

Определение положения объекта на изображении по фрагментам границы

И.С. Симакин

Ярославский государственный университет им. П.Г. Демидова,
Факультет информатики и вычислительной техники,
Ярославль, Россия
--sis--@mail.ru

Аннотация

В работе рассматривается задача распознавания объектов на изображении по фрагментам границы. Предлагаемый метод, основанный на машинном обучении, позволяет находить объекты, совпадающие с заданными эталонами, с учетом преобразований вращения, масштабирования и параллельного переноса. Метод не требует замкнутости границы, что позволяет ему работать с пересекающимися и частично закрытыми объектами. Ключевой особенностью данного метода является обучение процессу подбора положения объекта а не процедуре сравнения.

Ключевые слова: выделение объектов, распознавание по форме границы, контурный анализ, машинное обучение.

1. ВВЕДЕНИЕ

Распознавание объектов на изображении по форме границ заключается в определении положения объектов, каждый из которых совпадает с одним из заданных эталонов.

В литературе описаны эффективные методы решения данной задачи, работающие в случае, когда положение объектов определяется аффинными преобразованиями, включающими составляющие поворота, масштабирования и параллельного переноса, а граница распознаваемого объекта представляется в виде замкнутого контура.

В книге [1] описан метод, основанный на представлении контура границы в виде периодического комплексного сигнала инвариантного к параллельному переносу. Умножение такого сигнала на комплексный множитель приводит к повороту и масштабированию контура. В этом случае распознавание сводится к выбору начальной фазы сигнала при помощи вычисления линейной свертки сигнала выделенного контура и сигнала эталонного контура и определению коэффициента корреляции.

Другим распространенным вариантом решения данной задачи является метод, описанный в [2], когда положение выделенного контура сначала нормализуется, а затем представляется в виде вектора признаков, после чего решается задача классификации.

Подобные методы обладают рядом ограничений. Во-первых, для определения формы предполагаемого объекта, его требуется сначала выделить, что вызывает трудности при работе со сложными объектами. Во-вторых, поскольку такие методы работают с замкнутыми контурами, то они не позволяют распознавать пересекающиеся или частично закрытые объекты. В-третьих, не смотря на то, что подобные методы не очень чувствительны к шуму, они плохо справляются с объектами, у которых есть размытая часть с неустойчивой границей.

Предложенный в данной работе метод определения положения объекта по фрагментам границы, позволяет работать в приведенных выше случаях. Данный метод основан на машинном обучении процессу совмещения эталона с заданными фрагментами границы. Процедура подбора положения эталона сведена к трем подзадам: выбору начального приближения для положения эталона по паре фрагментов, локальной оптимизации положения эталона и применению приемов совмещения для преодоления проблем, возникающих в результате использования локальной оптимизации.

Предложенный метод рассмотрен на примере задачи распознавания альфа-треков на фотографиях плоских диэлектрических детекторов (см. рисунок 1) [3].

Оставшаяся часть работы организована следующим образом. Во втором разделе приведено описание модели, в том числе способ представления фрагментов границ. Третий раздел посвящен процессу подбора положения объекта по набору фрагментов. В четвертом разделе кратко описана процедура поиска объектов, основанная на предложенном методе. В пятом разделе приведены условия и результаты компьютерного эксперимента.

2. ОПИСАНИЕ МОДЕЛИ

Пусть входные изображения и эталоны представлены в виде наборов ломанных линий, описывающих границы. Такое представление можно получить из исходного растрового изображения при помощи алгоритма выделения границ [4]. На рисунке 1 представлен пример фотографии детектора альфа-частиц и соответствующий набор граничных линий. Набор эталонов представлен на рисунке 2.

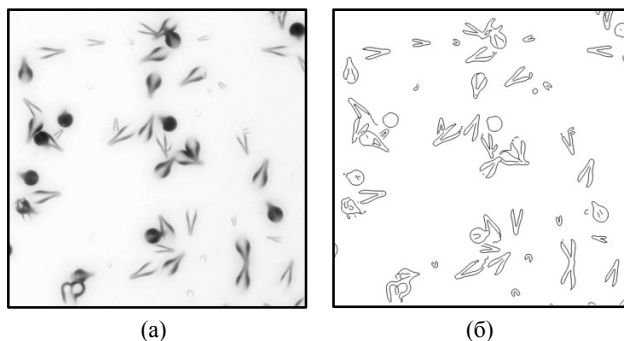


Рисунок 1: Растровое изображение поверхности детектора альфа-частиц (а) и соответствующие линии границ (б).



Рисунок 2: Набор из девяти эталонов.

Линия Γ представлена конечной последовательностью точек (Y_1, Y_2, \dots, Y_n) , заданных парами координат (Y_{x_i}, Y_{y_i}) , расставленными с шагом, равным 1.

Задача заключается в том, чтобы на входном изображении найти положения объектов, совпадающих с заданными эталонами. Положение каждого объекта определяется преобразованием координат T . В данной работе ограничимся рассмотрением аффинных преобразований, не содержащих составляющей перекоса, которые можно представить как суперпозицию вращения T_φ , равномерного по обеим осям координат масштабирования T_α с коэффициентом $\alpha > 0$ и параллельного переноса T_Δ [1]

$$T(\Delta_x, \Delta_y, \varphi, \alpha) = T_\varphi \cdot T_\alpha \cdot T_\Delta.$$

Расстояние между положениями T_1 и T_2 линии Γ определим как

$$d_T(T_1, T_2, \Gamma) = \sqrt{\frac{1}{n} \cdot \sum_{i=1}^n |Y_i \cdot T_1 - Y_i \cdot T_2|^2}.$$

