

УДК 004.932.2

Н.Л. Щеголева

## Разработка программных средств для моделирования биометрических систем распознавания человека по изображению лица

Сформулированы и обоснованы требования к средствам моделирования систем распознавания изображений лиц, принципы их построения и анализа. Предложена архитектура пакета, приведены состав подсистем и их назначение.

**Ключевые слова:** распознавание лиц, предметно-ориентированное моделирование.

В России биометрические системы стали появляться уже в середине девяностых годов. По ряду очевидных причин все они были чрезвычайно дороги. Однако совершенствование цифровых камер, давших возможность получать все более качественные изображения, вычислительной техники, обеспечившей возможность быстрой обработки значительных объемов данных, появление новых методов обработки и анализа, активное развитие мультимедийных и цифровых технологий дали резкий толчок для развития биометрических систем. На сегодняшний день такие системы существенно подешевели, что стало причиной для их широкого распространения. Но это только одна сторона, вторая же причина популярности биометрических систем – это возможность обеспечения безопасности офиса, дома, интеллектуальных и материальных ресурсов и т.д.

Среди существующих биометрических систем стоит особо отметить системы распознавания человека по изображению лица. Они являются наиболее удобными, поскольку эти технологии *ненавязчивы (распознавание происходит на расстоянии, не задерживая и не отвлекая человека), они, как правило, пассивны (не требуют специальных действий со стороны человека), они не ограничивают пользователя в свободе перемещений*» [1].

Однако в работе [2] справедливо указывается, что факторы, влияющие на процесс обработки и анализа изображений, могут быть весьма изменчивыми и неформализуемыми. Изображения лиц, несмотря на их известную структуру, расположение глаз, носа, рта, не являются исключением. Основные факторы, которые усложняют работу системы, – это трудные условия съемки, недостаток освещения, сложный текстурированный фон, сокрытие частей лиц, поворот влево/вправо, наклон головы, а также эмоции, которые испытывает человек в разные моменты времени.

Отсутствие общепринятого формализованного описания перечисленных факторов, вносящих разного рода неопределенности в процесс обработки, приводит к тому, что в отличие от многих других областей обработки сигналов говорить о существовании единственного оптимального алгоритма для решения той или иной задачи обработки изображений (а их огромное количество) практически невозможно.

Последние десять лет в области распознавания человека по изображению лица ведутся активные разработки и предложены различные методы распознавания, использующие такие подходы, как метод главных компонент, методы с использованием гистограмм, нейросети, байесовские сети, классификатор Viola-Jones, статистические методы и т.д. При этом ни один из методов не может быть признан предпочтительнее остальных. Иными словами, ни один из существующих методов не позволяет получить 100% результат.

Однако, как показал опыт, существует некоторая совокупность методов, позволяющая добиться более качественного результата [3]. Это приводит к важной проблеме выбора некоторой совокупности методов из существующих и определения последовательности их применения к имеющимся данным для обеспечения требуемого результата распознавания, т.е. построению системы распознавания человека по изображению лица (Face Recognition Systems – FaReS).

Поэтому дальнейшее повышение качества и эффективности биометрической идентификации может быть достигнуто за счет рационального сочетания методов, оптимизации архитектуры FaReS и максимизации производительности ее аппаратного и программного обеспечения (ПО).

Следует отметить, что в настоящее время основные исследования направлены на совершенствование методов обработки изображений лиц, однако чрезвычайно мало внимания уделяется вопросам, посвященным разработке и оптимизации FaReS как взаимосвязанной последовательности методов предобработки, экстракции/селекции признаков и классификации для решения конкретной задачи. Связано это, в частности, с очень малым количеством полноценных средств моделирования FaReS, что не позволяет добиться максимальной эффективности разрабатываемых систем.

Сложность решаемых задач биометрии требует разработки концепции построения средств моделирования, в рамках которых можно было бы исследовать перспективные структуры систем и их основные модули, ставить натурные эксперименты, получая эффективные решения.

