

## КЛАССИФИКАЦИЯ И ОЦЕНКА ГЕОМЕТРИЧЕСКИХ ПАРАМЕТРОВ СЛОЖНЫХ ПОВЕРХНОСТЕЙ ПО RGB-D ИЗОБРАЖЕНИЯМ

© 2022 Ю. И. Рубцова✉, Р. Н. Яковлев

*Санкт-Петербургский Федеральный исследовательский центр Российской академии наук, Санкт-Петербургский институт информатики и автоматизации Российской академии наук  
14-я линия В.О., 39, 199178 Санкт-Петербург, Российская Федерация*

**Аннотация.** При разработке систем управления формациями модульных робототехнических систем (РС) актуальной является задача классификации и оценки геометрических параметров сложных поверхностей. Для решения соответствующей задачи в работе был предложен авторский подход, отличающийся возможностью извлечения геометрических характеристик исследуемых поверхностей. Данный подход основан на применении камеры глубины Intel RealSense D435, реализующей оценку расстояний от объектива камеры до участков сцены в метрических единицах, предобученной сегментационной нейросетевой модели HRNet, осуществляющей выделение на изображении целевого участка сложной поверхности, а также ряде аналитических алгоритмов, реализующих оценку ключевых параметров поверхностей и классификацию сложных поверхностей, представляющих отдельные участки рельефа, на основе их геометрических характеристик. Апробация и оценка качества предложенного подхода производилась на основании тестового набора данных, включающего в себя 4500 изображений. Данный набор данных включает в себя изображения сцен — помещений, содержащих как минимум одну поверхность, по которой РС предстоит осуществлять движение. Усредненные показатели точности (accuracy, recall, precision) классификации по углу наклона и по типу перепадов высот для соответствующего набора данных составили {0.74, 0.68, 0.67} и {0.76, 0.68, 0.74} соответственно, при этом усредненная доля корректно классифицированных поверхностей по обоим классификациям составила 62,6 %. Согласно результатам проведенного тестирования, предложенное решение позволяет успешно классифицировать различные поверхности по типу угла наклона и по типу перепадов высот, а также осуществлять оценку геометрических параметров сложных поверхностей по RGB-D изображениям.

**Ключевые слова:** HRNet, Intel RealSense, классификация поверхностей, геометрия сложных поверхностей, RGB-D изображения, анализ изображений.

### ВВЕДЕНИЕ

В настоящее время одной из актуальных проблем в области управления модульными робототехническими системами (РС) является задача выбора оптимальных формаций модульных РС, исходя из специфики внешнего окружения. К особенностям внешнего окружения относится информация о наличии

препятствий на пути следования РС, а также физические свойства и геометрия поверхности, по которой РС предстоит осуществлять движение. На сегодняшний день существует ряд работ, посвященных классификации сложных поверхностей, в частности, широко представлены исследования, направленные на решение задачи классификации поверхностей на основе данных сенсорных устройств различного типа, таких как: IMU-сенсоры [1], камеры видимого спектра [2], а также тактильные пьезоэлектрические или емкостные датчики [3]. Как правило, авторы соответ-

✉ Рубцова Юлия Игоревна  
e-mail: [julia\\_rubik@mail.ru](mailto:julia_rubik@mail.ru)



Контент доступен под лицензией Creative Commons Attribution 4.0 License.

The content is available under Creative Commons Attribution 4.0 License.

ствующих исследований, разделяют задачи классификации поверхностей внутри помещений и на открытой местности. В контексте открытых пространств распространенной является следующая классификация поверхностей: гладкие (бетон, металлические пластины), крупнодисперсные (булыжники, гравий), среднелдисперсные (асфальт) и мелкодисперсные (песок, земля, трава) поверхности. В разрезе помещений может быть выделена классификация, предложенная в работе [4]. Авторы выделяют 6 типов поверхностей внутри помещений: керамогранитная плитка, лакированный паркет и ламинат, скользкий виниловый пол, ковровое покрытие высотой 0,5–1 мм, ковролин, а также мягкий ковер. Среди совмещенных классификаций отдельного внимания заслуживает исследование [5]. В данной работе авторы произвели классификацию поверхностей и выделили следующие типы поверхностей: деревянная лестница, деревянные бруски, пол с невысокими препятствиями, гладкий пол, асфальт, трава и земля.

Ключевым недостатком представленных выше решений является неполнота признакового описания поверхностей, в частности, практически нет исследований, акцентированных на классификации сложных поверхностей с точки зрения их геометрических характеристик, которые оказывают значительное влияние на эффективность механического движения РС.

