

Интеллектуальный анализ данных и распознавание образов

Метод распознавания объектов живописи в неконтролируемых условиях с обучением по одному примеру*

Н.С. СКОРЮКИНА^I, А.Н. МИЛОВЗОРОВ^{II}, Д.В. ПОЛЕВОЙ^{II,III,IV}, В.В. АРЛАЗАРОВ^{II,III,IV,V}

^I ООО «Смарт Энджинс Сервис», Москва, Россия

^{II} Национальный исследовательский технологический университет «МИСиС», Москва, Россия

^{III} Институт системного анализа Федерального исследовательского центра «Информатика и управление» Российской академии наук, Москва, Россия

^{IV} Московский физико-технический институт (ГУ), Долгопрудный, Россия

^V Институт проблем передачи информации им. А.А. Харкевича РАН, Москва, Россия

Аннотация. В статье рассмотрена задача идентификации картин на фотографиях, снятых в условиях музейной экспозиции при помощи мобильных устройств. Описан подход к решению, основанный на использовании созвездия особых точек и их дескрипторов в качестве компактного описания изображения. Приведена двухэтапная схема для определения лучшего соответствия изображению-запросу среди изображений-эталонов. На первом этапе сопоставление производится в модели «мешок признаков», на втором учитывается взаимное расположение точек. Метод предусматривает отказ в случае, когда соответствие не может быть установлено. Предложен способ геометрической нормализации изображения-запроса для достижения большей устойчивости алгоритма к кадрированию и смене ракурса. Рассмотрены преимущества подхода по сравнению с альтернативными современными решениями. Приведены результаты работы на открытом наборе данных WikiArt и оценки скорости для различных вычислительных платформ.

Ключевые слова: *распознавание живописи, особые точки, обработка изображений.*

DOI: 10.14357/20790279180501

Введение

В настоящее время в большинстве музеев и галерей используются различные инструменты для самостоятельного ознакомления с экспозицией. Наряду с классическими аудиогuidaми получили широкое распространение мобильные приложения, использующие методы обработки и анализа изображений. Некоторые из таких приложений распознают графические коды экспонатов

(штриховые или QR) [1], для других [2,3] входными данными являются фото- или видео- кадры с экспонатом, снятым крупным планом (Smartify, Artbit). «Мобильные гиды» последней категории удобнее для пользователя [4], чем решения с ручным вводом номера экспоната или распознаванием QR: номер и код достаточно малы, и их поиск и ввод требует дополнительных действий, не связанных с обзором экспозиции.

Алгоритмы анализа изображений произведения искусства не обязательно распознают конкрет-

* Работа выполнена при частичной финансовой поддержке грантов РФФИ (проекты 16-07-00616, 17-29-03236 и 15-29-06081).

ный экспонат. Так в работах [5-7] рассматривается задача определения авторства по изображению картины и проводится анализ качества результатов при использовании различных способов извлечения признаков, модели «мешка признаков» и методов статистического обучения. Задача определения художественного стиля картины рассмотрена в работах [8-10], подходы к решению которой аналогичны описанным выше. Использование нейронных сетей [11] позволяет установить одновременно стиль, жанр и автора, однако точность работы такой системы невысока.

С другой стороны, на сегодняшний день большинство музеев имеет общедоступные электронные каталоги, содержащие, помимо справочной информации, изображения экспонатов. Проиндексировав подобный каталог, можно получить полную справку о картине на снимке, установив для нее точное соответствие в базе. Подходы с использованием классифицирующих сетей нашли широкое применение [12-14] в решении данной задачи. Несомненным преимуществом таких методов является их применимость для классификации других объектов искусства (часы, скульптуры и т.д.) [14]. К недостаткам можно отнести сложность обновления – при изменении экспозиции требуется подготовка новых данных для обучения, не гарантируется достижение аналогичного качества классификации при увеличении количества классов, а общие затраты времени на переобучение достаточно высоки. Другим подходом к решению является комбинация модели «мешок локальных признаков» с алгоритмами аппроксимированного поиска ближайших соседей в пространстве дескрипторов. Подобный метод, предложенный в работе [15], потенциально может быть расширен на большее количество данных, однако не устойчив к смене

ракурса, а также подвержен межклассовым коллизиям, поскольку не учитывает взаимное расположение сопоставленных локальных особенностей. Отметим, что предложенные ранее методы разработаны для архитектур клиент-сервер, и, согласно заявленным оценкам быстродействия, не подходят для реализации полного цикла распознавания на клиентском устройстве.