**Структура FaReS.** Процесс обработки изображений лиц состоит из следующих этапов [3] (рис. 1): детекции области каждого лица на исходном цифровом изображении; экстракции признаков – представления изображения выделенного лица (всей выделенной прямоугольной области) в форме вектора исходных признаков; селекции некоторых признаков из полного набора исходных признаков или редукции исходного пространства признаков (СП/РПП); сравнения признаков нового образа (изображения лица) с признаками эталонов и принятия решений о принадлежности этого образа к одному из известных классов.

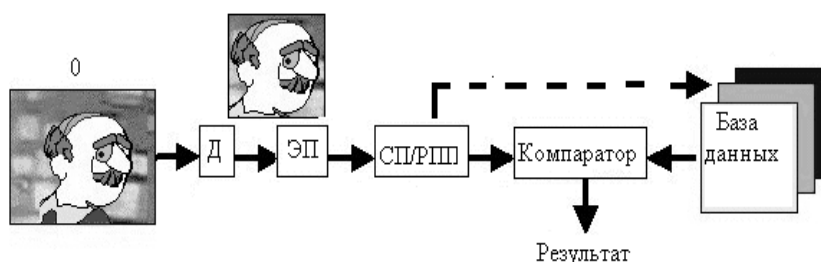


Рис. 1. Структура системы распознавания изображений лиц

Следует заметить, что каждый из перечисленных выше этапов обработки изображений может быть реализован различными способами, состоять из нескольких методов (каждый из которых зависит от ряда параметров), иметь параллельную или каскадную структуру. Все это определяет наличие очень большого количества вариантов FaReS, причем точность распознавания, вычислительная сложность, время обработки данных и т.д. полученных систем могут существенно отличаться. Поэтому основные усилия разработчиков будут сосредоточены на поиске оптимальной структуры FaReS (рис. 2).



Рис. 2. Алгоритм разработки системы [4]

Сложность построения структуры оптимальной FaReS, большое количество параметров и переменных, неполнота и недетерминированность исходной информации, разнообразие и случайный характер воздействий внешней среды – все это обуславливает необходимость моделирования.

**Современные средства моделирования.** Анализ существующего ПО с точки зрения возможности использования его для создания средств моделирования FaReS показывает, что его можно разделить на две группы [4]:

- универсальное ПО (среды программирования на языках высокого уровня – Microsoft Visual Studio, Piton, Eclipse и т.д.);
- специализированные пакеты программ – среды технического моделирования и вычислений (MATLAB, LabVIEW и т. д.) и специализированные программные средства с фиксированным количеством модулей (National Instruments Vision Assistant [5], FaReS-MOD (Face Recognition System Modeler) [6] и FaRetSys (Face Retrieval System) [7], Fabio [8]).

Очевидно, что использование универсального ПО будет сопряжено с большим количеством проблем: необходимо наличие грамотных специалистов по обработке изображений, времени, поскольку программирование всех необходимых алгоритмов обработки изображений лиц весьма трудоемкий процесс, написание огромного объема сервисного ПО для визуализации изображений, обеспечения интерактивного ввода параметров, векторной графики и т.п. Поэтому данный вариант представляется не самым лучшим решением.

Преимуществом большинства сред технического моделирования и вычислений являются универсальность и широкий набор сторонних библиотек математической статистики и цифровой обработки изображений, однако в них отсутствует возможность написания удобного и развитого интерфейса.

Альтернативой является использование готового специализированного ПО для моделирования FaReS. Рассмотрим существующие решения более подробно.

Среда Vision Assistant не является полноценным средством моделирования систем распознавания, т.к. ориентирована на решение задач пакетной обработки изображений, позволяет создавать в интерактивном режиме только линейные вычислительные модели без возможности классификации.

Среда FaReS-MOD является самой функциональной. Имеющийся в ней набор модулей позволяет проектировать системы, содержащие все этапы распознавания от предобработки до классификации, но только линейной структуры. В ней детально проработаны методология построения систем распознавания, настройка систем и получение результатов в процессе моделирования.

Среда FaRetSys во многом аналогична предыдущей, но позволяет реализовать каскадную и параллельную структуру для построения эффективных, в плане качества распознавания и быстродействия, систем на базе простых классификаторов. Возможно расширение функционала пользовательскими модулями. Недостатками среды являются малый базовый функционал, невозможность изменения параметров модулей системы.