Существует широкий ряд методов, способных реализовывать построение трехмерной карты окружения РС в динамической среде [6], а также ряд сегментационных алгоритмов и архитектур, позволяющих осуществлять распознавание отдельных участков внешнего окружения РС по RGB-D изображениям [7, 8]. Однако, согласно проведенному ранее исследованию [9], данные решения не предоставляют возможности по одновременной сегментации участков внешнего окружения РС с последующим предоставлением интерпретируемых данных в отношении пространственных характеристик участков поверхности движения РС.

Таким образом, настоящее исследование посвящено разработке подхода к класси-

фикации и оценке геометрических параметров сложных поверхностей движения, полученных по результатам сегментации внешнего окружения РС. Предлагаемое решение позволит реализовывать рекомендательные системы, а также полноценные системы управления формациями модульных РС в зависимости от пространственных характеристик поверхностей, по которым РС предстоит осуществлять движение.

## **1. ОБЗОР СОВРЕМЕННОГО СОСТОЯНИЯ ИССЛЕДОВАНИЙ**

Разработка подхода к классификации и оценке геометрических параметров сложных поверхностей по изображениям предполагает решение следующей группы основных задач: сегментация исследуемого изображения с целью выделения сегментов сложных поверхностей, по которым РС способна осуществлять движение; оценка пространственного положения исследуемого участка сложной поверхности; классификация и оценка параметров исследуемого участка сложной поверхности.

В настоящее время существует большое количество методов и подходов, направленных на решение задачи сегментации изображений [10–12]. Большинство современных методов сегментации основаны на применении искусственных нейронных сетей [7, 8, 13]. Наиболее подходящим решением в контексте разработки подхода к классификации и оценке параметров сложных поверхностей представляется нейросетевая модель HRNet [7], характеризующаяся высоким качеством сегментации (значение метрики IoU по данным первоисточника превышало 85 %) и способностью функционировать в режиме реального времени. В рамках предыдущих исследований [9] оригинальная нейросетевая модель была дообучена на наборе данных, включающем в себя такие типы сцен как: сцены с лестницами, сцены с ровными поверхностями, сцены с одиночными препятствиями, а также сцены с наборами препятствий. По результатам проведенных экспериментов дообученная модель HRNet продемонстрировала способность успешно сегментировать

участки сложных поверхностей на единичных RGB изображениях — качество сегментации в условиях хорошей освещенности сцены достигло 90,2 % по метрике IoU.

В контексте разработки подхода к классификации и оценке параметров сложных поверхностей по изображениям необходимо производить оценку пространственного положения исследуемого участка поверхности относительно РС. В случае, если объектив фиксирующей видеокамеры жестко зафиксирован на РС, то оценку пространственного положения участка сцены можно производить непосредственно относительно фиксирующей видеокамеры. В общем случае оценка пространственного положения некоторого участка трехмерной сцены относительно объектива видеокамеры может быть реализована на основе данных карты глубины соответствующей сцены. Современные методы построения карт глубины преимущественно могут быть отнесены к двум основным категориям: аппаратные методы [14–16] и методы, основанные на использовании нейросетевых моделей [17–19]. К недостатку методов, основанных на использовании нейросетевых моделей, можно отнести сравнительно низкий уровень точности работы в задаче построения карт глубины в сравнении с аппаратными методами. Также стоит отметить, что методы построения карт глубины, основанные на использовании нейросетевых моделей, не имеют возможности получать оценку расстояний от объектива камеры до участков сцены в метрических единицах, что исключает возможность использования подобных решений в контексте решения задачи по оценке пространственного положения участков исследуемой трехмерной сцены. В рамках настоящего исследования было принято решение использовать камеру глубины RealSense D435 [16] в связи с тем, что при решении задачи построения карт глубины данное устройство показывает высокую точность работы в сравнении с аналогами, а также позволяет осуществлять оценку расстояний от объектива камеры до исследуемых участков поверхностей на сцене в метрических единицах.

Таким образом, в рамках настоящего исследования будет предложен подход к классификации и оценке геометрических параметров сложных поверхностей по RGB-D изображениям.

