

Как видно из табл. 1, множество дескрипторов ORB для конкретной базы изображений с использо- ванием ретуширования и бинарной обработки не позволяет однозначно классифицировать множество ОТ каждого из эталонов.

Причина – число ОТ, отнесенных к определен- ному классу, не является максимальным для эталонов Z^1 , Z^3 и Z^4 . Ошибка классификации составляет $E = 0,74$ по формуле (2) и $\beta = 0,68$ по формуле (3).

Оценим и сравним теперь результативность клас- сификации с применением обсуждаемых трёх вариан- тов обучения сети, различающихся способом управления сетью из-за различного числа настраиваемых нейро- нов. Исходный набор данных содержит 244 бинарных вектора длины 256. Количество итераций равно коли- честву входных векторов. Матрица весов – матрица действительных чисел размером 244 на 256. На тестовом компьютере с процессором Intel Core i5-4440 об- работка выполняется в течении 3.8 - 4.3 с.

В табл. 2 показаны распределения дескрипторов ОТ по классам (варианты произвольного выбора цен- тров и вычисления по формуле (4)) для версии, когда в процессе обучения настраивается нейрон- победитель. Во второй половине табл. 2 наблюдается достаточно высокий уровень различения эталонов: максимумы значений находятся на диагоналях и зна- чительно превышают другие элементы в строках. Ошибки в случае произвольных центров составляют $E = 0,63$ и $\beta = 0,67$. В случае сформированных по (4) центров – $E = 0,38$ и $\beta = 0,34$. Как видим, примене- ние сети Кохонена в случае специального формиро- вания центров (4) приводит к значительному сниже- нию ошибки классификации, а следовательно, улуч- шается качество различения объектов по множеству значений их дескрипторов.

Таблица 2

Распределение ОТ эталонов в первом варианте обработки

Эталон	Произвольный выбор центров				
	Классы				
	1	2	3	4	5
Z^1	19	0	20	3	8
Z^2	6	14	18	5	7
Z^3	6	5	18	3	15
Z^4	6	1	14	21	8
Z^5	4	4	5	20	14

Эталон	Формирование центров (4)				
	Классы				
	1	2	3	4	5
Z^1	41	4	2	0	3
Z^2	20	26	1	0	3
Z^3	17	2	23	0	5
Z^4	14	2	2	30	2
Z^5	4	4	9	0	30

Кроме того, значения ошибок E и β оказались достаточно близки. Для варианта модификации трёх ближайших к победителю нейронов в случае произ- вольных центров ошибка составляет $E = 0,62$ и $\beta = 0,60$. Для сформированных в виде (4) центров $E = 0,30$ $\beta = 0,26$. Для варианта модификации всей сети нейронов ошибка в случае произвольных цен- тров составляет $E = 0,58$ и $\beta = 0,61$. Для сформиро- ванных в виде (4) центров $E = 0,22$ $\beta = 0,20$.

Как видим, с расширением числа настраиваемых нейронов, особенно для варианта специально сфор- мированных центров (4), ошибка классификации су-

щественно снижается. Например, динамика измене- ния значений β для трех вариантов обработки соста- вили 0,34, 0,26 и 0,20. Это подчеркивает возможно- сти адаптации сети Кохонена к произвольным дан- ным, особенно в случае успешного формирования первоначальных центров классов. Расширение набора адаптируемых нейронов от 1 до общего их числа по- вышает показатели удачной работы сети.

Выводы

Эффективность классификации визуальных объектов по множеству особых точек с использова- нием сети Кохонена непосредственно зависит от сле-

дующих ключевых взаимосвязанных факторов: распознаваемая база, метод предварительной обработки, способ формирования дескрипторов, начальный выбор центров, метрика для сравнения дескрипторов, размер кортежа подстраиваемых нейронов.

Исследование подтвердило универсальность и возможности успешной адаптации сети Кохонена к произвольным наборам визуальных данных, особенно в случае удачного формирования первоначальных центров классов. В ходе обучения сеть Кохонена как-бы «вытягивает» степень разрозненности данных, улучшая конечные показатели классификации. Это позволяет решать более сложный спектр задач классификации. Расширение набора адаптируемых нейронов в процессе обучения значимо совершенствует показатели действенной работы сети.

Наряду с путем непосредственного отнесения дескрипторов распознаваемого изображения по максимальному подобию к центрам в классификации могут быть применены способы долевого формирования подобия с весовым значением каждого из классов. Эти методы менее чувствительны к фактору равномерности распределения дескрипторов эталонов по классам. В плане влияния на результат классификации обладают преимуществом те методы формирования дескрипторов, которые имеют большую размерность дескриптора, поэтому BRISK имеет преимущество по отношению к ORB.

