

# Сравнение двух алгоритмов проективно-инвариантного распознавания плоских замкнутых контуров с единственной вогнутостью <sup>1</sup>

Притула Н.Е.  
ИППИ РАН  
pritulan@yandex.ru

Николаев П.П.  
ИППИ РАН  
nikol@iitp.ru

Шешкус А.В.  
ИСА РАН  
astdcall@gmail.com

## Аннотация

*В работе рассматриваются два алгоритма распознавания проективно преобразованных замкнутых плоских контуров, обладающих единственной вогнутостью. Для анализа изображения контуров моделируются с использованием ортогональной модели камеры с фиксированным фокусным расстоянием, что накладывает ограничение на размерность множества проективных преобразований. Из двух рассматриваемых алгоритмов один решает задачу в общем виде, а второй использует априорную информацию о модели камеры. Проверяется гипотеза об улучшенных точностных характеристиках второго алгоритма. Приводятся результаты 20000 численных экспериментов, подтверждающих гипотезу в среднем.*

Ключевые слова: *распознавание, гладкий контур, проективное преобразование, инвариантное отображение.*

## 1. Введение

Развитие и распространение компьютерной обработки информации привели к возникновению в середине XX века потребностей в технологиях, позволяющих машинам осуществлять автоматическое распознавание обрабатываемой ими информации. В результате получила развитие дисциплина "распознавание образов", решающая сложнейшие задачи из различных сфер человеческой деятельности: распознавание текста, речи, отпечатков пальцев.

Некоторые из этих задач хорошо изучены и решены в общем виде. Так, существует множество алгоритмов, решающих задачу распознавания двумерных объектов в двумерном пространстве, в частности оптическое распознавание символов. Если доступны чёткие

изображения текста, такие как сканированные печатные документы, то точность разработанных методов распознавания превышает 99%. [1].

Другие проблемы, напротив, представляют собой поле подзадач, не поддающихся стандартизации, либо требующих для однозначного анализа слишком большого вычислительного ресурса. Примером такой задачи может служить задача распознавания нежестких 3D моделей, в частности, распознавание лиц. Несомненно, решение этой задачи представляет большой интерес, особенно для области искусственного интеллекта.

В качестве промежуточного этапа на пути решения этой задачи можно выделить распознавание криволинейных двумерных объектов в трехмерном пространстве. Разработка алгоритмов в этой области послужит основой для решения задачи распознавания 3D моделей.

В некоторых случаях при решении задачи распознавания криволинейных двумерных объектов применяют алгебраические инварианты (такие как степень задающего полинома или набор параметров задающей кривой), однако ввиду нерегулярной структуры большинства объектов реального мира применение этих методов ограничено.

Дифференциальные инварианты также получили распространение при решении задач распознавания данного типа [2]. Достоинством такого подхода является многообразие объектов, для которых он применим. Тем не менее, ошибки, накапливающиеся при вычислении дифференциальных инвариантов высших порядков, затрудняют дальнейшую обработку данных. Таким образом, область применения данного подхода ограничена.

Так же, как и в ряде других работ [3,4], в данной работе для решения задачи распознавания применяется поиск опорных (проективно инвариантных) пар точек для формирования проективного базиса. Данный метод хорошо работает для кривых с большим числом особенностей.

<sup>1</sup>Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 13-01-12107-офи\_м

Задача, рассматриваемая в данной статье, представляет собой задачу классификации контуров криволинейных двумерных объектов в трехмерном пространстве. Приходящие на вход контуры сопоставляются с элементами коллекции эталонов. На основе сопоставления определяется, к какому эталону относится тестовый контур. При получении тестовых данных используется ортогональная модель камеры с фиксированным фокусным расстоянием. Это накладывает ограничение на множество тестовых контуров. Первый алгоритм не использует информацию о модели камеры и решает задачу в общем виде (как если бы проективное преобразование было произвольным). Второй алгоритм, напротив, решает задачу только на допустимом множестве проективных преобразований. В качестве тестовых данных используются замкнутые контуры, обладающие единственной вогнутостью ("рогалики" в классификации [5]). Выбор этого типа контуров в качестве тестовых данных обусловлен наличием у них четырех проективно устойчивых точек (две точки двойного касания и две точки перегиба), что позволяет легко построить базис проективного преобразования. Далее мы рассмотрим различные этапы проведения эксперимента, начиная с процесса создания тестовой выборки.

