

УДК 004.932.2

**В.И. Махов, Д.М. Бочаров**

Донецкий национальный технический университет, г. Донецк  
кафедра программного обеспечения интеллектуальных систем

## **ИССЛЕДОВАНИЕ МЕТОДОВ ОПРЕДЕЛЕНИЯ ОПОРНЫХ ТОЧЕК НА ИЗОБРАЖЕНИИ**

### *Аннотация*

*Махов В.И., Бочаров Д.М. Исследование методов определения опорных точек на изображении. Проведен анализ существующих алгоритмов для определения опорных точек на изображении. Исследован метод определения опорных точек в задачах поиска лиц на изображениях и рассмотрен алгоритм SIFT.*

*Ключевые слова:* опорные точки, особые точки, точка ядра, локализация.

**Постановка проблемы.** Обнаружение опорных точек на изображениях это важная цель во многих авторизированных системах. Алгоритмов обнаружения опорных точек уже существует достаточно много. Большинство из них могут эффективно обнаружить точку ядра, когда качество изображения хорошее. Но когда качество изображения плохое, эффективность этих алгоритмов значительно снижается.

На первый взгляд наборы соответствующих точек на изображениях дают довольно мало информации об изображениях и наблюдаемой сцене, но на самом деле это не так. Например, если рассматриваются несколько изображений одной сцены и наборы соответствующих точек на этих изображениях, то есть возможность определить настройки и положение камеры для каждого изображения. К сожалению, на данный момент нет универсального алгоритма качественно решающего данную задачу. Связано это с тем, что при съемке сцены изображения ее точек подвергаются различным искажениям, например проективные преобразования, связанные с перемещением камеры либо объектов сцены, изменения освещенности сцены и т.д.

**Анализ литературы.** Рассмотрена актуальность выделения опорных (особых) точек на изображениях [1-3], проведен анализ математических методов и алгоритмического обеспечения для локализации объектов на изображениях [3-5].

**Цель статьи** – исследование методов и обзор существующих алгоритмов для определения опорных (особых) точек на изображении.

**Постановка задачи.** Наиболее популярный на сегодня подход к решению задач локализации объектов на изображениях – это использование скользящих окон [1-2]. В этом случае, как правило, возникают сложности с выбором размера окна, т. е. определением масштаба локализуемых объектов. Также трудности возникают при поиске повернутых на некоторый неизвестный угол объектов. Для разрешения этих проблем часто используют несколько окон различного размера, сравнение содержимого с несколькими изображениями, на каждом из которых объект повернут на различные углы. Однако вычислительная сложность исходного алгоритма при этом возрастает настолько, что, как правило, приходится отказываться от поиска повернутых объектов, а часто также и от определения масштаба.

В рамках данной работы необходимо рассмотреть методы и алгоритмы локализации, использующую понятие особых точек для решения задач локализации повернутых и масштабированных графических объектов.

**Результаты исследований.** Рассмотрим преобразования изображений, называемые аффинными:

$$T_a(a_{11}, a_{12}, a_{13}, a_{21}, a_{22}, a_{23}): f_0(a_{11}x + a_{12}y + a_{13}, a_{21}x + a_{22}y + a_{23}) \quad (1)$$

Их удобно описывать так называемой матрицей аффинного преобразования с использованием однородных координат [3]:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{pmatrix} \\ T_a(A): f(z) = f_0(Az), z = (x, y, 1)^T. \quad (2)$$

Если известны координаты центров хотя бы трех объектов на искомом изображении и координаты центров этих же объектов на базовом изображении, то можно легко восстановить параметры преобразования координат, отображающего одно изображение на другое (рис. 1).

Алгоритм локализации на основе особых точек выглядит следующим образом:

- на искомом и базовом изображениях выделяются множества некоторых объектов (их центры – особые точки);
- находятся соответствия трех или более особых точек искомого изображения особым точкам базового изображения;
- вычисляются параметры преобразования координат.

