

## Распознавания с использованием активного сенсора. Алгоритмы и оценки.

Д.И. Медников, С.Ю. Сергунин, М.И. Кумсков

Московский Государственный Университет, Москва, Россия.

В работе описывается метод распознавания, использующий активный сенсор и приводятся теоретические оценки его вычислительной сложности.

### Постановка задачи и определения.

Задача распознавания объектов на полутоновых изображениях ставится следующим образом. Пусть дан набор полутоновых изображений проекций  $Pr_j(O_i)$  объектов  $O_i$ , набор полутоновых изображений эталонных фонов  $B_k$  и полутоновое изображение тестовой сцены  $Sc$  со сложным фоном. Требуется найти какие объекты, в каких проекциях находятся на сцене, и в какой части сцены они находятся, или же сообщить, что известных объектов на сцене нет, то есть построить функцию  $F(Sc, O_i) = \{e_i, pos_i, orient_i, scale_i\}$ , где  $e_i = \begin{cases} 0, & \text{если объект } O_i \text{ отсутствует на сцене } Sc \\ 1, & \text{если объект } O_i \text{ присутствует на сцене } Sc \end{cases}$ ,  $pos_i = (x_i, y_i)$  – координаты объекта  $O_i$  на сцене  $Sc$  (определено если  $e_i = 1$ ),  $orient_i = (\rho_i, \theta_i, \varphi_i)$  – ориентация проекции  $Pr(O_i)$  присутствующей на сцене  $Sc$  относительно эталонной системы координат объекта  $O_i$ ,  $scale_i$  – масштаб проекции  $Pr(O_i)$  присутствующей на сцене  $Sc$ , относительно эталонного масштаба объекта  $O_i$ .

В качестве объектов  $O_i$  в работе принимаются твердые трехмерные объекты. Под их проекциями  $Pr(O_i)$  понимаются изображения объектов с определенной точки на однородном фоне. В качестве особых точек рассматриваются характерные участки изображений, не имеющие в своей  $r$ - окрестности соседей. ОТ описываются своим положением на изображении и числовым маркером, указывающим тип их окрестностей.

Под конфигурацией особых точек  $Conf(O_i(Im))$  в области понимается вектор, описывающий некоторым образом их взаимное расположение в ней. Рейтингом особых точек относительно друг друга называется  $Rank(O_i, O_j)$  число, описывающее «похожесть» конфигураций их окрестностей.

Особенностью предлагаемого подхода к распознавания является введение понятия *активного сенсора*. При распознавании система получает данные об изображениях только через окно сенсора, и «забывает» о других частях изображения. При этом окно перемещается по

изображению и меняет свои размеры в зависимости от этапа распознавания и текущей проекции кандидата нахождение на сцене.

### Описание схемы распознавания.

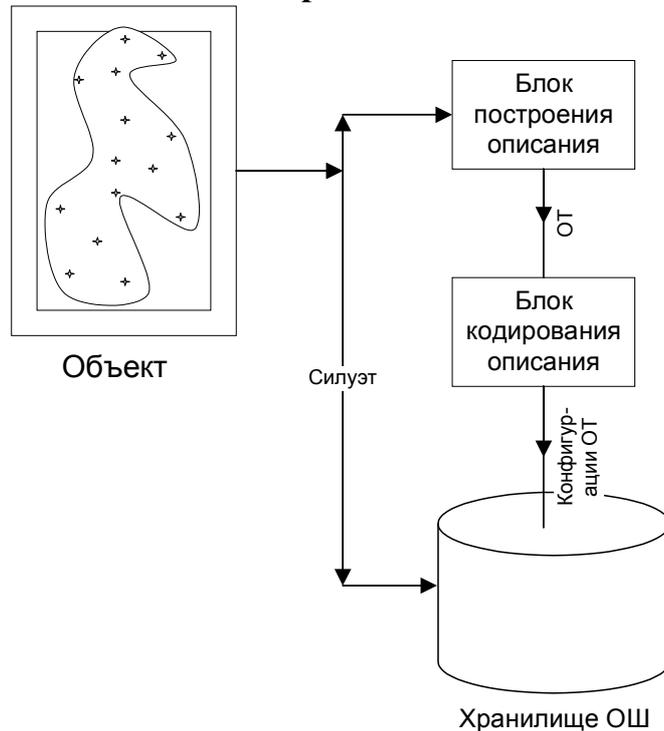


Рисунок 1. Схема наполнения базы ОШ.