Научная новизна исследования состоит в построении метода структурной классификации на основе модификаций сети Кохонена, что обеспечивает необходимую результативность различения объектов. **Практическая значимость** – получение прикладных многовариантных программных моделей для оценивания эффективности метода классификации и подтверждение ее эффективности для конкретных баз изображений. Дальнейшее улучшение результативности распознавания может быть достигнуто за счет подходов обучения с учителем, т.к. классы дескрипторов базы известны. Другим раз-

витием методов обучения на сетях Кохонена есть применение пакетной обработки, имеющей, как правило, более высокие показатели качества классификации из-за тщательного и многократного анализа данных. В целом это позволяет обучить сеть и достичь требуемых показателей классификации универсально для произвольной базы.

Список литературы

1. Осовский С. *Нейронные сети для обработки информации* / С. Осовский. – М.: Финансы и статистика, 2002. – 344 с.
2. Борисов Е. *Кластеризатор на основе нейронной сети Кохонена*. Режим доступа: <http://mechanoide.kiev.ua/neural-net-kohonen-clusterization.html>
3. Ethan Rublee, Vincent Rabaud, Kurt Konolige, Gary Bradski. *ORB: an efficient alternative to SIFT or SURF*. – *Computer Vision (ICCV), IEEE International Conference on IEEE*, pp. 2564 – 2571, 2011.
4. Гороховатский В.А. *Структурное распознавание изображений с применением моделей интеллектуальной обработки и самоорганизации признаков* / В.А. Гороховатский, А.В. Гороховатский, А.Е. Берестовский // *Радиоэлектроника, информатика, управление*. – 2016. – №3 (38). – С. 39–46.
5. Кохонен Т. *Самоорганизующиеся карты* / Т. Кохонен – М.: БИНОМ. Лаборатория знаний, 2013. – 655 с.
6. Gorokhovatsky V.A. *Efficient Estimation of Visual Object Relevance during Recognition through their Vector Descriptions* / V.A. Gorokhovatsky // *Telecommunications and Radio Engineering*. – 2016, Vol. 75, No 14. – P. 1271–1283.
7. Новотарський М.А. *Штучні нейронні мережі: обчислення*. / М.А. Новотарський, Б.Б. Нестеренко // *Праці інст. математики НАН України*. – 2004, т. 50. – 408 с.
8. Горбаченко В. И. *Сети и карты Кохонена*. [Электронный ресурс]. – Режим доступа: http://gorbachenko.self-organization.ru/articles/Self-organizing_map.pdf
9. *OpenCV Open Source Computer Vision*. [Электронный ресурс]. – Режим доступа: <https://docs.opencv.org/master/index.html>.
10. *OpenCV Image Thresholding* [Электронный ресурс]. – Режим доступа: http://docs.opencv.org/3.3.0/d7/44d/tutorial_py_thresholding.html

Надійшла до редколегії 11.02.2018

Рецензент: д-р техн. наук, проф. С.П. Пуятін, Харківський національний університет радіоелектроніки, Харків.

КЛАСИФІКАЦІЯ ЗОБРАЖЕНЬ ВІЗУАЛЬНИХ ОБ'ЄКТІВ ЗА МНОЖИНОЮ ДЕСКРИПТОРІВ ОСОБЛИВИХ ТОЧОК НА БАЗІ НЕЙРОННОЇ МЕРЕЖІ КОХОНЕНА

В.О. Гороховатський, Д.В. Пупченко

Вирішена задача розпізнавання візуальних об'єктів на основі класифікації дескрипторів зображень із застосуванням мережі Кохонена. Обговорюються варіанти та результати оброблення в плані різного числа нейронів, що налагоджуються. Аналізуються критерії оцінки рівня правильної класифікації. Приводяться результати програмного моделювання мережі з різними варіантами оброблення та формування центрів класів.

Ключові слова: комп'ютерний зір, особливі точки зображення, дескриптори особливих точок, детектор ORB, мережа Кохонена, програмне моделювання, кількість налагоджуваних нейронів, критерії правильної класифікації.

CLASSIFICATION OF IMAGES OF VISUAL OBJECTS BY A SET OF UNIQUE POINT DESCRIPTORS BASED ON THE KOHONEN NEURAL NETWORK

V.O. Gorokhovatsky, D.V. Pupchenko

The task of recognizing visual objects based on classification of image descriptors using the Kohonen network is solved. Variants and results of computation in terms of different numbers of adapted neurons are discussed. The criteria for assessing the level of the correct classification are analyzed. The results of the software modeling of the network with different variants of processing and formation of the centers of classes are given.

Keywords: computer vision, special image points, special point descriptors, ORB detector, Kohonen network, software simulation, number of tunable neurons, correct classification criterion.