

На правах рукописи

КУДРЯШОВ Павел Павлович

АЛГОРИТМЫ ОБНАРУЖЕНИЯ ЛИЦА ЧЕЛОВЕКА
ДЛЯ РЕШЕНИЯ ПРИКЛАДНЫХ ЗАДАЧ
АНАЛИЗА И ОБРАБОТКИ ИЗОБРАЖЕНИЙ

05.13.01 - Системный анализ, управление и обработка информации

АВТОРЕФЕРАТ

диссертации на соискание ученой степени
кандидата технических наук

Волгоград - 2007

Работа выполнена на кафедре "Системы автоматизированного проектирования и поискового конструирования" Волгоградского государственного технического университета.

Научный руководитель: доктор технических наук, профессор
Фоменков Сергей Алексеевич

Официальные оппоненты: доктор технических наук, профессор
Вагин Вадим Николаевич

доктор технических наук, профессор
Лукьянов Виктор Сергеевич

Ведущая организация: Институт проблем передачи информации
им. А. А. Харкевича (РАН)

Защита диссертации состоится 8 ноября 2007 г. на заседании диссертационного совета Д 212.028.04 при Волгоградском государственном техническом университете по адресу: 400131, Волгоград, пр-т. им. В. И. Ленина, 28.

С диссертацией можно ознакомиться в библиотеке Волгоградского государственного технического университета.

Автореферат разослан 3 октября 2007 г.

Ученый секретарь
диссертационного
совета

Водопьянов В. И.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы диссертации. В настоящее время в связи со стремительным развитием цифровой фотографии и цифрового видео очень перспективной является задача распознавания образов на цифровых изображениях.

Последние десять лет в области распознавания образов 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. (Kanade, Rowley, Schapire, Shneiderman, Viola, Загоруйко, Журавлев). Часть этих подходов распознавания инвариантны по отношению к объекту, другие используют такие априорные знания об объекте, как: форма, цветовая гамма, взаимное расположение частей и др.

Несмотря на то, что в реальном мире существует огромное количество различных объектов, значительный интерес представляет разработка алгоритмов обнаружения более узкого класса объектов - лиц человека. Конкретными практическими применениями алгоритмов распознавания лиц могут быть: системы автоматического учета числа посетителей; системы пропускного контроля в учреждениях, аэропортах и метро; автоматические системы предотвращения несчастных случаев; интеллектуальные интерфейсы "человек-компьютер" и др.

Задача обнаружения лица человека является сложной ввиду нескольких основных причин: высокая вариативность лиц человека, обусловленная анатомическими и фенотипическими особенностями индивидов; различные условия освещенности, определяющиеся типом, количеством и направлением источников света; необходимость обнаружения лиц, находящихся в произвольных положениях и наклонах.

Существующие потребности в создании подобных систем накладывают жесткие ограничения на скорость работы алгоритмов, которые должны работать в режиме близком к реальному времени. Однако даже наиболее быстрые из существующих подходов (Viola 2001, Lienhart 2002, Shneiderman 2004) позволяют обнаруживать лица в реальном времени лишь при вертикальном положении лица и непригодны для обнаружения лиц, повернутых в плоскости изображения под произвольным углом.

Для успешного функционирования системы обнаружения лиц обеспечение высокой скорости работы должно также сочетаться с малым количеством ложных обнаружений. В существующих системах при увеличении процента обнаружений свыше 90% наблюдается экспоненциальный рост числа ложных обнаружений, что затрудняет использование этих систем в случаях, когда требуется процент обнаружения близкий к 100%.

Таким образом, учитывая, что наиболее важными характеристиками систем обнаружения являются скорость и процент обнаружения, актуальной является задача разработки и внедрения новых, более быстрых и точных алгоритмов, способных обнаруживать лицо человека, наклоненное под произвольным углом в плоскости изображения.

Цель работы и задачи исследования. Целью диссертационной работы является повышение эффективности обнаружения лица человека на цифровых изображениях за счет разработки и применения новых и модифицированных алгоритмов. Под повышением эффективности понимается улучшение следующих характеристик: повышение процента обнаружения лиц человека, снижение процента ложных обнаружений, снижение времени обучения классификатора и времени обработки изображения по отношению к результатам существующих алгоритмов.

Для достижения данной цели были поставлены и решены следующие задачи:

- описать особенности и способы решения задачи обнаружения лица человека, провести обзор подходов и алгоритмов обнаружения, выявить их ограничения;
- разработать новые и модифицировать существующие алгоритмы для решения задачи обнаружения лица человека;
- реализовать предложенные алгоритмы в программном модуле обнаружения лица человека;
- провести проверку работоспособности и эффективности модуля при решении тестовых и практических задач;
- разработать и внедрить программный комплекс автоматической коррекции красных глаз.

Объект исследования. Процесс обнаружения лица человека на цифровых изображениях с использованием вычислительной машины.

Предмет исследования. Алгоритмы анализа цифровых изображений и обнаружения лица человека.

Методы исследования. В диссертации использованы методы системного анализа, распознавания образов, компьютерного зрения, математической статистики, параллельного программирования, объектно-ориентированного проектирования систем.

Научная новизна работы состоит в следующем:

- предложена модификация существующего алгоритма Viola-Jones с использованием нового типа признаков треугольной формы и матрицы объекта. Использование треугольных признаков позволяет добиться более точной аппроксимации наклонных черт лица. Использование матрицы объекта позволяет повысить скорость обучения при одновременном снижении объема базы

обучающих примеров. Использование треугольных признаков и матрицы объекта инвариантно к объекту распознавания (может использоваться при обнаружении любых других объектов);

- предложен новый алгоритм прогнозирования положения и размеров лица человека с использованием цветовой информации, позволяющий значительно сократить время на обнаружение лица, наклоненного на произвольный угол в плоскости изображения;
- предложен новый алгоритм подтверждения обнаружения лица с использованием механизма поворотов, позволяющий существенно снизить процент ложных обнаружений. Использование алгоритма подтверждения инвариантно к объекту распознавания;
- предложен гибридный алгоритм обнаружения лица человека, основанный на алгоритме Viola-Jones, алгоритме прогнозирования положения и размеров лица человека и алгоритме подтверждения обнаружения лица с использованием механизма поворотов.

Обоснованность и достоверность результатов, приведенных в диссертационной работе, обеспечиваются использованием зарекомендовавших себя методов системного анализа, компьютерного зрения и распознавания образов, подтверждаются результатами работы разработанного модуля быстрого обнаружения лиц на тестовой базе изображений, а также результатами его работы при решении практической задачи обнаружения лица человека в системе автоматической коррекции красных глаз "Red Eye Remove".