## 2. Создание тестовой выборки

Тестовые данные представлены коллекцией эталонов и поступающими на вход контурами (иначе тестовыми контурами). Элементы коллекции эталонов (Рис. 1) были получены с помощью кривых Безье, построенных на наборах базисных точек. Координаты этих точек выбираются вручную. Они расположены в квадрате  $[-40, 40] \times [-40, 40]$  условных единиц. Каждый эталон был получен путем "склейки" двух кривых Безье (Рис. 2). На рисунке изображены две кривых Безье, базисными точками которых являются наборы  $\{P_1, P_2, P_3, P_4, P_5\}$  и  $\{P_1', P_2', P_3', P_4', P_5', P_6', P_7'\}$ . Местами стыка двух кривых Безье являются точки двойного касания контура. Гладкость итоговой кривой достигается требованием совпадения касательных в местах стыка. Для всякой кривой Безье прямая, образованная первой и второй точкой из набора базисных точек ( $P_1$  и  $P_2$  соответственно), представляет собой касательную к кривой Безье в крайней базисной точке. Если расположить такие точки одной и другой кривой Безье ( $P_1, P_2, P_1', P_2'$ ) на одной прямой, то совпадение касательных в месте стыка обеспечит гладкость итоговой кривой (что и было сделано).

Для поступающих на вход контуров необходимо определить имеет ли место проективная эквивалентность с каким-либо элементом коллекции эталонов (задача классификации). Природа этих тестовых контуров не случайна. Все тестовые "рогалики" относятся к одному из  $k=10$  классов эквивалентности, порожденных элементами коллекции эталонов. Каждый эталонный элемент порождает один

класс эквивалентности, состоящий из  $n=100$  образов - проекций на различные плоскости под различными углами. Таким образом все тестовые данные представлены  $n \cdot k = 1000$  элементами. В ходе эксперимента каждый тестовый контур сопоставляется каждому эталону.

Каждый тестовый "рогалик" был построен при случайных параметрах камеры. Параметры камеры представляют собой тройку  $(C, N, \phi)$  - центр проецирования  $C$ , нормаль плоскости проецирования  $N$  и угол поворота базиса плоскости проецирования  $\phi$  относительно начального положения (рис 3). Координаты нормального вектора плоскости проецирования выбраны произвольно (однако длина вектора = 1). Угол поворота базиса на плоскости по часовой стрелке также представлен случайным числом от 0 до  $2\pi$ .

При построении тестовой выборки нами были предъявлены следующие требования:

Требования	Принятые меры
Гладкость контура эталона	Использование кривых Безье, склеенных гладким образом, для формирования контура
Отсутствие тестовых рогаликов, содержащих бесконечно удаленную точку, небольшие размеры тестового рогалика по сравнению с эталонами,	1) Координаты центра проецирования лежат в квадрате $[-30, 30] \times [30, 30]$ 2) Из потока сгенерированных данных выбираются контуры, диаметр которых не превышает 100
Существенность вогнутости контура	Отношение максимальных расстояний до прямой двойного касания со стороны внешней выпуклости и внутренней вогнутости не меньше 0.2. (рис 4)

Тестовые контуры подвергаются процедуре ресемплинга (передискретизации). Данная процедура проводится для эмуляции шумов, получаемых при работе с изображениями, полученными из камер. В нашей работе ресемплинг проводится с разным количеством итоговых точек (157, 207, 257, 307, 357, 407). Количество итоговых точек отражает степень зашумленности обрабатываемого изображения. Разное количество итоговых точек позволяет оценить зависимость качества работы алгоритмов, от степени зашумленности изображения.

## 3. Сравнение двух контуров

Расстояние между тестовым и эталонным контурами вычисляется следующим образом:

- 1) На вход поступают координаты тестового контура и эталона в виде двух массивов точек. Координаты этих контуров заданы в разных базисах
- 2) Алгоритм оценивает взаимное расположение плоскости эталона, плоскости тестового контура и камеры в предположении, что тестовый контур совпадает с эталонным. Затем осуществляется проецирование тестового контура на плоскость эталона. В результате получаем координаты спроецированного контура в базисе эталона.
- 3) С помощью выбранной метрики оцениваем близость двух контуров, полученное число - результат эксперимента.