## 2. РАЗРАБОТКА ПОДХОДА К КЛАССИФИКАЦИИ И ОЦЕНКЕ ПАРАМЕТРОВ СЛОЖНЫХ ПОВЕРХНОСТЕЙ ПО RGB-D ИЗОБРАЖЕНИЯМ

В рамках настоящего исследования предлагается подход к классификации и оценке параметров сложных поверхностей по RGB-D изображениям. Источником изображений служит камера Intel RealSense D435, закрепленная на модульной РС таким образом, чтобы оптическая ось объектива данной камеры располагалась в горизонтальной плоскости.

Предлагаемый подход к классификации и оценке параметров сложных поверхностей по RGB-D изображениям предполагает реализацию предварительного шага, который выполняется однократно в процессе развертывания соответствующего решения. На данном этапе устраняются эффекты дисторсии, а также перспективные искажения изображений в соответствии с алгоритмом, описанным в работе [20].

Разработанный подход к классификации и оценке параметров сложных поверхностей по RGB-D изображениям включает в себя следующие основные этапы:

1. Представление множества точек исследуемого участка сложной поверхности в трехмерном пространстве:

- a. Определение пространственного расположения областей сцены, соответствующих пикселям исследуемого RGB-D изображения;

- b. Сегментация сложных поверхностей с использованием дообученной нейросетевой модели HRNet [9];

2. Оценка геометрических параметров поверхностей;

3. Классификация сложных поверхностей, представляющих отдельные участки рельефа, на основе их геометрических характеристик;

а. Классификация поверхностей по углу наклона;

б. Классификация поверхностей по степени изменчивости высот.

Как можно заметить, предлагаемый подход содержит три ключевых этапа, каждый из которых направлен на решение отдельной группы подзадач. Рассмотрим каждый из этапов предложенного подхода более детально.

### 2.1. Представление множества точек исследуемого участка сложной поверхности в трехмерном пространстве

Для представления множества точек исследуемого участка сложной поверхности в трехмерном пространстве необходимо, в первую очередь, определить пространственное расположение областей сцены, которые соответствуют пикселям полученного с камеры Intel RealSense D435 изображения. Рассмотрим алгоритм определения пространственного расположения областей сцены более подробно.

Определим пространственную систему координат, связанную с объективом видеокамеры: ось  $Z$  ориентирована по направлению оптической оси фиксирующей видеокамеры, оси  $X$  и  $Y$  коллинеарны осям  $x$ ,  $y$  плоскости изображения. В таком случае, пространственные области, связанные с каждым пикселем изображения, отклонены от нормали к плоскости  $XY$  на некоторые плоские углы в плоскостях  $XZ$  и  $YZ$ .

Пусть размер полученного с камеры Intel RealSense D435 изображения в пикселях по горизонтальным и вертикальным осям равен  $img_{pix_x}$  и  $img_{pix_y}$  соответственно. Для каждого пикселя  $p_i$  на данном изображении с индексами по горизонтальной и вертикальной осям  $i_{pix_x}$  и  $i_{pix_y}$  соответственно вычислим угловое отклонение:

$$\varphi_{-x_{p_i}} = \left( i_{pix_x} - \frac{img_{pix_x}}{2} \right) \cdot \frac{\alpha}{img_{pix_x}};$$

$$\varphi_{-y_{p_i}} = \left( i_{pix_y} - \frac{img_{pix_y}}{2} \right) \cdot \frac{\alpha}{img_{pix_y}};$$

где  $\varphi_{-x_{p_i}}$  — угловое отклонение для пикселя  $p_i$  по оси  $x$ ;  $\varphi_{-y_{p_i}}$  — угловое отклонение для

пикселя  $p_i$  по оси  $y$ ;  $\alpha^\circ$  — угол обзора камеры по горизонтальной оси изображения (оси  $x$ );  $\beta^\circ$  — угол обзора камеры по вертикальной оси изображения (оси  $y$ ).

Поскольку для каждого пикселя  $p_i$  с использованием камеры RealSense D435 можно получить оценки расстояний  $D_i$  до центров фактических пространственных участков сцены, ассоциированных с пикселями  $p_i$ , то оценка  $R_i$  положения пространственного участка сцены, охватываемого пикселем  $p_i$ , относительно системы отсчета, связанной с объективом видеокамеры, может быть произведена следующим образом:

$$\left\{ \begin{array}{l} R_i = \frac{r_i}{|r_i|} * D_i \\ r = \left( tg(\varphi_{-x_{p_i}}); tg(\varphi_{-y_{p_i}}); 1 \right) \end{array} \right.,$$

где  $r$  — направляющий вектор, характеризующий положение пространственного участка сцены, охватываемого пикселем  $p_i$ , относительно системы отсчета, связанной с объективом видеокамеры.