Практическая значимость и внедрение.

- разработан модуль быстрого обнаружения человеческих лиц, позволяющий обнаруживать лица человека на цветных цифровых изображениях при произвольном угле наклона лица в плоскости изображения и отличающийся более высокой точностью и скоростью обнаружения по сравнению с существующими аналогами;
- разработана система полностью автоматической коррекции красных глаз, позволяющая корректировать цифровые изображения без участия человека. Система может быть использована для коррекции цифровых фотографий рядовых пользователей, в фотосалонах, фотоаппаратах, в составе цифровых комплексов высокопроизводительной печати фотолабораторий и в автоматических киосках фотопечати.

Система автоматической коррекции красных глаз "Red Eye Remove" внедрена в фотосалоне "DeaLux" в г. Волгограде. С использованием системы откорректировано свыше

5000 фотографий клиентов сети. Система автоматической коррекции красных глаз "Red Eye Remove" и система обнаружения лица человека "Rapid face detection" зарегистрированы в отраслевом фонде алгоритмов и программ ФГНУ "Государственный координационный центр информационных технологий" Федерального агентства по образованию РФ.

Положения, выносимые на защиту:

- модифицированный алгоритм Viola-Jones с использованием признаков треугольной формы и матрицы объекта;
- алгоритм прогнозирования положения и размеров лица человека на основе цветовой информации;
- алгоритм подтверждения обнаружения с использованием механизма поворотов;
- гибридный алгоритм обнаружения лица человека;
- модуль быстрого обнаружения лиц человека;
- автоматическая система коррекции красных глаз человека.

Апробация результатов работы. Основные положения диссертации докладывались и обсуждались на научных семинарах кафедры "САПР и ПК" ВолгГТУ, а также на Международных и Всероссийских научных и научно-практических конференциях: "Информационные технологии в науке, образовании, телекоммуникации и бизнесе (IT + SE)" (Гурзуф, Украина, 2007), "Интеллектуальные системы (AIS). Интеллектуальные САПР (CAD)" (Дивноморское, 2007), "Технологии Microsoft в теории и практике программирования" (Москва, 2007), "Новые информационные технологии" (Судак, 2007), "Информационные технологии в образовании, технике и медицине" (Волгоград, 2006), Региональная конференция молодых исследователей Волгоградской обл. (Волгоград, 2006).

Публикации. Основные положения диссертации отражены в 25 опубликованных работах. В том числе 4 статьи напечатаны в рецензируемых научных журналах центральной печати, из которых 3 статьи опубликованы в журналах, рекомендованных ВАК для публикации основных результатов диссертационных работ. В список опубликованных работ также входят два свидетельства об официальной регистрации программ для ЭВМ.

Структура и объем работы. Диссертационная работа состоит из введения, четырех глав, заключения, приложений, списка литературы. Общий объем диссертации - 121 страница, включая 44 рисунка, 11 таблиц, список литературы из 155 наименований.

СОДЕРЖАНИЕ РАБОТЫ

Во **введении** обоснована актуальность темы диссертации, сформулированы цели и задачи исследования, определена научная новизна, приводится перечень основных положений, выносимых на защиту, излагается краткое содержание глав диссертации.

В **первой главе** производится обзор работ связанных с подходами к решению задачи обнаружения лица на цифровых изображениях.

Задачу обнаружения лица человека на цифровом изображении можно сформулировать следующим образом. Пусть имеется цифровое изображение – двумерная матрица пикселей размером $W \times H$, в которой каждый пиксель представлен некоторым неотрицательным числом (0-255 в случае черно-белого изображения или 0-255³ в случае цветного изображения). Требуется определить список наклонных прямоугольников, описанных вокруг лиц людей, присутствующих на изображении:

$$LeanRect_i = \{x, y, width, height, angle\},$$

где x, y – координаты центра i -го прямоугольника, $width$ – ширина, $height$ – высота, $angle$ – угол наклона прямоугольника к вертикальной оси изображения.

Существует две основных категории алгоритмов, разработанных для решения задачи обнаружения лица: методы эмпирического распознавания и методы моделирования изображения лица человека.

Первая категория алгоритмов основывается на опыте человека в решении задачи распознавания лиц и пытается формализовать этот опыт. Так исследователи, использующие этот подход, стараются определить условия и признаки, по которым человек выносит решение, является ли данное изображение лицом или нет. При этом формируется ряд эвристик, на основе наличия и взаимного соответствия которых, автоматическая система может определить факт присутствия лица на изображении. К числу таких признаков относятся: лицо обладает определенным оттенком цвета "кожи"; лицо зачастую симметрично относительно вертикальной оси; на лице присутствуют отдельные черты (глаза, нос, рот); черты лица имеют строго определенные отношения размеров и взаимное положение.

Вторая категория алгоритмов не пытается формализовать процессы, происходящие в мозге человека, и основывается на аппарате математической статистики. Методы этой категории основываются на хорошо проработанной теории распознавания образов, или, еще в более общем смысле – на теории к решению задач классификации. При этом задача обнаружения лица рассматривается как частный случай задачи классификации с двумя предопределенными классами: "лицо" и "не лицо". Для определения принадлежности изображения к одному из классов используется вектор признаков, являющийся

отображением анализируемого изображения в пространство более высокой, чем исходная, размерности. Ввиду того, что общее количество признаков очень велико, используются различные процедуры сокращения размерности и выделения наиболее значимых признаков. Так большинство методов данной категории используют две стадии: стадию обучения (с учителем или без него) и стадию классификации. На стадии обучения из всего набора признаков выделяются существенные, также определяются пороговые значения признаков, используемые для принятия решения об отнесении изображения к тому или иному классу. На стадии классификации производится вычисление значений признаков и определение принадлежности изображения к классу. Ввиду того, что стадия обучения и выбора наиболее значимых признаков не требует вмешательства человека, результаты методов второй категории в значительной степени превосходят результаты эмпирических методов обнаружения.

Последние десять лет в области распознавания лица человека ведутся активные разработки и предложены различные методы распознавания, использующие такие подходы второй категории, как метод главных компонент [Shneiderman, Pentland], нейросети [Rowley, Kanade, Lin], алгоритм Viola-Jones [Viola, Lienhart, Whitehill], машины опорных векторов [Lienhart] и т.д.