Этап наполнения базы проходит при активном участии человека оператора, то есть является полуавтоматическим. В систему заносятся изображения объектов  $O_i$ . В **Блоке построения описания** на каждом изображении  $Im$  для каждого масштаба (количество масштабов используемых в работе является параметром системы) этого изображения при помощи того или иного алгоритма определяется положение особых точек  $\{SP_i(Im)\}$ , далее им присваиваются некоторые характеристики участка изображения, составляющего их ближайшую окрестность (например средняя яркость этого участка, или количество ветвей контуров объектов пересекающихся в особой точке), то есть каждая особая точка характеризуется своими координатами и вектором характеристик своей окрестности. Некоторые из них выделяются оператором как реперные, они должны располагаться в самых характерных участках объекта на изображении. После того как ОТ всех объектов выделены (они составляют множество  $\{SP_i\}$ ), в том же блоке проводится их маркировка на основе кластер анализа векторов характеристик, соответствующих особым точкам (в качестве маркера берется номер кластера  $i$  в множестве  $\{KL_i\}$ ). Таким образом, многомерное пространство векторов характеристик сводится к одномерному пространству маркеров. Полученные маркированные особые точки поступают в **Блок кодирования описания**, где описываются

векторами своих конфигураций  $K(SP_i)$ , то есть происходит обратный переход от одномерного пространства маркеров в многомерное пространство векторов конфигураций, но теперь описания ОТ зависят только от других особых точек, а не от соответствующих им участков изображения. Так получаются необходимые нам описания изображений  $D(Im_i)$ , абстрагированные от самих изображений, то есть двумерная матрица пикселей изображения переводится в плоский граф своих особых точек, таким образом, существенно уменьшается вычислительная сложность при дальнейшем распознавании. Полученные векторы записываются в **Хранилище объектных моделей** (базу данных), параллельно туда же для каждого объекта записывается его силуэт  $Sil$ . В хранилище вектора конфигураций особых точек индексируются для дальнейшего быстрого поиска похожих по конфигурациям ОТ по их индексам. Итак, в **Хранилище** оказываются все модели проекций объектов  $M(Pr_i(O_j))$ .

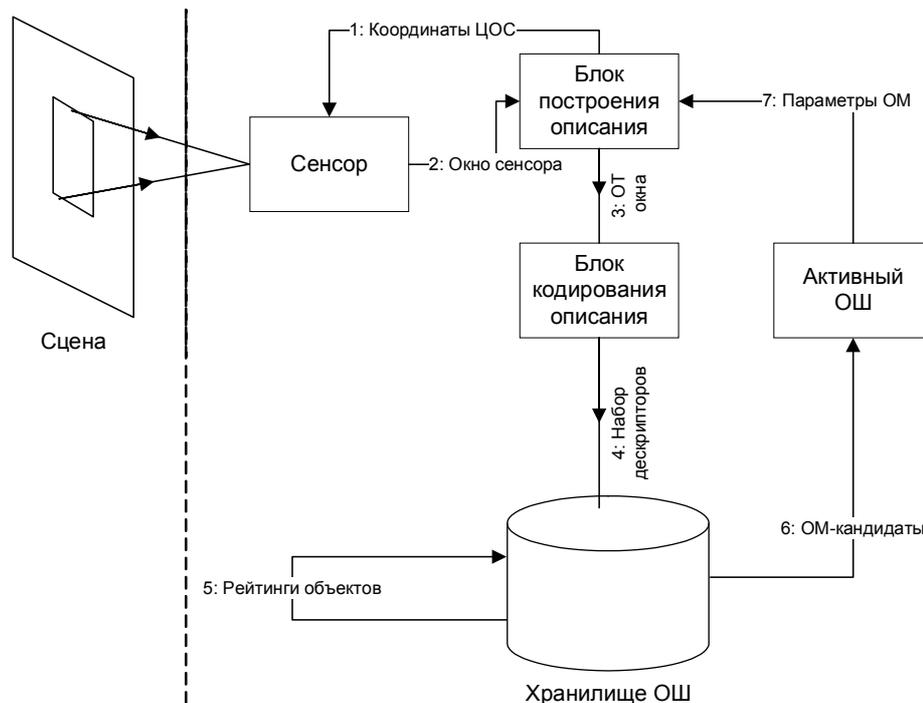


Рисунок 2. Выбор активной модели.