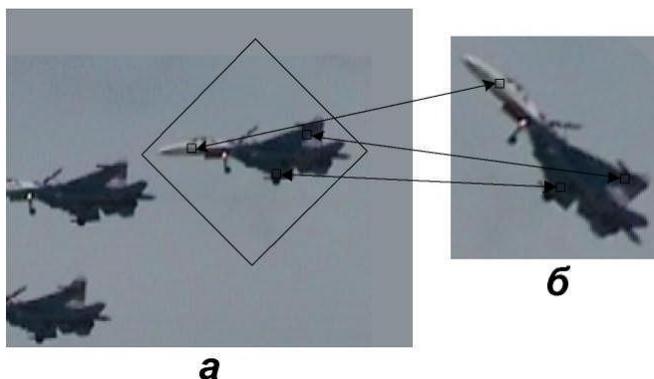


Рисунок 1 – Восстановление параметров преобразования координат:  
а – базовое изображение; б – искомое

Для практического решения проблемы выделения объектов необходимо выбрать такое их множество, для которого бы существовали эффективные процедуры их нахождения на произвольном изображении. Потому в качестве объектов часто используют локальные экстремумы функции яркости, функций составляющих цвета, локальные экстремумы градиента функции яркости и т.д. Центры таких объектов называют особыми точками. Главное свойство, которому должны удовлетворять особые точки, они должны сохранять своё взаимное расположение при всех допустимых преобразованиях изображений.

Взаимное расположение локальных экстремумов функции яркости будет сохраняться и при таких операциях, как увеличение масштаба, поворот, зеркалирование, не слишком сильное изменение яркости, контраста, гаммы. Однако проблемной остается операция уменьшения масштаба т. к. при этом две особых точки могут слиться в одну. Также наличие шума, например, при получении изображения с цифровой фотокамеры может привести к появлению большого числа ложных особых точек.

Для нахождения параметров преобразования на искомом изображении выбирают  $n \geq 3$  особых точек:  $(x_i^\omega, y_i^\omega)$   $i=1 \dots n$ , пусть на базовом изображении им соответствуют точки  $(x_i^0, y_i^0)$ , что они не лежат на одной прямой. Из этих точек составляются матрицы следующего вида:

$$U^{\omega 0} = \begin{pmatrix} x_1^{\omega 0} & x_2^{\omega 0} & \dots & x_n^{\omega 0} \\ y_1^{\omega 0} & y_2^{\omega 0} & \dots & y_n^{\omega 0} \\ 1 & 1 & \dots & 1 \end{pmatrix}; \quad U^\omega = \begin{pmatrix} x_1^\omega & x_2^\omega & \dots & x_n^\omega \\ y_1^\omega & y_2^\omega & \dots & y_n^\omega \\ 1 & 1 & \dots & 1 \end{pmatrix} \quad (3)$$

В случае идеальных изображений, к которым было применено исключительно преобразование координат, определяемое матрицей аффинного преобразования  $A$ , выполняется равенство:

$$U^\infty = AU^{\omega 0}. \quad (4)$$

В таком случае аффинное преобразование можно искать как решение оптимизационной задачи:

$$A^* = \arg \min_{A \in R^6} [U^\omega - AU^{\omega 0}] \quad (5)$$

где под символом  $[\ ]$  понимается сумма квадратов элементов матрицы:

$$[M] = \sum_{i=1}^{rows(M)} \sum_{j=1}^{cols(M)} M_{ij}^2. \quad (6)$$

Для случая  $n = 3$ :

$$A \approx U^\omega (U^{\omega 0})^{-1}, |U^{\omega 0}| \neq 0. \quad (7)$$

Решение задачи (7) для  $n > 3$  также может быть найдено аналитически, но трудоемкость его нахождения растет с  $n$ , потому для высоких  $n$  может потребоваться решение уже численными процедурами, что намного повышает вычислительную сложность задачи в целом, что часто крайне нежелательно.