Программное средство Fabio в значительной степени аналогично представленным выше средам, имеет модульную структуру, основной особенностью является использование для реализации графического редактора моделей и интерфейсной части Microsoft Visual Studio 2010, а для реализации функций, предназначенных для анализа изображений, – пакет Matlab. Возможно расширение функционала пользовательскими модулями, написанными на языке пакета Matlab. Недостатками являются ограниченный набор методов и линейная структура системы.

**Описание модели эксперимента.** Необходимо отметить, что в статьях, посвященных распознаванию изображений лиц, модели экспериментов авторами описываются различным образом, что часто затрудняет понимание, каким образом проводился эксперимент и какие параметры были выбраны для конкретных методов, что не позволяет исследователю повторить эксперимент и корректно сравнить результаты. В связи с этим необходимо унифицировать описание экспериментов.

Большинство исследователей предпочитают описание в форме таблицы, поэтому рассмотрим способ компактной и однозначной записи модели [9] и ее связь с табличным описанием. Пусть описание эксперимента выполнено в форме таблицы, дополнительно известно, что классификация выполнялась по минимуму расстояния, использовалась процедура кросс-валидации, размер изображения 112×92. Очевидно, что эта важная информация должна быть известна, но весьма часто авторы ее не указывают в описании эксперимента.

**Описание эксперимента**

База данных	Число классов (K)	Число образов в классе (Q)	Количество эталонов в классе (L)	Количество тестовых образов (Q–L)	Метод экстракции признаков	Метрика	Ранг
ORL	40	10	1	9	2DPCA	L1	1

Тогда модель эксперимента в соответствии с [9] будет иметь вид:

$$\text{ORL}(40/1/9/CV) \{2\text{DPCA}:112 \times 92 \rightarrow d \times d / \text{KMP}/L1/\text{rank}=1\},$$

где определены:

ORL(40/1/9/CV) – база лиц, структура данных и процедура кросс-валидации (CV);

2DPCA:112×92→d×d – использованный метод и размеры данных до и после редукции размерности пространства признаков;

KMP/L1/rank = 1 – тип классификатора, метрика и ранг результата.

Данное описание является более компактным и содержит все данные эксперимента.

**Требования.** Анализ работ посвященных разработке систем распознавания, показал, что требования, предъявляемые к программным средствам, предназначенным для моделирования FaReS, могут быть разделены на:

1. Общие:

- удобство использования,
- форматы используемых данных,
- возможность использования различных тестовых баз данных изображений,
- организация структуры программного средства,
- функциональный состав модулей,
- возможность интерактивно формировать из имеющихся готовых блоков любые системы обработки без использования какого-либо специального командного языка,
- описание модели эксперимента,
- сохранение структуры и параметров моделей систем и возможность их загрузки для повторения эксперимента или дальнейшей работы,
- возможность корректировки параметров используемых модулей после тестирования FaReS,
- возможность осуществлять автоматическую обработку данных в моделируемой системе при их обновлении,
- оценка качества разработанной системы.

2. Инженерные:

- скорость обработки данных,
- возможность использования видеозаписей,
- возможность оптимизации ранее разработанных систем,
- возможность получения готового программного продукта,
- построение гибкой структуры моделируемых FaReS.

3. Исследовательские:

- анализ состава и характеристик пространства признаков, представляющих информацию о лицах,
- поиск минимального размера пространства признаков для представления лиц конкретной базы данных,
- формирование отчетов в зависимости от целей исследования,
- возможность добавления новых функциональных модулей,
- сравнительный анализ различных структур FaReS,
- возможность использования разработанных FaReS для анализа изображений другой природы,
- автоматический выбор оптимальной структуры среди заданных пользователем.

**Обоснование требований.** Рассмотрим данные требования более подробно. Заметим, что удобство использования поставлено первым, т.к. разработка биометрической системы идентификации человека по изображению лица является весьма сложным процессом, требующим от разработчика максимум времени и внимания, а это означает, что эргономические показатели, предъявляемые к данному виду ПО, будут весьма высоки.

