

УДК 004.4

ОБЗОР ПРОГРАММНЫХ СИСТЕМ РАСПОЗНАВАНИЯ ИЗОБРАЖЕНИЙ

Головкин Михаил Евгеньевич

Российский государственный социальный
университет, г. Москва

Краснов Андрей Евгеньевич

Федеральное государственное автономное учреждение
«Государственный научно-исследовательский институт
информационных технологий и телекоммуникаций»

В статье представлена информация о наиболее широко используемых системах распознавания изображений. Приводится их описание, характеристики, сравнительный обзор с указанием категорий изображений, которые наилучшим образом распознаёт каждая система. В заключительной части обзора делается вывод о том, какая система является наиболее рациональной для распознавания изображений различных категорий, а также для фотограмметрии и трёхмерного моделирования.

Ключевые слова: программные системы, распознавание изображений, распознавание цветных изображений, распознавание чёрно-белых изображений, распознавание текста, распознавание лиц, скорость распознавания, качество распознавания, фотограмметрия, трёхмерное моделирование.

В настоящее время всё более широкое распространение получают информационные системы, функционирование которых напрямую связано с распознаванием изображений.

Значительный интерес представляет распознавание как отдельных фрагментов изображения, так и всей сцены в целом, причем инвариантно к группе преобразований изображения (сдвига, изменения масштаба, поворота) [1].

Системы распознавания можно разделить на несколько классов по категориям обрабатываемых изображений – цветные, чёрно-белые, текст, лица. Наиболее важные характеристики систем, а также критерии для их сравнения:

- 1) производительность
- 2) качество распознавания
- 3) способность распознавать цветные изображения, чёрно-белые изображения, текст, рисунки.

Ниже приведен обзор основных используемых в настоящее время программных систем распознавания изображений с описанием их основных характеристик и функциональности.

RealityCapture

RealityCapture – программное обеспечение фотограмметрии, которое позволяет создавать 3D модели из ряда

обычных изображений и/или из образов объектов, формируемых путем сканирования поверхностей объектов лазерным лучом [8].

RealityCapture превосходит другое аналогичное ПО благодаря функции быстрого выстраивания, позволяющей выстроить изображения за несколько секунд [3].

Это полнофункциональное приложение с богатым, простым в использовании интерфейсом и множеством функций. Наиболее важные из них показаны на рис. 1.



Рисунок 1 - Функциональность системы **RealityCapture**

Также в системе доступен сверхбыстрый режим предпросмотра на ноутбуке, который выравнивает изображения меньше чем за 10 минут на ноутбуке. RealityCapture позволяет объединить рисунки, изображения, полученные с помощью лазерного луча, а также изображения, полученные при аэрофотосъемке.

Agisoft Photoscan

Программа Agisoft Photoscan позволяет автоматически создавать высококачественные 3D модели объектов на основе цифровых фотографий [10].

Для реконструкции объекта в Photoscan достаточно загрузить фотографии, дополнительной информации не требуется.

Алгоритм работы программы приведен на рис. 2.

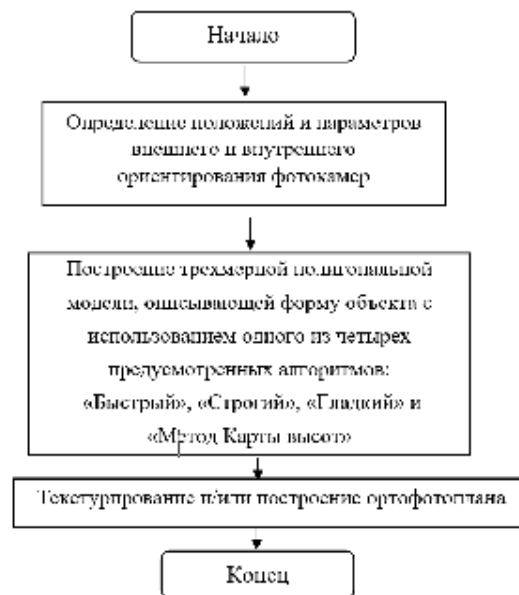


Рисунок 2 - Алгоритм работы системы **Agisoft Photoscan**

Pix4d
Pix4D Solutions – линейка программных продуктов, которая используется для обработки аэрофотоснимков и получения высокоточных ортофотопланов, построения 3D-моделей, карт отражений и карт индексов.

