

УДК 004.932.2:004.93'1



Н.В. Власенко

ХНУРЭ, г. Харьков, Украина, gorohovatskaja@gmail.com

ПОСТРОЕНИЕ ИНФОРМАТИВНЫХ КОМПАКТНЫХ ОПИСАНИЙ И КЛАССИФИКАЦИЯ ОБЪЕКТОВ ПУТЕМ ПРЕДСТАВЛЕНИЯ В ОРТОГОНАЛЬНОМ БАЗИСЕ

Рассматривается применение ортогонального базиса для построения сжатых информативных описаний изображений. Предложен эффективный метод построения компактных описаний на основе применения базиса функций Уолша. Приведены результаты сравнительных экспериментов предложенного метода с известными подходами.

СТРУКТУРНОЕ ОПИСАНИЕ ВИДЕО-ОБЕКТА, ФУНКЦИИ УОЛША, ОРТОГОНАЛЬНЫЙ БАЗИС, СЖАТИЕ, КЛАССИФИКАЦИЯ

Введение

Классификация – это отнесение данных к определенному классу путем выделения существенных признаков из общей массы данных. Архиважным здесь представляется формирование системы признаков, количество и качество которых необходимо и достаточно для осуществления эффективной классификации. Известные методы выявления точечных особенностей изображений [1, 2] позволяют сформировать описание видео-объекта в виде конечного множества векторов-дескрипторов – характерных признаков (ХП).

Модель сигнала изображения представляют как конечный набор Z ХП из пространства R^n векторов фиксированной размерности с действительными компонентами [1-4]. Описание видео-объекта – это подмножество $Z \subset R^n$. При этом векторы, элементы описания, полученного, например, с использованием методов SIFT или SURF [1], являются инвариантными к смещениям, масштабу и вращению. Нумерация элементов во множестве Z получается в процессе формирования описания.

Описание Z имеет в некотором смысле противоречивый характер с точки зрения требований к классификации. С одной стороны, оно является в некотором смысле однородным (векторы из одного видео-объекта). С другой стороны, возникают трудности из-за того, что векторы из описаний разных объектов могут быть эквивалентными. Кроме того, в практических задачах классификатору часто достаточно ограниченного числа элементов описания Z для принятия решения с тем же уровнем ошибки. Это порождает возможность постановки и решения проблемы уменьшения избыточности путем достижения компактности описаний для эффективного хранения и обработки данных [3].

Эффективное представление в пространстве векторов может быть связано с преобразованием к некоторому ортогональному пространству. Прикладной интерес по причине достаточно простой программной или аппаратной реализации вызывает применение в этих целях функций Уолша (ФУ) и преобразований, связанных с ними [3-5].

Для компактности опишем Z в виде матрицы $Z = \{z_{ij}\}, i = \overline{1, s}, j = \overline{1, n}$, строки которой содержат s векторов описания. Поставим задачу путем перехода в другое пространство на основе построения отображения $\Theta: Z \rightarrow U$ за счет применения ортогонального разложения по системе ФУ сформировать более компактное информативное описание, обладающее лучшими свойствами для классификации.

Цель статьи – разработка эффективного в плане быстродействия метода классификации изображений на основе построения сжатых описаний с использованием наиболее информативных компонентов спектра Уолша.

Задачи исследования включают формализацию и обоснование метода формирования компактного информативного описания и практическое сопоставление синтезированного метода с известными подходами при применении их для обработки изображений из реальных баз видео-информации.

2. Метод сжатия описания на основе ортогонального базиса

Применения двумерных преобразований в обработке изображений связаны с подчеркиванием отдельных свойств, кодированием изображений путем отбрасывания малых по величине коэффициентов преобразования, сокращением размерности описания для повышения быстродействия [6].

