

# **Вычислительно-эффективный метод поиска нечетких дубликатов в коллекции изображений**

© Пименов В.Ю.

Санкт-Петербургский Государственный университет,  
факультет Прикладной математики - процессов  
управления  
vitaly.pimenov@gmail.com

## **Аннотация**

В работе развивается метод решения задачи поиска нечетких дубликатов в коллекции изображений, основанный на выявлении и сопоставлении точечных особенностей изображений. Предложена методика повышения быстродействия за счет использования модели системы зрительного внимания. Точечные особенности выявляются с помощью методов теории масштабируемого пространства и представляются дескрипторами типа SURF. Сопоставление изображений производится путем сопоставления дескрипторов точечных особенностей, наиболее заметных в терминах модели зрительного внимания. Экспериментальные результаты показывают, что применение модели внимания обеспечивает существенный прирост быстродействия без ухудшения значений качественных показателей.

## **1. Введение**

Настоящая работа продолжает исследования [1-8], направленные на разработку методов поиска нечетких дубликатов изображений, основанных на выявлении и сопоставлении точечных особенностей изображений. Одним из ключевых факторов, препятствующих

широкому применению методов данного класса, является их высокая вычислительная сложность [2,3,7,8].

Цель настоящего исследования состоит в разработке методики повышения быстродействия методов выявления и сопоставления точечных особенностей изображений, а также в разработке программного комплекса, реализующего новую методику, и оценке ее производительности.

Поиск нечетких дубликатов изображений является одной из важных задач компьютерного зрения [1,9] и в настоящее время привлекает внимание исследователей в областях информационного поиска и робототехники. Это связано с тем, что использование информации о нечетких дубликатах позволяет разрабатывать эффективные алгоритмы решения задач поиска новостей [10], определения и отслеживания новостных сюжетов [11], поиска нарушений авторского права [4], автономной навигации мобильных роботов, обнаружения движущихся объектов, манипуляции объектами.

Внимание к задаче поиска нечетких дубликатов в рамках проблем информационного поиска возросло с возникновением в 2003 г. семинара TRECVID [12].

**Определение [7].** Два изображения, которые могут быть переведены друг в друга путем элементарных преобразований, таких как поворот, сдвиг, изменение угла обзора, изменение разрешения, изменение масштаба, изменение освещения, называются *нечеткими дубликатами*.

Нечеткими дубликатами являются фотоснимки и видеок cadры одной сцены, сделанные с различных ракурсов, при различном освещении, в различные моменты времени, при помощи различных регистрирующих устройств. Также нечеткие дубликаты возникают при редактировании изображений.

Задача поиска нечетких дубликатов тесно связана с классической задачей компьютерного зрения – задачей сопоставления изображений (совмещения изображений, регистрации изображений, *image registration, image matching*) [9,13]. Задача сопоставления изображений состоит в установлении соответствий между областями изображений, кроме того, зачастую необходимо установление в явном виде преобразования, связывающего изображения. При этом специфика прикладных задач обычно обуславливает ограничение возможных преобразований рамками некоторого априорно заданного класса.

Отличительная особенность задачи поиска нечетких дубликатов заключается в отсутствии каких-либо ограничений на допустимые преобразования, связывающие изображения. Этим обуславливается значительная сложность данной задачи, так как для ее решения необходимо представление изображений в форме, обеспечивающей ковариантность по отношению к неизвестным и потенциально сложным геометрическим и фотометрическим преобразованиям.

Задачи поиска нечетких дубликатов включают следующие частные задачи: поиск дубликатов по содержанию, сравнение пары изображений, кластеризация коллекции изображений. В настоящей работе рассмотрена задача кластеризации. Процесс решения задачи состоит из двух основных этапов: построение модели изображения; проведение кластеризации коллекции изображений.

## 2. Модель изображения

Анализ известных исследований показывает, что для построения модели изображения могут быть использованы как глобальные, так и локальные свойства изображения.

Существующие модели, основанные на глобальных свойствах изображения, в общем случае не гарантируют эффективности идентификации нечетких дубликатов. Так, использование цветовых свойств, к примеру матрицы изменения яркости, может приводить к ошибкам в случае изменения условий освещения сцены или использования определенных фильтров [14].

С другой стороны, использование характерных локальных признаков, устойчивых к геометрическим и фотометрическим преобразованиям, представляется более предпочтительной альтернативой. Одним из подходов, основанных на использовании локальных свойств изображений, являются *методы выявления точечных особенностей*. Фундаментом данных методов служит теория масштабируемого пространства [15,16], дающая математически строгое описание дифференциальной структуры многомерных сигналов. Методы выявления точечных особенностей изображений являются областью активных исследования (см., например, [1-9,14,17,18,20]).

