

УДК 004.9

В.А. Ковалев, И.В. Сафонов**ЭКСПЕРИМЕНТАЛЬНОЕ ИССЛЕДОВАНИЕ
КАЧЕСТВА РАСПОЗНАВАНИЯ ИЗОБРАЖЕНИЙ СЦЕН
В ЗАДАЧАХ МАШИННОГО ЗРЕНИЯ**

Проводится исследование качества распознавания цветных изображений сцен для дальнейшего использования наиболее эффективных приемов и методов в разрабатываемом программном комплексе Mobile Voiced Visual Assistant. Указанное ПО предназначается для помощи людям с нарушениями функций зрения. В ходе экспериментов для каждого из исследуемых методов определяются оптимальные параметры, применение которых позволяет добиться наилучшего качества в распознавании объектов реального мира.

Введение

Основная цель данного исследования – поиск оптимального метода распознавания и его параметров для дальнейшего использования в разрабатываемом авторами программном комплексе Mobile Voiced Visual Assistant (MVVA). Указанное ПО разрабатывается для применения на мобильных устройствах с целью обеспечения более уверенной ориентации в пространстве (в помещении или на улице) людей с нарушением функций зрения. В данной статье подразумевается использование версии MVVA, которая позволяет оценивать степень близости (похожести) предмета, находящегося в поле зрения камеры, к некоторому классу объектов и выражать ее в виде звукового сигнала соответствующей высоты в реальном масштабе времени. При этом применение какого-либо специального оборудования, кроме собственно мобильного вычислительного устройства, не предполагается.

В перспективе разрабатываемый программный продукт должен работать как приложение на любых портативных устройствах с камерой, к которым в первую очередь относятся мобильные телефоны. Отсюда вытекает ряд требований к алгоритмам и программному обеспечению, которое их реализует. Во-первых, приложение должно быть ориентировано на работу с изображениями низкого или среднего качества. Во-вторых, необходимо обеспечить максимальную устойчивость алгоритма к изменению условий съемки. К числу требований можно также отнести возможность работы алгоритма в реальном времени в условиях ограниченных аппаратных возможностей.

Известно, что в работе некоторых алгоритмов классификации и распознавания большую роль играет подбор оптимальных параметров. От их выбора сильно зависит качество распознавания изображений сцен. Поэтому важным этапом анализа эффективности того или иного метода является оценка его параметров. С учетом всех указанных выше факторов авторами разработан соответствующий план исследований и проведены эксперименты, позволяющие решить указанную задачу.

В статье представлены результаты вычислительных экспериментов с целью поиска эффективного классификатора. Рассматриваются оба случая: распознавание при заранее зафиксированном разбиении множества изображений на обучающую и тестовую выборки, а также формирование выборок случайным образом с последующим определением точности распознавания путем усреднения результатов 300 независимых тестов.

1. Инструментарий исследования и исходные данные

В качестве среды разработки и инструмента анализа данных был использован пакет статистических вычислений и графики R [1]. Исходные данные представляли собой кадры видеопоследовательностей – 24-битные изображения различных предметов размером 320×240 пикселей (рис. 1), полученные при помощи веб-камеры Logitech QuickCam Pro 9000. Эти изображения были разбиты на восемь классов, условно названных «портфель», «кресло», «кружка», «дверной проем», «цветок», «шкаф», «телефон» и «окно». Всего рассматривалось 73 кадра.



Рис. 1. Примеры изображений – представителей всех восьми классов

Дескрипторы изображений представляют собой вытянутые в вектор матрицы совместной встречаемости [2, 3]. Поскольку для 24-битных изображений использование таких дескрипторов в рамках нашей задачи представлялось нецелесообразным, матрицы совместной встречаемости строились не по оригинальным изображениям, а по изображениям, полученным из исходных путем уменьшения цветового пространства снимка до 8 бит (рис. 2). Эта операция резко сокращает объем памяти, необходимый для хранения матрицы совместной встречаемости, и значительно увеличивает скорость обработки кадра.