Обработка в Pix4D находит применение и для выявления изменений ландшафта, а также для ликвидации последствий чрезвычайных ситуаций [12].

Отличительной особенностью ПО Pix4D является возможность обработки мультиспектральных снимков [5].

Pix4Dmapper автоматически преобразует изображения, полученные с беспилотных летательных аппаратов (БПЛА) или изображения наземной съемки, и обеспечивает составление высокоточных карт и 3D-моделей с географической привязкой.

Один из основных компонентов системы - мультиспектральная камера Sequoia, принципы работы которой перечислены на схеме (рис. 3).



Рисунок 3 - Принципы работы и характеристики камеры **Sequoia**

ABBYY FineReader
ABBYY FineReader представляет собой программный продукт, поддерживающий работу со многими современными моделями сканеров, а также корректно распознающий любые снимки текста, сделанные цифровой камерой или

при помощи мобильного телефона [13]. Встроенная система распознавания может открывать файлы практически любых форматов [5]. Предусмотрена функция предварительной обработки изображения, исправляющая качество текстов. Другие

функциональные возможности ABBYY | FineReader перечислены на рис. 4.

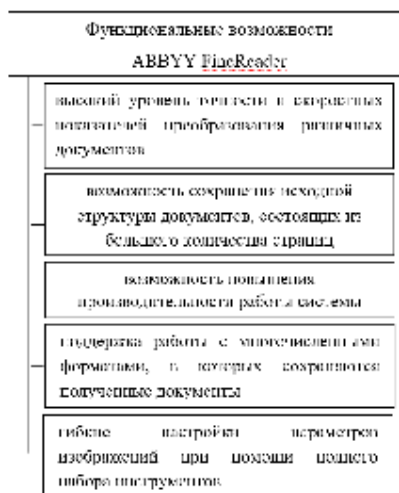


Рисунок 4 - Функциональные возможности ABBYY FineReader

Cunei Form

Cunei Form – это программная система оптического распознавания от российского разработчика Cognitive Technologies [14].

Для работы с системой достаточно отсканировать страницу с текстом, и через несколько секунд можно получить готовый результат в многофункциональном текстовом редакторе.

Cunei Form – шрифтонезависимая система. Алгоритмы, заложенные в

CuneiForm, исходят из правил написания букв и не требуют задания каких-либо эталонов, или обучения. В системе используется технология интеллектуального самообучения на базе адаптивного распознавания символов [2].

Распознаются любые печатные шрифты. Не распознается рукописный текст.

Характеристики системы перечислены на рис. 5.