Далее построим функцию $\zeta(P, \Gamma)$ степени расхождения (противоположную к качеству совмещения) эталона P с частью входного изображения Γ . Определим функцию ζ для одной точки γ линии таким образом, чтобы она вычисляла расстояние до ближайшей к ней точки эталона P :

$$\zeta(P, \gamma) = \min_{\rho \in P} |\rho - \gamma|.$$

Для всей части входного изображения Γ зададим ζ как

$$\zeta(P, \Gamma) = \sqrt{\frac{1}{n} \cdot \sum_{i=1}^n \zeta^2(P, \gamma_i)}. \quad (1)$$

2.1 Фрагменты линий

В основу рассматриваемой системы распознавания положена работа с небольшими фрагментами линий. Фрагмент F длины n линии Γ , начинающийся с точки под номером s , задается в виде

$$(f_1, f_2, \dots, f_n) = (Y_s, Y_{s+1}, \dots, Y_{s+n-1}), \\ ((x_{f_1}, y_{f_1}), \dots, (x_{f_n}, y_{f_n})) = ((x_{Y_s}, y_{Y_s}), \dots, (x_{Y_{s+n-1}}, y_{Y_{s+n-1}})).$$

Такое представление фрагментов будем называть точечным или координатным.

В параметрическом представлении форму и положение фрагмента будут определять пять величин, включая координаты вершины фрагмента ξ_x и ξ_y , угол наклона ξ_φ , кривизну ξ_c , и выраженность s-образного загиба ξ_s :

$$(\xi_x, \xi_y, \xi_\varphi, \xi_c, \xi_s).$$

Геометрический смысл параметрического представления фрагмента показан на рисунке 3.

Такое параметрическое представление хорошо описывает фрагменты гладких линий, то есть при переходе от точечного представления к параметрическому необходимая для распознавания объектов информация остается, а несущественная шумовая составляющая отбрасывается.

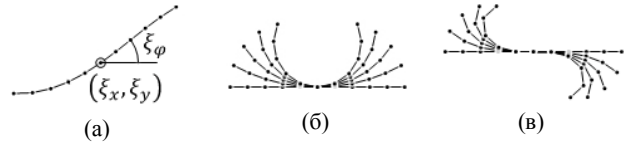


Рисунок 3: Фрагмент линии (а), изменение кривизны (б) и изменение s-образного загиба (в).

Процедура перехода от точечного задания фрагментов к параметрическому основана на преобразовании линии в угловое представление, интегрировании и последующем линейном разложении результата по специально выбранному ортонормированному базису.

К фрагментам, представленным в параметрическом виде, могут быть применены аффинные преобразования

$$\xi' = \xi \cdot T.$$

При этом параметры преобразованного фрагмента можно получить по формулам

$$(\xi'_x, \xi'_y) = (\xi_x, \xi_y) \cdot T, \\ \xi'_\varphi = \xi_\varphi + \varphi_T, \\ \xi'_c = \frac{\xi_c}{\alpha_T}, \\ \xi'_s = \xi_s.$$

Результат такого преобразования будет точным, если T не содержит составляющей масштабирования ($\alpha_T = 1$). В противном случае предложенные формулы будут являться хорошим приближением, достаточным для применения в рассматриваемой системе.

Функцию, вычисляющую расстояние между фрагментами ξ и ψ , заданными в параметрическом виде, обозначим как

$$d_\xi(\xi, \psi).$$

Она вычисляется через формулу корня квадратного из взвешенной суммы квадратов разностей специально построенных величин, заданных непосредственно через параметры фрагментов без перехода к точечному представлению.

Степень расхождения ζ эталона P и некоторого набора фрагментов изображения $\Xi = \{\xi_1, \xi_2, \dots, \xi_n\}$ определим через расстояние между фрагментами. Обозначим множество фрагментов эталона P как $\Psi(P)$. Зададим функцию ζ для одного фрагмента ξ , как расстояние от ξ до ближайшего к нему фрагмента эталона P :

$$\zeta(P, \xi) = \min_{\psi \in \Psi(P)} d_\xi(\xi, \psi).$$

После этого определим функцию ζ для всего набора фрагментов Ξ через формулу

$$\zeta(P, \Xi) = \sqrt{\frac{1}{n} \cdot \sum_{i=1}^n \zeta^2(P, \xi_i)}.$$

3. ПОДБОР ПОЛОЖЕНИЯ ОБЪЕКТА ПО ФРАГМЕНТАМ ГРАНИЦЫ

Рассматриваемый метод распознавания основан на процедуре подбора положения эталона по нескольким заданным фрагментам таким образом, чтобы достичь совпадения со всеми этими фрагментами. Это можно сформулировать в виде оптимизационной задачи на минимизацию степени расхождения набора из n фрагментов $\xi_1, \xi_2, \dots, \xi_n$ и эталона P :

$$\begin{aligned} \tau_n(P, \xi_1, \xi_2, \dots, \xi_n) &= \tau_n(P, \Xi) = T_{\min}: \\ \zeta(P \cdot T_{\min}, \Xi) &= \min_T \zeta(P \cdot T, \Xi), \\ \Xi &= \{\xi_1, \xi_2, \dots, \xi_n\}. \end{aligned} \quad (2)$$

Поскольку от процедуры подбора положения требуется достичь совпадения со всеми заданными фрагментами, то после решения оптимизационной задачи (2), нужно проверить, что степень расхождения каждого из них с эталоном не превосходит порог ζ_0

$$\forall \xi \in \Xi, \quad \zeta(P \cdot T_{\min}, \xi) \leq \zeta_0.$$

При невыполнении этого условия констатируется, что предложенные фрагменты несовместимы, то есть не могут принадлежать одному объекту.

Если выбрать размер фрагмента таким, чтобы он составлял не более трети всего эталона, то можно сделать допущение, что одного фрагмента не достаточно для уверенного определения положения объекта. Такое допущение позволяет избежать необходимости реализации процедуры подбора положения по одному заданному фрагменту.