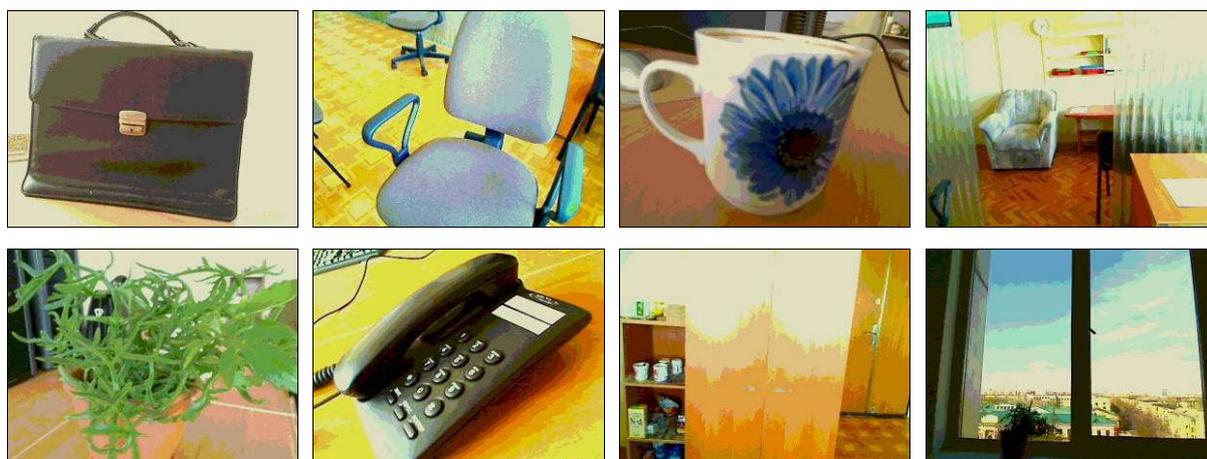


Рис. 2. 8-битные аналоги исходных изображений

В подходе с фиксированной обучающей выборкой изображения каждого класса были предварительно разделены на два подмножества. Первое подмножество содержало три характерных представителя данного класса и использовалось в дальнейшем в качестве обучающего. Снимки второго подмножества подавались на вход классификатора для распознавания. Использование выборок малых размеров было обусловлено спецификой решаемой задачи.

В случае вычисления среднего значения качества распознавания по результатам 300 независимых тестов список изображений класса перемешивался случайным образом, и элементы из первой половины списка определяли обучающую выборку, элементы из второй половины – тестовую. На заданных таким образом выборках определялась эффективность классификатора. На следующем шаге эти действия повторялись. Итоговое значение эффективности классификатора вычислялось как среднее по 300 итерациям.

Таким образом, все дальнейшие манипуляции проводились в пакете статистических вычислений R с 73 файлами, содержащими предварительно подсчитанные матрицы совместной встречаемости 8-битных изображений. Задачей классификатора являлось определение принад-

лежности тестового изображения к одному из восьми известных классов. Для подхода с фиксированным набором элементов в тестовой выборке обучение и распознавание на заданных параметрах проводились единожды. При случайном отборе изображений в тестовую и обучающую выборки на заданных параметрах проводились 300 последовательных итераций обучения и распознавания с переопределением элементов в выборках на каждом шаге.

Собственно распознавание проводилось с использованием двух различных методов:

- опорных векторов (SVM);
- ближайшего соседа (расстояние до центроида класса).

Метод опорных векторов рассмотрен как один из лучших методов классификации, существующих в настоящее время. Второй из представленных методов принят во внимание как не требующий больших вычислительных ресурсов, что хорошо подходит для вычислений на не слишком мощных процессорах мобильных устройств.

2. Этапы исследования

Будем рассматривать четыре этапа исследования:

- подбор оптимальных значений управляющих параметров в методе опорных векторов;
- распознавание через суммарное взвешенное расстояние до объектов класса;
- изменение типа дескрипторов (распознавание на дескрипторах, полученных совмещением дескриптора исходного изображения и дескриптора прореженного вдвое изображения);
- предварительная обработка дескрипторов цветных изображений (удаление значений по порогу, применение метода главных компонент).

Большинство из этих этапов представляют собой вычислительно емкие задачи. Для уменьшения пространства признаков и, следовательно, уменьшения времени работы алгоритмов проведена несложная предварительная обработка дескрипторов изображений. Те позиции, на которых находились нулевые элементы в каждом дескрипторе, из дальнейшего анализа исключались. Такой прием позволил намного уменьшить количество существенных признаков. Например, из потенциально возможного их количества 65 536 для исходных изображений после предварительной обработки дескрипторов оставалось не более 5000.

Далее подробно описаны только эксперименты для случая фиксированного распределения изображений между обучающей и тестовой выборками. Сущность экспериментов при подходе с многократным перемешиванием элементов класса не меняется, разница заключается только в количестве итераций. Поэтому для данного подхода приведены только результаты и комментарии к ним там, где это необходимо.