1. Постановка задачи

Задача распознавания произведений живописи на изображениях, рассмотренная в данной работе, может быть сформулирована следующим образом. Изображение-запрос Q должно быть отнесено к одному из классов

$$C = \{C_i\}_{i \in [0, N]},$$

где C_i – класс изображений i -го экспоната при $i \in [1, N]$; C_0 – класс прочих изображений, соответствующий значению «неизвестная картина».

Для каждого C_i задано изображение-эталон T_i .

Воспользуемся следующими предположениями:

1. Изображение-запрос Q получено с помощью камеры мобильного устройства неподготовленным пользователем во время экскурсии, следовательно:
 - а) может содержать визуальные дефекты – блики, расфокусированные области, шум;
 - б) ракурс, кадрирование, освещение и цветовой баланс неизвестны (рис. 1а);
 - в) на снимке могут присутствовать посторонние объекты, такие как декорации, картинные рамы, посетители (рис. 1б).
2. Изображение-эталон T представляет собой изображение высокого разрешения с фронтальной проекцией картины или ее цифровую репродук-



Рис. 1. Примеры изображений-запросов: а) картина сфотографирована издали при слабом освещении, б) в кадре посетитель



Рис. 2. Примеры изображений-эталонов для картин: а) Клод Моне «De Voorzaan en de Westerhem», б) Сальвадор Дали «La persistència de la memòria»

цию. Эталон не содержит посторонних объектов и визуальных дефектов. Эталонами могут служить, например, качественно отсканированные страницы художественных альбомов (рис. 2).

3. Картина (как на запросе, так и на эталоне) может относиться к любому стилю – реализм, импрессионизм, абстракционизм, фрактальная графика и т.д.
4. Количество классов N соответствует размеру коллекции и для одной галереи может достигать сотен тысяч экспонатов [16].

Заметим, что в рассматриваемом сценарии использование нестабильного интернет-соединения не должно влиять на функционирование клиентской части, а, значит, должна быть возможность выполнить полный цикл распознавания на мобильном устройстве. Из этого естественным образом вытекают следующие требования:

1. Время полного цикла распознавания на клиентском устройстве не должно превышать 1 с.
2. Размер необходимых для распознавания данных должен позволять загрузить их на клиентское устройство.

2. Предлагаемый подход

Итак, перед нами стоит классическая задача классификации на основе сравнения с эталоном. Качество ее решения зависит от функции невязки ρ , определенной для пар изображений, которую использует распознающий алгоритм. Существует несколько способов оценивания функций невязки. В нашем варианте постановки их можно сравнивать с помощью аналога индекса Девиса-Болдина [17]:

$$D = \frac{1}{N} \sum_{i=1}^N \max_{i \neq j} \frac{S_i + S_j}{\rho(T_i, T_j)},$$

где

$$S_i = \frac{1}{|C_i|} \sum_{X \in C_i} \rho(X, T_i), \quad T_i - \text{эталон класса } C_i.$$

Чем ниже значение D , тем лучше будет работать алгоритм, использующий данную ρ . Здесь учитывается, что центроиды классов нам не известны, но для разумно выбранных ρ эталоны должны быть близки к центроидам.

Представление изображения с помощью описания в метрическом пространстве позволяет использовать в качестве ρ расстояния (Евклида, Хэмминга и т.п.), для которых существуют эффективные структуры данных для поиска ближайшего соседа. К сожалению, на сегодняшний день не предложены метрические представления всего изображения, подходящие для условий решаемой задачи. При этом для малых участков изображений отображения в метрическое пространство, обеспечивающие высокое качество их классификации, известны. Решаемая в последнем случае задача проще, поскольку отсутствует проблема посторонних объектов, а проективные искажения могут быть аппроксимированы аффинными [18] или даже подобиями [19]. Устойчивый поиск наиболее информативных областей изображения – «особых» точек – позволяет представлять изображение как небольшой набор локальных метрических дескрипторов. Но, как было упомянуто ранее, модель «мешка признаков» недостаточно устойчива к межклассовым коллизиям. Применительно к живописи, в зависимости от выбранного алгоритма вычисления дескрипторов, к одному классу могут быть отнесены как изображения одних и тех же объектов с разных ракурсов (левый и правый профиль на портретах, отражения), так и произведения с типовым набором объектов (натюрморты). Повысить устойчивость классификации можно, учитывая информацию о взаимном расположении особых точек для оценки геометрического соот-

ветствия описаний, но для модели «созвездия признаков» неизвестно общее отображение в метрическое пространство. В [20] предложен двухэтапный подход, совмещающий достоинства обеих моделей. На первом шаге выполняется аппроксимированный поиск ближайших изображений в модели «мешок точек». На втором – оценивается геометрическое соответствие с каждым из отобранных на первом шаге кандидатов.