Рассмотрим подробнее этап 3. Результатом работы каждого из двух алгоритмов является проецирование контура тестового "рогалика" на плоскость эталона. Для численного определения степени близости двух контуров в работе используется расстояние Фреше (минимум среди максимумов расстояний между соответствующими точками при всех возможных монотонных соответствиях между узлами двух ломаных) [6] (Рис. 5). Выбор такой метрики обусловлен возможностью нахождения поточечной корреспонденции (монотонного соответствия). Под поточечной корреспонденцией мы понимаем установление условного соответствия (соответствующие точки вообще говоря не проективно эквивалентны) между точками двух контуров, представляемых приближенно в виде ломаных. Если тестовый контур совпадает с эталонным (с точностью до проективного преобразования), то расстояние Фреше в идеале равно нулю.

#### 4. Описание алгоритмов

Алгоритм решения задачи по 4 точкам основан на построении матрицы проективного преобразования, коэффициенты которой можно найти, решая систему линейных уравнений, полученных из уравнений, связывающих координаты опорных точек в старом и новом базисе. В качестве таких 4 опорных точек берутся две точки двойного касания и две точки перегиба (Рис. 6). После нахождения матрицы проективного преобразования мы проецируем контур пришедшего "рогалика" на плоскость элемента коллекции эталонов.

Алгоритм решения задачи по 3 точкам основан на решении "проективной проблемы 3 точек", которая подробно описана в [7]. В качестве 3 базисных точек мы берём точки двойного касания и точку пересечения касательных в точках перегиба (в качестве касательной мы берём звено ломаной). Предполагается, что в силу большей устойчивости прямых относительно ошибок и искажений при выделении контура проективно преобразованного объекта, алгоритм по 3 точкам будет работать стабильнее (мы заменили проективно неустойчивые точки перегиба точкой пересечения их касательных)

Как видно из описания двух алгоритмов, единственное различие между ними заключается в методе нахождения матрицы проективного преобразования. При этом в первом случае метод нахождения матрицы проективного преобразования достаточно ясен, тогда как во втором случае требуется предоставить некоторые пояснения.

Дальнейшие пояснения в этой части статьи представляют собой беглое изложение решения проблемы трех точек, представленное в [7]. Пусть нам даны координаты трех пар точек в двух различных системах координат в  $R^2$ . Требуется найти проективное преобразование ортогональной модели камеры с фиксированным фокусным расстоянием, переводящее каждую из 3 точек в парную ей. Иначе: найти взаимное расположение двух плоскостей в пространстве (плоскости объекта и плоскости изображения), таких что 3 прямые, проходящие через пары соответствующих точек пересекаются в 1 точке (положение камеры). Эта конфигурация представляет собой тетраэдр. При этом вершина тетраэдра должна быть расположена на заданном (фокусном) расстоянии от плоскости изображения и проекция вершины на эту плоскость попадает в заданную точку (начало координат базиса плоскости изображения) (рис 7). Решение задачи разделяется на 3 этапа.

- 1) Найти 3 расстояния от положения камеры до точек, лежащих в плоскости изображения и 3 плоскостных угла тетраэдра (мы можем это сделать сразу, так как известно взаимное расположение вершины тетраэдра и точек изображения).
- 2) Найти расстояния от положения камеры до точек плоскости объекта (на этом этапе потребуется решить уравнение 4 степени).
- 3) Найти координаты положения камеры и координаты точек изображения в системе координат объекта.

Рассмотрим этап 2 более подробно. Пусть  $R_{ab}$ ,  $R_{ac}$ ,  $R_{bc}$  - стороны треугольника, находящегося в плоскости объекта), а  $Q_{ab}$ ,  $Q_{ac}$ ,  $Q_{bc}$  - соответствующие углы тетраэдра. Требуется найти длины расстояний от вершины тетраэдра до вершин плоскости объекта -  $a$ ,  $b$ ,  $c$  (стороны тетраэдра). Решение может быть получено из системы уравнений:

$$R_{ab}^2 = a^2 + b^2 - 2ab \cdot \cos(Q_{ab})$$

$$R_{ac}^2 = a^2 + c^2 - 2ac \cdot \cos(Q_{ac})$$

$$R_{bc}^2 = b^2 + c^2 - 2bc \cdot \cos(Q_{bc})$$

Путем ввода новых переменных  $x$  и  $y$  ( $b = x \cdot a$ ,  $c = y \cdot a$ ) можно свести данную систему из 3 уравнений к двум уравнениям: уравнению 4 степени, определяющим переменную  $x$  и уравнению, определяющему  $y$  через  $x$ . Разрешив уравнение на  $x$  и  $y$  и вычислив соответственно  $a$ ,  $b$ ,  $c$  мы приходим к решению 2-ого этапа задачи.