Таким образом, пространственные координаты  $P_i^A$  участков сцены, связанных с пикселями изображения  $p_i$ , в глобальной системе координат  $A$  могут быть определены в соответствии со следующим выражением:

$$P_i^A = C^A + R_i^A,$$

где  $C^A$  — вектор, характеризующий пространственное положение фиксирующей видеокамеры, отложенный в глобальной системе отсчета,  $R_i^A$  — вектор, характеризующий положение пространственного участка сцены, охватываемого пикселем  $p_i$  в глобальной системе отсчета.

На следующем этапе дообученная модель нейронной сети HRNet выделяет на изображении сегменты, ассоциированные с участками сложных поверхностей. Обработывая изображение, данная нейросетевая модель возвращает количество найденных объектов и набор массивов (классы и маски). Маски представляют собой искомый набор пикселей изображения, принадлежащих исследуемому участку поверхности. Таким образом, зная целевой сегмент изображения, связанный с исследуемой поверхностью, представ-

ляется возможным определить множество точек  $P_i$ , принадлежащих исследуемой сложной поверхности:

$$P_i = (x_i, y_i, z_i).$$

Таким образом, исследуемый участок сложной поверхности формально может быть представлен в следующем виде:

$$Surf = \{P_i | i = 1, \dots, N\},$$

где  $N$  — число точек трехмерного пространства, ассоциированных с исследуемым участком сложной поверхности.

## 2.2. Оценка ключевых параметров поверхностей

Для отнесения некоторого исследуемого участка поверхности к одной из классификационных групп требуется произвести оценку параметров сложной поверхности. В рамках соответствующей оценки, в первую очередь, необходимо определить уравнение плоскости, наилучшим образом аппроксимирующей исследуемый участок поверхности  $Surf$ . В качестве такой плоскости примем плоскость  $O$ , для которой будет минимально усредненное расстояние между точками, принадлежащими участку поверхности  $Surf$ , и данной плоскостью. Общее уравнение искомой плоскости выглядит следующим образом:

$$Ax + By + Cz + D = 0; \quad x, y, z \in R, \quad (1)$$

где  $A, B, C, D$  — неизвестные параметры. Поиск значений данных параметров может быть представлен как задача параметрической оптимизации с ограничениями:

$$\frac{\sum_{i=1}^N \frac{|Ax_i + By_i + Cz_i + D|}{\sqrt{A^2 + B^2 + C^2}}}{N} \rightarrow \min_{A, B, C, D \in R}. \quad (2)$$

Таким образом, требуется определить такие значения параметров  $A, B, C, D$ , что при минимизации выражения (2) одновременно будет оставаться справедливым уравнение (1). Для решения оптимизационной задачи подобного рода может быть использован широкий спектр известных методов численной оптимизации [21], а также методов оптимизации на основе генетических алгоритмов [22].

Полученное уравнение плоскости (1) при известных параметрах будет являться оптимальным вариантом плоскостной аппроксимации пространственного распределения точек, принадлежащих исследуемому участку сложной поверхности  $Surf$ .

Затем осуществляется поиск 4 крайних точек  $P_1, P_2, P_3, P_4$ , для которых соответственно справедливы следующие выражения:

$$P_1 : \begin{cases} \max(x(P_i) | P_i \in Surf) == x(P_1) \\ \max(y(P_i) | P_i \in Surf) == y(P_1) \end{cases}$$

$$P_2 : \begin{cases} \max(x(P_i) | P_i \in Surf) == x(P_2) \\ \min(y(P_i) | P_i \in Surf) == y(P_2) \end{cases}$$

$$P_3 : \begin{cases} \min(x(P_i) | P_i \in Surf) == x(P_3) \\ \max(y(P_i) | P_i \in Surf) == y(P_3) \end{cases}$$

$$P_4 : \begin{cases} \min(x(P_i) | P_i \in Surf) == x(P_4) \\ \min(y(P_i) | P_i \in Surf) == y(P_4) \end{cases}$$

После того, как были определены точки  $P_1, P_2, P_3, P_4$ . Осуществляется определение точки пересечения двух прямых, построенных по точкам  $P_1, P_3$  и  $P_2, P_4$  соответственно. Полученную точку пересечения двух прямых обозначим как  $P_c$ .