3.1 Подбор положения объекта по двум фрагментам

Начнем с рассмотрения процедуры подбора положения по двум заданным фрагментам. Пусть даны два фрагмента ξ и ψ в параметрическом представлении

$$\begin{aligned} (\xi_x, \xi_y, \xi_\varphi, \xi_c, \xi_s), \\ (\psi_x, \psi_y, \psi_\varphi, \psi_c, \psi_s). \end{aligned}$$

Вычислим расстояние d между вершинами фрагментов

$$d = \sqrt{(\xi_x - \psi_x)^2 + (\xi_y - \psi_y)^2}.$$

Зададим нормализующее преобразование $T_N(\xi, \psi)$, таким образом, чтобы оно переводило вершины фрагментов в симметричные относительно начала координат точки $(-d/2, 0)$ и $(d/2, 0)$. Такое преобразование будет содержать составляющие вращения и параллельного переноса. После применения нормализующего преобразования параметры фрагментов примут вид

$$\begin{aligned} \xi_N &= \xi \cdot T_N(\xi, \psi) = (-\frac{d}{2}, 0, \xi_\varphi + \varphi_N, \xi_c, \xi_s), \\ \psi_N &= \psi \cdot T_N(\xi, \psi) = (\frac{d}{2}, 0, \psi_\varphi + \varphi_N, \psi_c, \psi_s). \end{aligned}$$

Пример применения нормализующего преобразования приведен на рисунке 4.

Если взять два фрагмента ξ и ψ , принадлежащих эталону P , то соответствующее нормализующее преобразование T_N будет определять некоторое положение эталона, в котором

эталон будет совмещен с нормализованными фрагментами ξ_N и ψ_N .

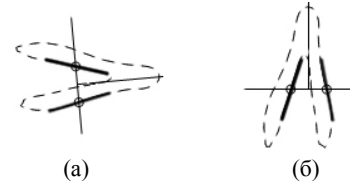


Рисунок 4: Нормализующее преобразование; (а) – положение эталона до нормализации (б) – положение эталона после нормализации по двум фрагментам.

Нормализованное положение инвариантно относительно параллельного переноса и вращения исходных фрагментов, то есть, каким бы образом ни перемещались и ни вращались фрагменты, после нормализации они примут одно и то же положение.

Пусть задана пара фрагментов исходного изображения ξ_Γ и ψ_Γ . Поскольку в рассматриваемой модели распознаваемые объекты являются преобразованными эталонами, то исходной паре соответствует некоторая пара фрагментов эталона ξ_P и ψ_P . Так как нормализованное положение инвариантно относительно преобразования вращения и параллельного переноса, то достаточно подобрать масштаб эталона α , чтобы после нормализации эти пары фрагментов совпали. Это можно условно записать в виде системы уравнений

$$\begin{cases} \xi_\Gamma \cdot T_N(\xi_\Gamma, \psi_\Gamma) \approx \xi_P \cdot T_\alpha \cdot T_N(\xi_P \cdot T_\alpha, \psi_P \cdot T_\alpha), \\ (\psi_\Gamma \cdot T_N(\xi_\Gamma, \psi_\Gamma) \approx \psi_P \cdot T_\alpha \cdot T_N(\xi_P \cdot T_\alpha, \psi_P \cdot T_\alpha). \end{cases} \quad (3)$$

Поскольку в случае реального распознавания пара ξ_P и ψ_P заранее не известна, то ее нужно выбрать из всех возможных пар фрагментов эталона P . Масштаб α , задающий преобразование T_α выбирается из диапазона допустимых значений. Таким образом, у системы есть три неизвестных ξ_P, ψ_P и α .

Оговорка, что система уравнений (3) записана условно, означает, что она не совсем корректна с математической точки зрения. Вследствие зашумленности изображений и других факторов эти уравнения не будут выполняться строго, поэтому их нужно решать приближенно. Однако такая форма записи, во-первых, более наглядна, а во-вторых, позволяет выполнять алгебраические преобразования уравнений.

Для достижения корректности формулировки воспользуемся метрикой сравнения фрагментов и заменим систему уравнений (3) на следующую оптимизационную задачу:

$$\begin{aligned} \chi(\xi_P, \psi_P, \alpha) &= \\ &= |\xi_\Gamma \cdot T_N(\xi_\Gamma, \psi_\Gamma) - \xi_P \cdot T_\alpha \cdot T_N(\xi_P \cdot T_\alpha, \psi_P \cdot T_\alpha)| + \\ &+ |\psi_\Gamma \cdot T_N(\xi_\Gamma, \psi_\Gamma) - \psi_P \cdot T_\alpha \cdot T_N(\xi_P \cdot T_\alpha, \psi_P \cdot T_\alpha)|, \\ (\xi_P, \psi_P, \alpha)_{\min} &= \min_{\substack{\xi_P \in P \\ \psi_P \in P \\ \alpha_{\min} \leq \alpha \leq \alpha_{\max}}} \chi(\xi_P, \psi_P, \alpha). \end{aligned} \quad (4)$$

Следует отметить, что значение функции χ в точке минимума равно значению степени расхождения ζ между нормализованной парой фрагментов $(\xi_\Gamma, \psi_\Gamma)$ и масштабированным нормализованным эталоном:

$$\chi(\xi_P, \psi_P, \alpha) = \zeta(\{\xi_\Gamma, \psi_\Gamma\} \cdot T_N(\xi_\Gamma, \psi_\Gamma), P \cdot T_\alpha \cdot T_N(\xi_P \cdot T_\alpha, \psi_P \cdot T_\alpha)).$$

Данную оптимизационную задачу можно решить при помощи перебора, если зафиксировать некоторый набор возможных значений масштаба α (фрагменты ξ_P и ψ_P выбираются из конечного множества фрагментов эталона P), однако, это неэффективно.