Выделив на изображениях некоторые множества объектов и определив их центры, необходимо найти соответствие особых точек базового изображения особым точкам искомого изображения. Существует несколько путей решения этой задачи.

В некоторых случаях удастся отнести каждый из объектов, выделенных на изображении, к отдельному классу. Так, например, объекты, выделенные на изображениях рис. 2, можно отнести к классам «нос самолета», «правое крыло самолета», «левое крыло самолета». Если каждый из объектов изображения имеет уникальный класс, то легко сопоставить объекты базового и искомого изображений. Однако на практике крайне редко удастся назначить каждому объекту уникальный класс, что является главным недостатком данного метода.

Этот подход заключается в переборе возможных соответствий групп из  $n$  особых точек искомого изображения группам из  $n$  особых точек базового изображения с оценением степени успеха такого сопоставления при помощи какого-либо критерия  $q(U^\infty, U^{\omega 0})$ .

Например, при  $n > 3$  в большинстве случаев не удастся найти ни одного соответствия групп из  $n$  особых точек, которое бы полностью удовлетворяло равенству (4), поэтому в качестве критерия можно выбрать степень расхождения решения с этим равенством:

$$q(U^\omega, U^{\omega 0}) = U^\omega - AU^{\omega 0}. \quad (8)$$

**Описание алгоритма SIFT.** Алгоритм SIFT разделяется на 5 частей [4]:

1) построение пирамиды гауссиан и их разностей. На этом шаге обеспечивается инвариантность к масштабированию;

- 2) определение экстремумов;
- 3) уточнение особых точек;
- 4) определение ориентаций особых точек (обеспечивается инвариантность к повороту);
- 5) построение дескрипторов (обеспечивается инвариантность к освещению, шуму, изменению положения камеры).

На первом шаге алгоритма *SIFT* строится масштабируемое пространство изображений – набор изображений, сглаженных фильтром Гаусса:

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (9)$$

где  $(x, y)$  – координаты точки,  $\sigma$  – радиус размытия.

По ним строится разность гауссиан  $D(x, y, \sigma)$  – попиксельное вычитание изображений в одной октаве с разным коэффициентом размытия. На этом шаге обеспечивается инвариантность к масштабированию. Затем определяются экстремумы, которые заносятся в список потенциальных особых точек.

Далее происходит уточнение особых точек, которое состоит из двух составляющих:

1) исключаются точки с малой контрастностью с помощью вычисления экстремума разности гауссиан;

2) исключаются граничные точки (точки, имеющие большой локальный изгиб вдоль границы и малый в перпендикулярном направлении).

На конечном шаге алгоритма *SIFT* для окрестности особой точки вычисляются изменения яркостей точек, по которым строится дескриптор. Дескриптор – это вектор из 64 чисел, позволяет получить инвариантность относительно положения камеры. Затем дескриптор нормализуется, за счёт чего достигается инвариантность относительно изменения освещения.

В качестве модели, удовлетворяющей вышеперечисленным требованиям, было рассмотрено описание изображения зрачка как центра локального минимума функции яркости [5]. Действительно, такая модель достаточно проста для того, чтобы можно было реализовать эффективные алгоритмы на ее основе, и в то же время достаточно «инвариантна», т.к. почти не зависит от размера зрачка. Помимо этого, с помощью нее хоть и с меньшей эффективностью, но все же могут быть описаны ситуации, когда глаза закрыты, т. к. обычно можно обнаружить по крайней мере один четко выраженный минимум на изображении ресниц (рис. 2).