Подробная схема выбора активной модели выглядит следующим образом:

1) Блок построения описания задает первоначальные координаты центра  $(x,y)$  окна сенсора  $WAS(x,y,s)$  и уровень масштаба изображения  $s$ . Сначала  $s$  выбирается максимальным, так чтобы все изображение попало в окно.

2) Сенсор делает «снимок» окна с заданными координатами и передает в Блок построения описания, где на этом участке изображения выделяются и маркируются особые точки.

3) Набор маркированных ОТ окна сенсора передается в Блок кодирования описания, где строится набор его дескрипторов (векторов конфигураций ОТ), т.е. представление набора в виде символьных строк.

4) Набор дескрипторов поступает в хранилище ОМ и каждая модель в хранилище сравнивает полученные дескрипторы с собственными, хранимыми внутри модели.

5) При этом вычисляется и накапливается рейтинг модели  $P(Im)$  – показатель, характеризующий количество совпавших дескрипторов, а тем самым, возможность присутствия на сцене объекта  $Im$ .

6) Модель, набравший максимальный рейтинг становится активной моделью - кандидатом. Остальные модели становятся моделями кандидатами (неактивными), только если их рейтинг превысит порог активации.

7) Активная модель передает свои параметры в Блок построения описания, для дальнейшего подтверждения.

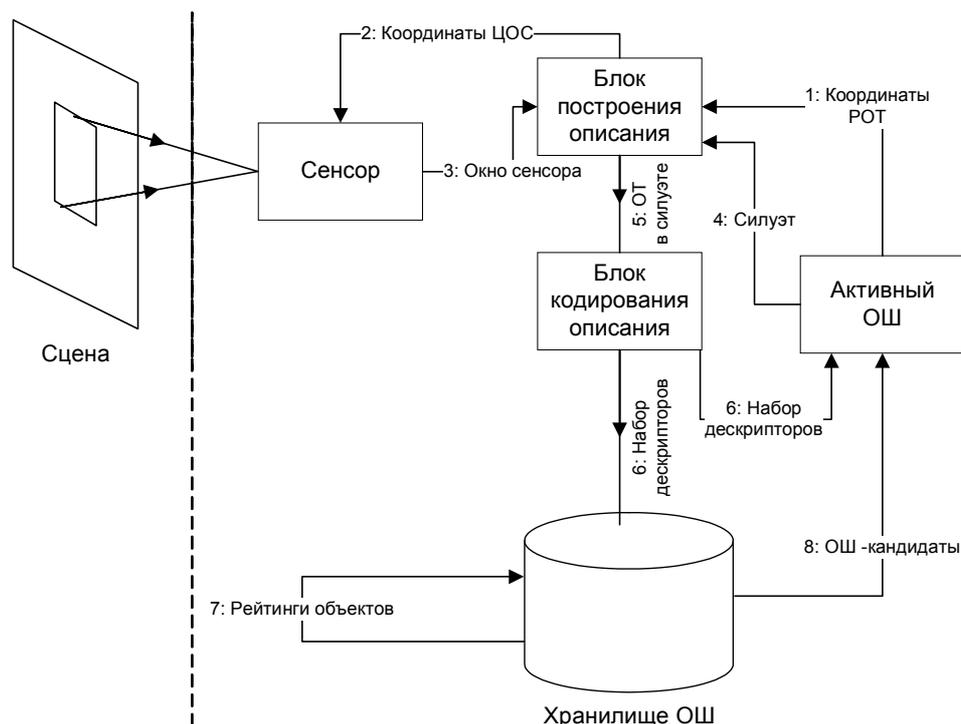


Рисунок 3. Подтверждение активной модели.

Второй этап процесса распознавания - подтверждение активной модели-кандидата:

1) Активный ОШ  $P(O)$  передает координаты своих реперных особых точек  $\{RSP_i(P)\}$  в Блок построения описания для установки в них  $WAS$ .