## 5. Численные эксперименты и их анализ

### 5.1. Вычисление внутриклассового расстояния

Внутриклассовые эксперименты представляют собой сопоставление эталона и всех тестовых объектов, входящих в класс эквивалентности эталона. В идеале внутриклассовые расстояния должны быть равны нулю. При проведении экспериментов внутриклассовое расстояние усреднялось при различных параметрах передискретизации, а именно 157, 207, 257, 307, 357 или 407 итоговых точек. Была заметна тенденция снижения внутриклассового расстояния при увеличении количества точек ресемплинга, что объяснимо повышением точности представления исходной кривой. В таблице 1 в качестве примера приведены данные по внутриклассовому расстоянию по классу №3.

### 5.2. Определение пороговых значений

Так как конечная прикладная задача - это задача классификации, т.е. определения по некоторому входному контуру, является ли он той или иной эталонной фигурой, то по имеющейся выборке был произведен расчет оптимального порогового значения для каждой эталонной фигуры (или класса), при превышении которого заданный контур классифицировался как не относящийся к этому классу.

Само пороговое значение имеет практическую значимость, но более интересен показатель качества такого порогового значения. Рассмотрим все входные контуры (1000 штук) с фиксированным количеством точек ресемплинга (307), для каждой эталонной фигуры рассчитаем долю входных контуров, классифицированных относительно нее неверно. Такую долю будем называть штрафом. В таблице 2 приведена информация по оптимальным пороговым значениям для рассматриваемой тестовой выборки и минимальным достижимым штрафам.

По таблице видно, что фигуры из седьмого и девятого класса тяжело отделяются от контуров других классов, оба алгоритма имеют на них ненулевой штраф. По всем остальным классам можно заметить, что, во-первых, комбинируя два алгоритма (используя для некоторых классов один из них, а для других - другой) можно добиться итогового нулевого штрафа, т.е. безошибочной классификации для каждого класса. Во-вторых, алгоритм "3 точки" имеет ошибки в меньшем числе классов, и суммарный по всем классам штраф алгоритма "3 точки" составляет 0,0137 против 0,0513 у алгоритма "4 точки", что говорит о лучшем качестве алгоритма "3 точки" для исследованного набора тестовых данных.

### 5.3. Сравнение распределений межклассовых расстояний

В результате проведенных экспериментов были определены медианные межклассовые расстояния,

которые были записаны в таблицы 3 и 4 (по одной для каждого алгоритма), в которых элемент  $i$ -ой строки и  $j$ -го столбца суть медиана расстояний при проецировании овалов класса  $j$  к элементу коллекции эталонов  $i$ . Затем эти таблицы были построено отнормированы так, чтобы диагональные элементы (т.е. внутриклассовое расстояние) стали равными единице.

По значениям всех ячеек данных таблиц, кроме диагональных, были построены гистограммы плотности распределения значений таблиц для каждого из алгоритмов. Они представлены на рисунке 8.

Нетрудно заметить, что моды распределения расстояний алгоритмов по 3 и 4 точкам расположены в отрезках [27-36] и [0-9] соответственно, на основании чего можно сделать вывод о лучшем качестве работы алгоритма по 3 точкам.

## 6. Выводы

В результате проведения экспериментов был достигнут следующий результат: подтвердилась гипотеза о лучших точностных характеристиках алгоритма, использующего априорную информацию о модели камеры. К недостаткам алгоритма можно отнести сложность реализации.

