

S. O. Shaposhnikov
Saint Petersburg Electrotechnical University «LETI»

E. Yu. Yatkina
National Research Tomsk Polytechnic University

ON FORMULATING AND CORRECTING DEGREE PROGRAM AIMS

Presents results of analysis of engineering degree programs aims description based on examples from Russian and foreign universities. Requirements to the description of program aims are formulated. A method is proposed for checking the correspondence between the program aims and the requirements to their description.

Engineering degree programs development, program aims and learning outcomes, public professional accreditation of engineering degree programs

УДК 004.932.2

Н. Л. Щеголева
Санкт-Петербургский государственный электротехнический
университет «ЛЭТИ» им. В. И. Ульянова (Ленина)

Концепция построения комплекса программных средств для моделирования систем поиска изображений лиц

Обсуждается проблема разработки систем распознавания изображений лиц. Приводится обзор состояния этой проблемы. Предложена концепция построения комплекса программных средств для моделирования FaRetSys, в рамках концепции определяется состав компонентов, их функции и методика использования комплекса при решении задач моделирования. Уточнена модель компактного описания компьютерного эксперимента, отражающая состав и параметры блоков моделируемых систем, отличающаяся возможностью описания параллельных, каскадных и комбинированных систем.

Распознавание лиц, моделирование систем поиска изображений лиц, программные средства

Развитие вычислительной техники и математических методов обработки изображений лиц (ИЛ) сделало возможным разработку систем поиска людей по ИЛ (FaRetSys – Face Retrieval System), обеспечивающих результат, близкий к 100 %. Однако существует несколько объективных факторов, определяемых сценарием использования FaRetSys, влияющих на результативность при обработке реальных изображений лиц. Перечислим основные:

– изменение (перемещение источника) освещения при движении человека в зависимости от времени суток, которое приводит к появлению теней на лице, меняя форму и пропорции лица и не позволяя точно указать расположение черт лица;

– ракурс съемки: захват изображения движущегося человека происходит с разных углов, что вызывает искажение или изменение пространственных соотношений между чертами лица;

– различные выражения лица, используемые человеком для выражения своих чувств и эмоций, изменение психофизиологического состояния человека;

– возрастные изменения лица;
– появление новых модных аксессуаров (очки, пирсинг, тату);

– макияж (камуфляж), изменяющий части лица, его пропорции;

– сокрытие части лица вследствие изменения имиджа (борода, усы, шляпа);

– постоянное увеличение объема анализируемых баз изображений;

– изменение/обновление баз изображений (динамические базы изображений);

– временные и аппаратные ограничения.

Сложность объекта исследования (изображения лица) и разнообразие сценариев использования привели к тому, что на сегодняшний день для

их обработки создано большое количество различных методов. Это ставит разработчиков перед необходимостью выбора одного подхода из множества существующих и оценки применимости каждого метода для решения конкретной задачи.

Наиболее значимой идеей, позволяющей в некоторой степени решить проблему выбора и повысить результативность FaReS, стало коллективное распознавание: использование нескольких классификаторов, каждый из которых принимает решение о принадлежности к классу одного и того же ИЛ, с последующим объединением (фузией) их решений с помощью некоторого алгоритма.

Начало исследований коллективного распознавания было положено работами Л. А. Растригина и Р. Х. Эренштейна [1], [2]. Исследование данной проблемы зарубежными учеными было начато позднее – в конце 90-х.

Исследователи считали, что объединение решений нескольких хороших классификаторов обеспечит наиболее точное решение, принимаемое системой в целом, т. е. необходима оптимизация классификаторов [3]. Однако практика опровергла данное предположение, и была поставлена задача «оптимизации в целом», которая стала развитием подхода, предложенного Л. А. Растригиным и Р. Х. Эренштейном.

Методы оптимизации в целом делятся на две группы [3]:

- выбора компетентного классификатора;
- объединения решений классификаторов.

В задачах поиска ИЛ выбор компетентного классификатора реализуется в известных методах «бут-страпинг» (bootstrap aggregation), или бэгинга (bagging), бустинга (boosting).

Значительно шире используются методы объединения решений классификаторов, а среди них – методы многоярусного обобщения. Популярность этих методов обусловлена следующими причинами [4]:

1. *Статистическая.* Классификационный алгоритм можно рассматривать как процедуру поиска в пространстве гипотез H наилучшей гипотезы. Обучаясь на ограниченном наборе данных, алгоритм может найти множество различных гипотез, одинаково хорошо описывающих обучающую выборку. Строя ансамбль (т. е. используя несколько алгоритмов), «усредняется» ошибка каждой индивидуальной гипотезы и уменьшается риск выбора неверного решения.