2) Блок построения описания задает координаты  $WAS$  и уровень пирамиды изображения.

3) Сенсор передает полученное окно в Блок построения описания

4) Активная модель передает силуэт объекта в Блок построения описания

5) Блок построения описания формирует набор особых точек изображения с учетом силуэта объекта, т.е. отсекая ОТ, соответствующие фону. Полученный набор поступает в Блок кодирования описания

6) Блок кодирования описания строит дескрипторы и передает их в Хранилище ОШ и в Активную модель  $P(O)$  для вычисления рейтингов.

7) В Хранилище ОМ каждая модель отдельно вычисляет свой рейтинг. Если после прохождения окном сенсора всех  $\{RSP_i(P)\}$  рейтинг активной модели не превышает порога подтверждения, то уровень пирамиды изображения понижается и происходит переход в шаг 1) для этого уровня. Если порог активации превышает в какой-то момент, то объект  $O$  считается найденным и подтверждение происходит для следующей по рейтингу модели-кандидата. Иначе считается, что в данной позиции на изображении объекта  $O$  нет, и подтверждение происходит для следующей по рейтингу модели-кандидата.

8) Возможно пополнение списка моделей кандидатов, если еще какие-то модели набрали рейтинг, больше порога активации.

При отсутствии моделей кандидатов производится поиск активного кандидата на более низком уровне пирамиды, чем в прошлый раз.

Итак, были рассмотрены основные виды работы РИАС системы, определена ее базовая модульную структуру и общий алгоритм работы. Были показаны основные отличительные особенности РИАС системы – ограничение области видимости системы при помощи окна активного сенсора, двухуровневая система распознавания, быстро находящая точки интереса на огрубленных изображениях и проверяющая их на ограниченных силуэтами объектов участках. Первая особенность позволяет сделать распознавание более интеллектуальным – каждая модель подтверждается по-своему, без каких либо усилий со стороны человека оператора, что улучшает его точность. Вторая – позволяет сохранить вычислительные ресурсы за счет необходимости проверки меньшего числа особых точек на первом «грубом» этапе и подробного исследования лишь отдельных участков изображения ограниченных силуэтами. Использование силуэтов так же повышает качество разделения объект\фон, т.к. при подробной верификации на втором этапе, вносящие ошибки, особые точки фона не используются.

**Теоретические оценки.**

Построим теоретические оценки сложности предлагаемого алгоритма.

*Лемма 1.* На прямоугольном изображении размеров  $M \times N$ , не может быть выделено более  $\frac{2M*N}{R^2}$  особых точек, если по определению ОТ в ее  $R$  окрестности не может быть других ОТ, в независимости от алгоритма выделения.

*Доказательство.* Покроем прямоугольник квадратами размеров  $\frac{R}{\sqrt{2}} \times \frac{R}{\sqrt{2}}$ . Их будет не более  $\frac{\sqrt{2}M}{R} * \frac{\sqrt{2}N}{R} = \frac{2M*N}{R^2}$  штук. Каждый такой квадрат имеет диагональ длины  $R$ , поэтому по определению ОТ, в нем может располагаться не более одной особой точки (при любом положении точки внутри квадрата, он целиком попадет в ее  $R$  окрестность). Таким образом, число ОТ будет не превышать число таких квадратов, то есть  $\frac{2M*N}{R^2}$ .

*Следствие 1.* При уменьшении масштаба изображения в 2 раза, максимальное число ОТ, которые могут на нем располагаться сократится в 4 раза.

*Доказательство.* Это утверждение является элементарным следствием Леммы 1. В случае уменьшения масштаба изображения в 2 раза, его линейные размеры тоже уменьшатся в 2 раза ( $M \times N \rightarrow \frac{M}{2} \times \frac{N}{2}$ ), таким образом, по Лемме 1 вместо  $\frac{2M*N}{R^2}$  особых точек, на нем максимально сможет поместиться  $\frac{M*N}{2*R^2} = \frac{1}{4} * \frac{2M*N}{R^2}$ .

*Лемма 2.* При одноуровневом распознавании без уменьшения масштаба, рассматриваемых изображений (классическая схема) потребуется  $\frac{4M^2*N^2}{R^4} * K$  операций сравнения. Где  $K$  – число изображений в базе,  $M \times N$  их размер,  $R$  – радиус пустой окрестности ОТ.