3. Построение компактного описания изображения

Вопрос выбора алгоритмов детектирования и метрического описания особых точек при использовании подхода Лоу [20] является открытым. Мы используем алгоритм детектирования YACIPE [21], поскольку он сочетает устойчивость к перепадам освещенности с высокой скоростью вычисления, а его использование не ограничено патентом. В качестве метрического описания предлагается использовать бинарный дескриптор RFD, потому что, с одной стороны, он отвечает требованиям по скорости вычисления на мобильных устройствах, а с другой – показывает отличные результаты на открытых наборах данных [22].

Из соображений уменьшения вычислительных затрат и увеличения устойчивости, изображения предварительно обрабатываются следующим образом:

- меньший из размеров изображения приводится к стандартному значению с сохранением пропорций (это компенсирует недостаточную инвариантность дескрипторов к масштабированию);
- высокочастотный шум подавляется фильтром Гаусса;
- производится цветоредукция путем взвешенного усреднения цветовых каналов.

На полученном изображении вычисляются особые точки и дескрипторы их областей, их количество искусственно ограничивается до M лучших по внутренней оценке YACIPE. Таким образом, описание изображения I можно представить как:

$$\omega = W(I) = \{ \langle p_i, f_i \rangle \}_{i \in [1, M]}$$

где

$p_i = \langle x_i, y_i \rangle$ – координаты i -й точки
 f_i – дескриптор окрестности i -й точки

4. Алгоритм обработки набора эталонов

Для каждого эталона T_i построим описание $\omega_i = W(T_i)$, а затем проиндексируем – для каждого $j \in [1, |\omega_i|]$ занесем запись вида $\langle i, f_j^i \rangle$ в рандомизированное иерархически кластеризующее поис-

ковое дерево [23], которое позволяет выполнять приближенный поиск ближайших соседей с существенным выигрышем по скорости в сравнении с линейным поиском.

5. Алгоритм обработки запроса

На изображении-запросе Q зона картины предварительно локализуется с использованием предположения о прямоугольности рамы. Поиск зоны осуществляется с помощью быстрого алгоритма поиска четырехугольников [24] со снятым ограничением на соотношение сторон. Изображение в такой зоне проективно нормализуется, после чего выполняются все шаги построения компактного описания, начиная с предобработки. Такой подход ранее уже показал хорошие результаты в решении задачи поиска соответствия для кадра на экране среди эталонных видеофрагментов [25].

Предварительная локализация позволяет предотвратить следующие проблемы:

- недостаточное описание области картины из-за посторонних объектов в кадре, на которых могут быть особые точки с наилучшими оценками;
- вычислительные затраты на сопоставление дескрипторов областей, расположенных вне картины;
- значимое несовпадение масштаба и угла наклона между эталоном в базе и изображением картины, что приводит к некорректному результату сопоставления дескрипторов.

Определим исправленное изображение как $Q^* = H^0(Q)$, где H^0 – проективное преобразование с кадрированием результата. После того, как построено компактное описание $\omega^* = W(Q^*)$, распознавание проходит в два этапа:

1. Для каждой особой точки $\langle p, f \rangle \in \omega^*$ в индексе производится аппроксимированный поиск ближайших соседей дескриптора f . К найденным дескрипторам применяется процедура голосования – дескриптор f_j^i добавляет голос эталону T_i . Затем отбираются K эталонов-кандидатов с наибольшим количеством голосов.
2. Для каждого из K выбранных вариантов с помощью алгоритма RANSAC [26] выполняется поиск проективного преобразования H , переводящего, в пределах заданной геометрической погрешности δ , точки запроса Q^* в точки эталона T . Пара точек $\langle p, p' \rangle$ с близкими дескрипторами считается верным сопоставлением, если:

$$|H(p) - p'| < \delta, \quad p \in \omega^*, \quad p' \in \omega^T.$$

В качестве окончательного результата выбирается эталон T_b , для которого число верных со-

поставлений $G(\omega^*, \omega_b)$ оказалось максимальным. При этом, если их количество меньше определенного порогового значения R , результатом будет ответ «неизвестная картина», чтобы избежать ложных срабатываний (например, на картинах, для которых в поисковой базе нет описания эталона). В случае, когда два кандидата имеют одинаковое количество сопоставлений, дополнительно рассчитывается оценка преобразования:

$$e_i = \frac{1}{G(\omega^*, \omega_i)} \sum_{j=0}^{G(\omega^*, \omega_i)} \frac{|H(p_j^*) - p_j^j|}{\delta}, \quad \delta > 0, p_j^* \in \omega^*, p_j^j \in \omega_i$$