## 7. Приложение

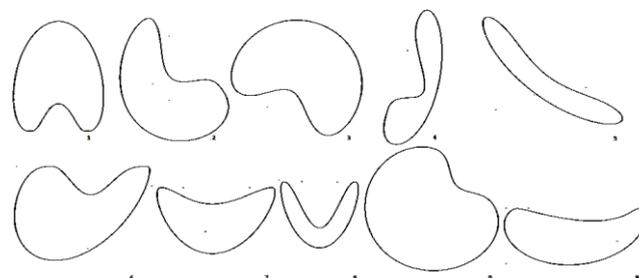
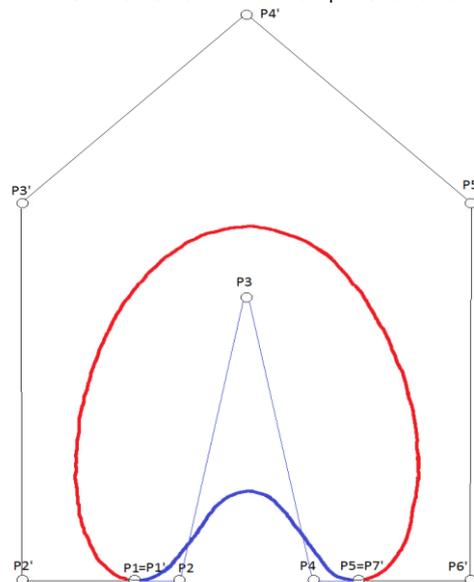
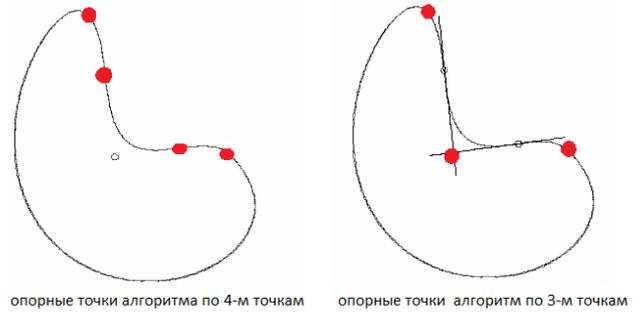
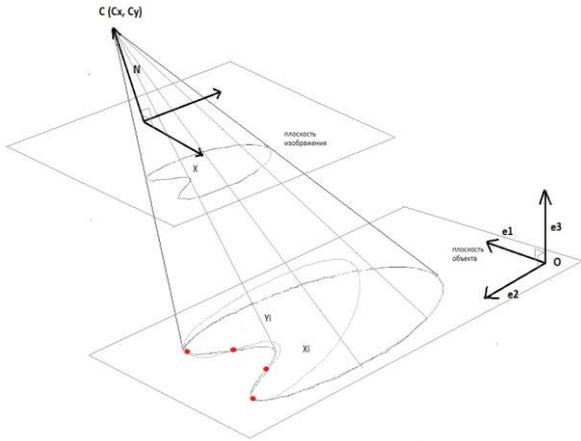