*Доказательство.* Пусть в базе имеется  $K$  изображений объектов размером  $M \times N$ . особых точек на них суммарно будет не более  $\frac{2M*N}{R^2} * K$ . На исследуемом изображении будет не более чем  $\frac{2M*N}{R^2}$  особых точек. Таким образом для поиска объектов будет проведено  $\frac{4M^2*N^2}{R^4} * K$  операций сравнения.

*Теорема 1.* При предложенном подходе к распознаванию потребуется не более чем  $\frac{1}{2^{2(l+1)}} * \frac{4M^2*N^2}{R^4} * K + \frac{\pi P^2 * M^2 * N^2 * S}{R^{6*2^{z-3}}}$  действий для распознавания объекта на изображении. Где  $K$  – число изображений в базе,  $M \times N$  их размер,  $R$  – радиус пустой окрестности ОТ,  $2^l$  масштаб огрубления на первом шаге алгоритма распознавания,  $P$  – радиус исследуемых окрестностей реперных ОТ,  $\frac{1}{2^z}$  доля реперных ОТ среди всех особых точек

изображения,  $S$  - число изображений кандидатов выбираемых в конце первого шага.

*Доказательство.* Пусть в базе имеется  $K$  изображений объектов размером  $M \times N$ . На первой стадии распознавания работа идет на масштабе, уменьшенном в  $2^l$  раз, поэтому особых точек на них суммарно будет не более  $\frac{1}{2^{2l}} * \frac{2M*N}{R^2} * K$ . Таким образом, для сравнения всех их особых точек с точками исследуемого изображения потребуется произвести  $\frac{1}{2^{2(l+1)}} * \frac{4M^2*N^2}{R^4} * K$  сравнений особых точек. На втором шаге сравниваются  $P$  окрестности реперных особых точек, которых не более  $\frac{1}{2^z}$  от числа ОТ исследуемого изображения. В каждой  $P$  окрестности особых точек не более чем  $\frac{2\pi P^2}{R^2}$  (вместо прямоугольника  $M \times N$  покрываем квадратиками  $\frac{R}{\sqrt{2}} \times \frac{R}{\sqrt{2}}$  круг радиуса  $P$ ). Значит, всего придется исследовать  $\frac{2\pi P^2}{R^2} * \frac{M*N}{R^2*2^{z-1}}$  особых точек на исследуемом изображении. На втором этапе выбирается  $S$  изображений кандидатов, которые рассматриваются уже в оригинальном масштабе, то есть на них содержится максимум  $\frac{2M*N}{R^2} * S$ . Значит, придется сделать  $\frac{\pi P^2 * M^2 * N^2 * S}{R^6 * 2^{z-3}}$ . Итого получаем  $\frac{1}{2^{2(l+1)}} * \frac{4M^2*N^2}{R^4} * K + \frac{\pi P^2 * M^2 * N^2 * S}{R^6 * 2^{z-3}}$  сравнений.

Таким образом, подбором параметров  $K, l, P, S$  можно добиться баланса между качеством распознавания и его скоростью (числом сравнений), сделав ее выше чем у классической схемы распознавания,

например, выбрав  $l > \frac{\log_2(\frac{\pi P^2 * M^2 * N^2 * S * R^4}{R^6 * 2^{z-3} * 4M^2 * N^2 * K})}{2} - 1$ .

### **Заключение.**

Таким образом, показаны широкие возможности уменьшение вычислительной сложности предложенной схемы распознавания при сохранении приемлемого качества распознавания. Вопросы эффективности работы предлагаемой схемы были рассмотрены авторами в предыдущих работах[1-2].

### **Список литературы.**

1) Mednikov D.I., Milovidov A., Sergunin S.Yu., Kumskov M.I. "Object identification on grayscale images with complicated background using the active sensor model", РОАИ – 2008, труды конференции.

2) D. I. Mednikov, A. Milovidov, S. Yu. Sergunin, M. I. Kumskov "Identification of stable description elements using an active sensor", Pattern Recognition and Image Analysis, Volume 19, Number 1 / Март 2009 г.