Рассмотрим преобразование, связанное с разложением по семейству прямоугольных базисных функций, в качестве которых рассмотрим дискретные ФУ, представляющие собой целочисленные векторы конечной размерности, состоящие из 1 и -1. В практических применениях исследуются ФУ с размерностью $n = 2^k$. Полный набор ФУ для фиксированного n образует ортогональную матрицу Адамара [3]. Преобразуем описание путем умножения матриц

$$U = Z * A, \quad (1)$$

где A – матрица Адамара размера $n \times n$, включающая n векторов ФУ w_1, \dots, w_n . Например, матрица Адамара размерностью 4 имеет вид:

$$A = \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix},$$

где коэффициент $1/2$ служит для обеспечения свойства ортонормальности. Ортонормальность при практическом применении ФУ играет не такую принципиальную роль, как, например, при использовании базиса Фрэя-Чена [2], т.к. нормировочные коэффициенты равны для всех ФУ одного базиса. Выражение (1) можно трактовать как одномерное преобразование Уолша, выполненное для каждого из векторов описания Z .

Метод синтеза сжатого описания через отображение $\Theta: Z \rightarrow U$ представим как последовательность шагов.

1. Применяя к исходному описанию преобразование (1), получим матрицу $U = \{u_{ij}\}$ размером $s \times n$. Вычислим квадраты элементов U и будем анализировать матрицу $U^2 = \{u_{ij}^2\}$. Значение u_{ij}^2 соответствует энергии ХП с номером i из описания Z , приходящейся на базисный вектор w_j (с номером j). Как известно, при представлении в ортонормированном базисе полная энергия сигнала равна сумме компонентов энергии, соответствующих всем базисным векторам [2].

2. В целях построения информативного кортежа из спектра применим анализ посредством дисперсионного критерия [3,4], вычисляя дисперсии элементов в каждом из n столбцов матрицы U^2 :

$$\sigma_i^2 = \frac{1}{s-1} \sum_{k=1}^s (u_{ik}^2 - u_i^2)^2,$$

где $u_i^2 = \frac{1}{s} \sum_{k=1}^s u_{ik}^2$, $i = \overline{1, n}$. Значения дисперсии характеризуют разброс величин откликов ХП для отдельных ФУ спектра. Далее отсортируем полученный массив дисперсий $\sigma_1^2, \dots, \sigma_n^2$ с сохранением номеров ФУ, для которых они вычислены.

3. Выберем фиксированное число $m \ll n$ ФУ с наибольшими дисперсиями. Выбранные ФУ w_1, \dots, w_m с их исходными номерами j_1, \dots, j_m будем использовать как наиболее информативную часть базиса для получения сжатого описания. Здесь применен подход [5], где в целях выделения и

фильтрации протяженных объектов используется ограниченный набор ФУ с наиболее подходящими свойствами.

4. Осуществим классификацию векторов описания Z на m классов в соответствии с выбранным информативным кортежем ФУ. Для этого проанализируем избранные m столбцов матрицы U^2 и для каждого из s векторов исходного описания реализуем оптимальное правило R_Z отнесения вектора $z_q \in Z$ к одному из классов k в диапазоне $1, \dots, m$:

$$R_Z: z_q \in k \mid \arg \max_{h=1, \dots, m} u_{qh}^2 = k, k = \overline{1, m}, q = \overline{1, s}. \quad (2)$$

Таким образом, каждый элемент исходного описания будет отнесен к тому из m классов, для которого квадрат отклика с соответствующей функцией Уолша будет максимальным. Правило R_Z реализует классификацию на основе оптимальной согласованной фильтрации элементов описания в ортогональном базисе кортежа ФУ. Здесь возможны варианты обработки с ограничениями и отсеиванием, когда вектор ХП не классифицируется ни к одному из классов в силу незначимости отклика.

5. В результате осуществления (2) получим распределение s элементов по m классам. Сформируем сжатое описание u^* в интегрированном виде как вектор целых чисел

$$u^* = (u_1, \dots, u_m), \sum_{j=1}^m u_j = s, \quad (3)$$

элементы u_j которого – количества векторов ХП, отнесенные к классу j . Учитывая, что эталоны классов могут содержать в своем составе разное число ХП, в практических задачах часто необходимо осуществить нормировку описания (3): $u_i = u_i / s$, что приводит к соотношению $\sum_{j=1}^m u_j = 1$. Схема построения метода приведена на рис. 1.