Основными недостатками методов данной категории являются: длительное время и необходимость наличия большой выборки на этапе обучения, обучение только на прямоугольных изображениях объектов, в подавляющем большинстве случаев - отбрасывание цветовой информации.

Кроме того, даже наиболее быстрый и точный в данной категории алгоритм Viola-Jones не позволяет обнаруживать лицо человека под произвольным углом поворота в режиме реального времени. Что в значительной мере затрудняет или делает невозможным его использование с учетом растущих потребностей современных производственных систем.

Таким образом, выявлена необходимость проведения научного исследования для модификации существующих и разработки новых более совершенных методик и алгоритмов обнаружения лица человека на цифровых изображениях.

Во **второй главе** описываются предложенные автором подходы к повышению эффективности решения задачи обнаружения лица человека.

Так по результатам обзора существующих подходов к обнаружению лица человека, проведенному в первой главе, было выявлено, что наилучшим алгоритмом на текущий момент является алгоритм, предложенный Viola и Jones в 2001 году. Данный алгоритм относится к семейству ассоциативных машин и способен классифицировать поступающие изображения на два класса: "объект" и "не объект". Алгоритм требует предварительного

обучения классификатора в режиме обучения с учителем. Эффективность алгоритма для классификации лиц человека подтверждена научными трудами различных авторов.

Для снижения временных затрат на этапе обучения и снижения необходимого количества образцов в обучающей выборке рассматривается подход с использованием матрицы объекта.

Предлагается введение матрицы объекта M_0 – прямоугольной бинарной матрицы, по размеру совпадающей с размером изображений в обучающей базе, в которой единицами обозначены пиксели, являющиеся пикселями объекта, а нулями – пиксели, являющиеся пикселями фона.

$$M_{0ij} = \begin{cases} 1, \forall k \text{ } Img_{kij} \in object \\ 0 \end{cases},$$

где M_{0ij} - пиксель с координатами i, j матрицы объекта, Img_{kij} - пиксель с координатами i, j k -го изображения обучающей выборки, *object* – объект классификации.

Фронтальное лицо человека имеет овальную форму и занимает от 75 до 90% площади ограничивающего прямоугольника в зависимости от условий кадрирования. А профильное лицо, в свою очередь, занимает от 70 до 85% площади прямоугольника. Использование матрицы объекта (рис. 1) позволяет исключить до 30% незначимых пикселей изображений обучающей выборки или до 40% незначимых признаков на этапе обучения классификатора лица.

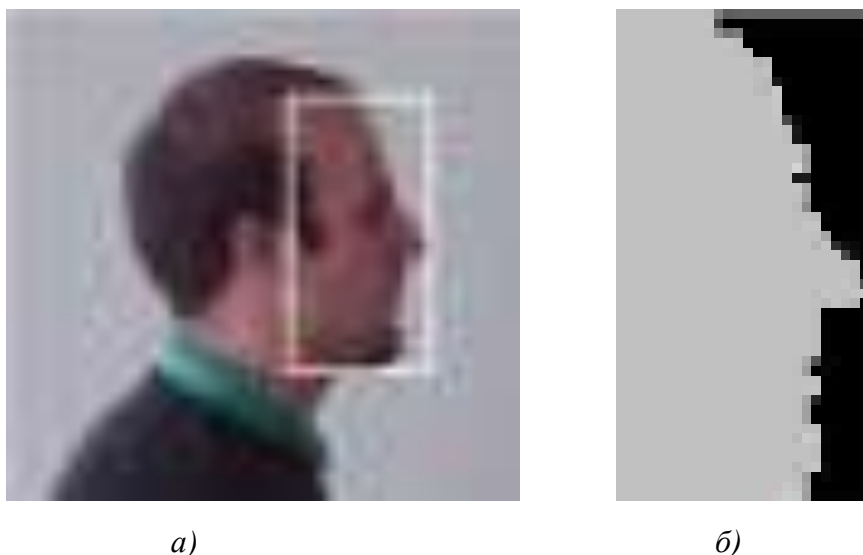


Рис. 1. Изображение профильного лица на этапе обучения (а) и матрица объекта с черными незначимыми пикселями (28% площади) (б).

Также использование матрицы объекта позволяет явным образом выделить области незначимого фона на изображениях в обучающей выборке. Это, в свою очередь, позволяет

априорно исключить из рассмотрения фоновые области изображения и уменьшить необходимое число изображений лиц на различном фоне, требуемых для обучения.

Для повышения репрезентативных возможностей алгоритма и качества аппроксимации наклонных черт лица предлагается использование нового типа треугольных признаков.

Признак, используемый для анализа изображения можно представить кортежем

$$Feature = \{T, O, S\},$$

где T – тип признака, O – координата левого верхнего угла признака, S – размер признака по горизонтали и вертикали в пикселях.

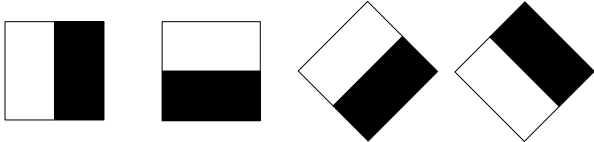
Значение признака рассчитывается по следующей формуле:

$$Feature_i = k_W \sum V_W - k_B \sum V_B,$$

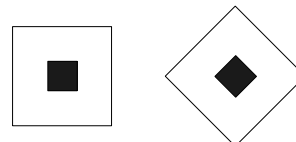
где $\sum V_W$ и $\sum V_B$ – суммы интенсивностей всех пикселей изображения в белых и черных областях признака соответственно, k_W и k_B – коэффициенты их нормировки по площади.

Предлагается использование нового типа признака – треугольных признаков, расширяющих набор существующих прямоугольных признаков. Рассматривается два вида треугольных признаков, представленных на рис. 2. На указанном рисунке черные области соответствуют отрицательным весам, а белые – положительным.

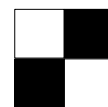
а. Граничные



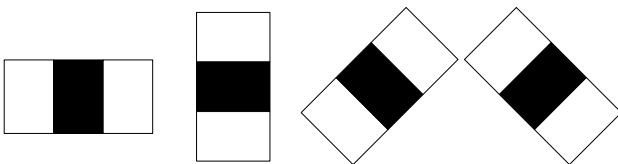
в. Центральные



г. Диагональные



б. Линейные



д. Треугольные

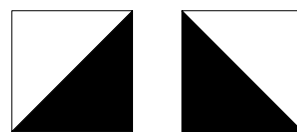


Рис. 2. Прямоугольные (а, б, в, г) и треугольные (д) типы признаков.