Точечной особенностью (в литературе также используются термины «характеристическая точка», «точка интереса», «контрольная точка», *interest point*) называется точка изображения, структура окрестности которой ковариантна заданным преобразованиям изображения.

Для выявления точечных особенностей разрабатываются методы-детекторы. Обзоры и сравнения этих методов представлены в работах [1,2,8,17,18,20].

Для точечных особенностей, найденных с помощью детектора, рассчитываются дескрипторы – вектора, описывающие структуру окрестности точечной особенности. Как правило, эти вектора формируются на основе набора значений первых и вторых производных изображения в точке. Кроме того, дескриптор рассчитывается таким образом, чтобы его значение было инвариантным по отношению к требуемым преобразованиям изображения (обычно это те преобразования, относительно которых ковариантна структура окрестности точечной особенности). Типичными преобразованиями являются поворот, растяжение, сжатие, монотонное изменение яркости, аффинные и проективные преобразования.

Набор дескрипторов составляет модель изображения, которая впоследствии используется для сопоставления изображений. Решение о том, являются ли сопоставляемые изображения нечеткими дубликатами или нет, принимается на основе анализа наборов дескрипторов – моделей изображений.

Сопоставление дескрипторов осуществляется в два этапа:

- 1) на первом этапе устанавливаются пары наиболее близких дескрипторов (тем самым устанавливаются соответствия между точечными особенностями);
- 2) на втором этапе на основе набора пар близких дескрипторов принимается решение о соответствии между изображениями.

Близость между дескрипторами чаще всего вычисляется как расстояние в евклидовом пространстве [3, 7] либо как расстояние Махаланобиса [6].

Для принятия решения о соответствии между изображениями можно ввести расстояние в пространстве изображений [7] и принимать решение на основе сравнения расстояния с пороговым значением. Также применяются методы статистической верификации, такие как RANSAC [19]; методы геометрической верификации [1,8], устанавливающие возможность построения геометрического преобразования, связывающего пары близких точечных особенностей; а также методы непосредственного установления данного преобразования [9,13].

### 3. Вычислительная сложность

Источник вычислительной сложности сопоставления изображений заключается в большом числе точечных особенностей. Для изображения размером  $640 \times 480$  точек результат работы распространенных методов-детекторов содержит в среднем около тысячи точек, в то время как для изображений размером  $1920 \times 1080$  точек – до десятков тысяч. Для каждой точечной особенности необходим расчет дескриптора, кроме того, алгоритмы сопоставления наборов дескрипторов имеют вычислительную сложность, зависящую от размера сопоставляемых наборов.

На практике время поиска точечных особенностей и расчета их дескрипторов при использовании распространенных методов, таких как SIFT и SURF, имеет порядок 0.3 – 1 с. [2]. Время сопоставления пары изображений имеет порядок 0.02 с. [3,7]. Подобные показатели быстродействия ставят под сомнение возможность эффективного применения этих методов в задачах с характерными ограничениями на быстродействие, в частности в задачах реального времени.

Одним из способов преодоления указанных недостатков является разработка эффективных численных методов поиска точечных особенностей, расчета дескрипторов, вычисления расстояний между дескрипторами, поиска ближайших дескрипторов. К таким методам относятся методы приближенного анализа дифференциальной структуры изображений на основе вейвлетов и интегральных изображений [2], методы индексирования дескрипторов [1,3,7], методы понижения размерности дескрипторов [14].

Также повышение быстродействия достигается за счет использования специализированного аппаратного обеспечения: интегральных схем, процессоров цифровой обработки сигналов, систем параллельных вычислений.

Однако данные подходы не устраняют изначальную причину высокой вычислительной сложности, так как число анализируемых точечных особенностей остается высоким. В то же время в работах, посвященных сравнительному анализу методов-детекторов [17,18,20], отмечается, что используемые детекторы имеют относительно невысокую степень воспроизводимости, т.е. точечные особенности, найденные в эталонном изображении, не всегда могут быть найдены в преобразованном изображении, хотя качество сопоставления изображений остается высоким. При этом число невозпроизводимых точечных особенностей может достигать 40% от общего числа найденных точечных особенностей.

Это наблюдение говорит о том, что, с точки зрения задачи сопоставления изображений, невоспроизводимые точечные особенности являются избыточными данными. В то же время в известных алгоритмах сопоставления изображений все точечные особенности обрабатываются единообразно. Таким образом, обработка избыточных данных представляет собой источник дополнительной вычислительной сложности, исключение которого не повлияет на результат сопоставления изображений.

В настоящей работе предлагается методика количественной оценки степени воспроизводимости точечных особенностей. Использование такой методики позволит понизить вычислительную сложность процедуры сопоставления изображений за счет исключения из нее точечных особенностей, обладающих низкой воспроизводимостью.

