

Выделение переднего плана на картах глубины с помощью скелетного представления силуэтов объектов*

Д.С. Белобородов, Л.М. Местецкий

dmitribeloborodov@yandex.ru | mestlm@mail.ru

Россия, Московский государственный университет имени М.В.Ломоносова

В данной статье рассматривается задача обнаружения переднего плана на картах глубины. Предложен метод решения этой задачи, основанный на построении скелета связанных областей на карте глубины. Передний план определяется как множество точек изображения, которые образуют связанные области и обладают меньшей глубиной, чем точки, окружающие эти области. Передний план соответствует объектам, находящимся ближе всего к камере. Обнаружение переднего плана является важной задачей при обработке изображений с окклюзиями. Приведённый в статье метод позволяет находить передний план на изображениях с несколькими объектами, которые могут перекрывать друг друга или самих себя, независимо от их глубины. Также демонстрируется работа предложенного метода на примере карт глубины, полученных с помощью устройства Microsoft Kinect.

Ключевые слова: передний план, изображения с окклюзиями, непрерывный скелет, серединные оси, сегментация, Kinect

1. Введение

При распознавании поз и жестов по RGB изображениям возникает проблема сегментации человеческих фигур. Она особенно сложна при выделении объектов на пёстром фоне. Использование камеры глубины даёт важную дополнительную информацию, которая существенно облегчает решение, поскольку появляется возможность более надёжного отделения фоновых пикселей от пикселей переднего плана. На границе объекта переднего плана имеется перепад глубины, который можно определить. Однако перепад глубины является локальным свойством, он позволяет увидеть малый фрагмент границы какой-то части объекта. Для сегментации фигуры человека, определения положения рук в случае сложной позы с окклюзиями необходимо произвести глобальный анализ всех таких участков границы. Тут важно отметить, что границы между элементами разной глубины не обязательно соотносятся как <объект-фон>, а могут представлять собой пару <объект-объект>. Это происходит, когда элемент переднего плана, например, рука, виден на фоне более удалённых элементов человеческой фигуры (лицо, туловище, другая рука). В этом случае корректное определение границы объекта переднего плана требует глобального анализа всего дальностного изображения.

Выделение переднего плана и объектов интереса на изображениях является важной задачей в области распознавания образов. В частности, такая задача встречается в распознавании жестов в системах дистанционного управления. Существует немало устройств, которые позволяют в реальном времени получать не только RGB изображение сцены, но и карту глубины, на которой каждой точке изоб-

ражения ставится в соответствие расстояние до камеры.

Задача заключается в том, чтобы обнаружить на изображении передний план — множество точек, соответствующих областям, которые находятся ближе к камере, чем окружающие их фоновые объекты. Эта процедура является начальным этапом во многих системах распознавания жестов.

Существует несколько разных подходов к решению этой задачи:

В одном из подходов обнаружение объектов интереса или выделение признаков, характеризующих их, производится с помощью обученных классификаторов, таких, как нейронные сети. Например, такой подход применяется в статье [4]. Недостатком этого подхода является необходимость в обучении классификатора, для которого требуется большая выборка размеченных данных.

Другой подход основывается на анализе нескольких подряд идущих кадров и оценке модели фона (например, смесью нормальных распределений). Такие методы рассматриваются в статьях [3, 5]. У этого подхода также есть недостаток: он не позволяет находить статичные объекты интереса.

Также существует подход, основанный на анализе графов. Такой подход рассматривается в [6] применительно к простым объектам.

В данной статье рассматривается метод, который позволяет в реальном времени находить передний план по единственной карте глубины изображения без использования методов машинного обучения. Метод также применим к изображениям с окклюзиями, на которых допустимы наложения силуэтов нескольких объектов или частей одного объекта. Предложенный метод основывается на выделении границ на карте глубины и построении скелета полученных связанных областей. Предлагается критерий для определения принадлежности вершин и

ветвей скелета к областям переднего плана, основанный на анализе глубины изображения на границе области.

Демонстрируется работа предложенного метода на примере карт глубины, полученных с помощью прибора Microsoft Kinect.

2. Описание метода

Ключевым моментом предложенного метода является построение скелета для всех объектов, находящихся в кадре, а также для областей фона, независимо от их глубины. В целом, для выделения переднего плана производятся следующие действия:

1. Выделение границ на карте глубины. Это необходимо для разделения перекрывающихся объектов и разрешения окклюзий.
2. Построение скелетов всех связных областей полученной карты глубины с границами.
3. Определение факта принадлежности к переднему плану для каждого узла построенных скелетов.
4. Восстановление областей, соответствующих переднему плану, по скелету.