Для построения более эффективного решения заметим, что в нормализованном положении пара фрагментов полностью описывается семимерным вектором b :

$$b(\xi_N, \psi_N) = (d, \xi_\varphi + \varphi_N, \psi_\varphi + \varphi_N, \xi_c, \psi_c, \xi_s, \psi_s).$$

Если считать, что каждая возможная пара фрагментов эталона определяет класс, то выбор пары фрагментов ξ_P и ψ_P эталона по заданному вектору b сводится к задаче многоклассовой классификации. При этом пару фрагментов не обязательно искать точно, так как для большинства случаев нахождения близкой пары фрагментов будет достаточно. В качестве функции ошибки классификации используем расстояние d_T между нормализующими преобразованиями определяемыми найденной (ξ_P, ψ_P) и оптимальной (ξ_P^*, ψ_P^*) парами фрагментов

$$err((\xi_P, \psi_P), (\xi_P^*, \psi_P^*)) = d_T(T_N(\xi_P, \psi_P), T_N(\xi_P^*, \psi_P^*), P).$$

Обучающее множество будем строить из набора всевозможных пар фрагментов эталона P , к которому применяются преобразования масштабирования с различными значениями α из допустимого диапазона. Количество обучающих примеров выбирается таким образом, чтобы их было достаточно для уверенной классификации.

Следует отметить, что для построения эффективной процедуры подбора положения по двум фрагментам требуется использовать быстрый метод классификации, обладающий меньшей трудоемкостью, чем полный перебор. В качестве примера такого метода можно привести [5].

Вернемся к процедуре подбора положения эталона по двум фрагментам. Будем считать, что пара фрагментов ξ_P и ψ_P уже определена (при помощи метода классификации), следовательно для решения оптимизационной задачи (4) осталось найти значение α . Вычислить его через отношение расстояния между вершинами фрагментов изображения ξ_Γ и ψ_Γ к расстоянию между выбранными вершинами фрагментов эталона ξ_P и ψ_P :

$$\alpha = \frac{d(\xi_\Gamma, \psi_\Gamma)}{d(\xi_P, \psi_P)}.$$

Далее найдем преобразование T , переводящее фрагменты эталона во фрагменты объекта, то есть задающее положение объекта. Выразим T из первого уравнения системы (2), применив к левой и правой части обратное преобразование $T_N^{-1}(\xi_\Gamma, \psi_\Gamma)$.

$$\xi_\Gamma \cdot T_N(\xi_\Gamma, \psi_\Gamma) \cdot T_N^{-1}(\xi_\Gamma, \psi_\Gamma) = \xi_P \cdot T_\alpha \cdot T_N(\xi_P, \psi_P) \cdot T_\alpha \cdot T_N^{-1}(\xi_\Gamma, \psi_\Gamma).$$

Так как суперпозиция прямого и обратного преобразования является тождественным преобразованием,

$$\xi_\Gamma = \xi_P \cdot (T_\alpha \cdot T_N(\xi_P, \psi_P) \cdot T_\alpha \cdot T_N^{-1}(\xi_\Gamma, \psi_\Gamma)).$$

Полученная формула осуществляет преобразование фрагмента эталона ξ_P во фрагмент объекта ξ_Γ . Легко проверить, что эта формула справедлива также для перехода от ψ_P к ψ_Γ . Следовательно, преобразование T , совмещающее эталон с парой фрагментов изображения ξ_Γ и ψ_Γ будет определяться формулой

$$\tau_2(P, \xi_\Gamma, \psi_\Gamma) = T = T_\alpha \cdot T_N(\xi_P, \psi_P) \cdot T_\alpha \cdot T_N^{-1}(\xi_\Gamma, \psi_\Gamma).$$

Таким образом, будем считать, что процедура подбора положения эталона по двум заданным фрагментам построена. Пример применения процедуры подбора положения эталона представлен на рисунке 5.

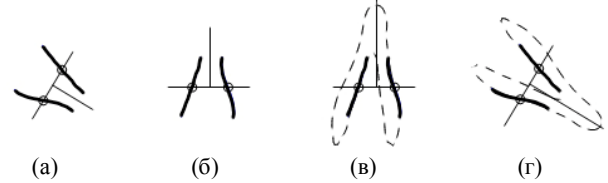


Рисунок 5: Стадии предсказания положения объекта по двум фрагментам; (а) – пара фрагментов; (б) – нормализованная пара фрагментов; (в) – предсказанное положение в нормализованной системе координат; (г) – предсказанное положение в исходной системе координат.

3.2 Локальная оптимизация положения эталона методом наименьших квадратов

Перейдем к рассмотрению процедуры подбора положения эталона по $n \geq 3$ фрагментам изображения. Итак, пусть заданы n фрагментов $\xi_1, \xi_2, \dots, \xi_n$ изображения Γ и эталон P . Как было определено в (2), процедура совмещения обозначается как

$$T = \tau_n(P, \xi_1, \xi_2, \dots, \xi_n).$$

Возьмем первые $n-1$ фрагментов и выполним их совмещение с эталоном P при помощи рекурсивного вызова процедуры подбора положения. Полученное положение назовем начальным и обозначим как T_0

$$T_0 = \tau_{n-1}(P, \xi_1, \xi_2, \dots, \xi_{n-1}).$$

Вспомним, что под выполнением процедуры τ_n подразумевается решение оптимизационной задачи с целевой функцией $\zeta(P \cdot T, \Xi)$ и со свободной переменной T . Будем пытаться решать эту задачу методом локальной оптимизации положения эталона с целевой функцией $\zeta(P \cdot T, \Xi)$. В качестве основы для построения метода будем использовать упрощенную версию модели, описанной в работе [6].

На каждой итерации используемого метода локальной оптимизации для каждого выделенного фрагмента изображения выбирается ближайший к нему фрагмент эталона. После этого система изменяет преобразование координат T так, чтобы “притянуть” сопоставленные пары фрагментов для достижения совмещения. Это сводится к приближенному решению некоторой системы линейных уравнений методом наименьших квадратов.