Понятие воспроизводимости может быть определено формально. Обобщая известные подходы к оценке воспроизводимости, можно ввести следующее определение.

**Определение.** Пусть в точке  $p$  изображения  $I$  детектором  $h$  найдена точечная особенность, и эта точка преобразованием  $M$  переводится в точку  $p'$  изображения  $M(I)$ . Пусть также существует достаточно малая окрестность  $U$  точки  $p'$  такая, что существует точка  $\bar{p} \in U$  изображения  $M(I)$ , в которой детектором  $h$  будет найдена точечная особенность. В этом случае точечная особенность в точке  $p$  воспроизводима детектором  $h$  по отношению к преобразованию  $M$ .

Из определения следует, что прямая оценка воспроизводимости требует анализа пары изображений  $I$  и  $M(I)$ , а также знания преобразования  $M$ . Такое определение удобно в задачах сравнительного анализа, где преобразование можно задать явно. С другой стороны, в задачах поиска нечетких дубликатов преобразование неизвестно. В связи с этим необходимо использовать косвенные методы оценки воспроизводимости.

Предлагаемый в работе метод основан на привлечении модели системы зрительного внимания. Предпосылкой использования модели зрительного внимания для оценки воспроизводимости явились результаты работы [20], в которой показано, что средняя воспроизводимости наиболее заметных (в терминах модели зрительного внимания) областей изображения по отношению к неизвестным заранее преобразованиям превышает 90%, в то время как воспроизводимость точечных особенностей, выявленных стандартными методами-детекторами, составляет порядка 60%.

Высокая степень воспроизводимости может быть интерпретирована как показатель того, что модель зрительного внимания сама по себе является качественным детектором и может быть использована напрямую. Однако следует уточнить, что современные вычислительные модели механизмов зрительного внимания в значительной степени основаны на эвристических подходах и имеют своей основной целью моделирование процессов зрительного восприятия в живых системах.

В настоящее время еще не создана формальная теория внимания, а использование существующих моделей внимания в качестве детекторов приводит к противоречиям между эвристическими принципами оценки заметности и формальными методами анализа дифференциальной структуры изображения. В частности расчет дескриптора точечной особенности производится в предположении, что структура изображения в окрестности особенности ковариантна ряду элементарных преобразований, в то время как выбор точечной особенности на основе степени заметности, рассчитанной по модели внимания, не гарантирует требуемых свойств ковариантности.

Методика, предлагаемая в работе, направлена на устранение данного противоречия и заключается в объединении формальных методов выявления и описания точечных особенностей с эвристическими методами моделирования зрительного внимания. Степень заметности, рассчитанная по модели внимания, используется в качестве косвенной оценки воспроизводимости точечных особенностей, выявленных стандартными методами. Тем самым, с одной стороны, процесс выявления точечных особенностей остается неизменным, что обеспечивает формальную корректность расчета дескрипторов; с другой стороны, появляется возможность фильтрации точечных особенностей до расчета дескрипторов.

**Основная гипотеза**, выдвигаемая в работе, состоит в следующем предположении: сокращение числа точечных особенностей на основе степени их заметности позволит существенно повысить быстродействие процесса сопоставления изображений без ухудшения значений его качественных показателей.

Для проверки сформулированной гипотезы был реализован метод выявления точечных особенностей и вычислительная модель зрительного внимания, детальному описанию которых посвящены последующие разделы статьи. Метод кластеризации коллекции изображений с целью поиска нечетких дубликатов был построен на основе метода, описанного в [7].

Оценка метода была проведена в соответствии с методикой семинара РОМИП на предоставленной организаторами семинара коллекции изображений. Результаты оценки были соотнесены с результатами эталонного метода, предложенного в работе [7]. Также была проведена оценка быстродействия эталонного и нового метода.

#### 4. Алгоритм выявления точечных особенностей

В работе использован алгоритм выявления точечных особенностей, в основе которого лежит теория о дифференциальной структуре многомерных сигналов – теория масштабируемого пространства [15,16,21]. Данная теория разрабатывалась, начиная с 1960-х гг., с целью построения формального описания структуры сигналов, содержащих информацию в разных масштабах, и методов обработки таких сигналов. Мотивацией для применения теории масштабируемого пространства в задачах компьютерного зрения является, с одной стороны, то, что изображения являются двумерными сигналами, содержащими информацию в разных масштабах, с другой стороны – потребность в математическом аппарате анализа изображений в условиях отсутствия априорной информации о масштабе представленных на них объектов. Эта необходимость обусловлена использованием при обработке изображений операций, зависящих от масштаба.

В рамках теории масштабируемого пространства разработано представление изображения во всех масштабах одновременно [15,16,21-24].