2. *Вычислительная.* Большинство обучающих алгоритмов используют методы нахождения экстремума некоторой целевой функции, который

может оказаться локальным. Ансамбли, осуществляющие поиск из различных начальных точек, могут обеспечить лучшую аппроксимацию неизвестной функции, чем один классификатор.

3. *Репрезентативная.* Искомая функция не может быть представлена ни одной из гипотез в H . Формирование взвешенной суммы гипотез позволит расширить пространство гипотез для представления функции.

Способ объединения классификаторов может быть каскадным, параллельным, иерархическим. Примером иерархического объединения является нейронная сеть (в данной статье не рассматривается). Каскад представляет собой последовательное соединение классификаторов. Параллельный способ предполагает одновременную работу нескольких классификаторов, решения которых объединяются посредством голосования, взвешенного голосования и т. д. [5], [6].

Результаты исследования эффективных методов построения некоторой совокупности (ансамбля) классификаторов [5], [6] показали, что необходимо использовать *разнообразные* классификаторы (Diversity in Classifier Ensembles), т. е. такие, которые обеспечивают некоррелированность ошибок классификации. Отсутствие общепринятого формального определения разнообразия стало причиной появления разного рода метрик для его оценки.

До 2003 г. проблема оценки разнообразия классификаторов достаточно активно обсуждалась в научной литературе и считалась одним из ключевых вопросов при построении ансамбля классификаторов. Однако в работе [5] были изучены 10 статистик измерения разнообразия бинарных классификаторов (Qstatistic, correlation, disagreement and the double fault, entropy of the votes, difficulty index, Kohavi-Wolpert variance, interrater agreement, generalized diversity, coincident failure diversity). Анализ полученных результатов привел к неожиданному выводу: несмотря на наличие связи между разнообразием классификаторов и точностью результата (в некоторых случаях), авторы *ставят под сомнение полезность применения мер разнообразия* при формировании ансамбля классификаторов для решения задач поиска и распознавания в реальных системах.

Здесь необходимо сделать 2 замечания: 1) в публикациях, посвященных обработке ИЛ, достаточно часто под классификатором понимается комбинация экстракции признаков (или селекции/редукции) и классификатора, что неверно; 2) для систем, использующих объединение реше-

ний классификаторов, авторами вместо описания структуры FaRetSys вводятся новые термины, например combination of multiple classifiers, classifier fusion, mixture of experts, voting pool of classifiers, committee machines, мультиклассификаторы и т. д.

Очевидно, что это связано с разными формулировками решаемых задач, предпосылок, на которых основано данное объединение классификаторов и их типов, способы получения итогового решения и т. д. Однако это порождает терминологическую проблему, что в свою очередь усложняет метаанализ и формулировку значимых выводов и рекомендаций.

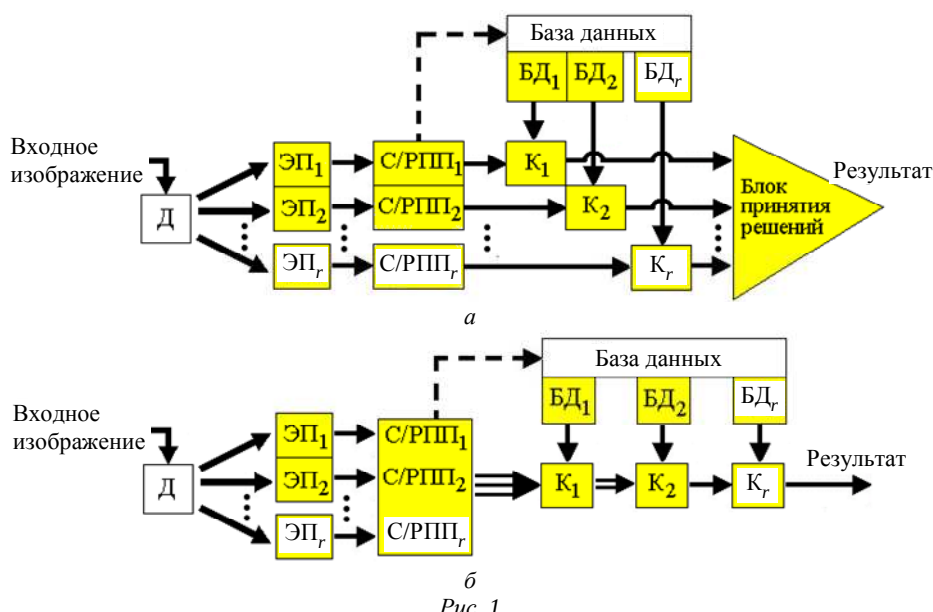
В связи с этим для исследования характеристик FaRetSys экстракция, селекция/редукция признаков и классификация должны рассматриваться как отдельные блоки, причем особый акцент следует делать на используемую структуру системы и ее влияние на результат решения задач лицевой биометрии.