2.1. Подбор оптимальных значений управляющих параметров в методе опорных векторов

Метод опорных векторов [4] – это набор эффективных алгоритмов, широко используемых для задач классификации. В нашем случае был выбран метод опорных векторов с ядром, основанным на радиальных базисных функциях:

$$K(x, y) = e^{-\gamma \|x-y\|^2}, \quad \gamma > 0.$$

Причины, по которым выбрано именно это ядро, подробно рассматриваются в работах [4, 5].

При использовании ядра, основанного на радиальных базисных функциях, важно правильно подобрать значения параметров C и γ . Здесь C – управляющий параметр метода, который позволяет находить компромисс между минимизацией суммарной ошибки и максимизацией разделяющей полосы при определении классов. Предсказать наиболее подходящие значения для этих параметров заранее невозможно, для различных наборов объектов они будут различаться. Поэтому одним из этапов анализа являлась оценка качества распознавания сцен на тестовых задачах для широкого диапазона всех сочетаний управляющих параметров C и γ . Сначала проводятся вычисления по сетке на большом диапазоне значений, потом выбирается область с наибольшей эффективностью и для этой области вычисления повторяются, но уже с меньшим шагом.

В качестве начального диапазона были выбраны следующие значения: $C = 2^{-5}, 2^{-4}, \dots, 2^{15}$, $\gamma = 2^{-100}, 2^{-99}, \dots, 2^3$. На этом диапазоне получаем зависимость эффективности распознавания от значений управляющих параметров. На рис. 3 эта зависимость представлена в виде изолиний одинакового значения эффективности.

Заметив, что область наибольшей эффективности находится в полосе значений $\gamma = 10^{-17} \dots 10^{-7}$, повторяем вычисления в этой полосе с уменьшенным шагом (рис. 4). По значениям, показанным на рис. 4, можно определить значения параметров C и γ , на которых достигается максимальная эффективность распознавания методом опорных векторов (65,3 %) на заданном наборе данных. В качестве таких значений можно, например, выбрать $C = 128$, $\gamma = 10^{-10}$.

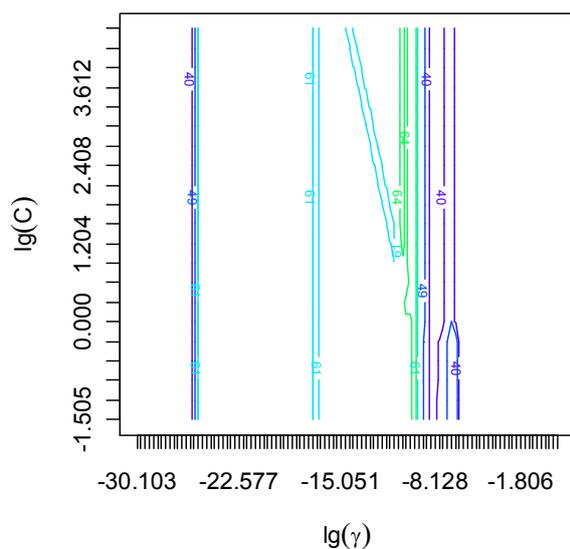


Рис. 3. Эффективность распознавания для начального диапазона управляющих параметров

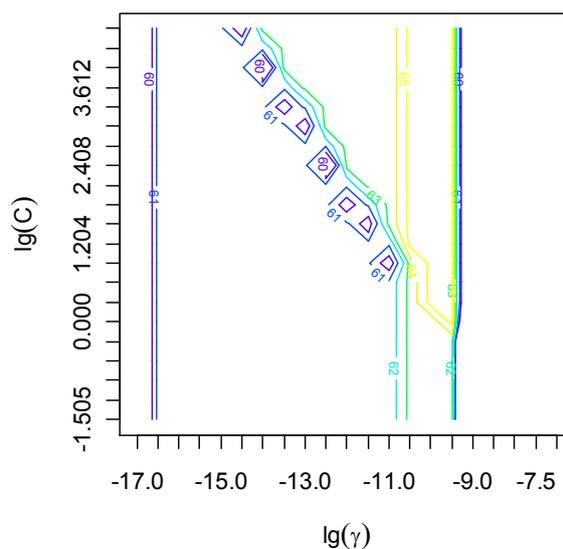


Рис. 4. Эффективность распознавания на уточненном интервале

2.2. Изменение типа матриц совместной встречаемости и их параметров

Эксперименты по распознаванию на первом этапе проводились для матриц совместной встречаемости цветов соседних пикселей. Разумно сделать предположение о том, что если в такие дескрипторы добавить информацию о соседстве цветов пикселей, лежащих на большем расстоянии друг от друга, то качество распознавания улучшится. Авторами была рассмотрена методика, состоявшая в добавлении к дескриптору изображения из первого этапа дескриптора прореженного вдвое изображения. Прореживание заключалось в удалении четных строк и четных столбцов оригинального изображения (рис. 5).