Также предлагается способ быстрого вычисления значений треугольных подобластей нового типа признаков. Для их вычисления используется понятие интегрального изображения, встречающееся в работах [Viola, Lienhart и др.]. Интегральное изображение – матрица, совпадающая по размерам с исходным изображением, в каждом элементе которой хранится сумма интенсивностей всех пикселей, находящихся левее и выше данного элемента, т.е. можно записать:

$$SAT(x, y) = \sum_{x' \leq x, y' \leq y} I(x', y'),$$

где $SAT(x, y)$ - значение элемента интегрального изображения с координатами x, y ,
 $I(x', y')$ - интенсивность пикселя исходного изображения с координатами x', y' .

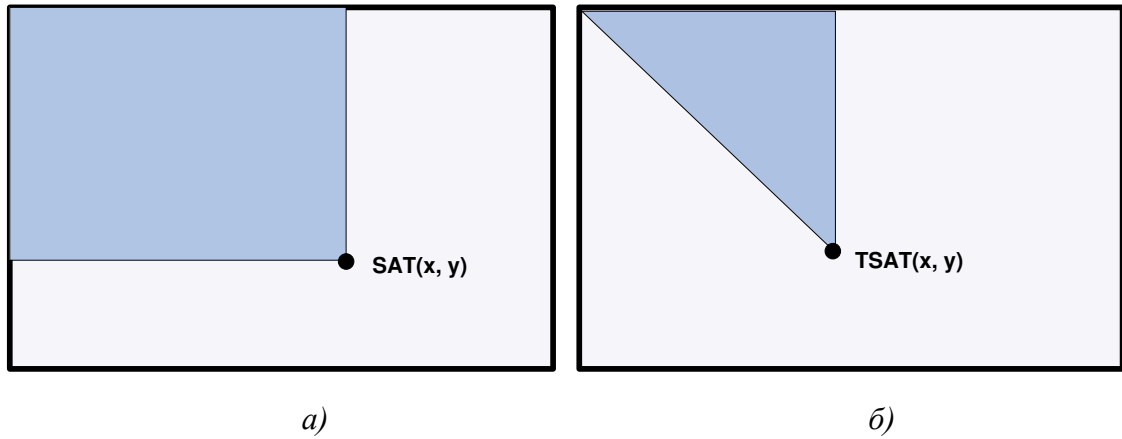


Рис. 3. Значение в точке интегрального изображения SAT (а) и $TSAT$ (б).

Используется дополнительное интегральное изображение $TSAT$, значение каждого элемента которого равно сумме интенсивностей всех пикселей, находящихся в прямоугольном треугольнике, расположенным левее и выше данного элемента (рис. 3):

$$TSAT(x, y) = \sum_{x' \leq x, y' \leq y, y' \leq y - |x - x'|} I(x', y')$$

Для быстрого вычисления значения каждого элемента интегрального изображения $TSAT$ за один проход цикла используется интегральное изображение SAT и формула:

$$TSAT(x, y) = TSAT(x - 1, y - 1) + SAT(x, y) - SAT(x - 1, y)$$

Используя вычисленное интегральное изображение $TSAT$ можно получить значение каждой треугольной подобласти $TrSum(t)$ за 4 обращения к SAT и $TSAT$ (рис. 4):

$$TrSum(t) = SAT(x - 1, y - 1) + TSAT(x + w - 1, y + w - 1) - SAT(x + w - 1, y - 1) - TSAT(x - 1, y - 1)$$

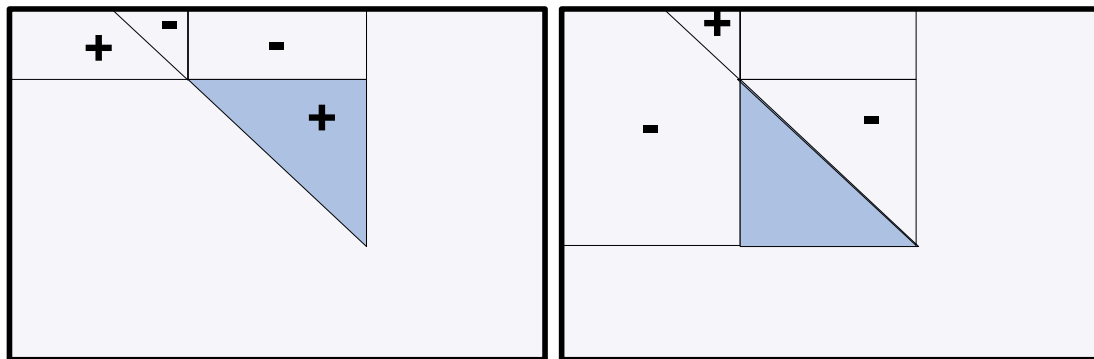


Рис. 4. Вычисление значения треугольной подобласти с использованием интегральных изображений SAT и $TSAT$.

Для вычисления значений двух треугольных подобластей одного треугольного признака требуется 6 обращений к интегральным изображениям *SAT* и *TSAT*.

Использование треугольных признаков позволяет повысить качество аппроксимации наклонных линий лица. Предложенный алгоритм быстрого вычисления значения треугольных признаков позволяет производить обнаружение без падения скорости по сравнению с базовым набором признаков.

Для снижения числа ложных обнаружений лиц предлагается алгоритм подтверждения обнаружения с использованием механизма поворотов.

Алгоритм Viola-Jones, обученный на распознавание лица человека достаточно устойчив (около 20 градусов) к поворотам изображения. При угле наклона свыше 20 градусов процент обнаружений резко падает. Однако при небольших углах наклона до 10 градусов, алгоритм продолжает надежно определять лицо человека. Случайные, ложные обнаружения, в отличие от истинных лиц, не обладают такой устойчивостью.

Используя данную особенность, предлагается следующий алгоритм подтверждения обнаружений. Для областей изображения, на которых были обнаружены лица, предлагается производить повторную проверку обнаружения на изображении, повернутом на 8 градусов по и против часовой стрелки. В случае если обнаружение ложное, с высокой степенью вероятности ни на одном из повернутых изображений оно не будет повторным.

Рассмотрим работу алгоритма Viola-Jones на базовом и повернутом изображениях при анализе 2-го граничного прямоугольного признака (рис. 2), состоящего из двух подобластей (рис. 5). При этом будем рассматривать темную подобласть, соответствующую глазам. Так на рис. 5 можно видеть, что на изображениях слева суммарная интенсивность пикселей подобласти глаз незначительно меняется при повороте ввиду того, что наиболее "весомые" для алгоритма элементы, обладающие большой площадью и низкой интенсивностью (глаза и брови) остаются в подобласти при повороте изображения.