Рис 1. Элементы коллекции эталонов

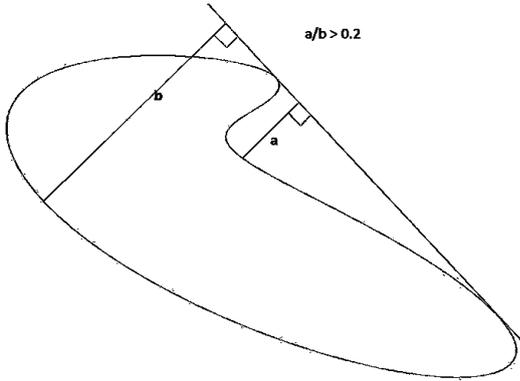


**Рис 2.** Склейка кривых Безье

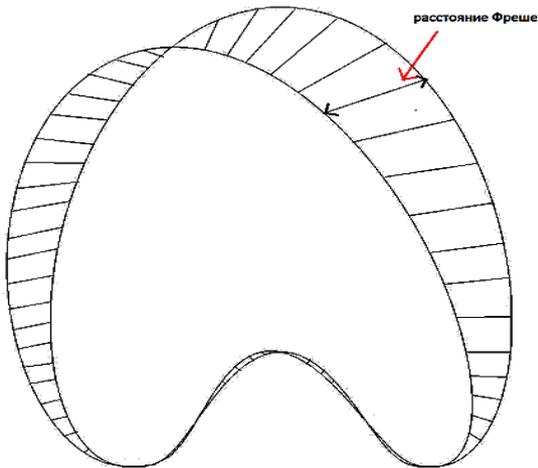


**Рис 6.** Выбор опорных точек

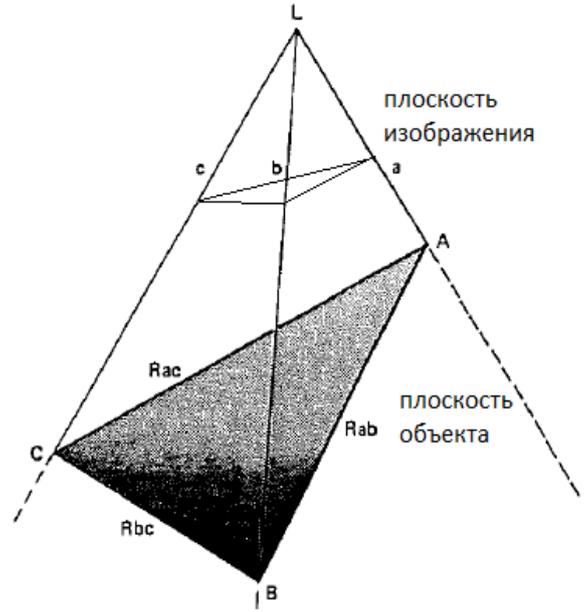
**Рис 3.** Проецирование изображения



**Рис 4.** Отношение внутреннего и внешнего максимального расстояния до прямой двойного касания



**Рис 5.** Расстояние Фреше



**Рис 7.** Проблема 3-х точек

**Таблица 1.** Среднее внутриклассовое расстояние и среднеквадратическое отклонение (класс №3)

Алгоритм\параметр ресемплинга	157	207	257	307	357	407
3 точки, среднее	0,025±0,015	0,018±0,01	0,014±0,0086	0,013±0,0078	0,013±0,0081	0,011±0,0067
4 точки, среднее	0,10±0,046	0,075±0,039	0,058±0,031	0,052±0,022	0,047±0,024	0,042±0,019

**Таблица 2.** Пороговые значения для классификации. Штрафы.

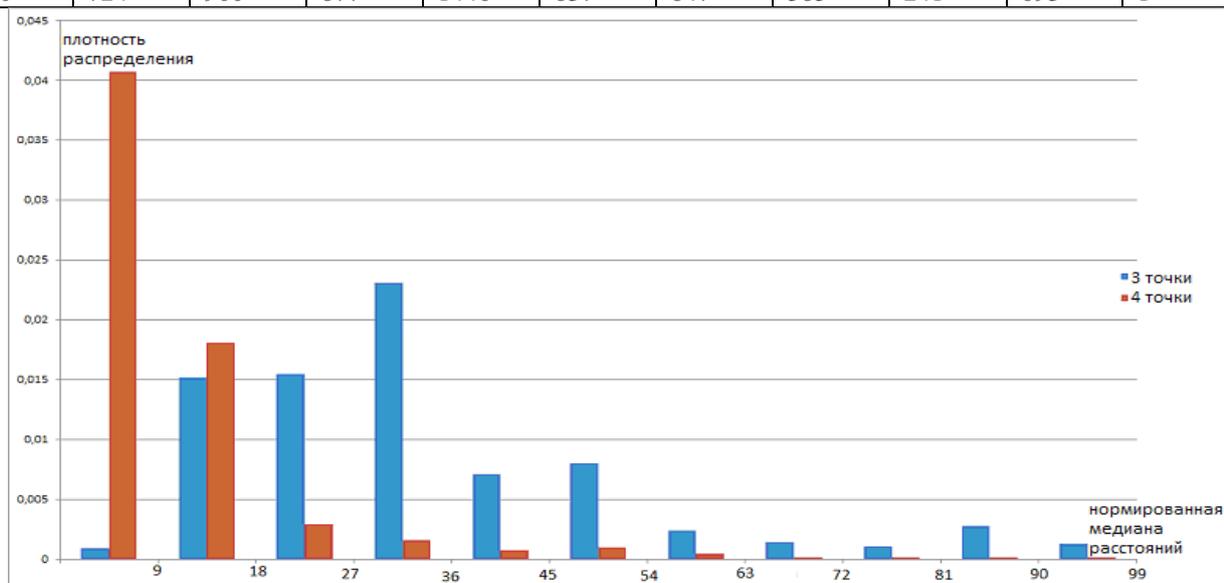
	1 класс	2 класс	3 класс	4 класс	5 класс	6 класс	7 класс	8 класс	9 класс	10 класс
3 точки, порог	0,0462	0,0445	0,0437	0,0368	0,0294	0,0669	0,0405	0,0442	0,0476	0,0256
3 точки, штраф	0	0	0	0	0,0027	0	0,0011	0	0,0099	0
4 точки, порог	0,0646	0,0737	0,1017	0,1038	0,07	0,1478	0,0745	0,0513	0,2766	0,1197
4 точки, штраф	0,0279	0,0029	0	0	0	0,0034	0,0075	0	0,0025	0,0071