Рис. 5. Оригинальное изображение и его прореженная в два раза копия

Таким образом, если $\|U_{ij}\|$, $i = \overline{1, n}$, $j = \overline{1, m}$, – матрица совместной встречаемости исходного изображения, то его дескриптором будет вектор $u_{image} = (u_{11}, u_{12}, \dots, u_{1m}, u_{21}, \dots, u_{2m}, \dots, u_{n1}, \dots, u_{nm})$. Аналогично для второго типа изображений: если $\|V_{ij}\|$, $i = \overline{1, k}$, $j = \overline{1, l}$, – матрица совместной встречаемости прореженного в два раза изображения, то в качестве его дескриптора нами принимался вектор

$$v_{image} = (v_{11}, v_{12}, \dots, v_{1l}, v_{21}, \dots, v_{2l}, \dots, v_{k1}, \dots, v_{kl}), \text{ где } k = \left\lceil \frac{n}{2} \right\rceil, \quad l = \left\lceil \frac{m}{2} \right\rceil.$$

Обучение классификатора и распознавание осуществлялись на нормализованных векторах вида $w = u_{image} \cup v_{image}$. Нормализация проводилась по общему количеству пар первого (305 522 пары) и второго (75 962 пары) изображений. Такая нормализация необходима для придания одинакового веса признакам исходного и уменьшенного в четыре раза изображения.

Как и в предыдущем случае, здесь тоже предварительно удалялись общие нулевые для каждого изображения элементы. Эксперименты показали, что для начального набора из 131 072 ($2 \times 65\,536$) элементов удаление общих незначимых признаков позволяет сократить размерность дескриптора до 7116 позиций, на которых хотя бы в одном дескрипторе находится ненулевой элемент.

Из обработанных таким образом дескрипторов был сформирован второй набор данных. На нем проводилась та же операция подбора управляющих параметров метода опорных векторов, что и на первом этапе. Как и ожидалось, максимальная эффективность распознавания повысилась и составила 69,4 %.

2.3. Предварительная обработка дескрипторов цветных изображений

Частично предварительная обработка дескрипторов уже проводилась на каждом из предыдущих этапов и заключалась в удалении общих нулевых элементов. Такая предобработка позволила убрать незначимые признаки из дальнейшего анализа и намного увеличить скорость обработки данных.

Кроме того, на любом изображении, полученном с цифровой камеры, присутствует шум, который не несет полезной информации и от которого нужно по возможности избавляться. Для этого нами была дополнительно применена процедура отбора элементов дескрипторов по порогу. Порог определялся в процентном отношении от общего количества пар соседних цветов в изображении. Нами были выбраны следующие процентные значения порогов: 0,1; 0,2; 0,5 %. Для исходных изображений с общим количеством пар 305 522 пороги в численном выражении составили 306, 611 и 1528 пар цветов. Для прореженных вдвое изображений с общим количеством пар 75 962 – 76, 152 и 380 пар соответственно. Пример работы фильтрации по порогу 20 показан на рис. 6. Разумеется, количество признаков после такой операции существенно уменьшится.

<i>img1</i>	24	...	60	12	6	657
<i>img2</i>	340	...	2	3	5	210
<i>img3</i>	253	...	0	18	10	148
<i>img4</i>	0	...	0	0	2	236
<i>img5</i>	0	...	0	0	8	251

→

<i>img1</i>	24	...	60	657
<i>img2</i>	340	...	2	210
<i>img3</i>	253	...	0	148
<i>img4</i>	0	...	0	236
<i>img5</i>	0	...	0	251

Рис. 6. Удаление признаков по порогу $T=20$

В табл. 1 указано количество признаков после фильтрации по различным порогам.

Таблица 1

Количество признаков после фильтрации по порогам

Дескрипторы	Начальное количество признаков	Пороги, %		
		0,1	0,2	0,5
исходного изображения	256 ²	563	423	245
прореженного изображения	256 ²	323	229	155
совмещенный	2×256 ²	886	652	400

После удаления признаков по порогу над преобразованными дескрипторами были проведены эксперименты для определения качества распознавания (рис. 7). Из рисунка видно, что наибольшая эффективность распознавания методом опорных векторов достигается на дескрипторах исходного изображения при фильтрации по порогу 0,5 %, что соответствует анализу всего 245 признаков (см. табл. 1). Совмещение двух типов дескрипторов прироста в качестве распознавания для данного набора данных не дает. Принимая во внимание еще и тот факт, что для построения и дальнейшего использования совмещенного дескриптора требуются дополнительные ресурсы компьютера (время на прореживание изображения, дополнительная память для хранения дескриптора такого изображения), в данном случае можно сделать вывод о нецелесообразности такой операции.