Таким образом

$$Q \in \begin{cases} C_i, & \text{if } G(\omega^*, \omega_i) \geq R \text{ and } G(\omega^*, \omega_j) + e_i > G(\omega^*, \omega_j) + e_j, i \neq j \\ C_0, & \text{if } G(\omega^*, \omega_i) < R \end{cases}$$

$$i, j \in [1, M]$$

Соответственно, функция невязки $\rho = \frac{1}{G}$:

$$\rho(Q^*, T_i) = \frac{1}{G(\omega^*, \omega_i)} - \text{невязка между нормализованным запросом } Q^* \text{ и эталоном } T_i.$$

ваным запросом Q^* и эталоном T_i .

6. Модификации алгоритма

При анализе результатов работы предложенного алгоритма было замечено, что алгоритм поиска четырехугольника часто выделяет зону картины вместе с картинной рамой. Дескрипторы окрестностей точек, найденных на резном багете, не вносят позитивного вклада в поиск. Количество точек ограничено сверху требованиями к быстродействию алгоритма, а зона картины может обладать менее выраженными, согласно выбранному детектору точек, локальными особенностями, чем рама. В результате, найденных на картине точек может быть слишком мало для описания. Такие случаи, ввиду большого разнообразия используемых кантов и багетов, затруднительно разрешать на стадии поиска четырехугольника.

Для коррекции таких случаев были протестированы две модификации базового алгоритма (Base).

Первая (Mod1) состоит в генерации для изображения-запроса покрытия равномерной прямоугольной сеткой и выборе в каждой ячейке фиксированного числа точек. Экспериментально были подобраны значения, которые позволяют сохранить точность и высокую скорость.

Вторая модификация (Mod2) опирается на наблюдение, что большая часть посторонних объектов находятся по краям картины. Предлагается для поиска использовать особые точки, расположенные в центральной части нормализованного запроса

са Q^* , а точки из краевой области (порядка 5% от линейных размеров) игнорировать.

7. Эксперименты

Для составления тестовых поисковых баз данных использовались соответствующие описанным выше требованиям изображения, которые находятся в открытом доступе [27]. Оригинальные картины относятся к 25 различным жанрам, таким как импрессионизм, кубизм, поп-арт, неоклассицизм и т.д.; общий объем использованной подвыборки составляет ~80 тыс. изображений. Для экспериментальной оценки зависимости скорости и точности поиска от размера поисковой базы данных случайным образом были составлены наборы различного объема: 100, 1000 и 10000 изображений.

Тестовый набор запросов составлен из пользовательских изображений, размещенных в открытых источниках (GooglePhoto, Instagram). Подборка содержит фотографии, снятые при помощи различных мобильных устройств в музеях и галереях, аналогичные представленным на рис. 1. На фотографиях присутствует не более одной картины одновременно. Набор данных состоит из двух частей:

- положительная выборка, картина на фотографии присутствует в базе – 170 изображений;
- отрицательная выборка, картина на фотографии в базе отсутствует – 30 изображений.

Положительная выборка содержит повторы – фотографии одной и той же картины, сделанные разными пользователями, всего 40 уникальных картин.

Эксперимент проводился на следующих устройствах:

- (desktop) – процессор AMD® A10-7850k radeon r7 3700 МГц, 8 Гб ОЗУ;
- (mobile) – процессор Qualcomm Snapdragon 625 MSM8953 2000 МГц, 4 Гб ОЗУ.

Время создания и размер поисковой структуры данных для эталонов картин представлены в табл. 1. Видно, что с увеличением числа эталонов значения на рассматриваемом участке возрастают приблизительно линейно.

Усредненное время одного поискового запроса для разных по количеству эталонов поисковых баз представлено в табл. 2.

На всей отрицательной выборке во всех экспериментах был получен ответ «неизвестная картина», т.е. ложноположительных срабатываний не было. Для примеров положительных запросов оценивался процент правильных ответов, когда корректно определен идентификатор картины.