Структуры FaRetSys. Впервые работы, акцентирующие внимание на структуре FaRetSys, появились примерно в 2000 г. В них предлагались каскадная и параллельная структуры, которые позже были широко использованы другими исследователями. Общая структура систем была представлена в [7], [8] и изображена на рис. 1: *а* – параллельная и *б* – каскадная. Здесь приняты следующие обозначения: Д – блоки детекции лиц; ЭП – блоки экстракции признаков; С/РПП – блоки, реализующие селекцию признаков и/или редукцию пространства признаков; К – классификатор; БД – блоки, хранящие образы (эталоны) для различных категорий признаков.

Каскадная структура представляет собой последовательно соединенные комбинации блока экстракции признаков для представления ИЛ и классификатора. На каждом этапе обработки, начиная со второго, выполняется верификация решений, принятых на предыдущем каскаде. Каскадная структура используется для сокращения объема анализируемых изображений за счет постепенного отсева изображений, не удовлетворяющих требованиям, и последующей обработки более сложными и точными методами наиболее вероятных объектов для получения требуемого результата. В англоязычных публикациях для обозначения используется термин Cascade of Rejection Classifiers [9]–[11].

Каскадная схема позволяет реализовать быстрый и эффективный поиск в больших БД, последовательно сокращая базу данных, используемую для анализа, тем самым делая поиск более быстрым и эффективным. Таким образом, каскадные системы более удобны для пользователя, поскольку, как правило, анализ требует меньшего количества времени по сравнению с параллельной структурой. Следует отметить, что недостатком является необходимость переобучения при изменении (добавлении/удалении) состава блоков системы.

К сожалению, в большинстве опубликованных работ полностью отсутствует обоснование выбора тех или иных методов ЭП и особенно их места в каскаде. Также не поясняется цель преобразований, какие полезные свойства приобретают данные после выполненных трансформаций. Из этого следует, что составляются данные каскады



чисто эмпирически, на основе опыта исследователей и, возможно, анализа результатов ряда экспериментов по использованию и статистической (в лучшем случае) оценке результатов их различных комбинаций.

Наиболее перспективными комбинациями, по мнению автора, являются последовательное применение методов представления ИЛ в ИПП или методов представления ИЛ в ортогональном базисе и методов представления ИЛ в подпространствах, позволяющих реализовать и редукцию пространства признаков (2D PCA, 2D LDA, 2D CCA, 2D PLS). Это обусловлено тем, что в данном случае понятен смысл выполняемых преобразований – для последующего анализа выбираются наиболее значимые информативные признаки и существенно сокращается размерность пространства признаков.

Параллельная структура. Системы с параллельной организацией структуры содержат несколько комбинаций блока, реализующего один из методов представления ИЛ и классификатора, одновременно выполняющих обработку изображений. Результат формируется в блоке принятия решений на основе мажоритарных принципов или некоторых логических правил.

Параллельная система в целом имеет более высокую точность, поскольку она использует больше данных – расширяется пространство представлений изображения; рассматривается большее количество гипотез, совместное использование которых делает возможным «усреднение» ошибки каждой отдельной гипотезы, уменьшая риск выбора неверного решения; поскольку большинство методов поиска основаны на нахождении экстремума некой целевой функции, который может оказаться локальным, в то время как поиск экстремума из различных начальных точек может обеспечить лучшие результаты. Параллельные системы позволяют использовать готовые решения, имеющие наилучшие показатели по точности и скорости, а также обеспечивают простоту реализации и трактовки полученных результатов.

Этот подход очень часто применяют в *мультимодальных биометрических системах*. Популярность данного подхода связана с использованием ряда коммерческих продуктов, разработанных по принципу COTS-технологии (Commercial Off-The-Shelf – готовые к использованию). Использование таких продуктов позволяет более гибко выбирать производителя, увеличить количество, а также сложность/научеваемость разрабатываемых приложений для решения разных задач

биометрии, одновременно сокращая финансовые и временные затраты, однако большинство предлагаемых в литературе систем имеют параллельную структуру, потому что основной целью разработчиков было стремление снизить процент ошибок биометрических систем, а не увеличить пропускную способность системы и/или сократить время обработки. В связи с этим системы данной структуры могут быть менее удобны для пользователя.