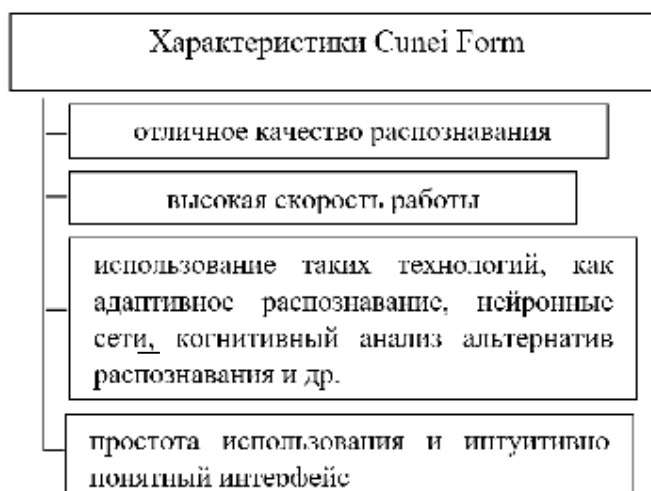


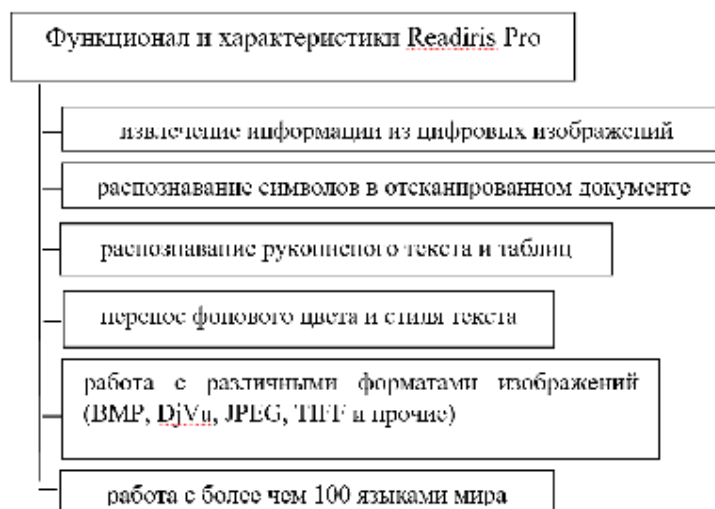
Рисунок 5 - Характеристики Cunei Form

Readiris Pro

Readiris Pro – программа, позволяющая преобразовывать документы в любые форматы. Приложение быстро

конвертирует формат изображения в txt, xml, pdf или docx документ [15].

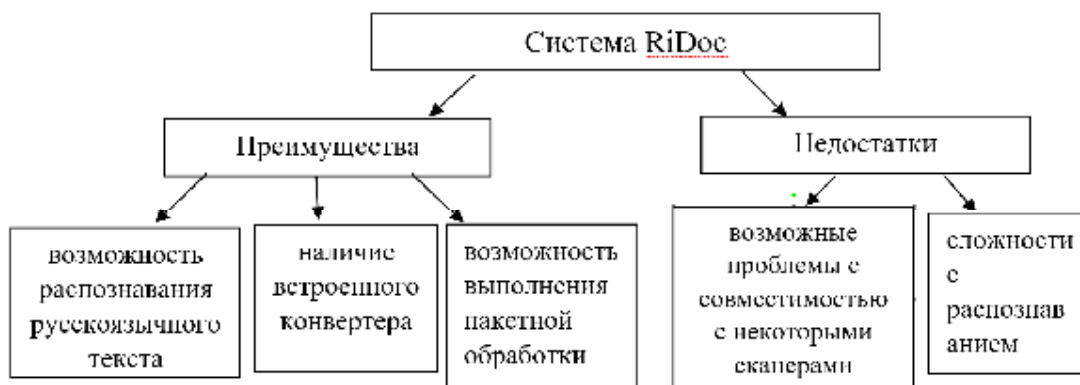
Краткое описание функционала и характеристик системы приведено на рис. 6

Рисунок 6 - Функционал и характеристики **Readiris Pro****RiDoc**

Приложение, которое предназначено для сканирования и уменьшения размера файла с сохранением качества [16]. Утилита может распознавать текст на разных языках. Кроме того, приложение может

конвертировать текстовые документы в изображения. Также есть возможность добавлять водяные знаки.

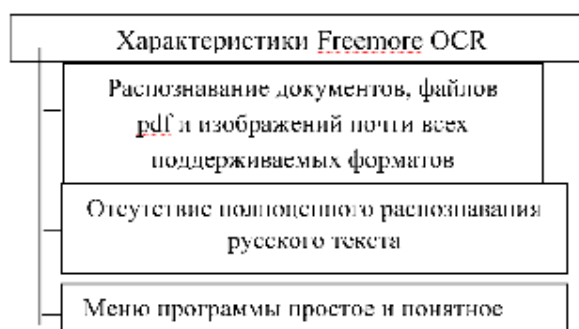
Преимущества и недостатки системы RiDoc схематически отображены на рис. 7.