**Определение [21].** Масштабируемое пространство изображения  $f: \mathbf{R}^2 \rightarrow \mathbf{R}$  есть семейство  $\{L(x, y, t): \mathbf{R}^2 \times \mathbf{R}_+ \rightarrow \mathbf{R}\}$  постепенно сглаживаемых, упрощающихся версий исходного изображения.

Доказано [16, 23], что единственным непрерывным линейным масштабируемым пространством является Гауссово масштабируемое пространство, для которого

$$L(x, y, t) = g(x, y, t) * f(x, y),$$

где  $*$  – операция свертки,  $g(x, y, t)$  – гауссова функция ядра

$$g(x, y, t) = \frac{1}{2\pi t} e^{-(x^2+y^2)/2t}$$

со стандартным отклонением  $t = \sigma^2$ , называемым *масштабом*. Масштабируемое пространство также является решением уравнения теплопроводности



$$\partial_i L = \frac{1}{2} \nabla^2 L$$

с начальным условием  $L(x, y, 0) = f(x, y)$ . Это позволяет формально ввести операцию дифференцирования изображения [16,21]. Таким образом, дифференциальная структура изображения в окрестности выбранной точки может быть описана с помощью разложения Тейлора. Набор коэффициентов ряда Тейлора до порядка  $N$  (обычно  $N = 2$ ) включительно может использоваться в качестве дескриптора рассматриваемой точки.

На основе значений частных производных до второго порядка включительно можно сконструировать величины, инвариантные по отношению к элементарным геометрическим преобразованиям – простейшие дифференциальные инварианты.

В рамках настоящей работы рассматривается определитель матрицы Гессе, величина, инвариантная по отношению к линейным преобразованиям и поворотам:

$$\det \mathbf{HL} = L_{xx}L_{yy} - L_{xy}^2.$$

Экстремальные свойства этой функции сохраняются при изменении масштаба [15], поэтому точки экстремума могут рассматриваться как точечные особенности.

Расчет значения определителя требует вычисления вторых производных и, следовательно, выполнения операций свертки. Прямые вычисления имеют высокую сложность, в связи с чем на практике обычно прибегают к использованию аппроксимаций. Наиболее часто свертку полных изображений заменяют сверткой изображений, уменьшаемых соответственно масштабному параметру  $t$ . В настоящей работе применен альтернативный метод, предложенный в [2], в рамках которого свертка осуществляется не гауссовой функцией, а ее приближением.

## 5. Модель системы зрительного внимания

После того как выявлены точечные особенности – точки экстремума определителя матрицы Гессе – производится оценка их заметности на основе модели системы зрительного внимания. Точечные особенности, имеющие низкую степень заметности, не используются в дальнейшем анализе.

Предлагаемая модель системы зрительного внимания основана на модели VOCUS [25], которая была использована в работе [20] при оценке воспроизводимости детекторов.

Ключевым понятием в современных вычислительных моделях зрительного внимания является понятие карты заметности (карты внимания, *saliency map*). Впервые предложенное в работе [26], понятие карты заметности означает «топографически упорядоченную карту, отражающую визуальную заметность соответствующих областей изображения» [27].

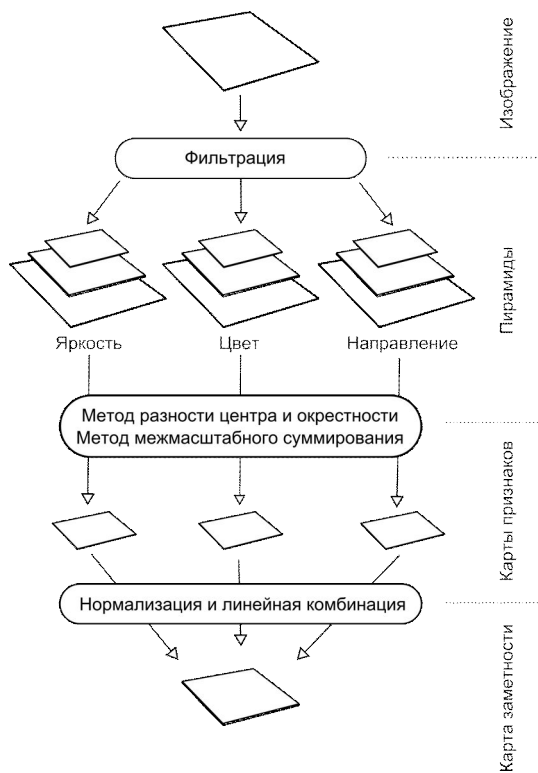


Рисунок 1. Блок-схема процедуры расчета карты заметности.