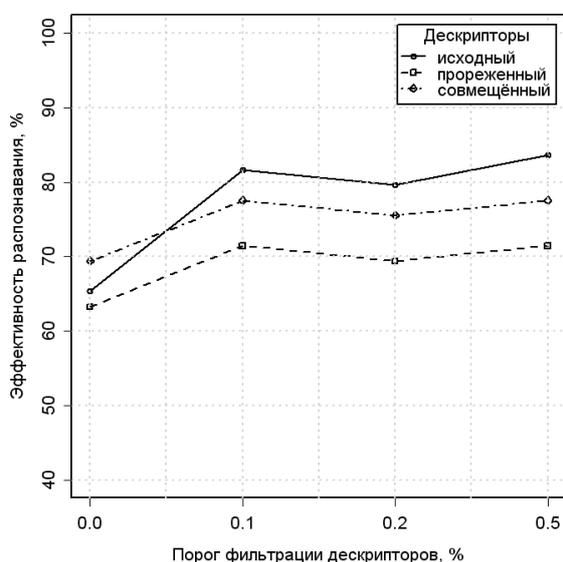


Рис. 7. Качество распознавания при различных порогах

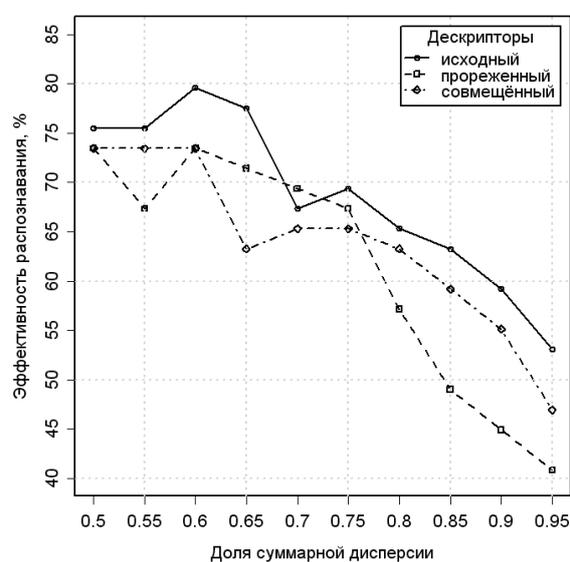


Рис. 8. Зависимость качества распознавания от доли суммарной дисперсии при фиксированной тестовой выборке

Другой возможный подход к предварительной обработке матриц выборки заключается в том, чтобы подвергнуть исходные дескрипторы функциональному преобразованию (с целью отбора информативных признаков), при котором размерность исходной задачи уменьшается и гарантируется линейная независимость новых признаков. Одним из таких способов является метод главных компонент [6, 7], широко применяемый в области распознавания образов и компьютерного зрения.

Анализируя изменение относительной доли дисперсии, вносимой первыми n главными компонентами, в зависимости от числа этих компонент, можно определить число компонент, которое целесообразно оставить в рассмотрении. Иногда в научной литературе принято сокращать размерность пространства как минимум до тех пор, пока добавление остальных главных компонент не будет повышать суммарную характеристику рассеяния более чем на 10 %.

Авторами проведен эксперимент по оценке качества распознавания в зависимости от количества анализируемых главных компонент. Для трех наборов данных, соответствующих различным дескрипторам, последовательно бралось количество главных компонент с относительной долей суммарной дисперсии от 0,5 до 0,95 с шагом 0,05. Выделенные главные компоненты далее использовались для распознавания методом опорных векторов.

Рассматривая рис. 8, можно сделать вывод о том, что наилучшее качество распознавания достигается для дескрипторов исходного изображения при относительной доле суммарной дисперсии, равной 0,6. Для дескриптора исходного изображения, на котором достигается максимальная эффективность 79,59 %, такая доля суммарной дисперсии соответствует всего лишь 14 главным компонентам (признакам).

2.4. Распознавание через суммарное взвешенное расстояние до объектов класса

Следующий этап исследования – распознавание по минимальному взвешенному расстоянию до объекта класса.