Табл. 1

Характерная зависимость времени создания поисковой базы данных от количества эталонов

Количество эталонов в БД, шт	Время создания индекса БД, с (desktop)	Размер структуры поиска, Мб
100	9	1.13
1000	90	12.4
10000	911	125.3

Табл. 2

Зависимость времени выполнения одного запроса от объема поисковой базы. Время усреднено по 200 запросам

Количество эталонов в БД, шт	Время обработки одного запроса, с					
	(desktop)			(mobile)		
	Base	Mod1	Mod2	Base	Mod1	Mod2
100	0.069	0.056	0.068	0.280	0.238	0.275
1000	0.070	0.057	0.070	0.313	0.260	0.307
10000	0.081	0.060	0.073	0.361	0.274	0.340

Табл.3

Качество распознавания (точность как % правильных ответов) положительной выборки

Количество эталонов в БД, шт	Точность, %		
	Base	Mod1	Mod2
100	95.8	98.2	98.8
1000	94.7	97.0	97.6
10000	91.7	94.1	96.4

Итоговое сравнение точности распознавания на положительной выборке для базового подхода и его модификаций представлено в табл. 3.

По результатам экспериментов можно сделать вывод, что версия алгоритма Mod2 является

наиболее удачной по соотношению скорости работы и качества распознавания.

Предложенная реализация имеет определенные ограничения. Так, в ходе эксперимента были выявлены картины, на которых выбран-

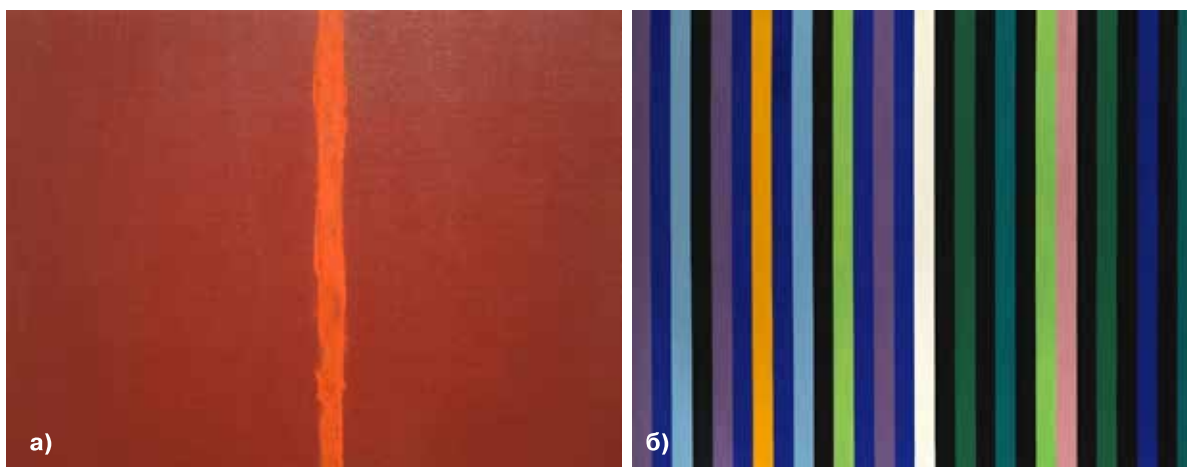


Рис. 3. Пример сложных для выделения особых точек изображений: а) Барнетт Ньюман «Onement Iii»; б) Джин Дэвис «Two Part Blue»

ный алгоритм поиска не находит особых точек. Примеры таких картин представлены на рис. 3. Стоит отметить, что изображения такого вида вообще не содержат локальных особенностей типа «точка».

Заключение

В работе показана эффективность алгоритма поиска изображения в базе данных с использованием компактного описания и специальных структур данных для решения задачи распознавания живописи. Метод протестирован на реальных данных из открытых источников и показал точность выше 95% и время работы 0.24-0.36 секунды на стандартном для 2017 года мобильном устройстве, что по точности сопоставимо с современными решениями, а по скорости превосходит. Отметим, что подход, изначально разработанный для идентификации телевизионного контента, при минимальных изменениях нашел применение и продемонстрировал высокое качество в другой области, что показывает его потенциальную универсальность в задачах классификации изображений различной природы. Установление геометрического соответствия с помощью RANSAC для 3D-координат [28] допускает использование предложенного метода, в том числе, для распознавания трехмерных объектов (скульптур, часов и прочих предметов искусства). Исходя из полученных результатов, дальнейшее развитие работы может быть ориентировано как на исследование свойств метода на других классах распознаваемых объектов, так и на улучшение его характеристик за счет использования более точных геометрических моделей и оценок центроидов целевых классов.