Далее осуществляется исследование четырех отрезков:  $L_1 = [P_c; P_1]$ ,  $L_2 = [P_c; P_2]$ ,  $L_3 = [P_c; P_3]$ ,  $L_4 = [P_c; P_4]$ . Каждый рассматриваемый отрезок  $L_j$  разделяется на  $K$  участков равной длины. После чего для каждого участка  $L_{jk}$  осуществляется расчет расстояния между ним и плоскостью  $O$ . При расчете расстояния  $d_{jk}$  до плоскости  $O$ , каждый участок  $L_{jk}$  аппроксимируется некоторой точкой  $P_{jk}$ , координаты которой задаются как усредненные координаты концов такого отрезка. Расстояние  $d_{jk}$  между некоторой точкой  $P_{jk}$  и плоскостью  $O$  определяется в соответствии со следующим выражением:

$$d_{jk} = \frac{|Ax(P_{jk}) + By(P_{jk}) + Cz(P_{jk}) + D|}{\sqrt{A^2 + B^2 + C^2}}.$$

Таким образом, каждому участку  $L_{jk}$  соответствует некоторое расстояние  $d_{jk}$  до плоскости  $O$ . Полученный набор данных в отношении параметров исследуемой поверхности позволяет провести ее дальнейшую классификацию.

### 2.3. Классификация сложных поверхностей, представляющих отдельные участки рельефа, на основе их геометрических характеристик

При оценке типов поверхностей с точки зрения их геометрических характеристик может быть предложено множество вариантов классификации типов поверхностей по различным наборам критериев. В рамках данной работы было принято решение использовать фасетный метод классификации. По результатам предварительного анализа были выделены следующие ключевые классификации сложных поверхностей: по углу наклона поверхности, по степени изменчивости высот. Рассмотрим каждый вариант классификации более подробно.

С точки зрения угла наклона исследуемого участка поверхности для формирования непересекающихся классификационных группировок была предложена следующая иерархическая классификация:

1. Поверхности с условно постоянным наклоном;

а. Горизонтальные поверхности;

б. Наклонные поверхности;

2. Поверхности с переменным наклоном;

На основании полученных на предыдущем этапе наборов дискретных данных могут быть построены зависимости между расстояниями от точки  $P_c$  в направлениях  $P_1, P_2, P_3, P_4$  и соответствующими расстояниями до плоскости  $O$ . Полученные зависимости могут быть представлены в следующем виде:

$$d_{jk} = f_j \left( |P_{jk} - P_c| \right), P_{jk} \in L_j.$$

Каждая полученная зависимость  $f_j$  ( $j=1, \dots, 4$ ) может быть аппроксимирована линейной функцией вида:

$$d_{jk} = a_j \cdot |P_{jk} - P_c| + b_j.$$

Коэффициенты  $a_j, b_j$  при аппроксимирующих функциях могут быть найдены с использованием метода наименьших квадратов. На основе значений данных коэффициентов может быть сделан однозначный вывод о классификационной подгруппе, к которой относится исследуемый участок сложной поверхности:

1. Если для всех зависимостей  $f_j$  коэффициент  $a_j$  линейной аппроксимирующей функции по модулю меньше, чем некоторый  $\varepsilon_1$ , то исследуемый участок сложной поверхности относится к типу поверхностей с условно постоянным наклоном.

а. Если при этом параметры  $A$ , в плоскости  $O$  по модулю меньше, чем некоторый  $\varepsilon_2$ , то исследуемый участок сложной поверхности может быть отнесен к классу горизонтальных;

б. В противном случае исследуемый участок сложной поверхности относится к классу наклонных.

2. В случае, если хотя бы для одной зависимости  $f_j$  коэффициент  $a_j$  соответствующей линейной аппроксимирующей функции по модулю превосходит некоторый  $\varepsilon_1$ , то исследуемый участок сложной поверхности однозначно может быть отнесен к классу поверхностей с переменным наклоном.

При оценке типов сложных поверхностей с точки зрения изменчивости высот для формирования непересекающихся классификационных группировок была предложена следующая классификация:

1. Ровная поверхность;

2. Поверхность с малыми перепадами высот;

3. Поверхность со значительными перепадами высот.

Для отнесения некоторого исследуемого участка поверхности к одной из представленных выше классификационных групп могут быть использованы результаты, полученные в процессе анализа наклона исследуемого участка поверхности.