Для распознаваемого объекта вычислялось взвешенное суммарное расстояние до всех изображений каждого из восьми классов. По минимальному из полученных восьми расстояний делались выводы о принадлежности анализируемого снимка к тому или иному классу. Для вычисления расстояния между дескрипторами брались четыре различные метрики: евклидова, Чебышева, манхэттенская и Канберра. Для векторов $X = (x_1, x_2, \dots, x_n)$ и $Y = (y_1, y_2, \dots, y_n)$ эти метрики определяются следующим образом:

$$- \text{евклидова: } d_{euclidean}(X, Y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2};$$

$$- \text{Чебышева: } d_{chebyshev}(X, Y) = \max_i |x_i - y_i|;$$

$$- \text{манхэттенская: } d_{manhattan}(X, Y) = \sum_{i=1}^n |x_i - y_i|;$$

$$- \text{Канберра: } d_{canberra}(X, Y) = \sum_{i=1}^n \frac{|x_i - y_i|}{x_i + y_i}.$$

Известно, что метрика Канберра очень чувствительна к малым отклонениям в окрестности нуля, а также к шуму в исходных данных. В нашем случае эта особенность не оказывает негативного влияния на результаты эксперимента в силу самого вида дескрипторов. Кроме того, на большинстве этапов исследования проводится дополнительная фильтрация признаков по порогу, что приводит к еще более значительному снижению шума в дескрипторах.

Результаты эксперимента приведены в табл. 2.

Таблица 2

Эффективность распознавания для различных дескрипторов изображения и различных метрик, %

Условия эксперимента	Метрики											
	евклидова			Чебышева			манхэттенская			Канберра		
Дескриптор исходного изображения	57,14			51,02			71,43			73,47		
Дескриптор прореженного вдвое изображения	57,14			32,65			69,39			81,63		
Совмещенный нормированный дескриптор	61,22			42,86			73,47			85,71		
Исходное изображение с удалением признаков по порогам (0,1; 0,2; 0,5 %)	57,14	57,14	57,14	51,02	51,02	51,02	71,43	71,43	71,43	65,31	63,27	61,22

Прореженное изображение с удалением признаков по порогам (0,1; 0,2; 0,5 %)	57,14	57,14	57,14	32,65	32,65	32,65	69,39	71,43	71,43	77,55	73,47	77,55
Совмещенный нормированный дескриптор с удалением признаков по порогам (0,1; 0,2; 0,5 %)	61,22	61,22	59,18	42,86	42,86	42,86	73,47	73,47	73,47	75,51	71,43	71,43

Максимальная эффективность распознавания в этом случае составила 85,71 %, и достигалась она для совмещенных нормированных дескрипторов с метрикой Канберра. Использование этой метрики оправдано и для других типов дескрипторов. Только в одном случае, когда использовался дескриптор исходного изображения с удалением признаков по порогу, применение манхэттенской метрики принесло лучшие результаты, чем вычисление в метрике Канберра.

3. Вычисление средних значений эффективности

Как уже упоминалось выше, в данном случае при проведении экспериментов в последовательности этапов ничего принципиально не меняется. Этот подход отличается от уже изложенного только тем, что на заданных параметрах тренировка и обучение классификаторов проводились не один раз, а 300, каждый раз с новым независимым распределением изображений между обучающей и тестовой выборками. Поэтому в данном разделе приводятся только результаты вычислений с небольшими комментариями.

На этапе подбора оптимальных значений управляющих параметров в методе опорных векторов максимальная эффективность распознавания составила 84,89 % при $C = 2^3$, $\gamma = 2^{-33}$. В отличие от предыдущей серии экспериментов здесь применение метода опорных векторов к совмещенному дескриптору не привело к повышению эффективности распознавания. Максимальное ее значение равнялось 84,67 %.

Основываясь на данных рис. 9, можно сделать вывод о том, что по сравнению со случаем фиксированного задания обучающей выборки (см. рис. 7) эффективность распознавания повысилась и составила значение, близкое к 90 % (фактическая величина – 88,69 %). Однако общая тенденция сохраняется: изменение типа дескриптора не дает ожидаемого прироста в качестве распознавания, хуже всего распознавание происходит на дескрипторах прореженных вдвое изображений, при совмещении дескрипторов происходит небольшое улучшение в распознавании, но самым эффективным остается применение дескриптора исходного изображения с прореживанием по порогу 0,5 %.