2.1 Выделение границ

Для выделения границ на карте глубины применяется оператор Собеля. Считается, что пиксель является точкой границы, если полученная величина превосходит некоторый заданный порог. Далее строится бинарное изображение, на котором черными точками помечаются пиксели, лежащие на границах, а также пиксели с неизвестной глубиной (области с неизвестной глубиной могут появляться по краям объектов или на заднем плане из-за особенностей работы Kinect); белыми точками помечаются все остальные пиксели. В результате получается карта глубины с выделенными границами, на которой разделены все области, внутри которых глубина меняется достаточно плавно.

2.2 Построение скелетов областей

Скелет плоской фигуры представляет собой множество центров вписанных в эту фигуру окружностей. Будем называть сайтами многоугольной плоской фигуры ее вершины и ребра. Скелет многоугольной фигуры может быть представлен в виде графа, каждому узлу которого ставятся в соответствие его координаты, радиус вписанной окружности с центром в этом узле, а также два или три сайта фигуры, которых касается эта окружность. Ребра графа представляют собой непрерывные кривые, соединяющие узлы. Любая точка на ребрах скелета является центром некоторой окружности, вписанной в данную фигуру.

Построение скелета производится методом, описанным в [2]. Данный метод позволяет эффективно

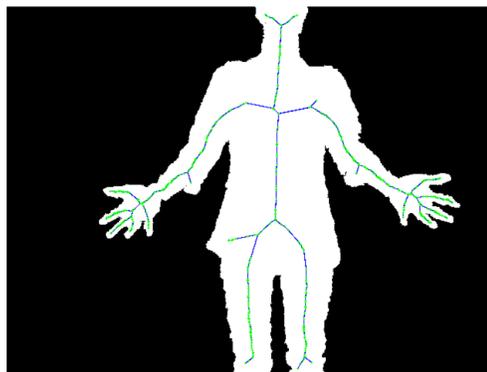


Рис. 1: Пример скелета бинарного изображения.

строить скелет для карт глубины, что позволяет обрабатывать данные в реальном времени.

На этом этапе сначала производится аппроксимация белых связных областей изображения, полученных на предыдущем шаге, многоугольными фигурами (метод описан в [2]). Далее для этих многоугольных фигур строится скелет. В дальнейшем скелетное представление фигур используется для определения принадлежности к переднему плану.

2.3 Принадлежность узла скелета к переднему плану

Далее предлагается для каждого узла каждого полученного скелета определить, принадлежит ли он области, соответствующей переднему плану. Метод основан на идее, впервые описанной в [1].

Определим понятие регулярного узла скелета. Обозначим за $d(\nu)$ глубину в точке ν . Пусть данному узлу с координатами (x_0, y_0) соответствует вписанная в многоугольную фигуру окружность радиуса R с центром в этом узле. Эта окружность касается границ фигуры по меньшей мере в двух точках. Обозначим множество точек, лежащих внутри этой окружности, как C_R . Построим еще одну окружность с центром в (x_0, y_0) и радиусом $R + \varepsilon$, где ε – некоторое маленькое положительное число. Обозначим множество точек, лежащих внутри этой окружности, как $C_{R+\varepsilon}$.

Границы многоугольной фигуры делят множество $C_{R+\varepsilon}$ на несколько множеств A_0, A_1, \dots, A_k , ровно одно из которых целиком содержит в себе C_R (поскольку окружность вписанная). Пусть $C_R \subseteq A_0$, $\bigcup_{i=1}^k A_i = A$.

Будем считать узел регулярным, если при малых ε можно найти не зависящее от ε положительное число h такое, что выполняется неравенство:

$$\min_{\nu \in A} d(\nu) > \max_{\nu \in C_R} d(\nu) + h$$

Данное неравенство означает, что на участках границы, которых касается данная вписанная окружность, существует некоторый перепад глубин. При

этом при движении через границу в точке касания изнутри вписанной окружности глубина увеличивается.

Регулярным ребром будем считать ребро, у которого хотя бы один из инцидентных узлов является регулярным.

Предлагается проверять условие регулярности узла следующим образом:

Сначала находятся все точки касания соответствующего максимального круга границы фигуры. Для этого нужно перебрать все сайты, в которых круг касается границы. Если сайт является вершиной, то её координаты и будут задавать точку касания. Если же сайт является ребром, то ищется точка, в которой окружность касается этого ребра.