Рисунок 2 – Центры локальных минимумов функции яркости в случае полуоткрытых и закрытых глаз

Для ускорения процедуры поиска можно воспользоваться еще некоторой априорной информацией, например, если известно, что видеочкамера имеет определенное положение и люди стоят либо сидят, то можно предположить, что изображение лица наклонено не слишком сильно, например, не более, чем на  $\pm 45^\circ$ . Также может быть известно, что люди находятся не дальше и не ближе какого-либо расстояния от видеочкамеры, тогда размер лица на изображении будет достаточно определенным, в данном случае можно наложить дополнительные ограничения на возможный масштаб изображения (фактически на расстояние между особыми точками).

Также некоторое количество особых точек может быть отсеяно по величине заданного заранее порога.

Все эти эмпирические предположения позволяют наложить дополнительные ограничения на множество возможных преобразований  $A$ , а, следовательно, и сократить количество анализируемых фрагментов изображений, существенно сократив время работы алгоритма (рис.3).

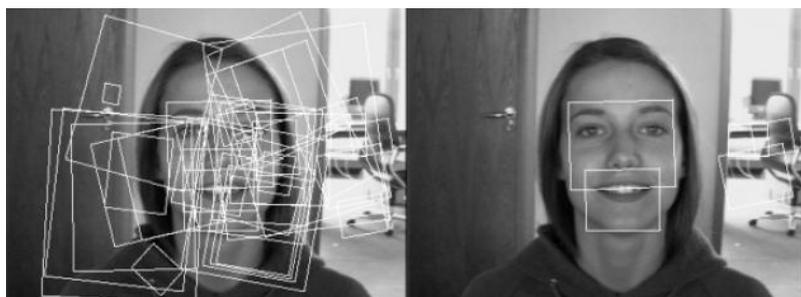


Рисунок 3 – Множество фрагментов до и после наложения ограничений на возможный масштаб и наклон головы

**Выводы.** Алгоритмы, использующие понятие особых точек, могут быть применены для решения широкого круга задач, связанных с локализацией изображений. Они могут давать значительно лучшие результаты в случае наличия масштабных преобразований и поворотов изображений по сравнению с алгоритмами, построенными по принципу скользящего окна.

Использование алгоритмов, основывающихся на переборе, выглядит наиболее перспективным, однако, их вычислительная сложность достаточно велика. Тем не менее, как показал анализ применения алгоритма SIFT в задаче локализации изображения лица, использование дополнительной эмпирической информации является эффективным путем снижения вычислительной сложности.

### Список литературы

1. Методы преобразования 2D-фотографий помещения в 3D-модель / Д.М. Бочаров, Р.А. Сорокин // Сучасні інформаційні системи і технології : матеріали другої міжнародної науково-практичної конференції (Суми, 21-24 травня 2013 року) / Сумський державний університет. Суми, 2013. – С. 150-151.
2. Предварительная обработка 2D-изображений при 3D-реконструкции / Д.М. Бочаров, Р.А. Сорокин // Искусственный интеллект. Интеллектуальные системы ИИ-2013 : материалы международной научно-технической конференции (пос. Кацивели АР Крым, 23 – 27 сентября 2013 года) / Донецкий национальный технический университет. Донецк, 2013. – С. 70-72.
3. Васильев В. Н. Математические методы и алгоритмическое обеспечение анализа и распознавания изображений в информационно-телекоммуникационных системах / В. Н. Васильев, И. П. Гуров, А. С. Потапов // Всероссийский конкурс обзорно-аналитических статей по приоритетному направлению «Информационно-телекоммуникационные системы» [Электронный ресурс] – Режим доступа: URL: <http://www.itc.edu.ru/itkonkurs2008/>
4. Построение SIFT дескрипторов и задача сопоставления изображений [Электронный ресурс] – Режим доступа: URL: <http://habrahabr.ru/post/106302/>
5. Krestinin I. A. Peculiar Point Technique For Object Detection in Image Analysis / I. A. Krestinin, O. S. Seredin // 8th International Conference on Pattern recognition and Image analysis: New information technologies. – PRIA-8 - 2007, Yohkar-Ola. 113 с. – 2007 у.