Предложенный метод преобразует исходное описание в виде матрицы размера $s \times n$ к вектору из m числовых значений, причем можно выбрать $m \ll n$, что существенно снижает вычислительные затраты особенно при больших значениях числа векторов s , которое в практических задачах достигает 200-400. При этом преобразование эталонных описаний к сжатому виду выполняется на предварительном этапе и не влияет на общее время

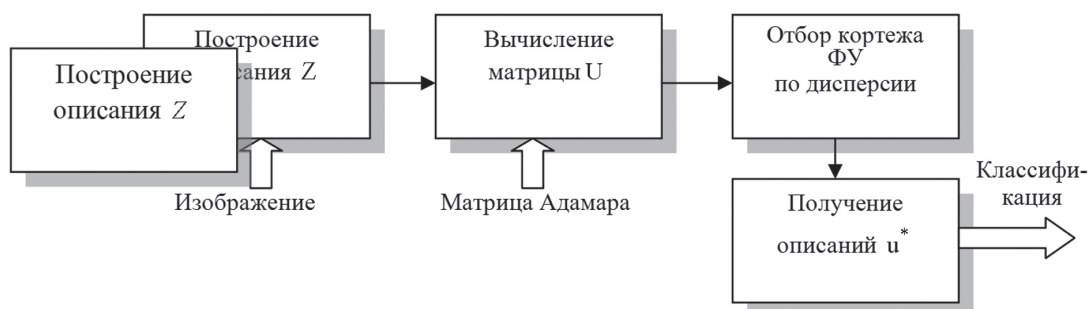


Рис. 1. Схема построения метода

классификации. Отбрасывание части коэффициентов разложения не вызывает существенных ошибок или потери нужных свойств. Размером объема описания (параметр m) можно управлять, повышая характеристику достоверности.

Выбор функций Уолша в качестве ортогонального базиса, кроме простоты аппаратной реализации, дополнительно обоснован их высокими фильтрационными свойствами относительно действия аддитивных флуктуационных помех, которые характерны для прикладных задач обработки изображений [5].

Важным частным случаем рассмотренного метода есть вариант $m = n$, т.е. все имеющиеся ФУ участвуют в классификации. Для этой ситуации сжатие описания возможно на основе анализа значимости элементов представления (3) путем отбрасывания малых или выделения m наибольших по величине среди u_j . Однако наши исследования показывают, что энергия представления в базисе ФУ для векторов ХП на практике распределяется очень неравномерно по спектру. Три ФУ (с номерами 1, 3, 49) в значительной мере доминируют над остальными, что приводит к классификации подавляющего числа ХП к этим трем классам. За счет этого фактора происходит чрезмерное сжатие информации, и в результате ухудшаются показатели качества классификации, особенно при действии помех. Такая картина наблюдается для нескольких исследуемых баз видеоданных, что говорит об объективности отмеченных спектральных свойств и полученных выводов.

2. Классификация на основе сжатого информативного описания

Будем использовать сжатое описание (3) в задаче классификации путем поиска эталона с наибольшим значением меры подобия по правилу [8]:

$$c(u) = \arg \underset{q \in \{1, \dots, K_Q\}}{\text{opt}} \vartheta[u, u(q)], \quad (4)$$

где $c(u)$ – функция классификации описания u ; $\vartheta[u, u(q)]$ – мера подобия между векторами в пространстве R^m ; $u(q)$ – вектор описания эталона с номером q ; K_Q – множество классов. При этом в процессе классификации описание объекта строится с учетом ФУ, синтезированных для видео-базы, что обеспечивает дополнительные согласование и помехозащищенность. В результате представления с учетом сжатия формируется множество K_p классов признаков, которое включает m классов и основано на выделенном кортеже ФУ.

За счет представления координатного векторного пространства ХП в ортогональном базисе ФУ получаем новое векторное пространство, переходя от представления объекта в виде множества векторов к одному вектору, отражающему спектральное представление для наиболее значимых (информативных)

элементов базиса. Теперь классификация сводится к представлению объекта в виде вектора u з заданным для базы множеством классов K_p .