Рис. 5. Устойчивость алгоритма к небольшим поворотам изображения:
истинное лицо(а) и ложное обнаружение (б).

Другая ситуация происходит при ложном обнаружении. При повороте "весомый" элемент изображения (темная одежда), обладающий большой площадью и низкой интенсивностью, на повернутом изображении выходит за границы подобласти, вследствие чего значение признака резко уменьшается, в результате чего принимается решение об отбросе данного окна поиска. Использование описанного механизма поворотов позволяет с высокой степенью уверенности отсекал случайные ложные обнаружения.

Варьируя угол подтверждения проверки можно изменять соотношение между количеством ложных обнаружений и пропусков. Эмпирически при обнаружении фронтальных лиц человека оптимальным значением угла подтверждения является значение в 8 градусов.

Для создания быстрого и надежного способа определения вероятных областей лица человека с целью ускорения обработки на дальнейших этапах обнаружения, предлагается алгоритм определения размеров и положения лица человека на основании цветовых характеристик изображения.

Алгоритм использует ограничения (априорные знания) задачи обнаружения человеческих лиц. Так лица людей на цифровых изображениях имеют так называемый цвет "кожи", который достаточно компактно описывается в цветовых пространствах RGB, HSV и других.

С использованием априорных метрик для классификации кожи в пространстве HSV выделяется маска кожи. Маска кожи представляет собой однобитное изображение, где единицами помечены те пиксели изображения, которые могут быть кожей.

$$M_{S_{ij}} = \begin{cases} 1, IsSkin(Img_{ij}) \\ 0 \end{cases},$$

где $M_{S_{ij}}$ - пиксель с координатами i, j маски кожи, Img_{ij} - пиксель с координатами i, j исходного изображения, $IsSkin(H, S, V)$ – функция принадлежности пикселя к цвету "кожи", заданная в цветовом пространстве HSV и определяемая соотношением:

$$IsSkin(H, S, V) = V \geq 40 \ \& \ H \leq -0.4V + 75 \ \& \ 10 \leq S \leq (-H - 0.1V + 110) \\ \& \ S \leq \begin{cases} 0.08(100 - V)H + 0.5V, H \geq 0 \\ S \leq 0.5H + 35, H < 0 \end{cases}$$

После вычисления полученная маска кожи обрабатывается алгоритмом обрамляющих эллипсов, обнаруживающим контуры эллипсов вокруг границ компактного расположения пикселей маски. На этой стадии алгоритм выдает значительное количество эллипсов, большинство из которых не являются контурами человеческих лиц.

Далее происходит обработка полученных эллипсов. При этом отсекаются слишком маленькие эллипсы, которые соответствуют цветовому шуму. Также удаляются эллипсы со слишком большим отношением большого и малого радиусов: $R_L/R_S > 5$. Такие вытянутые эллипсы часто соответствуют конечностям человека, различным деревянным панелям мебели и др. По итогам получается маска кожи, обведенная эллипсами, каждый из которых обводит предполагаемое лицо человека (рис. 6а).

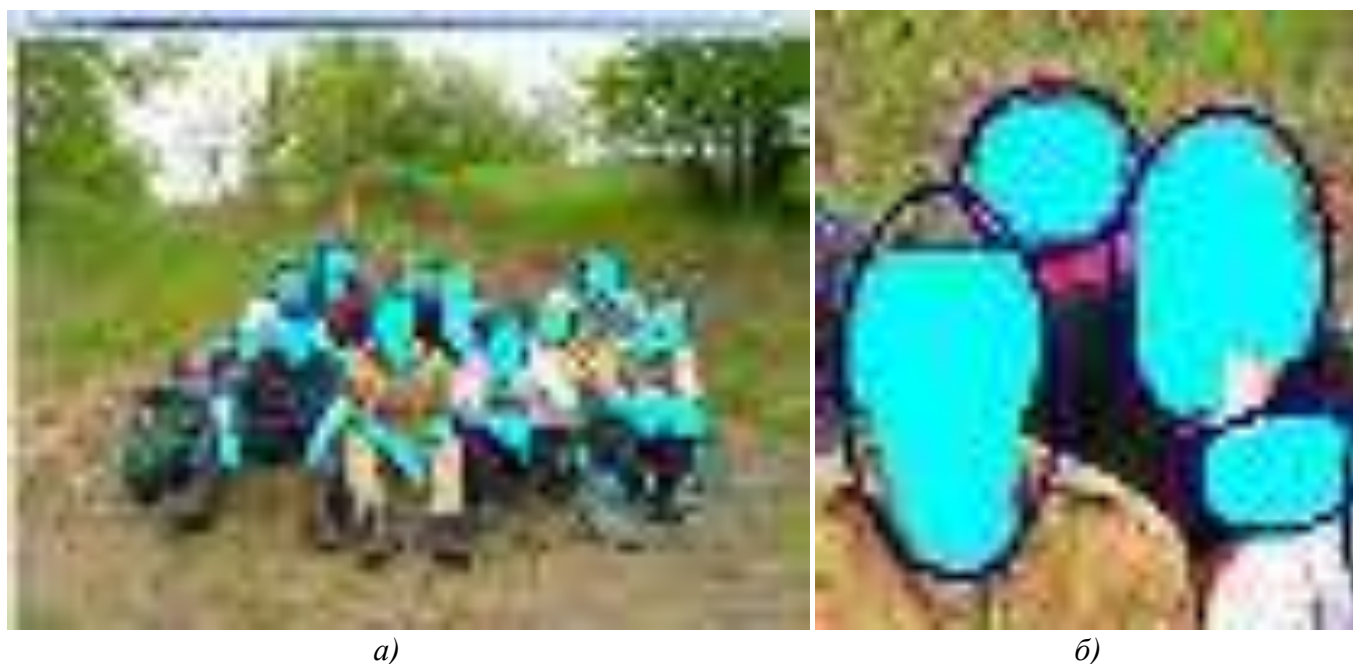


Рис. 6. Области цвета "кожи", обведенные эллипсами слева (а) и уточненные регионы лиц справа (б).