**Таблица 3.** Нормированные медианные расстояния между классами, алгоритм «3 точки»

Класс	1	2	3	4	5	6	7	8	9	10
1	1	30,22	35,6	53	5521	35,45	32,9	52,5	35,34	46,8
2	62,5	1	14,56	32,75	39	9,35	24,37	33,35	20,26	26,3
3	52	21,76	1	42,3	38,1	29,8	17,8	43,5	16,6	35,5
4	90,39	21,34	28,57	1	29,23	31,23	28	26	36,5	23,63
5	2624	5218	1432	14,69	1	1432	45,6	1432	5332	7,17
6	31	9,83	11,35	19,86	25,58	1	15,87	19,1	12,1	17,77
7	900	153	222	355	15	1930	1	86,5	60,6	16,96
8	95,7	35,8	40,8	16,15	39469	29,58	50,4	1	54,18	18,2
9	2673	33,18	87,8	77,94	24,11	810	22,3	32,5	1	29,2
10	215	68	47	463	20	1069	15	34	33	1

**Таблица 4.** Нормированные медианные расстояния между классами, алгоритм «4 точки»

Класс	1	2	3	4	5	6	7	8	9	10
1	1	2	401	3,83	6,29	335,4	7,34	9,17	340	9,5
2	2,19	1	279,2	6,79	7	245,8	9,44	10,78	1210	11,67
3	7,83	7,56	1	8,45	9,34	4,84	9,77	10,7	15,34	10,94
4	7,18	5,08	8,9	1	9,96	9,38	11,65	10,74	18,19	12,22
5	21,56	30,85	3661	48,4	1	2841	5,06	5,85	3173	7,97
6	5,43	5,18	3,69	5,35	6,62	1	7,13	7,42	19,39	7,79
7	810	1051	889	1126	7,54	1000	1	6,23	841	7,05
8	208	259	5264	126	57,03	4068	35,54	1	4470	15,49
9	4,82	4,76	3,5	4,92	5,09	3,89	5,17	5,35	1	5,38
10	724	900	877	1446	837	847	585	241	693	1



**Рис. 8.** Гистограммы для рассматриваемых алгоритмов

## Литература

- [1] В.Л. Арлазаров, П.А. Куратов, О.А.Славин Распознавание строк печатных текстов // Сб. трудов ИСА РАН "Методы и средства работы с документами". – М.: Эдиториал УРСС. – 2000
- [2] А.С. Горбань, А.Н. Каркищенко Дифференциальные и интегральные инварианты в задачах обнаружения аффинных симметрий контура // Искусственный интеллект. 2005.
- [3] Николаев П.П., Николаев Д.П. Проективно инвариантное распознавание плоских контуров на примере кривых с симметриями // Труды ИСА РАН. 2009. Т. 45. с.209-221.
- [4] Николаев П.П. Распознавание проективно преобразованных плоских фигур. I. Анализ и инвариантное отображение составных овалов // Сенсорные системы. 2011. Т. 25. N 2. С. 118-137.
- [5] Николаев П.П. Модели константного зрительного восприятия. I. Проективно инвариантное представление гладких объектов // Сенсорные системы. 1995
- [6] Chambers, Erin Wolf and Colin de Verdière, Éric and Erickson, Jeff and Lazard, Sylvain and Lazarus, Francis and Thite, Shripad (2010) Homotopic Fréchet distance between curves or, walking your dog in the woods in polynomial time. Computational Geometry, 43 (3). pp. 295-311. ISSN 0925-7721
- [7] Martin A. Fischler and Robert C. Bolles (June 1981). «Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography». Comm. Of the ACM 24: 381–395.
- [8] S. Startchik, R. Milanese, T. Pun Projective and illumination invariant representation of disjoint shapes // IVC(16), No. 9-10, July 1998, pp. 713-723.
- [9] Полевой Д.В., Самойлов О.С. Задача контроля качества при создании и развитии систем оптического распознавания печатного текста. // Технологии программирования и хранения ия данных / Под ред.В. Л. Арлазарова, Н. Е. Емельянова, Труды Института системного анализа РАН; Т. 45. — М.: ЛЕНАНД, 2009.
- [10] Kalle Astorm Fundamental Difficulties with Projective Normalization of Planar Curves // Applications of Invariance in Computer Vision, pp. 199-214, 1993