Карта заметности строится поточно, следуя работам [25,28]. Для каждой точки входного изображения заметность рассчитывается по трем признакам, используемым в зрительной системе человека [29]: яркости, цвету и ориентации. После этого производится интеграция признаков и строится карта внимания.

## 6. Расчет признаков заметности

Расчет признаков заметности основан на эвристических процедурах, предложенных для модели зрительного внимания VOCUS [25].

**Яркость.** Карты яркости рассчитываются путем применения метода разности центра и окрестности (*center-surround mechanism*) к изображениям гауссовой пирамиды. Пирамида содержит 5 изображений  $(s_0, \dots, s_4)$ , построенных из исходного изображения, преобразованного в полутоновое, с помощью последовательного гауссовского сглаживания и уменьшения масштаба. Яркость центра и яркость окрестности [25,29] рассчитываются отдельно для каждой точки путем сравнения яркости точки с усредненной яркостью ее окрестности. Подобный способ вычисления согласуется с современными представлениями об организации обработки информации в коре головного мозга человека [30]. Усредненное значение яркости вычисляется по квадратной области  $\Omega$  размера  $7 \times 7$  точек вокруг рассматриваемой точки. Из соображений эффективности вычисления производятся на основе интегральных изображений [31].

Карты яркости центра и яркости окрестности рассчитываются для изображений гауссовой пирамиды  $(s_2, s_3, s_4)$ . Изображения  $(s_0, s_1)$  исключаются из рассмотрения в целях повышения устойчивости к возможному шуму в исходном изображении [28]. Полученные карты объединяются с помощью метода межмасштабного суммирования (*across-scale addition*) [25].

**Цвет.** Карты цветовой заметности вычисляются как карты яркости четырех базовых цветов: красного, зеленого, синего и желтого. Выбор базовых цветов соответствует современным представлениям об обработке информации о цвете зрительной системой человека [32]. Исходное RGB изображение преобразуется в изображение в универсальном цветовом пространстве CIE LAB [32], чтобы обеспечить соответствие евклидоваго расстояния между цветами их зрительному восприятию. Полученное LAB изображение используется для построения гауссовой пирамиды в соответствии с описанной выше процедурой. На каждом уровне пирамиды LAB изображение используется для расчета четырех карт яркости цвета. Таким образом для каждого базового цвета строится пирамида карт яркости. Карты цветовой заметности строятся на основе пирамид яркости цветов (аналогично построению карты заметности яркости) путем применения метода разности центра и окрестности и метода

межмасштабного суммирования. Детальное описание способа построения карт цветовой заметности приведено в [25].

**Направление.** Карты заметности ориентации строятся для четырех базовых направлений:  $0^\circ, 45^\circ, 90^\circ, 135^\circ$  – в различных масштабах. Для каждого базового направления карта рассчитывается с помощью применения соответствующего преобразования Габора [32] к изображениям пирамиды разности гауссианов (*DoG, difference of Gaussians*) [1]. Результирующие изображения имитируют отклик нейронов зрительной коры мозга человека, отвечающих за обработку информации о направлении [29]. Метод межмасштабного суммирования применяется для получения карт заметности ориентации из преобразованных пирамид.

Построенные карты заметности признаков нормализуются и объединяются в карту заметности. Способ нормализации заимствован из [25]: карта заметности каждого признака умножается на коэффициент  $1/\sqrt{m}$ , где  $m$  – число локальных максимумов карты, превышающих заданный порог, значение которого было экспериментально выбрано равным 0.65% от глобального максимума карты. Подобное преобразование применяется для акцентирования ярко выраженных пиков заметности. Нормализованные карты заметности признаков суммируются с одинаковыми весами, результатом чего становится карта заметности. Равные веса используются в целях наиболее точного воспроизведения модели VOCUS, хотя известно, что использование переменных весов более предпочтительно [33].

Степень заметности точки характеризуется ее яркостью на карте внимания – значением от 0 до 255. Отсевание точечных особенностей производится простейшим пороговым методом: точечные особенности, которым соответствуют значения заметности меньшие, чем порог, выбранный равным 100, исключаются из дальнейшей обработки. При таком значении порога среднее число точечных особенностей, сохраняемых для обработки, составляет порядка 1/10 от их общего числа.

## 7. Алгоритм расчета дескрипторов

В работе использован алгоритм SURF, предложенный в [2]. Целью разработки этого алгоритма было повышение быстродействия

процедуры расчета дескриптора в сравнении с алгоритмом SIFT [1]. Для этого представление структуры окрестности точечной особенности с помощью набора значений градиента изображений в точках окрестности заменяется представлением, основанным на расчете отклика вейвлетов Хаара, которое может быть эффективно реализовано с применением интегральных изображений [31]. Подробное описание алгоритма SURF и его сравнение с рядом альтернативных алгоритмов приведено в [2].