Далее производится уточнение найденных обрамляющих эллипсов. Критерием необходимости уточнения является процент заполнения эллипса пикселями маски кожи, попадающими в область эллипса. В случае, если отношение суммарной площади пикселей маски кожи, попадающих в эллипс к площади эллипса слишком мало, производится операция уточнения эллипсов.

$$S_{skin}/S_{Ellipse} < 0.6$$

Операция уточнения состоит из двух этапов: морфологических операций эрозии и дилиации маски кожи с ядром операции равным 1/10 размера эллипса и повторного запуска алгоритма обрамляющих эллипсов на уточненной маске (рис. 6). Данная операция последовательно производится для всех эллипсов до тех пор, пока не выполнится соотношение площадей маски кожи и эллипса указанное выше или размер обнаруженных на очередной стадии эллипсов не станет слишком мал.

Для дополнительного повышения скорости работы алгоритма на этапе обнаружения пикселей цвета "кожи" и обрамляющих их эллипсов используется предварительное уменьшение размеров изображения в 2-4 раза. При этом происходит квадратичное снижение объема вычислений, связанных с анализом цветовых характеристик пикселей. Также при этом не требуется удаление слишком маленьких обрамляющих эллипсов, так как подобные области цветового шума исчезнут в результате операции понижения разрешения. В дальнейшей обработке обнаруженные эллипсы лиц масштабируются в систему координат исходного изображения.



Рис. 7. Приоритетные направления обнаружения лица.

Последним этапом алгоритма является определение угла наклона лица человека в плоскости изображения. При этом используется тот факт, что лицо человека представляет собой эллипс с отношением большей и меньшей осей равным $2 / 1$ для фронтального лица и $4 / 1$ для профильного лица. Следовательно, используя угол наклона большей оси обрамляющего эллипса, считаем, что с наибольшей вероятностью лицо человека расположено в одном из двух направлений, соосных большей оси эллипса и расположенных под углом 180 градусов друг к другу (рис. 7). Причем, ввиду того, что в подавляющем большинстве случаев изображения расположены не вверх-ногами, более приоритетным является то направление, которое имеет меньший угол с направлением "вверх" изображения.

На основании алгоритмов, описанных ранее во второй главе, предлагается гибридный алгоритм обнаружения лица человека

Последовательность шагов гибридного алгоритма следующая:

1. уменьшение изображения в 2-4 раза;
2. вычисление маски кожи;
3. получение и уточнение обрамляющих маску кожи эллипсов и направлений их главных осей;

4. масштабирование положений центров и размеров полученных эллипсов в координаты исходного изображения;
5. получение прямоугольных участков изображения, в которых содержатся полученные эллипсы;
6. последовательная обработка полученных участков с запуском модифицированного алгоритма Viola-Jones. Выполняется обнаружение лица человека на участках, повернутых под всеми углами с шагом в 20 градусов, начиная от наиболее приоритетных направлений, полученных в п. 3. При этом выполняется до $360 / 20 = 18$ поворотов изображения. В случае обнаружения лица, наклоненного под каким-то определенным углом, обработка текущего участка прекращается и осуществляется переход к следующему участку;
7. выполняется подтверждение обнаружения с использованием механизма поворотов.

Гибридный алгоритм обладает следующими преимуществами по сравнению с базовым алгоритмом Viola-Jones:

- более высокая точность обнаружения за счет использования треугольных признаков и подтверждения обнаружения с использованием механизма поворотов;
- более высокая скорость обнаружения за счет:
 - отсекающие областей изображения, не являющихся лицами на основании цветовых характеристик изображения;
 - априорного знания о приблизительном размере лица для поиска алгоритмом Viola-Jones на основании анализа цветовых характеристик, и, следовательно, более низкого количества поисковых окон, подлежащих обработке;
 - априорного знания о наиболее вероятном наклоне лица к оси изображения и исключения необходимости проверять всех 18 углов поворота.

В **третьей главе** описан созданный программный модуль быстрого обнаружения лица человека, реализующий усовершенствованные и новые алгоритмы. Модуль имеет следующий интерфейс:

Вход: цветное изображение для обнаружения лиц

Выход: массив наклонных прямоугольников, описанных вокруг обнаруженных лиц людей, $LeanRect_i = \{x, y, width, height, angle\}$,

где x, y – координаты центра i -го прямоугольника, $width$ – ширина, $height$ – высота, $angle$ – угол наклона прямоугольника к вертикальной оси изображения.

Модуль состоит из трех основных подсистем:

- подсистема предварительного определения положения и размеров лица
- подсистема обнаружения лица
- подсистема подтверждения обнаружения лица

Базируясь на предложенных ранее алгоритмах, разработанный модуль обладает некоторыми ограничениями. В условиях недостаточного освещения и недоступности цветовой информации, требуемой для работы подсистемы предварительного определения положения и размеров лица, модуль будет работать медленнее, но надежнее, чем стандартный алгоритм Viola-Jones за счет использования предложенного механизма поворотов. Также, несмотря на то, что размеры и углы наклона около 85% лиц, обнаруженных на тестовых примерах, были корректно определены алгоритмом определения размеров и положения лица и обнаружены алгоритмом Viola-Jones на первом же повернутом изображении, остальные 15% лиц были определены в среднем за 5 поворотов изображения. Однако даже в этом случае гибридный алгоритм не пропускает такие лица, работая с более низкой скоростью, чем возможна в "идеальном" случае.

Для проверки эффективности работы алгоритма и сравнения скорости были проведены тестовые испытания на тестовых базах CMU Frontal, CMU PIE и UCD Colour, специально разработанных для сравнения алгоритмов обнаружения лица человека. В таблице 1 представлены результаты сравнения точности работы базового алгоритма Viola-Jones, реализованного в свободно распространяемой библиотеке Intel OpenCV, с модифицированным алгоритмом, использующим треугольный тип признаков и механизм подтверждения обнаружений с использованием поворотов.

Таблица 1.

Сравнение точности работы гибридного алгоритма.

Алгоритм / БД	CMU Frontal		CMU PIE		UCD Colour	
	TP	FP	TP	FP	TP	FP
OpenCV Viola-Jones	83%	12	93%	28	65%	11
Viola-Jones + треуг. признаки	83%	10	94%	25	69%	10
Viola-Jones + механизм поворотов	83%	9	92%	17	65%	6
Viola-Jones + треуг. пр. + мех. пов.	83%	7	94%	14	68%	6

В таблице 2 приведены результаты сравнения скорости работы базового алгоритма Viola-Jones, реализованного в Intel OpenCV, с гибридным алгоритмом обнаружения. Для обнаружения лиц, наклоненных под произвольным углом, с использованием базового алгоритма, каждое изображение обрабатывалось под 18 углами в диапазоне 0 – 340 градусов. Для определения влияния размера изображения на результаты работы алгоритмов было

произведено тестирование на трех базах, содержащих изображения различного размера: UCD Colour, CVL Face и базе из 100 фотографий пешеходных переходов г. Волгограда.