Пусть C – центр максимального круга, D – точка касания. Далее измеряются значения глубины в двух точках A и B , которые лежат на прямой CD и находятся на фиксированном небольшом расстоянии d от точки D . Пусть точка A лежит внутри максимального круга, а точка B – снаружи. Обозначим координаты точек A , B , C , D как (x_A, y_A) , (x_B, y_B) , (x_C, y_C) , (x_D, y_D) соответственно. Тогда координаты точек A и B будут считаться по следующим формулам:

$$k_{out} = 1 + \frac{d}{\sqrt{(x_D - x_C)^2 + (y_D - y_C)^2}}$$

$$k_{in} = 1 - \frac{d}{\sqrt{(x_D - x_C)^2 + (y_D - y_C)^2}}$$

$$x_A = k_{in}x_D + (1 - k_{in})x_C$$

$$y_A = k_{in}y_D + (1 - k_{in})y_C$$

$$x_B = k_{out}x_D + (1 - k_{out})x_C$$

$$y_B = k_{out}y_D + (1 - k_{out})y_C$$

Обозначим глубину в точках A и B как Z_{in} и Z_{out} соответственно, а величину $Z = Z_{out} - Z_{in}$ назовем перепадом глубин для данной точки касания. Тогда если перепад глубин для всех точек касания данного узла скелета больше нуля, то узел считается регулярным. В противном случае он считается нерегулярным.

Ясно, что регулярные узлы соответствуют областям, которые находятся ближе к камере, чем окружающий фон.

Если величина перепада точки касания близка к нулю, то изменение глубины связано либо с шумом, либо с локальными неровностями объекта. Если же

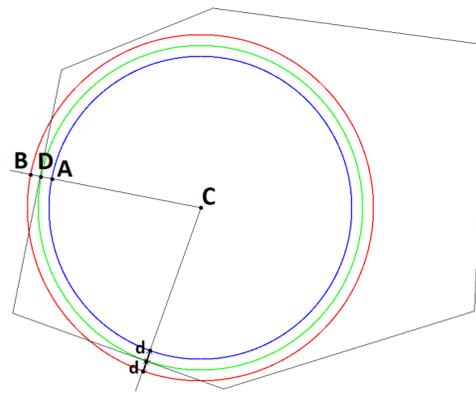


Рис. 2: Иллюстрация к алгоритму определения регулярности узла.

перепад достаточно велик, то точка касания находится на границе объекта с глобальным фоном или сильно удаленным объектом.

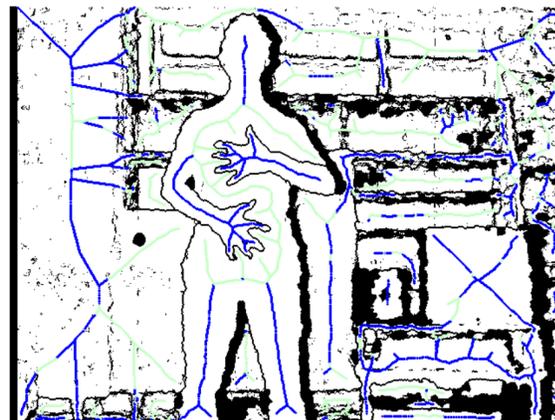


Рис. 3: Регулярные ребра (синий) и нерегулярные (зеленый).

Выберем два значения порогов глубины: $0 \leq Z_{low} < Z_{high}$. Будем относить к переднему плану только те узлы скелета, у которых перепады всех точек касания удовлетворяют неравенству $Z_{low} < Z < Z_{high}$. Регулируя значения этих порогов, можно выбирать разные объекты переднего плана. Параметр Z_{low} используется для устранения шума и повышения устойчивости алгоритма определения переднего плана. При уменьшении параметра Z_{high} объекты, находящиеся на фоне удаленных объектов, перестают считаться передним планом. Это свойство может использоваться для обнаружения окклюзий близко расположенных друг к другу объектов (таких, как руки человека на фоне тела) и отсеивания всех остальных объектов.

Составим скелет переднего плана следующим образом: в качестве узлов возьмем все узлы скелетов, для которых в каждой точке касания выполнено неравенство $Z_{low} < Z < Z_{high}$. Ребрами скелета пе-

реднего плана будут все ребра, инцидентные хотя бы одному из выбранных узлов.

2.4 Восстановление переднего плана

Далее требуется по скелету переднего плана восстановить силуэты объектов переднего плана, соответствующих этому скелету. Используем ребра скелета переднего плана для восстановления множества точек, соответствующих переднему плану.

Пусть ребро инцидентно двум узлам с координатами (x_a, y_a) и (x_b, y_b) , которым соответствуют вписанные окружности радиуса R_a и R_b соответственно. Каждой точке этого ребра также соответствует некоторая вписанная окружность. Будем считать, что радиус таких окружностей меняется линейно от R_a до R_b вдоль ребра.