Все перечисленные выше специализированные пакеты используют визуальное программирование для моделирования FaReS, процесс построения которой [3] состоит в выборе функциональных блоков (рис. 3), реализующих методы обработки, выборе их параметров, установлении между ними связей (рис. 4). Это позволяет уделять основное внимание вопросам, связанным с разработкой и оптимизацией систем распознавания как взаимосвязанной последовательности методов предобработки, экстракции/селекции признаков и классификации. Интерактивный характер разработки должен обеспечивать возможность анализа системы, а также модификации проекта FaReS – его структуры и параметров произвольно выбранных блоков.

Сохранение в проекте FaReS иерархии блоков и соединений модели, включая их настройки и информацию для отображения, может быть реализовано с помощью файла, имеющего специальную структуру [6–8].

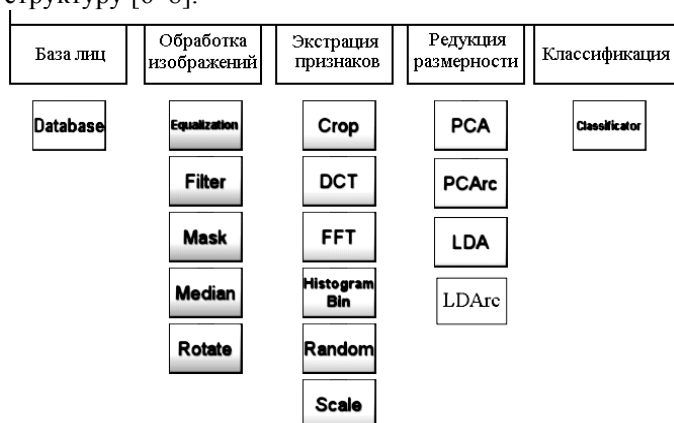


Рис. 3. Функциональные блоки, используемые в пакете FaReS-MOD

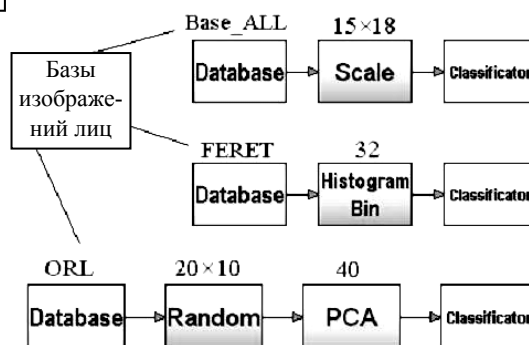


Рис. 4. Примеры FaReS, построенных в среде FaReS-MOD

Программное средство для моделирования должно обеспечивать возможность обработки наиболее популярных файловых форматов, загрузку, обработку, просмотр и сохранение бинарных, полутоновых и цветных изображений, возможность использования различных тестовых баз данных изображений, а также любых двумерных и одномерных числовых массивов, как целочисленных, так и вещественных.

В работе [4] обосновывается необходимость использования модульной структуры (рис. 5).