Итерации выполняются до тех пор, пока происходит существенное уменьшение значения целевой функции. В итоге система приходит в точку локального минимума. Результат локальной оптимизации T^* положения эталона P

для множества выделенных фрагментов изображения Ξ , выполненной из начального положения T_0 , обозначим как

$$\tau^*(P, \Xi, T_0) = T^*.$$

Примеры применения локальной оптимизации положения эталона приведены на рисунке 6.

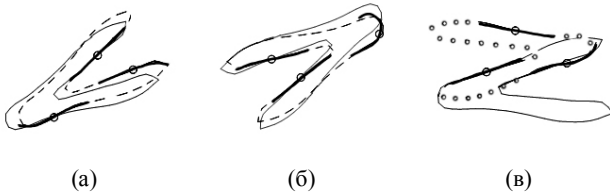


Рисунок 6: Локальная оптимизация положения; (а) и (б) – примеры удачного применения; (в) – пример, когда локальная оптимизация не позволяет достичь совмещения.

Может оказаться, что полученная точка локального минимума будет сильно отличаться от искомой точки глобального минимума. В этом случае эталон будет совмещен с изображением неправильно. Пример попадания в неправильную точку локального минимума приведен на рисунке 6 (в). Данная проблема в рамках предложенного метода локальной оптимизации не решается.

3.3 Приемы совмещения

Если внимательно проанализировать работу метода локальной оптимизации положения эталона на конкретных входных данных, то можно заметить, что многие случаи попадания в “неправильную” точку локального минимума похожи между собой. Следует подчеркнуть, что в данном случае речь идет исключительно о реально совместимых фрагментах, принадлежащих одному объекту.

Можно проклассифицировать такие случаи и для каждого класса придумать простой способ, который позволил бы выбраться из точки локального минимума, чтобы после этого достичь правильного совмещения.

Для реализации таких способов выхода из “неправильных” точек локального минимума будем использовать приемы совмещения. Под применением приема η будем понимать умножение текущего положения T_0 эталона P на некоторое преобразование T_η .

Использование приема η из начального положения T_0 будем считать успешным в случае, если выполняются два условия. Во-первых, необходимо, чтобы в результате применения преобразования T_η и последующей локальной оптимизации уменьшилось значение целевой функции ζ :

$$\zeta(P \cdot \tau^*(P, \Xi, T_0 \cdot T_\eta), \Xi) < \zeta(P \cdot \tau^*(P, \Xi, T_0), \Xi). \quad (5)$$

Во-вторых, потребуем, чтобы положение локального минимума, полученное после использования приема, значительно отличалось от положения, полученного без использования приема. В качестве порога “значительности” отличия положений прием половину расстояния между соответствующими начальными положениями (до локальной оптимизации)

$$d_T(\tau^*(P, \Xi, T_0 \cdot T_\eta), \tau^*(P, \Xi, T_0), P) > \frac{1}{2} \cdot d_T(T_0 \cdot T_\eta, T_0, P). \quad (6)$$

Рассмотрим, использование набора приемов $H = \{\eta_1, \eta_2, \dots, \eta_m\}$ в процессе совмещения эталона P с заданными фрагментами изображения $\Xi = \{\xi_1, \xi_2, \dots, \xi_{n-1}\}$ из начального положения T_0 . Будем последовательно пробовать выполнять приемы $\eta_1, \eta_2, \dots, \eta_m$ и проверять условия успешности их применения (5) и (6).

При обнаружении первого успешного приема η_j , процесс проверки прекращается и выполняется переход к новому положению T_1

$$T_1 = T_0 \cdot T_{\eta_j}.$$

Далее положение T_1 объявляется начальным для следующей итерации и процесс попытки применения приемов повторяется.

Если успешных приемов среди $\eta_1, \eta_2, \dots, \eta_m$ найти не удастся, то положение локального минимума $\tau^*(P, \Xi, T_0)$ объявляется конечным результатом совмещения

$$\tau_n(P, \Xi) = \tau^*(P, \Xi, T_0).$$

В итоге может выстроиться целая цепочка последовательного применения приемов, заканчивающаяся выполнением локальной оптимизации

$$\tau_n(P, \Xi) = \tau^*(P, \Xi, T_0 \cdot T_{\eta_{j_1}} \cdot T_{\eta_{j_2}} \cdot \dots \cdot T_{\eta_{j_k}}).$$

Пример набора приемов для нескольких эталонов представлен на рисунке 7.

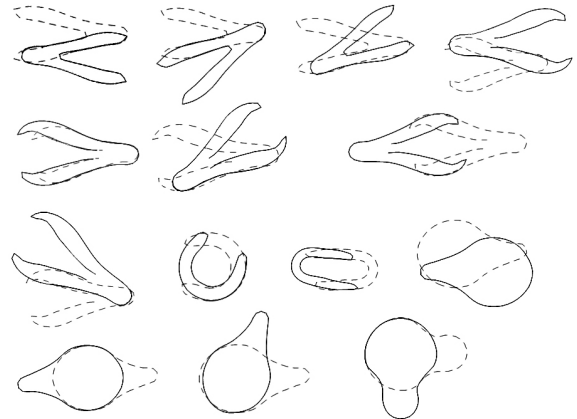


Рисунок 7: Пример набора приемов совмещения для нескольких эталонов.

Перейдем к вопросу о том, каким образом можно получить множество приемов совмещения для заданного эталона. Поскольку приемы используются для преодоления точек локального минимума, то можно предположить, что их следует строить на основе примеров попадания в локальные минимумы, полученных при обработке реальных данных. Сконструируем процедуру построения множества приемов совмещения как метод машинного обучения на примерах.