Литература

1. *Pérez-Sanagustín M. et al.* Using QR codes to increase user engagement in museum-like spaces // *Computers in Human Behavior*. – 2016. – Т. 60. – С. 73-85.
2. *Антошук С.Г., Годовиченко Н.А.* Анализ точечных особенностей изображения в системе «Мобильный виртуальный гид» // *Pratsi*. – 2013. – №. 1 (40). – С. 67-72.
3. *Andreatta C., Leonardi F.* Appearance based paintings recognition for a mobile museum guide // *International Conference on Computer Vision Theory and Applications, VISAPP*. – 2006.
4. *Leonard Wein.* 2014. Visual recognition in museum guide apps: do visitors want it?. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 635-638.
5. *Ivanova K. et al.* Features for art painting classification based on vector quantization of mpeg-7 descriptors // *Data Engineering and Management*. – Springer, Berlin, Heidelberg, 2012. – С. 146-153.
6. *Keren D.* Painter identification using local features and naive bayes // *Pattern Recognition, 2002. Proceedings. 16th International Conference on*. – IEEE, 2002. – Т. 2. – С. 474-477.
7. *Jou J., Agrawal S.* Artist identification for renaissance paintings. – 2012.
8. *Zujovic J. et al.* Classifying paintings by artistic genre: An analysis of features & classifiers // *Multimedia Signal Processing, 2009. MMSp'09. IEEE International Workshop on*. – IEEE, 2009. – С. 1-5.
9. *Arora R.S.* Towards automated classification of fine-art painting style: A comparative study : дис. – Rutgers University-Graduate School-New Brunswick, 2012.
10. *Lecoutre A., Negrevergne B., Yger F.* Recognizing Art Style Automatically in painting with deep learning // *Asian Conference on Machine Learning*. – 2017. – С. 327-342.
11. *Tan W.R. et al.* Ceci n'est pas une pipe: A deep convolutional network for fine-art paintings classification // *Image Processing (ICIP), 2016 IEEE International Conference on*. – IEEE, 2016. – С. 3703-3707.
12. *Hong Y., Kim J.* Art Painting Identification using Convolutional Neural Network // *International Journal of Applied Engineering Research*. – 2017. – Т. 12. – №. 4. – С. 532-539.
13. *Taverriti G. et al.* Real-time Wearable Computer Vision System for Improved Museum Experience // *Proceedings of the 2016 ACM on Multimedia Conference*. – ACM, 2016. – С. 703-704.
14. *Zhang R., Tas Y., Koniusz P.* Artwork Identification from Wearable Camera Images for Enhancing Experience of Museum Audiences // *arXiv preprint arXiv:1806.09084*. – 2018
15. *Ruf B., Kokipoulou E., Detyniecki M.* Mobile museum guide based on fast SIFT recognition // *International Workshop on Adaptive Multimedia Retrieval*. – Springer, Berlin, Heidelberg, 2008. – С. 170-183.
16. *Третьяковская галерея*, <https://www.tretyakovgallery.ru/collection/>
17. *Davies D.L., Bouldin D.W.* A cluster separation measure // *IEEE transactions on pattern analysis and machine intelligence*. – 1979. – №. 2. – С. 224-227.
18. *Yu G., Morel J.M.* A fully affine invariant image comparison method // *Acoustics, Speech and*

- Signal Processing, 2009. ICASSP 2009. IEEE International Conference on. – IEEE, 2009. – С. 1597-1600.
19. Bay H., Tuytelaars T., Van Gool L. Surf: Speeded up robust features //European conference on computer vision. – Springer, Berlin, Heidelberg, 2006. – С. 404-417.
20. Turcot P., Lowe D.G. Better matching with fewer features: The selection of useful features in large database recognition problems //Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on. – IEEE, 2009. – С. 2109-2116.
21. Лукоянов А.С., Николаев Д.П., Коноваленко И.А. Модификация алгоритма YAPE для изображений с большим разбросом локального контраста //Информационные технологии и нанотехнологии. – 2018. – С. 1193-1204.
22. Fan B. et al. Receptive fields selection for binary feature description //IEEE Transactions on Image Processing. – 2014. – Т. 23. – №. 6. – С. 2583-2595.
23. Muja M., Lowe D.G. Fast matching of binary features //Computer and Robot Vision (CRV), 2012 Ninth Conference on. – IEEE, 2012. – С. 404-410.
24. Skoryukina N., Nikolaev D.P., Sheshkus A., Polevoy D. (2015, February). Real time rectangular document detection on mobile devices. In Seventh International Conference on Machine Vision (ICMV 2014) (Vol. 9445, p. 94452A). International Society for Optics and Photonics.
25. Skoryukina N. et al. Snapscreen: TV-stream frame search with projectively distorted and noisy query //Ninth International Conference on Machine Vision (ICMV 2016). – International Society for Optics and Photonics, 2017. – Т. 10341. – С. 103410Y.
26. Fischler M.A., Bolles R.C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography //Communications of the ACM. – 1981. – Т. 24. – №. 6. – С. 381-395
27. WikiArt <http://www.wikiart.org/>
28. Karpenko S. et al. UAV control on the basis of 3D landmark bearing-only observations //Sensors. – 2015. – Т. 15. – №. 12. – С. 29802-29820.