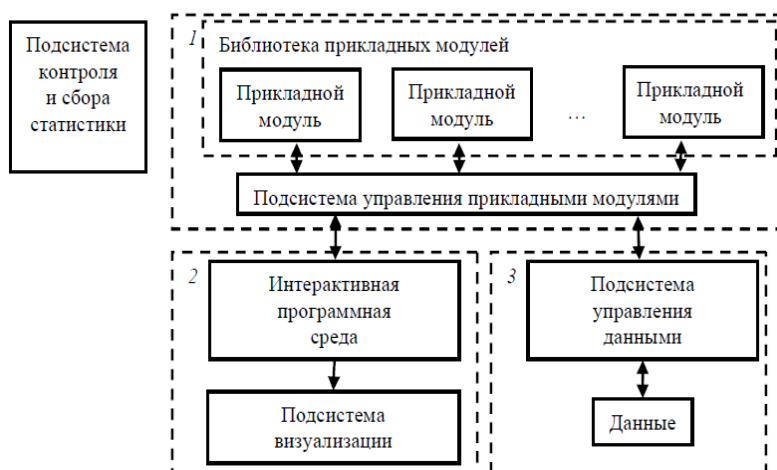


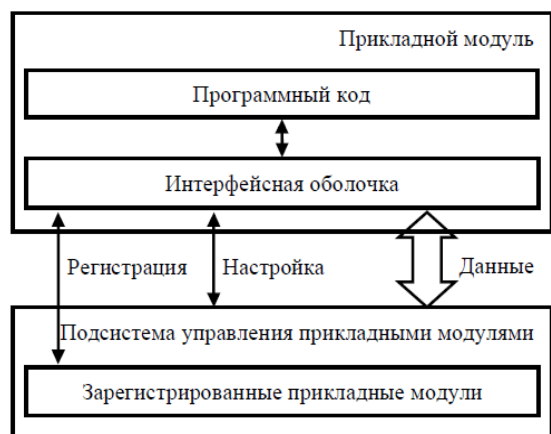
Рис. 5. Модульная структура программного средства [4]

В предложенной архитектуре выделяются три основные части, что позволяет разделить между собой систему ввода/вывода 3, интерактивную программную среду 2 и библиотеку прикладных модулей 1, что улучшает масштабируемость, сопровождаемость и надежность системы в целом. Модульная архитектура обеспечивает гибкость и расширяемость системы без существенного увеличения трудозатрат на ее реализацию.

Учитывая специфику предметной области, авторы выделяют следующие группы модулей: модули для работы с базами изображений; модули предобработки изображений; модули экстракции/селекции признаков; модули классификации; модули отображения результатов.

Все прикладные модули реализуют единообразный интерфейс взаимодействия кода модуля с другими подсистемами (рис. 6). Подсистема управления прикладными модулями осуществляет настройку модулей и предоставляет доступ к их функционалу, взаимодействует с интерактивной программной средой.

Очень важным при проектировании FaReS является оценка качества ее работы на конкретных данных. Два наиболее важных показателя: FAR (False Acceptance Rate) – процентный порог, определяющий вероятность того, что один человек может быть принят за другого (коэффициент ложного доступа) и FRR (False Rejection Rate) – вероятность того, что человек может быть не распознан системой (коэффициент ложного отказа в доступе). Однако для более полного анализа протокол результата моделирования должен сохранять информацию о структуре исходных данных, визуальную и текстовую информацию о проведенном процессе распознавания, реакцию системы (результат распознавания) на каждый предъявленный тестовый образ. Для визуализации данных должны исполь-



зоваться таблицы, 3D-графика, различного рода проекции, позволяющие оценить влияние используемого признакового пространства, его размерности, структуры системы на качество распознавания. Обязательно наличие временных оценок, которые позволяют сделать вывод о возможности работы в режиме реального времени.

Рис. 6. Интерфейс взаимодействия кода модуля с другими подсистемами [4]

Всех пользователей средств для моделирования FaReS можно условно разделить на три категории: студенты, которым необходимо изучить основы систем идентификации, инженеров, которые занимаются разработкой реальных систем, предназначенных для обеспечения безопасности, и ученых, основной задачей которых является разработка новых более эффективных и быстрых алгоритмов.

С этой точки зрения для инженеров наиболее актуальным будет следующее: возможность использования при моделировании видеозаписей, возможность получения готового программного продукта на основе смоделированной системы, т.е. фактически для них необходим модуль, который формирует уникальное приложение для отдельного заказчика и разработанного для конкретной базы данных с учетом всех ее особенностей – освещения, положения камеры и т.д.

Все FaReS можно разделить на простые, использующие быстрые алгоритмы экстракции признаков, не требующие больших объемов вычислений, и сложные – использующие методы редукции размерности пространства признаков, требующих существенных затрат вычислительных и временных ресурсов.

В настоящее время основным направлением обеспечения безопасности становится применение новых систем наблюдения, использующих значительное количество (от 30 и более) цифровых видеокамер. В средствах массовой информации регулярно появляются сообщения о разработке или испытании систем городского видеонаблюдения – сети из сотен/тысяч камер, установленных в больших городах на улицах и площадях, в аэропортах и на вокзалах, а также в общественном транспорте.