Точка на изображении с координатами (x, y) принадлежит переднему плану, если существует вписанная окружность с центром на ребре скелета переднего плана такая, что эта точка попадает внутрь окружности. Обозначим за ρ расстояние от точки (x, y) до ближайшей точки данного ребра. Тогда принадлежность точки (x, y) к переднему плану будет определяться следующим образом:

$$t = \frac{(x_a - x)(x_a - x_b) + (y_a - y)(y_a - y_b)}{(x_a - x_b)^2 + (y_a - y_b)^2}$$

$$R = \begin{cases} (1 - t)R_a + tR_b, & t \in [0, 1] \\ R_a, & t < 0 \\ R_b, & t > 1 \end{cases}$$

Здесь величина t параметризует положение проекции точки (x, y) на ребро, а R – радиус вписанной окружности для этой проекции. Тогда (x, y) относится к переднему плану, если $\rho \leq R$.

Для определения принадлежности к переднему плану точки (x, y) достаточно проверить выполнение этого неравенства относительно этой точки для всех ребер скелета переднего плана.

На самом деле, для каждого ребра можно рассмотреть множество точек (x, y) :

$$\min(x_a - R_a, x_b - R_b) \leq x \leq \max(x_a + R_a, x_b + R_b)$$

$$\min(y_a - R_a, y_b - R_b) \leq y \leq \max(y_a + R_a, y_b + R_b)$$

и проверить выполнение неравенства $\rho \leq R$ относительно данного ребра только для этих точек. Все точки из этого множества, для которых выполняется неравенство, помечаются как принадлежащие переднему плану.

3. Реализация алгоритма

Данный алгоритм был реализован и применен к последовательностям карт глубины, получаемых с

помощью устройства Microsoft Kinect. Тестирование алгоритма показало, что обработка поступающих данных может производиться практически в реальном времени со скоростью до 12 кадров в секунду.

В частности, метод был опробован в задаче сегментации изображения человека при наличии окклюзий. Одна из особенностей данного метода заключается в том, что он восстанавливает области переднего плана по скелету, поэтому области могут восстанавливаться неточно. Также метод иногда обнаруживает ложные узлы, не относящиеся к переднему плану. Частично эти проблемы связаны со «слепыми зонами» датчика Kinect, из-за которых на картах глубины могут присутствовать области с неопределенной глубиной.



Рис. 4: Выделенный передний план относительно человеческого тела.

Данный метод может быть использован в системах дополненной реальности, где требуется совместить реальное изображение, полученное с камеры, с синтетическими изображениями некоторых объектов. Особенно этот метод может быть полезен при визуализации виртуальных объектов на реальном изображении с окклюзиями. Примером такой задачи может быть визуализация виртуальной одежды на теле человека в «виртуальной примерочной».

4. Заключение

В статье предложен новый метод выделения переднего плана на изображении по карте глубины, который позволяет выделять элементы объекта разного уровня, как относительно глобального фона, так и относительно локальных объектов. В целом, метод показал свою работоспособность, а также возможность обрабатывать данные в реальном времени.

Предложенный метод не использует алгоритмов машинного обучения и позволяет эффективно оценить передний план сцены по одному кадру. Алгоритм требует ряд усовершенствований, направленных на повышение устойчивости и эффективности

его работы, на что и будут направлены дальнейшие исследования.

Распознавание динамических жестов на основе медиального представления формы изображений.

Литература

- [1] Куракин А.В. Распознавание динамических жестов в системе компьютерного зрения на основе медиального представления формы изображений, Кандидатская диссертация. Московский физико-технический институт, 2012.
- [2] Непрерывная морфология бинарных изображений. Фигуры, скелеты, циркуляры. — М.: ФИЗМАТЛИТ, Москва, 2009.
- [3] Enrico Bondi, Lorenzo Seidenari, Andrew D. Bagdanov, Alberto Del Bimbo. Real-time people counting from depth imagery of crowded environments. University of Florence, 2014.
- [4] Saurabh Gupta, Ross Girshick, Pablo Arbelaez, Jitendra Malik. Learning Rich Features from RGB-D Images for Object Detection and Segmentation. European Conference on Computer Vision (ECCV), 2014.
- [5] Michal Kepski, Bogdan Kwolek. Person Detection and Head Tracking to Detect Falls in Depth Maps. University of Rzeszov, 2014.
- [6] Giorgio Toscani, Stefano Rosa. Fast Graph-Based Object Segmentation for RGB-D Images. Politecnico di Torino.