Мера подобия ϑ в (4) может быть вычислена либо путем голосования, либо путем векторного представления объекта и сопоставления описаний как векторов. Понятно, что описание может при этом содержать элементы помех и ложных фоновых объектов, что приводит к искажениям вектора (3). В частности, можно использовать метрику в пространстве мультимножеств:

$$\rho(A, B) = \frac{\sum_i w_i |u_A(i) - u_B(i)|}{\sum_i w_i \max[u_A(i), u_B(i)]}, \quad (5)$$

где A, B – сопоставляемые векторы; $u_A(i), u_B(i)$ – значения компонент векторов с номером i ; w_i – весовые коэффициенты.

Кроме свойств различимости и помехозащищенности, важным моментом остается оценка выигрыша в быстродействии предложенной обработки по сравнению с традиционным голосованием. Проведем ее путем подсчета числа вычислительных операций. Обозначим t_E – время вычисления квадрата евклидова расстояния или скалярного произведения между векторами ХП размером n , t_c – время сравнения и суммирования двух чисел при поиске минимума расстояния и инкрементации числа голосов, s, s_1 – мощности описаний эталона и объекта. Тогда оценка времени при вычислении подобия в виде числа голосов выглядит как $T_v = s s_1 (t_E + t_c)$. Время для вычисления подобия в соответствии с обработкой (1)-(3) можно оценить как $T_m = s_1 [m(t_E + t_c) + t_c] + (t_E m) / n$, где выражение $(t_E m) / n$ соответствует вычислению подобия между векторами размером m . Отношение оценок имеет вид

$$\alpha = \frac{T_v}{T_m} = \frac{s s_1 (t_E + t_c)}{s_1 [m(t_E + t_c) + t_c] + (t_E m) / n}. \quad (6)$$

Детальный анализ (6) показывает, что в наибольшей степени величина α зависит от соотношения s / m , т.е. пропорционально увеличению s и уменьшается с увеличением m . Отсутствие зависимости от n можно объяснить тем, что оба варианта включают анализ всего описания классифицируемого объекта, включающего s_1 векторов одинаковой размерности n .

Конкретно при $m = 8$, $s = 100$ значение $\alpha \approx 12,5$, а при $s = 200$ – $\alpha \approx 25$. Как видим, преимущества предложенного метода в быстродействии для практических значений параметров могут достигать 10-20 раз.

3. Результаты компьютерного моделирования

Эксперименты проводились для базы из 12-ти эталонных изображений домашних животных (кошки) [8]. Предварительная обработка базы состояла в следующем. Для каждого эталона с

использованием метода SURF получено структурное описание как набор векторов размерностью $n=64$, при этом число векторов для эталонов базы колеблется от 66 до 262. Для перехода к спектральному пространству Уолша вычислены скалярное произведение каждого вектора с множеством ФУ и матрица U^2 для всей базы. Значения элементов u_{ij}^2 есть базовыми данными для проведения классификации.

Основой для применения метода сжатия описания является предположение о неравнозначности его элементов, а также предположение о том, что для осуществления классификации с высоким уровнем помехозащищенности достаточно ограниченного числа компонентов в базисе ФУ. Для отбора информативного кортежа ФУ размером $m=8$ элементов применен дисперсионный критерий. Наибольшие 16 значений дисперсии квадратов откликов представлены в табл. 1.

Как видим из табл. 1, функции 1, 3, 49 дают наиболее значимые отклики, которые существенно выделяются среди откликов остальных ФУ. Использование при классификации этих трех ФУ, как показало моделирование, приводит к тому, что подавляющее число элементов описания будет отнесено к одному из этих трех классов, что в результате значительно ухудшает показатели различимости объектов.

По этой причине исключим из рассмотрения ФУ с номерами 1, 3, 49 и выберем из табл. 1 в качестве информативного сокращенного базиса следующие восемь ФУ, на базе которых построим классификацию.

Выбранные 8 ФУ позволяют классифицировать векторы описания Z на классы и, как результат, получить для каждого эталона новое описание в виде вектора – кортежа из 8 коэффициентов по правилу (2). В результате получим новое информативное описание в виде векторов, содержащих количества ХП исходного описания, отнесенных к сформированному классу с использованием информативных ФУ.

Синтезированные описания представлены в табл.2. Для обеспечения сравнимости описаний из разных классов выполнена нормировка значений табл. 2 на число векторов в описании изображения/эталона в соответствии с выражением $u_i = u_i / s$.