К сожалению, в настоящее время основные исследования направлены на совершенствование методов обработки ИЛ, однако чрезвычайно мало внимания уделяется вопросам, посвященным разработке и оптимизации системы в целом. Это приводит к тому, что более точное решение задачи обеспечивается за счет *повышения сложности системы*, перенося, таким образом, решение проблемы точности на другой уровень. В связи с этим авторы [6], [11] отмечают, что опасность заключается в том, что ряд исследователей может вместо поиска метода и набора его параметров, наилучшим образом решающих задачу обработки изображений, пытаться найти лучший набор методов, а затем их лучшее сочетание.

В связи с этим правильным является поиск системы, имеющей *наиболее простую структуру* и решающей задачу с максимальной точностью, а *разнообразие необходимо обеспечивать за счет некоррелированности, независимости признаков*, полученных в блоках экстракции признаков. В противном случае постоянное усложнение системы без изучения всех характеристик, возможностей и ограничений реализованного метода, обоснования его использования, места в структуре системы и полученных преимуществ приводит к бесконечному возрастанию сложности системы и временных затрат на разработку, тестирование и модификацию системы.

Все изложенное делает разработку эффективной FaRetSys нетривиальной задачей, требующей не только хорошего знания методов обработки ИЛ, но большого опыта в проектировании и программировании такого рода систем. Чтобы решить эту проблему, необходима разработка программной системы для моделирования FaRetSys.

Концепция построения комплекса программных средств для моделирования FaRetSys. В основе концепции построения комплекса программных средств для моделирования FaRetSys лежат следующие базовые принципы:

1. Вербальное описание эксперимента, включающее исходные предпосылки и цели эксперимента.

2. Описание использованной базы данных и выбранных параметров базы.

3. Запись эксперимента в упрощенной форме, представляющей модель FaRetSys.

4. Представление результатов выполненных экспериментов (скриншоты, таблицы, рисунки).

5. Интерпретация полученных результатов.

Рис. 2 иллюстрирует предлагаемую концепцию, учитывающую требования [12] и современные тенденции развития биометрических систем и компьютерных технологий.

Архитектура и состав компонентов предлагаемого комплекса программных средств для моделирования FaRetSys показаны на рис. 3. Здесь приняты следующие обозначения: Э и СП – блок, реализующий экстракцию и селекцию признаков; ПиР – блок, реализующий редукцию простран-

ства признаков. База эталонов хранит образы (эталонные) для различных категорий признаков. При этом методы проекции и РППП могут быть реализованы, в том числе, в рамках облачных технологий, а в простых системах блок ПиР может отсутствовать.

Построение FaRetSys представляет собой итерационный процесс модификации заданной структуры системы для определения ее наилучших параметров на основе решения вариационной задачи и сравнительной оценки результативности разных вариантов FaRetSys. Методику использования программного средства поясняет рис. 4.

При этом метод проекции в собственные подпространства выбирается в зависимости от числа наборов исходных данных, структуры каждого набора, цели решаемой задачи, выбранного критерия и способа реализации метода проекции (1D, 2D, параллельный, каскадный).