Рисунок 7 - Преимущества и недостатки системы **RiDoc****Freemore OCR**

Используемая в приложении технология Side-by-Side помогает приложению отделять картинки от символов, чтобы не захватывать лишние

блоки. Распознанные документы сохраняются в формате doc, txt или pdf [17].

Характеристики Freemore OCR перечислены на рис. 8

Рисунок 8 - Характеристики **Freemore OCR**

Windows Hello

Windows Hello — одна из интересных функций операционной системы Windows 10.

Обеспечивает способ входа в систему с помощью распознавания лица: как только пользователь оказывается перед своим компьютером, автоматически происходит вход [18].

Система распознает и сопоставляет множество точек, описывающих определенные черты лица: глаза, губы, нос и т. п. Она описывает точное положение этих точек для каждого человека и назначает этому массиву точек определенный уникальный тег [7].

Когда текущий пользователь выходит из системы, система автоматически пытается распознать лица людей, оказывающихся перед экраном и сравнить полученные данные с хранящимися в локальной базе данных Windows Hello. При совпадении автоматически загружается профиль соответствующего пользователя.

С фотографиями или изображениями лица эта функция не работает.

Blink

Luxand Blink Pro – программа, которая использует для входа в систему в качестве пароля изображение лица пользователя с видеокамеры компьютера. Чтобы войти в свой профиль на компьютере,

достаточно подключенной веб-камеры [19]. Используя подобные передовые биометрические технологии распознавания лиц, Blink практически не делает ошибок, чему способствует опыт достигнутый в результате многолетних исследований. При правильной настройке и хорошем оборудовании ложные срабатывания практически исключены.

Программа Luxand Blink Pro действует почти аналогично сканеру отпечатков пальцев, за исключением того, что вместо ручного ввода пароля, для распознавания используется видеоскан лица пользователя [8].

KeyLemon

Программное обеспечение KeyLemon – это специализированное приложение для замены пароля, обычно набираемого с клавиатуры, изображением лица пользователя: система создает уникальную модель лица, которая применяется для входа в систему Windows либо выхода в Интернет [20]. Такой подход KeyLemon позволяет отказываться от многочисленных паролей, которые приходится запоминать и вводить ежедневно.

Основные характеристики системы перечислены на рис. 9.

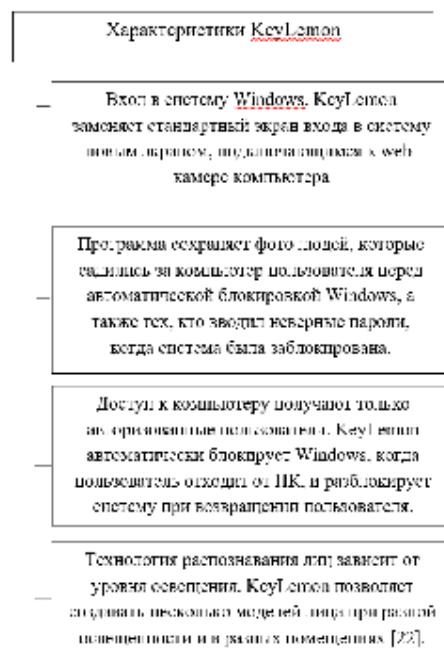


Рисунок 9 - Характеристики KeyLemon

Преимущества и недостатки системы схематически показаны на рис. 10.