В целом, можно сделать вывод, что скорость работы гибридного алгоритма при обнаружении лица человека, повернутого под произвольным углом в плоскости изображения, до 10 раз превышает скорость работы базового алгоритма.

Таблица 2.

Сравнение скорости работы гибридного алгоритма.

Алгоритм / БД	UCD Colour	CVL Face	Volgograd
Средний размер фото, пиксели	~400 x 300	640 x 480	2592 x 1944
Количество фото	94	798	100
Поиск Viola-Jones, с	30	6244	14238
Поиск гибридный алгоритм, с	28	1082	1340

Программный модуль, реализующий гибридный алгоритм, написан на языке MS Visual C++ 2005 с использованием библиотеки компьютерного зрения Intel OpenCV в виде динамически подключаемой dll и работает под управлением Windows 2000/XP.

Областями применения разработанного модуля могут быть: системы интеллектуального взаимодействия "человек – компьютер", системы теле- и видеоконференций, автоматические системы слежения и безопасности, системы контент-ориентированного сжатия и поиска, различные прикладные программ обработки мультимедиа (от систем коррекции красных глаз для домашних пользователей до систем учета числа посетителей в супермаркетах).

В **четвертой главе** приведены конкретные практические области применения разработанного модуля, а также результаты работы созданной с его использованием системы.

С использованием модуля обнаружения лица человека было разработано две программных системы: прототип системы контент-ориентированного сжатия изображений и система автоматической коррекции красных глаз "Red Eye Remove".

Прототип системы контент-ориентированного сжатия изображений позволяет кодировать отдельные части изображения JPEG с различным качеством. При этом система обнаруживает на изображениях лица людей и сжимает их с лучшим качеством, чем остальные области изображения. Ввиду того, что основным фокусом внимания человека на изображениях являются лица людей, система позволяет сохранять изображения в файлы меньшего размера при сохранении визуального качества изображения. Или, с другой стороны, при сохранении изображения в файл заданного размера, адаптивное сжатие системой позволяет получать визуально более качественные изображения.

Система является тестовым прототипом, однако результаты, полученные при ее реализации, позволяют предположить, что подобная технология контент-ориентированного сжатия может применяться в случаях, когда место для хранения изображений ограничено, например, в мобильных цифровых устройствах. Также подобный подход может быть применен в системах видеоконференций для снижения объема передаваемых данных. Система реализована на Microsoft Visual C++ 2005 и использует библиотеки Intel OpenCV и LibJpeg.

Разработанный модуль также вошел в состав полностью автоматической системы коррекции красных глаз "Red Eye Remove". Система состоит из двух основных частей: ядра системы, производящего все операции обработки и коррекции изображения, и графического интерфейса пользователя (рис. 8).



Рис. 8. Графический интерфейс пользователя программы "Red Eye Remove".

Общий алгоритм работы системы следующий:

- получение маски "красных" пикселей;
- сегментация красных пикселей и выделение эллиптических областей;
- отбрасывание недостаточно заполненных и слишком вытянутых областей;
- проверка каждой из найденных областей статистическими верификаторами (наличие вокруг областей цвета "кожи", радужной оболочки, блика, белка, бровей и др.);
- проверка окрестностей найденных областей верификатором с использованием гибридного алгоритма обнаружения лица человека и сопоставление найденных лиц человека и областей глаз, соответствующих им;

- вычисление общей оценки на основании оценок, полученных всеми верификаторами;
- в случае превышения общей оценкой определенного значения считается, что данная красная область является красным (засвеченным) глазом;
- пиксели каждого красного глаза корректируются особым образом, сохраняя естественный вид глаза после коррекции.

Система позволяет корректировать изображения в формате Jpeg без повторного пережатия всего изображения, при этом изменяются только небольшие области изображения, соответствующие откорректированным глазам. Сравнительный анализ результатов работы разработанной системы с аналогичными системами сторонних производителей представлен в таблице 3.

Таблица 3.

Сравнение работы системы Red Eye Remove.

Критерий	HP RedBot	FotoNation	RedEyeRemove
Обработка изображения 5МПикс ¹ , с	30 ²	2	4
Обнаружение "красных" глаз ³ , %	92%	68%	94%
Обнаружение "красных" глаз (лицо скрыто), %	12%	68%	76%
Устойчивость к повороту лица	45°	360°	360°
Обнаружение "белых" глаз ⁴	-	-	+
Настольная система	-	+	+
Автоматическая система	-	-	+

Система написана на Microsoft Visual C++ 2005 и использует библиотеки Intel OpenCV, LibJpeg, HTMLLayout, Boost, ZLib, OpenExif и др. Система доступна как в виде пользовательского приложения под операционной системой Windows, так и в виде автономной библиотеки, которую можно использовать в составе другого программного обеспечения, в автоматических системах высокоскоростной печати для фотолабораторий, в фотокиосках, в фотоаппаратах и др. Система внедрена в фотосалоне "DeaLux" в г. Волгограде. С использованием системы откорректировано свыше 5000 фотографий клиентов сети.

¹ Среднее время обработки изображения в наборе

² Ориентировочное время обработки, основанное на описании алгоритма работы системы

³ Результат получен на наборе из 1100 фотографий, полученных из разных источников

⁴ Тестовая версия

ОСНОВНЫЕ РЕЗУЛЬТАТЫ РАБОТЫ

- проведен сравнительный анализ подходов к решению задачи обнаружения лица человека. В качестве наиболее перспективного выбран подход с использованием алгоритма Viola-Jones;
- предложен модифицированный алгоритм Viola-Jones, использующий новый тип треугольных признаков и матрицу объекта. Треугольные признаки и матрица объекта могут быть использованы инвариантно к объекту распознавания;
- предложен гибридный алгоритм обнаружения лица человека, основанный на методах: прогнозирования положения и размеров лица человека с использованием цветовой информации и подтверждения обнаружения лица с использованием механизма поворотов. Алгоритмы подтверждения обнаружения инвариантны к объекту распознавания;
- предложенные алгоритмы реализованы в виде модуля быстрого обнаружения человеческих лиц. Результаты тестирования модуля на тестовых базах показали на 40% более низкий уровень ложных обнаружений и до 10 раз большую скорость работы при обнаружении лиц в произвольных положениях;
- разработанный модуль обнаружения человеческих лиц использован в составе автоматической системы коррекции засвеченных глаз человека Red Eye Remove;
- результаты тестирования системы Red Eye Remove показали превосходство системы в точности обнаружения над системами компаний HP (на 5%) и FotoNation (на 15%) на тестовом наборе из 200 фотографий.