Скорюкина Наталья Сергеевна. ООО «Смарт Энджинс Сервис», г. Москва, Россия. Научный сотрудник-программист. Количество печатных работ: 8. Область научных интересов: анализ изображений, компьютерное зрение. E-mail: skleppy.inc@gmail.com

Миловзоров Артемий Николаевич. Национальный исследовательский технологический университет «МИСиС», г. Москва, Россия. Студент. Область научных интересов: анализ изображений. E-mail: milovzorov.artemiy@yandex.ru

Полевой Дмитрий Валерьевич. Институт системного анализа Федерального исследовательского центра «Информатика и управление» Российской академии наук, г. Москва, Россия. Старший научный сотрудник. Кандидат технических наук. Национальный исследовательский технологический университет «МИСиС», г. Москва, Россия. Доцент. Количество печатных работ: более 20. Область научных интересов: цифровая обработка изображений, компьютерное зрение, искусственный интеллект. E-mail: dvpsun@gmail.com

Арлазаров Владимир Викторович. Институт проблем передачи информации им. А.А. Харкевича Российской академии наук, г. Москва, Россия. Ведущий научный сотрудник, кандидат технических наук. Количество печатных работ: 30. Область научных интересов: искусственный интеллект, машинное обучение, системы распознавания, информационные технологии. E-mail: vva777@gmail.com

Paintings recognition in uncontrolled conditions using one-shot learning

N.S. Skoryukina^I, A.N. Milovzorov^{II}, D.V. Polevoy^{III,IV}, V.V. Arlazarov^{II,III,IV,V}

^I Smart Engines Service LLC, Moscow, Russia

^{II} National University of Science and Technology "MISIS", Moscow, Russia

^{III} Institute for Systems Analysis, Federal Research Center "Computer Science and Control" of Russian Academy of Sciences, Moscow, Russia

^{IV} Moscow Institute of Physics and Technology, Dolgoprudny, Russia

^V Institute for information transmission problems (Kharkevich Institute) RAS, Moscow, Russia

Abstract. The paper considers the problem of paintings identification in photos acquired with mobile devices under the conditions of museum exhibition. The proposed approach is based on compact description of an image with a constellation of keypoints and corresponding local descriptors. Two-step comparison scheme is described for finding the best reference image matching the query. Bag-of-features approach is used as a first step, then mutual disposition of points is analyzed. Rejection of the query is performed if no suitable matches are found. Geometrical normalization of the query image is proposed to achieve higher robustness against scale and viewpoint variations. Advantages of the described approach over state-of-the-art solutions are considered. The results of the experiments conducted on the open WikiArt dataset are presented along with processing times for different hardware platforms.

Keywords: *paintings recognition, feature points, image processing.*

DOI: 10.14357/20790279180501

References

1. Pérez-Sanagustín M. et al. Using QR codes to increase user engagement in museum-like spaces // Computers in Human Behavior. – 2016. – T. 60. – С. 73-85.
2. Antoschuk S.G., Godovichenko N.A. Image local features analysis for "Mobile virtual guide" system. //Pratsi. – 2013. – №. 1 (40). – С. 67-72.
3. Andreatta C., Leonardi F. Appearance based paintings recognition for a mobile museum guide //International Conference on Computer Vision Theory and Applications, VISAPP. – 2006.
4. Leonard Wein. 2014. Visual recognition in museum guide apps: do visitors want it?. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14). ACM, New York, NY, USA, 635-638.
5. Ivanova K. et al. Features for art painting classification based on vector quantization of mpeg-7 descriptors //Data Engineering and Management. – Springer, Berlin, Heidelberg, 2012. – С. 146-153.
6. Keren D. Painter identification using local features and naive bayes //Pattern Recognition, 2002. Proceedings. 16th International Conference on. – IEEE, 2002. – T. 2. – С. 474-477.
7. Jou J., Agrawal S. Artist identification for renaissance paintings. – 2012.
8. Zujovic J. et al. Classifying paintings by artistic genre: An analysis of features & classifiers // Multimedia Signal Processing, 2009. MMSP'09. IEEE International Workshop on. – IEEE, 2009. – С. 1-5.
9. Arora R.S. Towards automated classification of fine-art painting style: A comparative study : diss. – Rutgers University-Graduate School-New Brunswick, 2012.
10. Lecoutre A., Negrevergne B., Yger F. Recognizing Art Style Automatically in painting with deep learning //Asian Conference on Machine Learning. – 2017. – С. 327-342.
11. Tan W. R. et al. Ceci n'est pas une pipe: A deep convolutional network for fine-art paintings classification //Image Processing (ICIP), 2016 IEEE International Conference on. – IEEE, 2016. – С. 3703-3707.
12. Hong Y., Kim J. Art Painting Identification using Convolutional Neural Network //International Journal of Applied Engineering Research. – 2017. – T. 12. – №. 4. – С. 532-539.
13. Taverri G. et al. Real-time Wearable Computer Vision System for Improved Museum Experience //Proceedings of the 2016 ACM on Multimedia Conference. – ACM, 2016. – С. 703-704.
14. Zhang R., Tas Y., Koniusz P. Artwork Identification from Wearable Camera Images for Enhancing Experience of Museum Audiences //arXiv preprint arXiv:1806.09084. – 2018
15. Ruf B., Kokipoulou E., Detyniecki M. Mobile museum guide based on fast SIFT recognition // International Workshop on Adaptive Multimedia Retrieval. – Springer, Berlin, Heidelberg, 2008. – С. 170-183.
16. The State Tretyakov Gallery, <https://www.tretykovgallery.ru/collection/>