Пусть дан пример входного изображения Γ , и эталон P . Первым делом построим на изображении Γ множество возможных положений частичного совмещения K . Для этого будем много раз применять процедуру подбора положения по тройке (или по паре) случайно выбранных попарно совместимых неперекрывающихся фрагментов. Поскольку

множество приемов для эталона P еще не построено, то при подборе положений ограничимся предсказанием начального положения по двум фрагментам и локальной оптимизацией

$$\kappa = \tau^*(P, \{\xi_1, \xi_2, \xi_3\}, \tau_2(P, \xi_1, \xi_2)).$$

Кроме того, обеспечим, чтобы множество возможных положений частичного совмещения состояло из заметно отличающихся друг от друга элементов, расстояние между которыми превосходит заданный порог

$$\forall \kappa_i, \kappa_j \in K, \quad i \neq j, \quad \delta_T(\kappa_i, \kappa_j, P) > \delta_{\min}.$$

В результате в окрестности каждого объекта на изображении Γ будет содержаться по одному или несколько различных положений частичного совмещения из множества K .

После этого выполняется стадия обучения, заключающаяся в построении набора возможных приемов. Для этого много раз выполняется попытка совмещения по тройке случайно выбранных попарно совместимых неперекрывающихся фрагментов $\Xi_3 = (\xi_1, \xi_2, \xi_3)$.

Как и в процессе обычного совмещения, первым делом по двум фрагментам вычисляется начальное положение T_0

$$T_0 = \tau_2(P, \xi_1, \xi_2).$$

После этого применяется локальная оптимизация из положения T_0

$$T_0^* = \tau^*(P, \Xi_3, T_0).$$

Далее выполняется проверка на попадание в “неправильную” точку локального минимума. Для этого сначала выбирается подмножество положений частичного совмещения K_{Ξ_3} , которые могут иметь отношение к выбранной тройке фрагментов Ξ_3 . Положение частичного совмещения κ_j будем считать имеющими отношение к Ξ_3 , если степень расхождения эталона P с фрагментами будет меньше некоторого порогового значения ζ_k

$$\zeta(P \cdot \kappa_j, \Xi_3) < \zeta_k \Rightarrow \kappa_j \in K_{\Xi_3}.$$

После этого из множества положений частичного совмещения, имеющих отношение к тройке Ξ_3 , выбирается положение κ_{\min} , обеспечивающее самое лучшее совмещение после применения локальной оптимизации

$$\begin{aligned} \kappa_{\min}: \zeta(P \cdot \tau^*(P, \Xi_3, T_0 \cdot \kappa_{\min}), \Xi_3) \\ = \min_{\kappa \in K_{\Xi_3}} \zeta(P \cdot \tau^*(P, \Xi_3, T_0 \cdot \kappa), \Xi_3). \end{aligned}$$

Положение локального минимума, достижимое из κ_{\min} , назовем конечным и обозначим как $T_{\kappa_{\min}}^*$:

$$T_{\kappa_{\min}}^* = \tau^*(P, \Xi_3, T_0 \cdot \kappa_{\min}).$$

Далее проверим четыре условия. Во-первых, конечное положение $T_{\kappa_{\min}}^*$ должно обеспечивать лучшее совмещение, чем начальное положение, подвергнутое локальной оптимизации.

$$\zeta(P \cdot T_{\kappa_{\min}}^*, \Xi_3) < \zeta(P \cdot T_0^*, \Xi_3).$$

Во-вторых, необходимо, чтобы было достигнуто совмещение для каждого из фрагментов тройки Ξ_3

$$\forall \xi_i \in \Xi_3, \quad \zeta(P \cdot T_{\kappa_{\min}}^*, \xi_i) < \zeta_0.$$

Второе условие позволяет исключить обучение приемам, применение которых хотя и улучшает значение целевой функции, однако не приводит в результате к правильному совмещению.

В-третьих, для того, чтобы получить нетривиальный прием, конечное положение $T_{\kappa_{\min}}^*$ должно существенно отличаться от начального положения T_0 :

$$d_T(T_{\kappa_{\min}}^*, T_0, P) > \delta_\eta.$$

В-четвертых, необходимо, чтобы конечное положение $T_{\kappa_{\min}}^*$ заметно отличалось от результата локальной оптимизации T_0^* начального положения T_0 . Отсутствие проверки этого условия привело бы к обучению приемам, работающим в ситуациях, с которыми и так справляется метод локальной оптимизации. Для проверки данного условия воспользуемся выражением, построенным по аналогии с (6):

$$d_T(T_{\kappa_{\min}}^*, T_0^*, P) > \frac{1}{2} \cdot d_T(T_{\kappa_{\min}}^*, T_0, P).$$

В случае нарушения любого из этих условий, будем считать, что построение приема из начального положения T_0 не удалось, и следует перейти к другой тройке случайных фрагментов для продолжения обучения.

Если же данные условия выполняются, то на основе положений T_0 и $T_{\kappa_{\min}}^*$ можно построить новый прием η . Преобразование T_η , задающее прием η , выбирается таким образом, чтобы оно переводило эталон P из начального положения T_0 в найденное конечное положение $T_{\kappa_{\min}}^*$:

$$T_0 \cdot T_\eta = T_{\kappa_{\min}}^*.$$

Отсюда следует, что

$$T_\eta = T_0^{-1} \cdot T_{\kappa_{\min}}^*.$$

Построенный прием η добавляется в список приемов N эталона P , после чего происходит переход к следующей попытке.

В процессе обучения приемам учитываются ранее построенные приемы. Если множество приемов N не пусто, то при выполнении попытки обучения новому приему, сначала проверяется возможность применения существующих приемов. Если из начального положения T_0 удастся успешно выполнить хотя бы один из этих приемов, то происходит переход к следующей попытке обучения с новой тройкой фрагментов.

После построения множества возможных приемов выполняется их ранжирование, то есть упорядочивание по убыванию полезности. Более полезными считаются приемы, применимые в большем количестве ситуаций.

Ранжирование выполняется следующим образом: сначала ищется прием, применимый в самом большом количестве ситуаций. После этого ищется прием, который лучше всего помогает в ситуациях, с которыми не справляется первый

прием. Далее – лучший прием, для ситуаций, с которыми не справились первые два приема, и так далее.