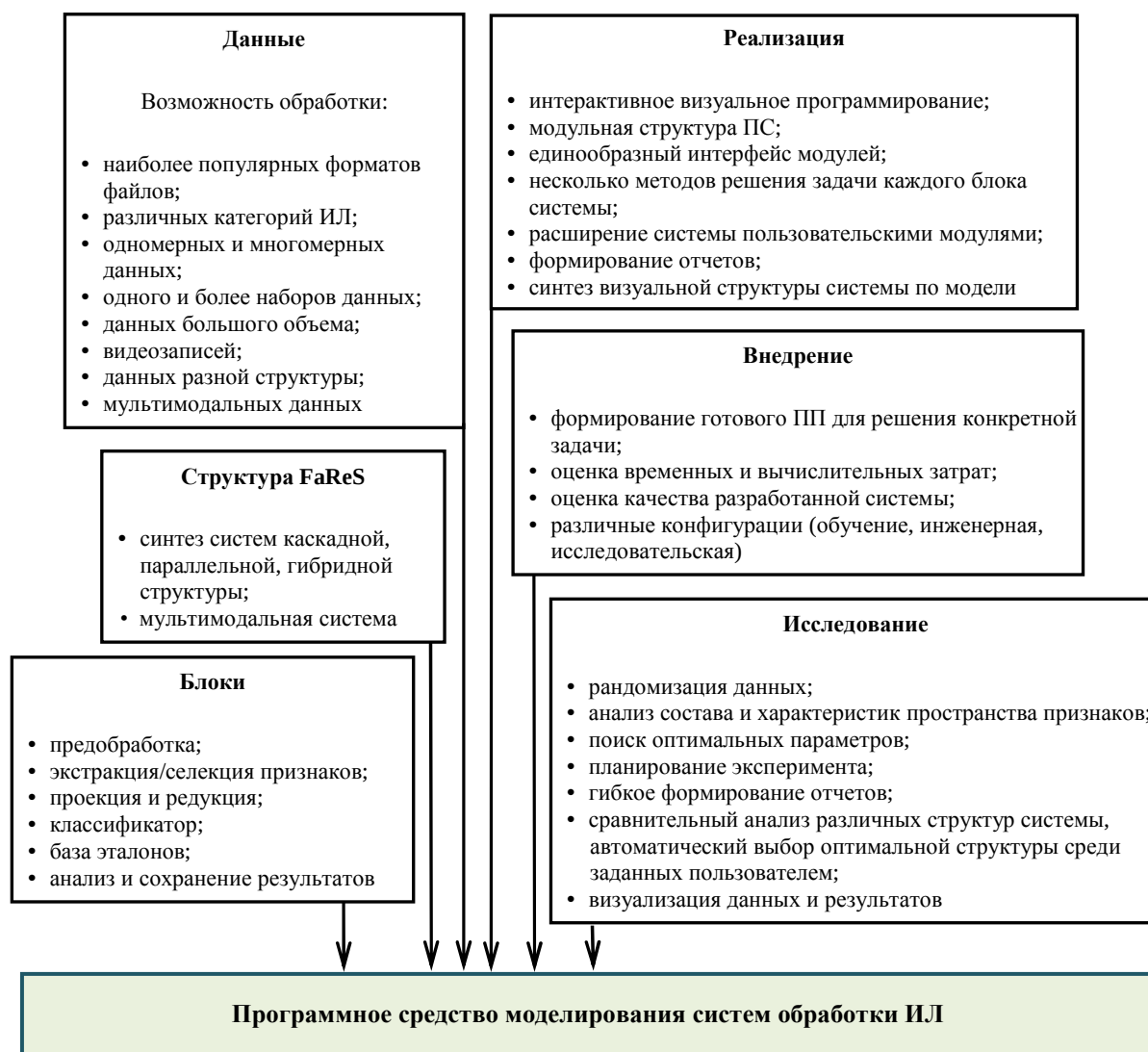


Рис. 2

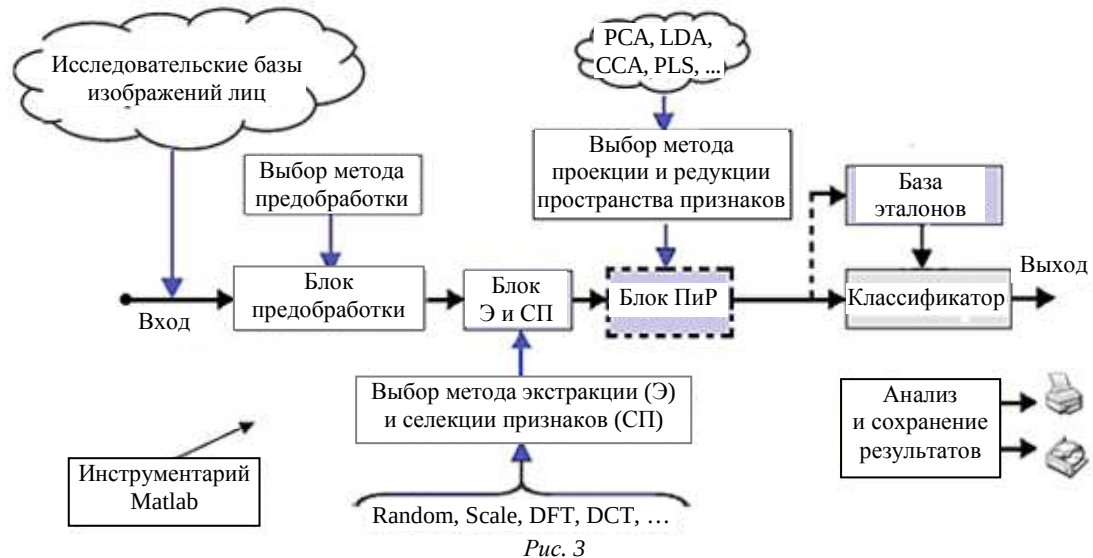


Рис. 3

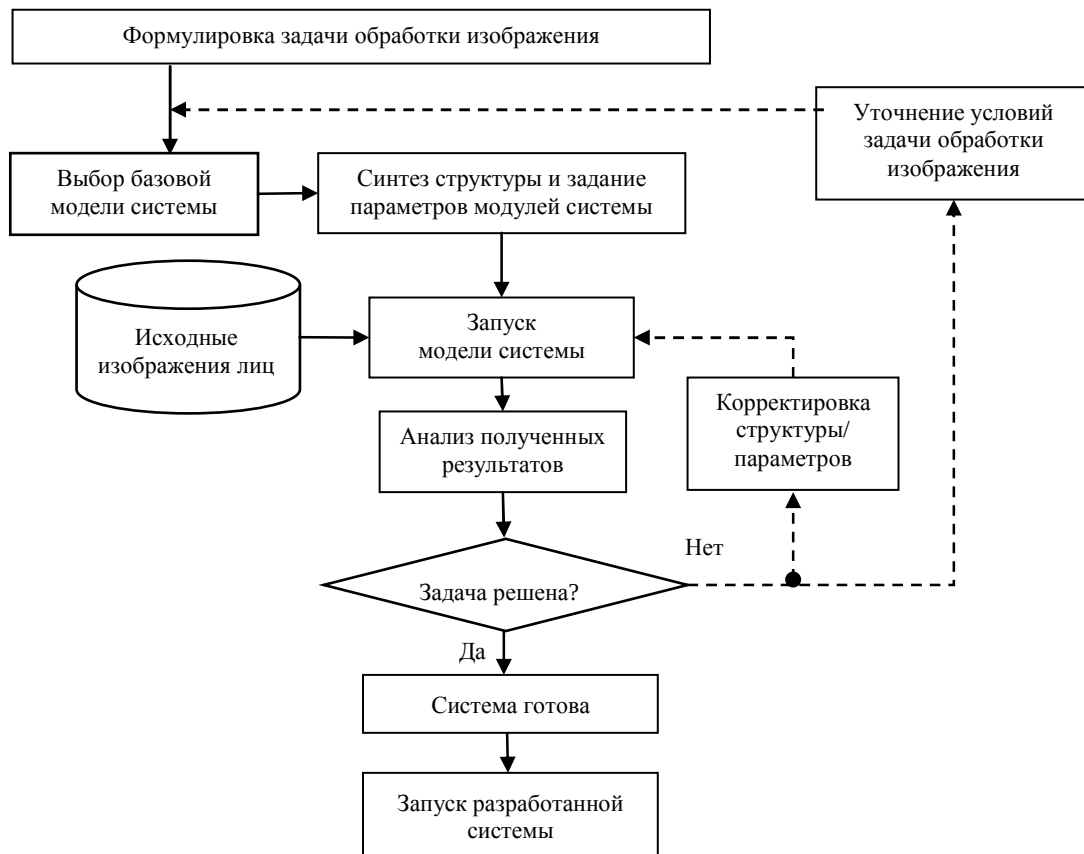


Рис. 4

Определяющим в выборе метода проекции является частота обновления исходных данных. При редких обновлениях данных проекция реализуется методами PCA, LDA, CCA и PLS. При большой динамике обновлений проекция реализуется в универсальных базах с учетом параметров моделей ИЛ в собственных базах.

Компактная форма записи компьютерного эксперимента. В работе [12] отмечается, что в настоящее время выбор оптимальной структуры

системы для решения практической задачи осуществляется на основе комбинации различных алгоритмов обработки и настройки их параметров. Однако в статьях, посвященных ИЛ, эксперименты описываются авторами по-разному, что часто затрудняет понимание, каким образом проводился эксперимент и какие параметры были выбраны для конкретных методов. Это не позволяет исследователю повторить эксперимент и корректно сравнить результаты. Первые работы,

затрагивающие вопросы унификации описания экспериментов, появились недавно. В них предлагалась частная модель описания рассматриваемой в экспериментах FaRetSys и не предлагалось обобщить данный способ описания на эксперименты с системами другой (произвольной) структуры. Далее предложена модель компактного описания компьютерного эксперимента (МКОЭ), отражающая состав и параметры блоков моделируемых систем, отличающаяся возможностью описания параллельных, каскадных и комбинированных систем.