Для анализа информации, поступающей с большого количества камер, необходима разработка новых быстродействующих систем распознавания, которые смогут обеспечить работу в режиме online, ориентированных на большие объемы исходных данных, а также баз с быстрой динамикой изменения их состава.

Реализация данного рода систем возможна только при условии применения аналитики на видеокамере [10], т.к. стандартный средний поток, выдаваемый IP-камерой, 2 мбит/с. Если камера мегапиксельная, то уже оперируют потоками более чем в 10 мбит/с, причем количество камер может быть и 100, и 1000, а сетевое оборудование не может работать со 100%-й загрузкой (для видеопотоков максимум 50%). Существует еще одна причина, по которой видеоаналитика должна быть установлена на источнике видеосигнала (на IP-камере или на кодере), – любая компрессия с целью сокращения передаваемого объема данных ухудшает качество изображения, а значит, и качество работы системы идентификации.

Поскольку вычислительные ресурсы видеокамер на данный момент ограничены, то реализовать вычислительно сложные алгоритмы в самих видеокамерах не представляется возможным, однако использование в них простых алгоритмов позволит существенно уменьшить загрузку сети, используя для более точного и сложного анализа только критически важную информацию, что позволит не только повысить уровень безопасности объекта, но и снизить стоимость системы. Следует отметить, что в данном случае FaReS (или некоторая ее подсистема) будет иметь параллельную структуру.

Объем информации, хранящейся в базах данных, постоянно растет, и поиск в них нужной информации становится все более сложным, поэтому и в этом случае необходимо использовать более гибкую структуру системы – комбинацию простых и сложных методов, поскольку время поиска становится все более значительным. Решение данной проблемы предложено в [11].

Применение унимодальных систем показало, что у них есть ряд недостатков: увеличение сложности вычислений, падение качества распознавания при росте количества распознаваемых людей, относительно слабая защищенность против фальсификаций. Поэтому наиболее перспективным решением ближайшего будущего является применение мультимодальных биометрических систем. Использование различной информации об одном и том же лице человека «Multisampling / Multi-imaging» (представление изображения лица одного и того же человека через отдельные его фрагменты) [12] и «Multirepresentation» (представление изображения лица одного и того же человека в форме различных признаков: яркостных, гистограммных, спектральных или в форме изображений, полученных различными способами: в видимом свете, при инфракрасном освещении и т.д.) [13, 14] – один из самых привлекательных способов повышения эффективности биометрических систем. Однако за рост качества распознавания разработчикам приходится платить усложнением структуры системы.

Все вышеизложенное делает очень важным возможность программного средства для моделирования FaReS проектировать системы, имеющие каскадные, параллельные или комбинированные структуры, а также комплексировать информацию на разных уровнях. Отсюда же вытекает и необходимость наличия нескольких классификаторов для моделирования FaReS – с помощью функций расстояния, Байесовский, Ada-Boost, голосование и т.д.

Заметим, что исследователи, занимающиеся разработкой новых методов идентификации, при использовании моделирования более других категорий пользователей заинтересованы в наличии богатого инструментария для анализа результатов экспериментов – им необходимы самые продвинутые средства визуализации: возможность построения проекций, трехмерных графиков, таблиц. Кроме того, им необходим модуль, который освободит их от рутинных операций по непрерывному изменению параметров в некоторых пределах с целью поиска оптимальных значений, т.е. должна быть возможность описания вычислительного эксперимента для заданной структуры и определения критерия качества полученных результатов.

**Заключение.** Анализ существующих специализированных средств с точки зрения изложенных выше требований показывает, что в наибольшей степени им удовлетворяет пакет FaReS-MOD, который позволяет, не прибегая к классическому программированию, быстро поставить эксперимент, исследовать систему распознавания лиц и сравнить ее результативность с другими, ранее созданными системами. Однако необходимо дополнение его рядом модулей, в частности, позволяющих строить FaReS, имеющих более сложную и гибкую структуру.