## 8. Алгоритм кластеризации

Для проведения кластеризации был использован алгоритм, представленный в [7]. В его основе лежит алгоритм кластеризации  $QT$  [34], атомарная операция которого – вычисление расстояния между двумя векторами.

Расстояние между изображениями рассчитывается как

$$\rho(I_n, I_m) = \frac{1}{\|OOS(I_n, I_m)\|}, \quad (1)$$

где  $OOS(I_n, I_m)$  – множество пар точечных особенностей, соответствующих рассматриваемым изображениям, которые являются ближайшими соседями друг друга. Это означает, что пара точечных особенностей  $(P, Q)$  будет включена в множество  $OOS(I_n, I_m)$  тогда и только тогда, когда  $P \rightarrow Q$  и  $Q \rightarrow P$ , где  $\rightarrow$  обозначает отношение однозначного ближайшего соседства, т.е.

$$\rho(P, Q) < \rho(P, Q'), \quad \forall Q' \in I_m,$$

$$\rho(Q, P) < \rho(Q, P'), \quad \forall P' \in I_n,$$

где  $\rho(P, Q)$  – евклидово расстояние между векторами дескрипторами  $P$  и  $Q$ .

В работе [3] представлена подробная аргументация, обосновывающая преимущества использования симметричного взаимно-однозначного соответствия по сравнению с другими подходами: соответствием многие-ко-многим, многие-к-одному, один-ко-многим, один-к-одному без симметрии.

Таким образом, в центре атомарной операции алгоритма  $QT$  находится процедура построения множества  $OOS(I_n, I_m)$ . В свою очередь эта процедура основана на процедуре поиска для данного дескриптора точечной особенности  $P \in I_n$  ближайшего соседа

$Q \in I_m$ . Поскольку число точечных особенностей, выявленных в одном изображении, может быть высоким, процедура поиска ближайшего соседа может оказаться вычислительно сложной. Для понижения вычислительной сложности прибегают к использованию различных способов предварительной фильтрации множества дескрипторов. В настоящей работе использована индексная структура, разработанная в [7], основанная на нормализации и дискретном хешировании дескрипторов.

Процедура кластеризации с помощью алгоритма  $QT$  может быть представлена следующими этапами:

- 1) Определение максимального диаметра кластера.
- 2) Выбор изображения  $I$  из коллекции.
- 3) Для каждого изображений  $I'$  из коллекции, отличного от  $I$  :
  - a) построение множества  $OOS(I, I')$ ;
  - b) вычисление расстояния в соответствии с формулой (1);
  - c) если расстояние меньше максимального диаметра, изображение  $I'$  добавляется к кластеру с центром  $I$ .
- 4) Удаление всех изображений, добавленных к кластеру с центром  $I$  из коллекции.
- 5) Переход к шагу 2).

При проведении экспериментов значение максимального диаметра кластера полагалось равным 0.1, т.е. изображения  $I'$ , добавляемые в кластер с центром  $I$ , должны иметь по крайней мере 10 пар дескрипторов в множестве  $OOS(I, I')$ .

## 9. Постановка эксперимента

Для оценки производительности разработанного метода был проведен эксперимент. Предложенный метод поиска нечетких дубликатов изображений был реализован в виде набора программных модулей, разработанных на языке Java. Для проведения эксперимента была использована коллекция изображений, предоставленная организаторами семинара РОМИП. Коллекция содержит 38 000 изображений различного размера и качества. Количество точечных особенностей для одного изображения варьируется от 0 до 11985. В среднем одно изображение содержит 652 точечные особенности. После отсеивания точечных особенностей низкой заметности их среднее число сокращается до 70-80 для одного изображения.

Для оценки результатов эксперимента вычислялись значения четырех метрик: уровень ошибок первого рода, уровень ошибок второго рода, точность и полнота. Оценка проводилась организаторами семинара РОМИП. Методология оценки и описания метрик описаны в [35].

Кроме этого, были проведены измерения быстродействия метода. Измерения проводились на компьютере с процессором Intel Core 2 Duo 1.83ГГц с объемом оперативной памяти 2Гб.

## 10. Анализ результатов эксперимента

В таблице 1 приведены результаты эксперимента в сравнении с результатами эталонного метода, описанного в [7], оценка которого была проведена в идентичных условиях.

Приведенные результаты показывают, что значения качественных показателей изменились незначительно, что соответствует основной гипотезе, выдвинутой в данной работе.

Таблица 1. Результаты оценки качества кластеризации коллекции.