17. *Davies D.L., Bouldin D.W.* A cluster separation measure //IEEE transactions on pattern analysis and machine intelligence. – 1979. – №. 2. – С. 224-227.
18. *Yu G., Morel J.M.* A fully affine invariant image comparison method //Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on. – IEEE, 2009. – С. 1597-1600.
19. *Bay H., Tuytelaars T., Van Gool L.* Surf: Speeded up robust features //European conference on computer vision. – Springer, Berlin, Heidelberg, 2006. – С. 404-417.
20. *Turcot P., Lowe D.G.* Better matching with fewer features: The selection of useful features in large database recognition problems //Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on. – IEEE, 2009. – С. 2109-2116.
21. *Lukoyanov A.S., Nikolaev D.P., Konovalenko I.A.* Modification of YAPE keypoint detection algorithm for wide local contrast range image // Information technologies and nanotechnology. – 2018. – С. 1193-1204.
22. *Fan B. et al.* Receptive fields selection for binary feature description //IEEE Transactions on Image Processing. – 2014. – Т. 23. – №. 6. – С. 2583-2595.
23. *Muja M., Lowe D.G.* Fast matching of binary features //Computer and Robot Vision (CRV), 2012 Ninth Conference on. – IEEE, 2012. – С. 404-410.
24. *Skoryukina N., Nikolaev D.P., Sheshkus A., Polevoy D.* (2015, February). Real time rectangular document detection on mobile devices. In Seventh International Conference on Machine Vision (ICMV 2014) (Vol. 9445, p. 94452A). International Society for Optics and Photonics.
25. *Skoryukina N. et al.* Snapscreen: TV-stream frame search with projectively distorted and noisy query //Ninth International Conference on Machine Vision (ICMV 2016). – International Society for Optics and Photonics, 2017. – Т. 10341. – С. 103410Y.
26. *Fischler M.A., Bolles R.C.* Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography //Communications of the ACM. – 1981. – Т. 24. – №. 6. – С. 381-395
27. *WikiArt* <http://www.wikiart.org/>
28. *Karpenko S. et al.* UAV control on the basis of 3D landmark bearing-only observations //Sensors. – 2015. – Т. 15. – №. 12. – С. 29802-29820.

N.S. Skoryukina. LLC “Smart Engines Service”, Moscow, Russia. Researcher-programmer. Number of publications: 8. Scientific interests: image processing, computer vision. E-mail: skleppy.inc@gmail.com

A.N. Milovzorov. National university of science and technology “MISiS”, Moscow, Russia. Student. Number of publications: 0. Scientific interests: image processing. E-mail: milovzorov.artemiy@yandex.ru

D.V. Polevoy. Institute for Systems Analysis, Federal Research Center “Computer Science and Control” of Russian Academy of Sciences, Moscow, Russia. Ph.D. National university of science and technology “MISiS”, Moscow, Russia. Senior researcher. Associate professor. Number of publications: more than 20. Scientific interests: image processing, computer vision, artificial intelligence, deep learning. E-mail: dvpsun@gmail.com

V.V. Arlazarov. Institute of information transmission problems (Kharkevich institute) of Russian Academy of Sciences, Moscow, Russia. Lead researcher, Ph.D. Number of publications: 30. Scientific interests: artificial intelligence, machine learning, recognition systems, information technology. E-mail: vva777@gmail.com