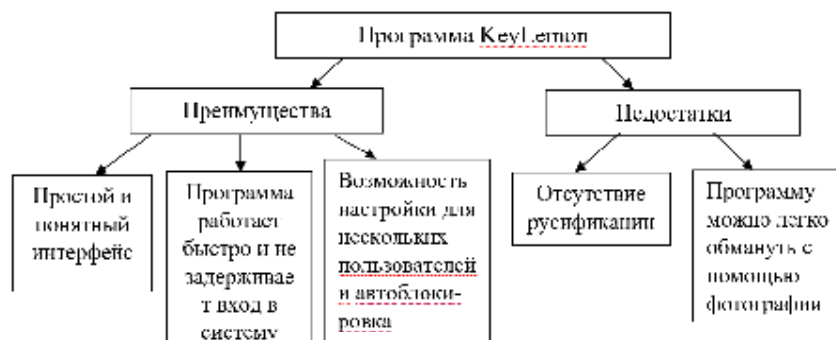


Рисунок 10 - Преимущества и недостатки KeyLemon

VeriFace Lenovo

Veriface Lenovo – это программа, которая дает возможность пользователям защитить собственные данные современным способом. Работа программы заключается в распознавании лиц, поэтому устройство обязательно должно быть оснащено камерой [21]. Поскольку ПО выпущено компанией

Lenovo, оно предназначено для устройств именно этого бренда.

В основе работы программы лежит специальная технология, которая «читает» черты лица.

Преимущества и недостатки VeriFace Lenovo показаны на рис. 11.



Рисунок 11 - Преимущества и недостатки VeriFace Lenovo

Большинство программных систем наилучшим образом подходят для распознавания изображений какой-либо определённой категории. Сравнение

программных систем с указанием распознаваемого типа изображений приведено в таблице 1.

Таблица 1 - Сравнение программных систем распознавания изображений

Система Тип изобра- жения	RealityCap- ture	Agisoft Photoscan	Pix4d	ABBYY FineReader	CuneiForm	Readiris Pro
Цветное	Работает	Работает	Работает	Не работает	Не работает	Не работает
Чёрно-белое	Работает	Работает	Работает	Не работает	Не работает	Не работает
Текст	Не работает	Не работает	Не работает	Работает	Работает	Работает
Лица	Работает	Не работает	Не работает	Не работает	Не работает	Не работает
Цветное	Не работает	Не работает	Работает	Работает	Работает	Работает
Чёрно-белое	Не работает	Не работает	Не работает	Не работает	Не работает	Не работает
Текст	Работает	Работает	Не работает	Не работает	Не работает	Не работает
Лица	Не работает	Не работает	Работает	Работает	Работает	Работает

Чем больше типов изображений может распознавать система, а также чем выше скорость и эффективность распознавания, тем более обоснованно её применение с технической и экономической точки зрения.

Как видно из таблицы, система RealityCapture может работать как с цветными и чёрно-белыми изображениями, так и с изображениями лиц, формируемыми фотокамерой. Кроме того, программа обладает множеством функций и характеризуется высокой скоростью

построения качественных трёхмерных моделей объектов при распознавании изображений.

Заключение

На основании приведённых характеристик систем и сравнительного анализа их функциональных возможностей следует сделать вывод о том, что наиболее рациональной программной системой для распознавания изображений различных категорий является RealityCapture.

В то же время, системы, подобные RealityCapture, широко применяются для трёхмерного моделирования. Они могут использоваться в более узком кругу специализированных задач, например, для фотограмметрии.

Также важно отметить, что во всех рассмотренных системах отсутствует возможность распознавания изображений, подвергнутых одновременным преобразованиям сдвига, поворота и масштабирования.