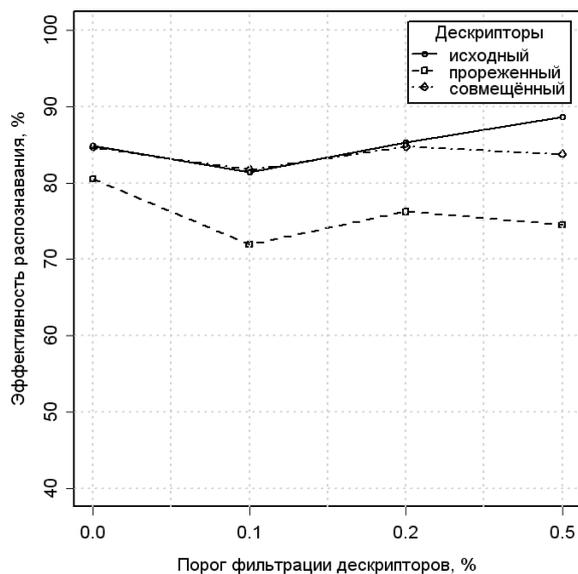


Рис. 9. Качество распознавания для редуцированных дескрипторов при вычислении эффективности в среднем

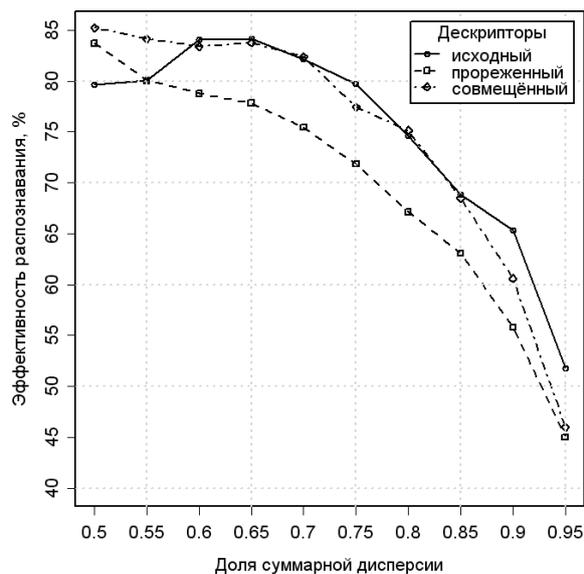


Рис. 10. Зависимость качества распознавания от доли суммарной дисперсии при вычислении среднего значения эффективности

На рис. 10 показаны результаты работы классификатора SVM после применения к матрицам совместной встречаемости метода главных компонент и использования полученных компонент (10–50 шт.) в качестве признаков. Неожиданным результатом является то, что максимальная эффективность распознавания 85,23 % достигается на совмещенном дескрипторе при суммарной доле дисперсии 0,5. Эта доля суммарной дисперсии соответствует анализу лишь 12 главных компонент (признаков). При дальнейшем же увеличении количества анализируемых главных компонент качество распознавания ухудшается. Вероятно, этот эффект можно объяснить тем, что после подобного редуцирования начального пространства признаков остаются только наиболее информативные. Признаки, отвечающие за второстепенные объекты, и другие не относящиеся к распознаваемому объекту детали изображения не рассматриваются.

Таблица 3

Эффективность распознавания для различных дескрипторов изображения и различных метрик при вычислении эффективности в среднем, %

Условия эксперимента	Метрики											
	евклидова			Чебышева			манхэттенская			Канберра		
Дескриптор исходного изображения	59,15			55,94			63,14			67,64		
Дескриптор прореженного вдвое изображения	57,04			45,85			55,43			62,43		
Совмещенный нормированный дескриптор	59,64			49,53			58,94			65,20		
Исходное изображение с удалением признаков по порогу (0,1; 0,2; 0,5 %)	59,89	59,48	59,44	56,35	56,44	56,19	63,01	62,73	62,23	66,71	66,57	66,25
Прореженное изображение с удалением признаков по	56,55	56,16	57,03	45,63	45,11	45,93	55,05	54,77	55,41	61,22	60,78	61,26

порогу (0,1; 0,2; 0,5 %)												
Совмещенный нормированный дескриптор с удалением признаков по порогу (0,1; 0,2; 0,5 %)	59,36	60,04	60,24	48,72	49,81	49,58	58,23	58,72	58,31	63,86	64,07	63,79

Использование в качестве классификатора суммарного взвешенного расстояния до объектов класса становится неоправданным в случае, когда речь идет о средних значениях эффективности. Доказательством этому служат данные табл. 3, особенно если их рассматривать в сравнении с данными табл. 2. Хотя метрика Канберра остается для этого случая самой эффективной, общее падение эффективности дает основания для того, чтобы в дальнейшем не рассматривать такой классификатор.