В статье сформулированы и обоснованы требования к средствам моделирования систем распознавания изображений лиц, принципы их построения и анализа. Предложена архитектура пакета, приведены состав подсистем и их назначение, что позволит синтезировать системы распознавания человека по изображению лица, имеющие различную архитектуру; существенно сократить затраты и снизить требования к квалификации экспертов при проектировании систем для решения многих практических и исследовательских задач биометрии; осуществлять тестирование и параметрическую оптимизацию систем распознавания изображений лиц с использованием стандартных и специальных баз данных; выполнять тестирование новых методов обработки изображений и сравнение их эффективности с существующими; исследовать перспективные варианты построения систем и оценивать или прогнозировать их точностные характеристики.

#### *Литература*

1. Пентланд А. Распознавание лиц для интеллектуальных сред // Открытые системы. – 2000. – № 3. – С. 17 – 20.
2. Обработка и анализ изображений в задачах машинного зрения / Ю.В. Визильтер, С.Ю. Желтов, А.В. Бондаренко и др. – М.: Физматкнига, 2010. – 672 с.
3. Кухарев Г.А. Системы распознавания человека по изображению лица / Г.А. Кухарев, Н.Л. Щеголева. – СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2006. – 176 с.
4. Хомяков М.Ю. Принципы построения пакета программ для моделирования систем распознавания изображений лиц / М.Ю. Хомяков, Г.А. Кухарев // Известия СПбГЭТУ «ЛЭТИ». – 2010. – № 7. – С. 41–46.
5. NI Vision Assistant Tutorial [Электронный ресурс]. – Режим доступа: <http://www.ni.com/pdf/manuals/372228a.pdf>, свободный (дата обращения: 21.01.2013).

6. Kukharev G. An environment for recognition system modeling / G. Kukharev, A. Kuzminski // Enhanced methods in computer security, biometric and artificial intelligence systems. – New York; USA: Kluwer Academic Press, Springer, 2005. – P. 157–164.
7. Kukharev G. Strategia budowy systemow klasy «Face Retrieval» / G. Kukharev, M. Mikiasz, B. Taudal // Metody Informatyki Stosowanej. – 2007. – Т. 12. – S. 61–72.
8. Щеголева Н.Л. Программный комплекс для моделирования систем распознавания изображений лиц // Матер. XIV Междунар. конф. по мягким вычислениям и измерениям (SCM'2011). – СПб. – 2011. – С. 271–274.
9. Kukharev G. FaRetSys – the environment for modeling Face Retrieval tasks / G. Kukharev, M. Miklasz, B. Tauda // Metody Informatyki Stosowanej. – 2007. – Т. 12, № 2. – S. 61–73.
10. Видеоаналитика: потребности и возможности // Журнал ТЗ. – 2010. № 2 [Электронный ресурс]. – Режим доступа: <http://www.tzmagazine.ru/jpage.php?uid1=516&uid2=543&uid3=546>, свободный (дата обращения: 21.01.2013).
11. Кухарев Г.А. Поиск изображений лиц в больших базах данных // Мир измерений. – 2009. – № 4. – С. 22–30.
12. Price J.R. Face recognition using direct, weighted linear discriminant analysis and modular subspaces / J.R. Price, T.F. Gee // Pattern Recognition. – 2005. – № 38. – P. 209–219.
13. Rajagopalan A.N. Face recognition using multiple facial features / A.N. Rajagopalan, K. Srinivasa Rao Karri, Y. Anoop Kumar // Pattern Recognition Letters. – 2007 – № 28. – P. 335–341.
14. Kamenskaya E. Some aspects of automated recognition of person's psychological characteristics from the facial image / E. Kamenskaya, G. Kukharev // Metody informatyki stosowanej. – 2008. – Vol. 15, №2. – P. 29–37.

---

**Щеголева Надежда Львовна**

Канд. техн. наук, доцент каф. математического обеспечения ЭВМ СПбГЭТУ, г. Санкт-Петербург

Тел.: 8 (812) 234-26-82

Эл. почта: NLSchegoleva@etu.ru

Shchegoleva N.L.

**Development of software for face recognition system modeling**

In the work we considered functions, composition and structure of a software tool designed to graphical modeling of face recognition systems. We developed a software tool which allows you to investigate methods, algorithms and structures of a processing system.

**Keywords:** face recognition, domain-specific modeling.

---