Метрика	Эталонный метод	Новый метод
Ошибки первого рода	$1.97 \times 10^{-4}$	$8.42 \times 10^{-5}$
Ошибки второго рода	0.430	0.462
Точность	0.721	0.739
Полнота	0.569	0.537

Относительно низкое значение точности объясняется спецификой алгоритма *QT*: условием попадания изображения в кластер является наличие не менее десяти пар симметрично ближайших соседей с точечными особенностями центра кластера. Эксперимент показывает, что этого условия не достаточно для получения высокой точности кластеризации. Для преодоления схожих трудностей в литературе предложены разнообразные методы статистической и геометрической верификации множества пар ближайших точечных особенностей (см., например, [1,8]). Подобные методы не были использованы в настоящем исследовании, так как преследовалась цель сохранить все свойства, присущие эталонному алгоритму, за исключением быстродействия.

Быстродействие метода измерялось двумя показателями: временем поиска точечных особенностей и расчета их дескрипторов

(для одного изображения) и временем сопоставления пары изображений (после расчета дескрипторов). Результаты измерения быстродействия представлены в таблице 2.

Таблица 2. Результаты оценки быстродействия.

<b>Метрика</b>	<b>Эталонный метод</b>	<b>Новый метод</b>
Время поиска (мс)	205	194
Время сопоставления (мс)	25	0.5

Использование модели системы зрительного внимания позволило сократить среднее время сопоставления изображений в 50 раз. Время поиска сократилось незначительно, однако следует отметить, что для нового метода во время поиска включено время, затрачиваемое на расчет карты заметности, которое составляет порядка 70% от общего времени поиска.

## 11. Заключение

В работе предложен метод поиска нечетких дубликатов изображений, основанный на выявлении и сопоставлении точечных особенностей изображений. Стандартный метод-детектор, основанный на расчете определителя матрицы Гессе, дополнен расчетом степени заметности точечных особенностей с помощью модели системы зрительного внимания. Новая методика позволила в 50 раз сократить среднее время сопоставления изображений за счет исключения из обработки точечных особенностей низкой заметности без ухудшения значений качественных показателей: точности, полноты, уровней ошибок первого и второго рода.

В числе недостатков разработанного метода наиболее существенны следующие: выбранный алгоритм кластеризации коллекции изображений не обеспечивает высокой точности; использованная модель системы зрительного внимания в значительной степени основана на эвристических подходах к задаче моделирования механизмов зрительного внимания живых систем.

В связи с этим наиболее важными являются следующие направления развития полученных результатов:

- экспериментальный анализ возможности применения альтернативных методов кластеризации;
- разработка модели системы зрительного внимания, имеющей строгое математическое обоснование.



## Литература

- [1] D. Lowe. Distinctive Image Features from Scale-Invariant Keypoints. In *International Journal of Computer Vision*, volume 60, pages 91-110, 2004.
- [2] H. Bay et al. SURF: Speeded Up Robust Features. *Comput. Vis. Image Underst.*, number 110 (3), pages 346-359. 2008.
- [3] W.-L. Zhao, C.-W. Ngo, H.-K. Tan and X. Wu. Near-Duplicate Keypoint Identification with Interest Point Matching and Pattern Learning. In *IEEE Transactions on Multimedia*, volume 9(5), pages 1037-1048, 2007.
- [4] Y. Ke, R. Suthanakar, L. Huston. Efficient Near-Duplicate Detection and Sub-Image Retrieval. In *ACM Multimedia Conference 2004*, pages 869-876, 2004.
- [5] А. П. Кудряшов. Извлечение и сопоставление точечных особенностей. Электронный научный журнал *Исследовано в России*, т.10, 2007. <http://zhurnal.ape.relarn.ru/articles/2007/104.pdf>
- [6] K. Mikolajczyk, C. Schmid. An affine invariant interest point detector. In *Proceedings of European Conference on Computer Vision (ECCV)*, pages 128-142, 2002.
- [7] В. Ю. Пименов. Метод поиска нечетких дубликатов изображений на основе выявления точечных особенностей. *Труды ПОМИИ 2007-2008*, с. 145-158. СПб.:НУ ЦСИ, 2008.
- [8] M. Awrangjeb, G. Lu. Techniques for efficient and effective transformed image identification. *J. Vis. Commun.* 2009.
- [9] B. Zitová, J. Flusser. Image registration methods: a survey. *Image and Vision Computing*, volume 21, pages 977-1000. 2003.
- [10] S. F. Chang, W. Hsu, L. Kennedy, L. Xie, A. Yanagawa, E. Zavesky and D.-Q. Zhang. Columbia University TRECVID-2005 Video Search and High-Level Feature Extraction. In *TRECVID*, 2005.
- [11] X. Wu, C.-W. Ngo and Q. Li. Threading and Autodocumenting News Videos. In *Signal Processing Magazine*, 2006.
- [12] TREC Video Retrieval Evaluation (TRECVID) Web Site. <http://www-nlpir.nist.gov/projects/trecvid>
- [13] L. G. Brown. A survey of image registration techniques. *ACM Computing Surveys*, volume 24, pages 325-376. 1992.
- [14] Y. Ke, R. Suthanakar. PCA-SIFT: A More Distinctive Representation for Local Image Descriptors. In *Computer Vision and Pattern Recognition*, volume 2, pages 506-513, 2004.
- [15] T. Lindeberg. Scale-space theory in computer vision. Kluwer/Springer. 1994.