Таким образом, процедура $\tau_n(P, \Xi)$ подбора положения эталона P по n фрагментам $\Xi = \{\xi_1, \xi_2, \dots, \xi_{n-1}\}$ изображения была сведена к процедуре τ_{n-1} подбора начального положения по $n - 1$ фрагменту, применению подготовленных на основе обучения приемов совмещения $H = \{\eta_1, \eta_2, \dots, \eta_m\}$ и локальной оптимизации $\tau^*(P, \Xi, T_0)$.

4. ПРОЦЕДУРА ПОИСКА ОБЪЕКТОВ

Процедура поиска объектов отвечает, во-первых, за выделение наборов фрагментов, по которым выполняются попытки подбора положения эталона, а, во-вторых, за выбор лучшего набора положений при построении конечной расстановки объектов на изображении, выдаваемой в качестве результата распознавания.

Поскольку построение процедуры поиска объектов рассматривалось при выполнении исследований как второстепенная задача, то для ее решения был выбран наиболее простой алгоритм, достаточный для проверки работоспособности и перспективности метода подбора положения эталона.

Итак, из множества фрагментов $\Psi(\Gamma)$ изображения Γ по порядку выбираются фрагменты ψ_i . На каждом шаге ψ_i обозначается как первый выделенный фрагмент ξ_1

$$\xi_1 = \psi_i \in \Psi(\Gamma).$$

После этого строится подмножество Ψ_{ξ_1} множества фрагментов $\Psi(\Gamma)$, соседних с ξ_1 , то есть расположенных на расстоянии, не превосходящем диаметр эталона P . После этого из Ψ_{ξ_1} случайным образом выбирается второй фрагмент ξ_2 , совместимый и не перекрывающийся с ξ_1 . Положение k , полученное в результате совмещения, добавляется во множество возможных положений объектов K , обозначенное по аналогии с множеством возможных положений частичного совмещения из раздела 3.3:

$$k = \tau_2(P, \xi_1, \xi_2), k \rightarrow K.$$

Далее из Ψ_{ξ_1} случайным образом выбирается фрагмент ξ_3 , попарно совместимый и не перекрывающийся с фрагментами ξ_1 и ξ_2 . Если вся тройка фрагментов оказывается совместимой, то результат совмещения k добавляется во множество K

$$k = \tau_3(P, \xi_1, \xi_2, \xi_3), k \rightarrow K.$$

Добавление новых фрагментов к выборке выполняется до тех пор, пока удается достигать совмещения. Попытки построения случайных наборов фрагментов $\{\xi_1, \xi_2, \dots\}$ выполняются для каждого ψ_i заданное количество раз.

Поскольку одному реальному объекту на изображении может соответствовать несколько различных элементов построенного множества возможных положений K , то необходимо выбрать лучшую расстановку объектов, таким образом, чтобы одному реальному объекту соответствовало одно положение эталона.

В качестве целевой функции для выбора лучшей расстановки объектов используется суммарное качество расстановки объектов. Значение этой величины зависит от качества совмещения с изображением каждого объекта, входящего в

расстановку, и от меры неразличимости каждой пары выделенных объектов.

Величина, характеризующая качество совмещения эталона P с частью изображения Γ для заданного положения k , строится на основе функции $\zeta(P \cdot k, \Gamma)$ (1).

Мера неразличимости двух объектов $\theta(P_1 \cdot k_1, P_2 \cdot k_2)$, заданных положениями k_1 и k_2 эталонов P_1 и P_2 , принимает значения от 0 до 1 и характеризует правдоподобность того, что оба объекта могут быть одновременно различимы на изображении. Если два объекта не пересекаются, то они легко различимы друг от друга, следовательно, для них значение функции θ будет равно 0. Если же два объекта полностью совпадают, то они являются заведомо неразличимыми, поэтому

$$\theta(P \cdot k, P \cdot k) = 1.$$

Для выбора лучшей расстановки объектов используется простой жадный алгоритм, выбирающий на каждом шаге из множества возможных положений K лучшее положение k , максимально увеличивающее суммарное качество расстановки. Выбор объектов осуществляется до тех пор, пока происходит увеличение суммарного качества расстановки.

Описанная выше реализация процедуры поиска объектов характеризуется повышенной вероятностью появления ошибок распознавания в сложных случаях и относительно низким быстродействием, недостаточным для работы с большими объемами данных в режиме пакетной обработки и в режиме реального времени.

5. КОМПЬЮТЕРНЫЙ ЭКСПЕРИМЕНТ

Для выполнения компьютерного эксперимента были использованы пять фотографий поверхности детектора альфа-частиц, масштабированные до размера 784×588 пикселей. Эти фотографии были разрезаны на куски размером 256×256 пикселей. Каждый из 34 полученных кусков использовался как входное изображение. Пример входного изображения представлен на рисунке 1.

Набор из девяти эталонов (см. рисунок 2) был построен вручную. Каждый эталон был вырезан из исходного изображения и повернут в горизонтальное положение.

После этого выполнялось обучение классификаторов для предсказания положения объектов по двум фрагментам, по одному на каждый эталон. На обучение одного классификатора требовалось около 5-7 минут.

После этого была выполнена процедура обучения приемам совмещения на части входных изображений, пока суммарное количество приемов не достигло 50.

Далее все входные изображения были обработаны системой распознавания. На обработку одного изображения размером 256×256 пикселей требовалось около 1 минуты. Результаты распознавания были проверены визуально. Типичный пример распознанного изображения приведен на рисунке 8. На рисунке 9 показано множество лучших положений и лучшая расстановка объектов для того же самого изображения.

Мелкие объекты остались нераспознанными в следствие того, что для них не были заданы эталоны. Недостаточная тщательность распознавания (пропуск) крупных объектов объясняется, во-первых, упрощенным устройством процедуры поиска объектов, а во-вторых, несовершенством механизма вычисления меры неразличимости объектов,

используемым в процессе выбора лучшей расстановки. Эти проблемы являются устранимыми и не имеют непосредственного отношения к самому метод определения положения объектов по фрагментам границы.