Введем понятие *элементарной системы* (ЭС) для обработки ИЛ, состоящей из БД, модуля экстракции/селекции признаков и классификатора.

Используем следующие обозначения: BD – наименование базы данных; K_s – количество классов в БД; Q – количество образов в классе; L – количество эталонов, представляющих класс; $Q - L$ – количество тестовых образов. Тогда описание базы данных системы можно представить как

$$BD(K_s/L/Q - L).$$

В процессе исследования с целью оценки робастности системы, устойчивости алгоритмов к разного рода искажениям, как правило, выполняется некоторая модификация тестовых образов, поэтому запишем $(Q - L) \rightarrow MOD$, где MOD обозначает тип выполняемого искажения – зашумление, сдвиг и т. д.

Способ формирования тестового набора (например, кроссвалидация) обозначим следующим образом: $Q - L : TV$. Тогда общее описание БД будет иметь вид

$$BD(K_s/L/(Q - L) \rightarrow MOD: TV). \quad (1)$$

Пример «Для анализа использована база данных ORL, состоящая из 40 классов, каждый из которых состоит из 10 изображений, при этом 1 изображение используется в качестве эталона, а 9 изображений – тестовые, для формирования тестового набора использована процедура кроссвалидации (CV)» записывается следующим образом:

$$ORL(40/1/9/CV).$$

Очевидно, что данная запись понятная, существенно короче, но при этом не потеряла информативности того обозначения, которое заменила.

Для описания экстракции/селекции признаков используем следующие обозначения: M – высота изображения в пикселях; N – ширина изображения в пикселях; TF – метод экстракции/селекции признаков; $d1 \times d2$ – размерность эталона; PO – предобработка; SF – выбор признаков. Тогда блок экстракции/селекции признаков будет иметь следующее описание:

$$\{PO / TF : M \times N \rightarrow d1 \times d2 / SF\}. \quad (2)$$

Пример «Размер исходного изображения составляет 112×92 пикселя, редукция размерности пространства признаков выполнена с помощью двумерного метода главных компонент 2DPCA, размерность эталона – матрица с размерами 11×11 » запишется следующим образом:

$$\{2DPCA : 112 \times 92 \rightarrow 11 \times 11\}.$$

Характеристики классификатора зависят от следующих параметров: Cls – используемый классификатор; Met – используемая метрика; rank – ранг результата.

Используя данные обозначения, описание классификатора будет иметь вид

$$[Cls/Met/rank]. \quad (3)$$

Пример «Использован классификатор по минимуму расстояния, метрика L1, ранг = 1» запишется следующим образом:

$$[KMP/L2/rank = 1].$$

Учитывая (1)–(3), МКОЭ для ЭС можно представить в виде рис. 5. Здесь BD – база данных; K_s – количество классов в БД; L – количество эталонов; Q – количество образов в классе; $Q - L$ – количество тестовых образов; MOD – тип искажения (зашумление, сдвиг); TV – способ формирования тестового набора (например, кроссвалидация); PO – вид предобработки; TF – метод экстракции признаков; $d1 \times d2$ – размерность эталона; M – высота изображения в пикселях; N – ширина изображения в пикселях; SF – способ выбора признаков; Cls – классификатор; Met – метрика; rank – ранг результата.

Предложенная МКОЭ отражает состав и параметры блоков моделируемых систем, отличается возможностью описания параллельных, каскадных и комбинированных систем.

База данных	Экстракция/селекция признаков	Классификатор
ES = BD($K_s/L/(Q - L) \rightarrow MOD: TV$)	{PO/TF: $M \times N \rightarrow d1 \times d2/SF$ }	[Cls/Met/rank]

Рис. 5

Пример. Описание системы, состоящей из модулей, представленных ранее, будет иметь вид

$$ES = ORL(40/1/9/CV) \{2DPCA : 112 \times 92 \rightarrow 11 \times 11\} [KMP/L2/rank = 1].$$

В качестве примера приведем модель системы для распознавания ИЛПО, представленной в [13]:

$$YaleB(38/1/2414) \{G\&Log / 2DDCT: 192 \times 168 \rightarrow d \times d\} [\Phi + KMP/L1/rank = 1].$$