СПИСОК ПУБЛИКАЦИЙ ПО ТЕМЕ ДИССЕРТАЦИИ

1. Кудряшов П. П. Гибридный алгоритм обнаружения человеческих лиц / П. П. Кудряшов, С. А. Фоменков // Информационные технологии. - 2007. - №10. - С.20-23.
2. Кудряшов П. П. Быстрый поиск человеческих лиц на изображениях / П. П. Кудряшов, С. А. Фоменков // Вестник компьютерных и информационных технологий. - 2007. - №7. - С.14-17.
3. Кудряшов П. П. Гибридный алгоритм распознавания человеческих лиц на цифровых изображениях / П. П. Кудряшов, С. А. Фоменков // Изв. ВолгГТУ. Серия "Актуальные проблемы управления, вычислительной техники и информатики в

технических системах": межвуз. сб. науч. ст. / ВолгГТУ.- 2007.- Вып. 2, №2.- С.43-46.

4. Кудряшов П. П. Быстрое распознавание человеческих лиц на цифровых изображениях / П. П. Кудряшов, С. А. Фоменков //Современные технологии автоматизации. - [Электронный ресурс]. - [2007]. - Режим доступа: http://www.cta.ru/online/online_progr-digital.htm
5. Кудряшов П. П., Алгоритм обнаружения лица человека для решения прикладных задач анализа и обработки изображений / П. П. Кудряшов, С. А. Фоменков // Интеллектуальные системы (IEEE AIS'07) и Интеллектуальные САПР (CAD-2007): Сб. тр. Международных научно-технических конференций. - М.: Изд-во физико-математической литературы, 2007. - Т.2. - С.132-134.
6. Кудряшов П. П. Автоматическая система коррекции эффекта "красных глаз" / П. П. Кудряшов, С. А. Фоменков //Открытое образование: прилож. к журн.: по матер. XXXIV междунар. конф. и дискус. науч. клуба, Ялта-Гурзуф, 20-30.05.07: Инф. технол. в науке, образ., телеком. и бизнесе (IT+SE '07).- 2007.-[Б/н]. - С.153-154.
7. Кудряшов П. П. Алгоритм распознавания лица человека / П. П. Кудряшов, С. А. Фоменков //Инновационные технологии в управлении, образовании, промышленности "АСТИНТЕХ-2007": матер. всерос. науч. конф., 18-20 апреля 2007 г. / Астрахан. гос. ун-т и др. - Астрахань, 2007. - Ч.2. - С.116-120.
8. Кудряшов П. П. Гибридный алгоритм поиска областей "красных глаз" / А.М. Герасимов, П. А. Колчин, П. П. Кудряшов, С. А. Фоменков //Технологии Microsoft в теории и практике программирования: тр. IV всерос. конф. студ., аспири. и мол. уч., 2-3 апр. 2007: Центральный регион, Москва / Моск. авиац. ин-т (гос. техн. ун-т) [и др.]. - М., 2007. - С.98-99
9. Кудряшов П. П. Двухступенчатый алгоритм обнаружения лица человека / А.М. Герасимов, П. А. Колчин, П. П. Кудряшов, С. А. Фоменков //Технологии Microsoft в теории и практике программирования: тр. IV всерос. конф. студ., аспири. и мол. уч., 2-3 апр. 2007: Центральный регион, Москва / Моск. авиац. ин-т (гос. техн. ун-т) [и др.]. - М., 2007. - С.97-98
10. Кудряшов П. П. Быстрый алгоритм обнаружения засвеченных глаз / П. П. Кудряшов, С. А. Фоменков //Технологии Microsoft в теории и практике программирования: межвуз. конкурс-конф. студ., аспири. и мол. уч. Северо-Запада: матер. [конф.], 13-14 марта 2007 г. / СПбГПУ [и др.]. - СПб., 2007. - С.169
11. Кудряшов П. П. Использование обрамляющих эллипсов для ускорения работы алгоритма Нааг при распознавании лиц / П. П. Кудряшов, С. А. Фоменков

//Технологии Microsoft в теории и практике программирования: межвуз. конкурс-конф. студ., аспири. и мол. уч. Северо-Запада: матер. [конф.], 13-14 марта 2007 г. / СПбГПУ [и др.]. - СПб., 2007. - С.170

12. Кудряшов П. П. Повышение скорости распознавания лица человека алгоритмом Нааг на цветных изображениях / П. П. Кудряшов //Информационные технологии, системный анализ и управление: сб. тр. IV всерос. науч. конф. мол. ученых, аспирантов и студ. / Таганрог. гос. радиотехн. ун-т [и др.]. - Таганрог, 2006. - С.6-7
13. Кудряшов П. П. Повышение точности алгоритма Нааг при распознавании объектов на цифровых изображениях / П. П. Кудряшов //Информационные технологии, системный анализ и управление: сб. тр. IV всерос. науч. конф. мол. ученых, аспирантов и студ. / Таганрог. гос. радиотехн. ун-т [и др.]. - Таганрог, 2006. - С.5-6
14. Кудряшов П. П. Препроцессинг как средство повышения качества распознавания объектов на цифровых изображениях / П. П. Кудряшов, С. А. Фоменков //Информационные технологии в образовании, технике и медицине: матер. междунар. конф., Волгоград, 23-26 октября 2006 г. /ВолгГТУ и др.- Волгоград, 2006. - С.158-160
15. Кудряшов П. П. Автоматическое обнаружение эффекта "красных глаз" на цифровых изображениях / П. П. Кудряшов, С. А. Фоменков //Информационные технологии в образовании, технике и медицине: матер. междунар. конф., Волгоград, 23-26 октября 2006 г. /ВолгГТУ и др.- Волгоград, 2006. - С.157-158
16. Свид. об офиц. рег. программы для ЭВМ № 8782. Система распознавания лиц человека на цифровых изображениях "Rapid face detection" - №8782; Заяв. 17.07.2007 опубл. 17.07.2007. ОФАП, 2007.
17. Свид. об офиц. рег. программы для ЭВМ № 8783. Автоматическая система коррекции "красных глаз" на цифровых фотографиях "Red Eye Remove" - №8783; Заяв. 17.07.2007 опубл. 17.07.2007. ОФАП, 2007.
18. 8 публикаций в сборниках тезисов региональных и международных конференций.