Список использованных источников

1. Головкин М. Е., Краснов А. Е. Методы выделения инвариантных признаков изображений. // Актуальные проблемы современной науки – 2016. – №4 (89). – с. 209-212.
2. Ипатов Ю.А., Кревецкий А.В. Методы обнаружения и пространственной локализации групп точечных объектов // Кибернетика и программирование. — 2014. – №6. – с. 17-25.
3. Черногорова Ю. В. Методы распознавания образов // Молодой ученый. – 2016. – №28. – с. 40-43.
4. Исаев А. Л., Газаров Д.А., Евсеев С.Д. Распознавание лиц по изображениям. // Символ науки. – 2017. – Т. 2, №4. – с. 70-76.
5. Исаев А. Л., Андросова Е. Е. Компьютерное моделирование комбинации из трехмерных объектов. // Символ науки. – 2016. – №10-2. – с. 44-49.
6. Финогеев А.Г., Четвергова М.В. Методика распознавания изображений на основе случайных деревьев в системах автоматизированного проектирования расширенной реальности // Современные проблемы науки и образования. – 2012. – №5.
7. Козлов В. Н. О распознавании аффинно разных дискретных изображений // Интеллектуальные системы. – 1998. – Т. 3, №3-4. – с. 95-122.
8. Цветков А. А., Шорох Д. К., Зубарева М. Г., Юрсков С. В., Шуклин А. В., Хамуш А. Л., Ануфриев И. Б. Алгоритмы распознавания объектов [Текст] // Технические науки: проблемы и перспективы: материалы IV Международной научной конференции. – СПб.: Свое издательство, 2016. – с. 20-28.
9. Гричуха К. RealityCapture. – Режим доступа: https://grinikkos.com/view_post.php?id=394 (дата обращения: 03.06.2018)
10. Agisoft PhotoScan Pro 1.1.5 Portable – создание 3D моделей из фото. – Режим доступа: <http://portable4pro.ru/foto/3d-modeling/agisoft-photoscan-pro.html> (дата обращения: 03.06.2018)
11. <http://portable4pro.ru/foto/3d-modeling/agisoft-photoscan-pro.html> (дата обращения: 03.06.2018)
12. Программное обеспечение Pix4D. – Режим доступа: <http://unmanned.ru/software/pix4d.htm> (дата обращения: 04.06.2018)
13. Эбби файн. – Режим доступа: <https://soul-car.ru/fajl/ebbi-fajn.html> (дата обращения: 04.06.2018)
14. Распознавание текста - OCR CuneiForm. – Режим доступа: <http://pro-spo.ru/text/341--ocr-cuneiform> (дата обращения: 05.06.2018)

15. Readiris Pro: что это за программа? – Режим доступа: <http://itfaqs.ru/readiris-pro-cto-eto-za-programma> (дата обращения: 04.06.2018)
16. RiDoc – программа для сканирования документов. [Электронный ресурс]. – Режим доступа: <http://computerologia.ru/ridoc-programma-dlya-skanirovaniya-dokumentov> (дата обращения: 05.06.2018)
17. Программы для распознавания текста. – Режим доступа: <http://softcatalog.info/ru/obzor/programmy-dlya-raspoznavaniya-teksta> (дата обращения: 06.06.2018)
18. Использование Windows Hello на устройствах с предварительной версией Windows. [Электронный ресурс]. – Режим доступа: <https://software.intel.com/ru-ru/articles/how-to-get-working-windows-hello-on-actual-windows-10-insider-preview> (дата обращения: 06.06.2018)
19. Luxand Blink Pro v 2.3 + Rus. [Электронный ресурс]. – Режим доступа: <https://winzoro.net/2011/12/26/luxand-blink-pro-v-23-rus.html> (дата обращения: 06.06.2018)
20. KeyLemon. [Электронный ресурс]. – Режим доступа: <https://store.softline.ru/keylemon/keylemon> (дата обращения: 06.06.2018)
21. Veriface Lenovo: что это за программа? [Электронный ресурс]. – Режим доступа: <https://prolenovo.ru/faq/veriface-lenovo-cto-eto.html> (дата обращения: 06.06.2018)
22. Видеоаналитика: распознавание лиц, детектор очередей, поиск объектов на видео. [Электронный ресурс]. – Режим доступа: <https://geektimes.com/company/ivideon/blog/286168> (дата обращения: 06.06.2018)

OVERVIEW OF SOFTWARE IMAGE RECOGNITION SYSTEMS

The article presents information about the most widely used image recognition systems. Their description, characteristics, a comparative survey with the categories of images, which are best recognized by each system, are given. The final part of the review concludes which system is the most rational for recognizing images of different categories, as well as for photogrammetry and three-dimensional modeling.

Index terms: software systems, image recognition, color image recognition, black and white image recognition, text recognition, face recognition, recognition speed, recognition quality, photogrammetry, three-dimensional modeling.

*Головкин Михаил Евгеньевич,
Краснов Андрей Евгеньевич, 2018*