На основе ЭС теперь достаточно просто записывается модель каскадной системы CS, представляющая собой последовательное соединение KS элементарных систем:

$$CS = ES_1 \rightarrow ES_2 \rightarrow \dots \rightarrow ES_{KS}.$$

Для записи модели параллельной системы необходим еще один параметр Fuz – комплексирование (фузия) решений, принятых KS элементарными системами, которые соединены параллельно, при этом ES_{Itog} является системой, которая принимает окончательное решение PS:

$$PS = \left\{ \begin{matrix} ES_1 \\ ES_2 \\ \vdots \\ ES_{KS} \end{matrix} \right\} : Fuz \ ES_{Itog}.$$

В отличие от моделей, представленных в [13], в предложенной модели ЭС (рис. 5) учитываются

проводимые в процессе исследований возможные модификации данных, особенности формирования тестового набора, а общий вид модели позволяет на ее основе описывать системы, имеющие каскадную, параллельную структуры, а также системы, сочетающие в себе их возможные комбинации.

Предложена концепция построения комплекса программных средств для моделирования FaRetSys, а в рамках концепции определен состав компонентов, их функции и методика использования комплекса при решении задач моделирования. Это позволит разрабатывать средства для моделирования FaRetSys с учетом всех требований и современных технологий. Предложенная расширенная МКОЭ представляется не только более наглядной и точно описывающей структуру системы распознавания, но и позволяет более просто и эффективно сравнивать структуры моделируемых систем, оценивать сложность их реализации и влияние последовательности выполняемых процедур на результаты работы, а также использовать полученные выводы в рамках метаанализа. Набор параметров, входящих в модель, достаточен для выполнения аналогичного эксперимента другими исследователями или другими средствами (платформами), а также для сравнительного анализа при тестировании моделируемых систем.

СПИСОК ЛИТЕРАТУРЫ

1. Растринин Л. А., Эренштейн Р. Х. Обучение коллектива решающих правил // Адаптивные системы. 1974. Вып. 4. С. 8–20.
2. Растринин Л. А., Эренштейн Р. Х. Принятие решений коллективом решающих правил в задачах распознавания образов // Изв. АН СССР. Автоматика и телемеханика. 1975. № 9. С. 134–144.
3. Городецкий В.И., Серебряков С. В. Методы и алгоритмы коллективного распознавания: обзор // Тр. СПИИРАН. 2006. Т. 1, вып. 3. С. 139–171.
4. Dietterich T. G. Ensemble Methods in Machine Learning // Int. Workshop on Multiple Classifier Systems, Lecture Notes in Computer Science. Springer-Verlag. 2000. Vol. 1857. P. 1–15.
5. Kuncheva, L. I., Whitaker C. J. Measures of Diversity in Classifier Ensembles and Their Relationship with the Ensemble Accuracy // Machine Learning. 2003. № 51. P. 181–207.
6. Kuncheva L. I. Combining Pattern Classifiers: Methods and Algorithms. John Wiley & Sons, Inc., 2004. 376 p.
7. Кухарев Г. А., Щеголева Н. Л. Системы распознавания человека по изображению лица. СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2006. 176 с.
8. Кухарев Г. А., Щеголева Н. Л. Методы построения систем распознавания человека по изображению лица: учеб. пособие. СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2007. 94 с.
9. Elad M., Hel-Or Y., Keshet R. Pattern Detection Using a Maximal Rejection Classifier // Pattern Recognition Letters. 2002. Vol. 23, № 12. P. 1459–1471.
10. Gama J., P. Brazdil Cascade Generalization // Machine Learning. 2000. №. 41(3). P. 315–342.
11. Ho T. K. Multiple classifier combination: Lessons and the next steps. In A. Kandel and H. Bunke, editors, Hybrid Methods in Pattern Recognition. World Scientific Publishing, 2002. P. 171–198.
12. Щеголева Н. Л. Разработка программных средств для моделирования биометрических систем распознавания человека по изображению лица // Докл. ТУСУРа. 2013. №1 (27). С. 37–44.
13. Методы обработки и распознавания изображений лиц в задачах биометрии / под ред. М.В. Хитрова. СПб.: Политехника, 2013. 388 с.

N. L. Shchegoleva

Saint Petersburg Electrotechnical University «LETI»

CONCEPTION OF SOFTWARE FOR FACE RETRIEVAL SYSTEMS MODELING

The problem of developing of software for face retrieval systems modeling discusses. Review of the status of the problem is presented. The conception of construction of software tools for modeling FaRetSys proposed. Within the concept defined components, their functions and methods of using in solving complex simulation tasks. Model of a compact description of computer simulation is clarified. Model contains the elements and parameters of the blocks modeled system. Characterized by the ability to describe the parallel, cascade and combined systems.

Retrieval face, software for face retrieval systems modeling, system structure