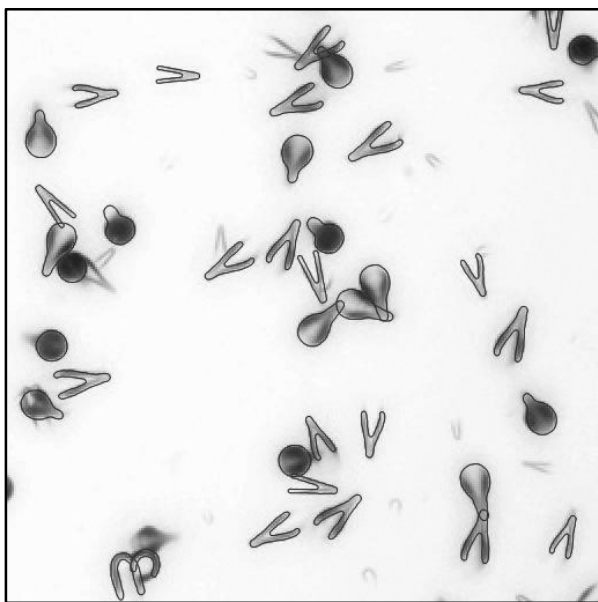


Рисунок 8: Результат распознавания.

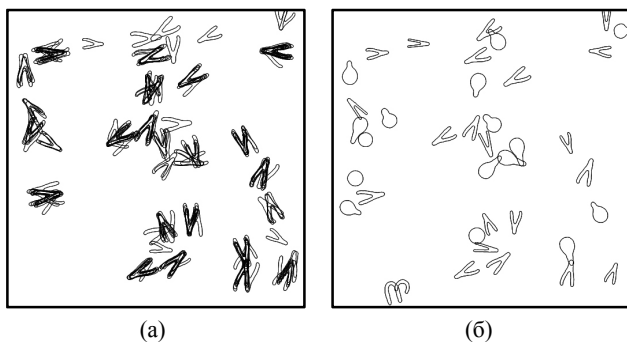


Рисунок 9: Множество возможных положений для одного из эталонов (а) и лучшая расстановка (б).

6. ЗАКЛЮЧЕНИЕ

Ключевой идеей предложенного метода можно считать использование приемов совмещения, получаемых при помощи процедуры обучения на примерах. Приемы позволяют добиться эффективной работы с эталонами с учетом самоподобия их частей.

Рассмотренная система распознавания на некоторых этапах работы использует машинное обучение на примерах. Важной особенностью, на которую следует обратить внимание, является использование процедуры обучения **процессу совмещения** эталонов с объектами, тогда как большинство описанных в литературе методов распознавания, основанных на обучении на примерах, осуществляет обучение **процессу узнавания** эталонов по набору признаков.

Не смотря на то, что процедура поиска объектов рассматривалась в данной работе лишь как часть системы, необходимая для исследования метода подбора положения эталона по фрагментам, и была построена в максимально

упрощенном виде, она, тем не менее, обеспечила возможность распознавания входных изображений, содержащих реальные данные. Результаты компьютерных экспериментов подтвердили работоспособность и перспективность идей, заложенных в основу системы.

В заключение следует отметить, что предложенный метод поиска объектов можно применять не только к изображениям, представленным в виде наборов линий. Вместо фрагментов границ можно было бы использовать, например, особые точки изображения.

7. ЛИТЕРАТУРА

- [1] Фурман Я.А. Введение в контурный анализ; приложения к обработке изображений и сигналов. 2-е изд., исп. – М.: ФИЗМАТЛИТ, 2003. – 592 с. ISBN 5-9221-0374-1.
- [2] Giannakou V., Tzouveli P., Avrithis Y. and Kollias S. Affine Normalization Curve Matching Method Using Curvature Scale-Space and Normalization. Sixth International Workshop on Content-Based Multimedia Indexing (CBMI 2008), London, 2008
- [3] Roussetski A.S. CR-39 Track Detectors in Cold Fusion Experiments: Review and perspectives. Condensed Matter Nuclear Science. Proceedings of the 11th International Conference on Cold Fusion, 2004.
- [4] Devernay F. A Non-Maxima Suppression Method for Edge Detection with Sub-Pixel Accuracy. Research report 2724, INRIA Sophia-Antipolis, 1995.
- [5] Andoni A., Indyk P. Near-Optimal Hashing Algorithms for Approximate Nearest Neighbor in High Dimensions. Communications of the ACM, vol. 51, no. 1, 2008.
- [6] Cootes T. F., Taylor C. J., Cooper D. H., and Graham J. Active shape models - their training and application. Computer Vision and Image Understanding, Jan. 1995.

Благодарности

Автор выражает благодарность сотрудникам лаборатории дозиметрии и радиоактивности окружающей среды химического факультета МГУ Ю.А. Сапожникову, С.Н. Калмыкову и И.Э. Власовой за поставленную задачу распознавания альфа-треков и предоставленные фотографии детекторов альфа-частиц.

Об авторе

Симакин Иван – аспирант Ярославского государственного университета им. П.Г. Демидова, факультет информатики и вычислительной техники, email: --sis--@mail.ru.

ESTIMATION OF OBJECT POSITION ON IMAGE USING EDGE FRAGMENTS

Abstract

In this paper we consider a problem of object recognition using edge fragments. The proposed method is based on machine learning. It permits to detect objects matching specified templates regardless of various rotation, scaling and translation transforms. The proposed method doesn't require closed edge contours therefore it allows working with intersecting and covered objects. Key feature of this method is learning the object position estimation process instead of learning the comparing procedure.

Keywords: object detection, contour analysis, shape recognition, machine learning.