- [16]L. M. J. Florack. Image structure. Kluwer/Springer. 1997.
- [17]K. Mikolajczyk, C. Schmidt. A Performance Evaluation of Local Descriptors. *IEEE Trans. Pattern Anal. Mach. Intell.*, volume 27 (10), pages 1615-1630. 2005,
- [18]K. Mikolajczyk et al. A Comparison of Affine Region Detectors. *Int. J. of Comput. Vis.* volume 65 (1-2), pages 43-72, 2005.
- [19]M. A. Fischler, R. C. Bolles. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Communications of ACM*, volume 24(6), pages 381-395. 1981.
- [20]S. Frintrop. The High Repeatability of Salient Regions. *Proc. of ECCV Workshop «Vision in Action: Efficient Strategies for Cognitive Agents in Complex Environment»*. 2008.
- [21]T. Lindeberg. Scale-space. *Encyclopedia of Computer Science and Engineering*, volume IV, pages 2495-2504. John Wiley and Sons. 2009.
- [22]A. P. Witkin. Scale-space filtering. In *Proceeding of International Joint Conference on Artificial Intelligence*, pages 1019-1022, 1983.
- [23]J. J. Koenderink. The structure of images. In *Biological Cybernetics*, volume 50, pages 363-396, 1984.
- [24]А. В. Скурихин. Применение методов масштабируемого пространства в обработке сигналов. <http://www.spiiras.nw.ru/rus/conferences/ict/Skurihin110604.ppt>
- [25]S. Frintrop. VOCUS: A Visual Attention System for Object Detection and Goal-Directed Search. Ph.D. Thesis Rheinische Friedrich-Wilhelms-Universität Bonn. 2006.
- [26]C. Koch, S. Ullman. Shifts in Selective Visual Attention: Towards the Underlying Neural Circuitry. *Human Neurobiology*, volume 4, pages 219-227. 1985.
- [27]E. Niebur. Saliency map. *Scholarpedia*, volume 2 (8), page 2675. 2007
- [28]C. Koch, L. Itti, E. Niebur. A Model of Saliency-Based Visual Attention for Rapid Scene Analysis. *IEEE Trans. Pattern Anal. Mach. Intell.*, volume 20, pages 1254-1259. 1998.
- [29]S. E. Palmer. *Vision Science, Photons to Phenomenology*. MIT Press, Cambridge. 1999.
- [30]N. D. B. Bruce, J. K. Tsotsos. Spatiotemporal Saliency: Towards a Hierarchical Representation of Visual Saliency. *Proc. 5th Int. Workshop on Attention in Cognitive Systems*, pages 98-111. 2008.
- [31]P. Viola, M. Jones, Rapid Object Detection Using a Boosted Cascade of Simple Features. *Proc. of the 2001 IEEE Computer Society Conf.*

on *Computer Vision and Pattern Recognition*, volume 1, pages 511-518, 2001.

- [32] Д. Форсайт, Ж. Понс. Компьютерное зрение. Современный подход. Вильямс. 2004.
- [33] Н. Hügli, А. Bur. Adaptive Visual Attention Model. Proc. of Image and Vision Computing New Zealand, pages 233-237. 2007.
- [34] L. J. Heyer, S. Kruglyak and S. Yooseph. Exploring Expression Data: Identification and Analysis of Coexpressed Genes. In *Genome Research*, volume 9, pages 1106-1115, 1999.
- [35] Результаты для дорожки поиска нечетких дубликатов в коллекции изображений. Приложение. – Труды российского семинара РОМИП'2009.

## **Computationally Efficient Near-Duplicate Image Detection Method**

Vitaly Pimenov

Saint-Petersburg State University, Faculty of Applied Mathematics  
and Control Processes  
vitaly.pimenov@gmail.com

The paper develops a method for near-duplicate image detection by local interest point detection and matching. New technique based on visual attention model is proposed to improve speed efficiency. Interest points are detected in image scale-space and represented with SURF descriptors. Image matching is performed by matching descriptors corresponding to interest points that are most salient in terms of visual attention model. Experimental results show that new technique significantly improves computational efficiency without sacrificing quality of near-duplicate image detection.