Далее экспериментально исследовалась классификация объектов с использованием компактных описаний при $m=8$, в результате получена оценка вероятности правильной классификации в зависимости от уровня аддитивных помех.

В качестве критерия близости преобразованных описаний использована мера (5).

Уровень действия аддитивной помехи, действующей на компоненты векторов, характеризовался дисперсией σ^2 , которая напрямую отражает соотношение сигнал-шум. Описания подвергались действию помехи, после чего проводилось сравнение с эталонами по минимуму (5).

На вход системы классификации поступало описание Z^* , подверженное действию помехи в виде $Z^* = Z + \xi$, где ξ – матрица шума с независимыми нормально распределенными элементами

Таблица 1

Наибольшие 16 значений дисперсии для базы домашних животных

№ п\п	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
№ФУ	3	1	49	33	2	50	4	35	34	36	17	51	13	11	52	19
дисперсия	66,26	51,85	13,7	8,63	7,37	6,82	5,48	5,35	3,18	3,048	2,729	2,207	2,102	1,94	1,699	1,614

Таблица 2

Состав эталонных классов по числу классифицированных ХП

Номер класса ХП/ номер ФУ								Число векторов эталона	Класс эталона
1/2	2/4	3/17	4/33	5/34	6/35	7/36	8/50		
25	11	8	35	13	23	13	18	146	1
3	8	1	14	6	27	1	6	66	2
56	41	8	57	16	57	17	10	262	3
21	2	5	38	12	18	3	18	117	4
35	27	2	41	12	31	15	15	178	5
33	19	9	43	13	35	15	12	179	6
43	31	6	52	9	31	15	9	196	7
31	11	10	36	16	20	8	14	146	8
40	15	4	29	9	25	1	5	128	9
30	12	11	37	12	13	13	24	152	10
46	25	6	42	11	41	13	13	197	11
27	11	2	13	7	9	9	11	89	12

(матожидание – 0, дисперсия – σ^2). Вероятность правильной классификации вычисляется как отношение числа правильных ответов к общему числу экспериментов. В результате 100 экспериментов получены зависимости вероятности правильной классификации от уровня помехи σ^2 (рис. 2).

Проведен сравнительный эксперимент по классификации для базы домашних животных, в котором описание представлено тремя вариантами: в пространстве 8 отобранных по дисперсионному критерию ФУ (с отбрасыванием ФУ 1, 3, 49); в пространстве 8 ФУ с наименьшими дисперсиями; в пространстве полного базиса Уолша (64 ФУ). Для сопоставления проведены также эксперименты для метода голосования путем подсчета числа голосов, порог эквивалентности двух ХП выбран равным 0,16 (1% от максимума значения евклидовой метрики).

Как видим и рис. 2, предложенный в статье метод построения информативного описания имеет несомненное преимущество в плане помехозащищенности перед традиционным голосованием. До уровня $\sigma^2 = 0,02$ эффективность методов примерно одинакова (вероятность правильного распознавания выше 0,94), далее с увеличением σ^2 кривая вероятности для разработанного метода выше, и при $\sigma^2 = 0,03$ вероятности составляют 0,8 и 0,6 соответственно. Высокая устойчивость к шуму обусловлена прежде всего интегральными свойствами ФУ. В то же время и быстродействие разработанного метода существенно выше: сравнительное время классификации одного изображения составило 0,03 сек и 0,95 сек соответственно, т.е. более чем в 30 раз лучше (компьютер Intel Core2 Duo 2,67 ГГц).

Другие модификации предложенного метода, как видно из рис. 2, имеют несколько худшие показатели помехозащищенности: если для предложенного варианта при $\sigma^2 = 0,01$ вероятность равна 0,99, то для методов в пространстве 8 ФУ с наименьшими значениями дисперсии и в пространстве

полного базиса Уолша из 64 ФУ вероятности составили 0,53, что неприемлемо на практике.