Заключение

В статье подробно рассмотрены несколько подходов к распознаванию цветных изображений (кадров видеопоследовательностей) сцен реального мира, полученных с помощью веб-камеры среднего качества. Проведена оценка влияния различных факторов на качество распознавания сцен, рассмотрены некоторые из них.

Выполнено исследование эффективности распознавания для двух различных подходов к заданию тестовой и обучающей выборок. Для подхода с фиксированным набором элементов в тестовой выборке обучение и распознавание на заданных параметрах проводились один раз. В другом из рассмотренных подходов формирование выборок происходило случайным образом с последующим определением точности распознавания путем усреднения результатов 300 независимых тестов.

Анализ эффективности двух подходов позволил выявить некоторые закономерности. Так, в методе опорных векторов изменение типа дескриптора не дает ожидаемого прироста в качестве распознавания. Хуже всего распознавание происходит на дескрипторах изображений, прореженных вдвое. При совмещении дескрипторов происходит небольшое улучшение в распознавании, но самым эффективным остается использование дескриптора исходного изображения с прореживанием по порогу 0,5 %. Кроме того, применение метода опорных векторов для случая среднего значения эффективности позволяет добиться максимально возможного качества распознавания 88,69 %.

В случае, когда реализуется подход с фиксированным набором элементов в выборках, наибольшее качество распознавания (85,71 % для восьми возможных классов) достигается при распознавании через суммарное взвешенное расстояние до объектов класса с использованием метрики Канберра и совмещенного нормированного дескриптора, но резкое падение показателей эффективности этого классификатора при переходе к средним значениям эффективности говорит о том, что в общем случае распознавание с использованием расстояния до центра не является перспективным подходом. Такой классификатор сильно зависит от начальных условий, от того, какие изображения из класса будут выбраны пользователем в обучающую выборку.

Для дальнейшего более подробного изучения и реализации в Mobile Voiced Visual Assistant принят метод опорных векторов как самый эффективный из рассмотренных в этой работе.

Список литературы

1. R Development Core Team. R: A Language and Environment for Statistical Computing / R Foundation for Statistical Computing // The R project for Statistical Computing [Electronic resource]. – Vienna, Austria, 2008. – Mode of access : <http://www.r-project.org>. – Date of access : 02.02.2009.

2. Ковалев, В.А. Анализ структуры трехмерных медицинских изображений / В.А. Ковалев. – Минск : Белорусская наука, 2008. – 263 с.
3. Kovalev, V. Color Co-Occurrence Descriptors for Querying-by-Example / V. Kovalev, S. Volmer // Int. Conf. on Multimedia Modelling. – Lausanne, Switzerland : IEEE Comp. Society Press, 1998. – P. 32–38.
4. Vapnik, V.N. The nature of statistical learning theory / V.N. Vapnik. – N.Y. : Springer – Verlag, 1995. – 188 p.
5. Keerthi, S.S. Asymptotic behaviors of support vector machines with Gaussian kernel / S.S. Keerthi, C.-J. Lin // Neural Computation. – 2003. – № 15 (7). – P. 1667–1689.
6. Прикладная статистика. Классификация и снижение размерности / С.А. Айвазян [и др.] ; под ред. С.А. Айвазяна. – М. : Финансы и статистика, 1989. – 607 с.
7. Jolliffe, I.T. Principal Component Analysis / I.T. Jolliffe ; Series : Springer Series in Statistics. – 2nd ed. – N.Y. : Springer, 2002. – 502 p.

Поступила 20.03.09

*Объединенный институт проблем
информатики НАН Беларуси,
Минск, Сурганова, 6
e-mail: vassili.kovalev@gmail.com
igorsafonov@gmail.com*

V.A. Kovalev, I.V. Safonov

**AN EXPERIMENTAL STUDY
OF THE RECOGNITION QUALITY OF REAL-WORLD SCENES
FOR COMPUTER VISION PURPOSES**

This paper presents results of an experimental study of recognition quality of color images of real-world scenes. The goal was to assess the implemented recognition techniques and to tune control parameters of image frame recognition employed in so-called Mobile Voiced Visual Assistant, which is aimed to assist visually-impaired people in recognition and audio-interpretation of real-world scenes and surrounding objects in real time. As a result, for each recognition algorithm the optimal set of control parameters was identified and suitable modes were chosen for further implementation in the introduced Mobile Assistant. The use of these parameters allows achieving the highest recognition rates possible in the framework of actual limitations.