Выводы

Предложенный в работе метод построения информативных компактных описаний на основе анализа свойств структурного представления в ортогональном базисе Уолша позволяет проводить классификацию объектов с высоким быстродействием без снижения уровня помехозащищенности. Проведенные вычислительные эксперименты подтвердили предположение о возможности существенного сокращения описания путем отбрасывания малых по величине коэффициентов преобразования.

При сравнении уровня помехозащищенности с традиционным голосованием установлено, что метод построения компактных описаний в пространстве кортежа выделенных ФУ не уступает по помехозащищенности голосованию, что обусловлено интегральными свойствами базиса Уолша. Основным достоинством предложенного подхода является значительный выигрыш в быстродействии: теоретически – в 10-20 раз, практически – в 30 раз по сравнению с голосованием.

В результате исследования разработан новый метод построения сжатых информативных описаний изображений на основе отбора наиболее значимых компонентов спектра преобразования Уолша, обеспечивающий высокое быстродействие в сравнении с традиционным голосованием без снижения показателя помехозащищенности.

Практическая ценность работы состоит в экспериментальном подтверждении эффективности синтезированного метода на прикладной базе видео-данных домашних животных.

Перспективное направление дальнейших исследований состоит в разработке адаптивной процедуры выбора функций Уолша на основе исследования свойств эталонов базы.



Рис. 2. Зависимость качества классификации от уровня помехи

Список литературы: 1. *Lowe D.* Distinctive Image Features from Scale-Invariant Keypoints/ D. Lowe // In International Journal of Computer Vision. – 2004. – vol. 60. – P. 91–110. 2. *Шапиро Л.* Компьютерное зрение [Текст] / Л. Шапиро, Дж. Стокман: пер. с англ.— М.: БИНОМ. Лаборатория знаний, 2006. – 752 с. 3. *Ахмед Н.* Ортогональные преобразования при обработке цифровых сигналов [Текст] / Н. Ахмед, К.Р. Рао: пер. с англ. – М.: Связь, 1980. – 248 с. 4. *Залманзон Л. А.* Преобразование Фурье, Уолша, Хаара и их применение в управлении, связи и других областях [Текст] / Л. А. Залманзон. – М.: Наука, 1989. – 496 с. 5. *Ересько Ю.Н.* Локализация изображений в автоматических визирах / Ю.Н. Ересько. – М.: Компания Спутник+, 2002. – 357 с. 6. *Pratt W. K.* Digital Image Processing / W. K. Pratt. – New York: John Wiley and Sons Inc., 2001. – 723 p. 7. Прикладная статистика: Классификация и снижение размерности: Справ. изд. / С.А. Айвазян, В.М. Бухштабер, И.С. Енюков, Л.Д. Мешалкин; под ред. С.А. Айвазяна. – М.: Финансы и статистика, 1989. – 607 с. 8. *Власенко Н.В.* Анализ кортежей наиболее значимых соответствий между описаниями объектов в задачах метрического распознавания [Текст] / Н.В. Власенко, В.П. Машталир // Реестрация, зберігання і обробка даних.— 2012. – Т. 14, № 3. – С. 20-32.

Поступила в редколлегию 16.01.2013

УДК 004.932.2:004.93'1

Побудова інформативних компактних описів та класифікація об'єктів шляхом представлення в ортогональному базисі / Н.В. Власенко // Біоніка інтелекту: наук.-техн. журнал. – 2013. – № 1 (80). – С. 93–98.

У роботі запропоновано ефективний метод класифікації зображень на основі побудови стиснених описів із використанням найбільш інформативних компонентів спектру Уолша. Представлено результати експериментального порівняння синтезованого методу з відомими підходами при використанні їх для обробки реальних баз відео-інформації.

Табл. 2. Лл. 2. Бібліогр.: 8 назв.

UDC 004.932.2:004.93'1

Building a compact informative descriptions and classification of objects by presenting in an orthogonal basis / N.V. Vlasenko // Bionics of Intelligence: Sci. Mag. – 2013. – № 1 (80). – P. 93–98.

In this paper we propose an effective method for image classification based on constructing a succinct description of using the most informative parts of the spectrum Walsh. The results of an experimental comparison of the synthesized method with known approaches when they are used to process live video data bases are presented.

Fig. 2. Tab. 2 Ref.: 8 items