Значение выражения (2), полученное в процессе оценки параметров  $A, B, C, D$  аппроксимирующей плоскости  $O$ , характеризует усредненное расстояние между точками, принадлежащими исследуемому участку поверхности **Surf**, и данной плоскостью. На основе значений данного выражения (2) может быть сделан однозначный вывод о классификационной подгруппе, к которой относится исследуемый участок сложной поверхности:

1. В случае, если результирующее значение выражения (2) меньше, чем некоторый  $\alpha_1$ , то

исследуемый участок сложной поверхности может быть отнесен к классу ровных;

2. В случае, если результирующее значение выражения (2) больше, чем некоторый  $\alpha_1$ , но не превосходит некоторый  $\alpha_2$ , то исследуемый участок сложной поверхности может быть отнесен к классу поверхностей с малыми перепадами высот;

3. В случае, если результирующее значение выражения (2) превосходит  $\alpha_2$ , то исследуемый участок сложной поверхности может однозначно быть отнесен к классу поверхностей со значительными перепадами высот.

Параметры  $\alpha_1$  и  $\alpha_2$  задаются эмпирически и представляют собой коэффициенты масштаба, прямо пропорционально зависящие от масса-габаритных характеристик используемых модульных РС.

Предложенная классификация сложных поверхностей, по которым РС предстоит осуществлять движение, нацелена на обеспечение максимального уровня обобщения и универсальности. Посредством данной классификации могут быть описаны различные типы сложных поверхностей, такие как: лестницы, труднопроходимые местности, поверхности предполагаемого движения РС со значительным количеством препятствий и др. При этом предложенная классификация позволяет унифицировать признаковое описание сложных поверхностей с сохранением в каждом конкретном случае данных о специфических особенностях исследуемой поверхности. В частности, в рамках предложенной классификации лестничный марш будет классифицирован как наклонная поверхность со значительными перепадами высот. Дополняя классификацию данными о протяженности поверхности, об ориентации аппроксимирующей плоскости, а также о средней глубине перепадов высот, можно сохранить большую часть информации о структуре сложной поверхности, без введения для каждой семантической категории поверхностей специфических характеристик и параметров.

Далее перейдем к оценке реализации предложенного подхода к классификации и оценке параметров сложных поверхностей по RGB-D изображениям.

### 3. ТЕСТИРОВАНИЕ РАЗРАБОТАННОГО ПОДХОДА К КЛАССИФИКАЦИИ И ОЦЕНКЕ ПАРАМЕТРОВ СЛОЖНЫХ ПОВЕРХНОСТЕЙ

В рамках настоящего исследования оценка качества разработанного подхода к классификации и оценке параметров сложных поверхностей производилась на основании тестового набора данных, включающего в себя 4500 изображений, на которых представлено 6000 поверхностей различного типа с полным покрытием обеих представленных выше классификационных группировок: 3 класса поверхностей по степени изменчивости высот и 3 класса по углу наклона. Тестовый набор данных был сформирован с использованием камеры глубины RealSense D435. Данный набор данных включает в себя изображения сцен — помещений, содержащих как минимум одну поверхность, по которой РС предстоит осуществлять движение. Примеры изображений из тестового набора данных представлены на рис. 1.

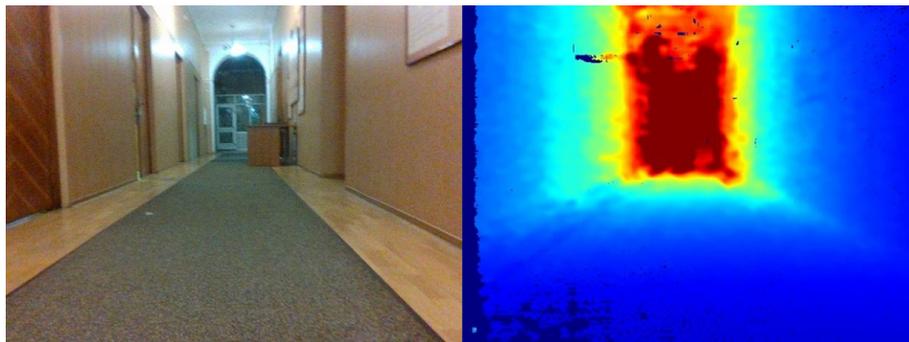
На рис. 1 представлены примеры RGB изображений из тестового набора данных и соответствующих данным изображениям карт глубины, полученных с использованием камеры глубины Intel Real Sense D435 [16]: сцен с ровной поверхностью (рис. 1.а), где в качестве такой поверхности выступает ровный пол; сцен с поверхностью с малыми перепадами высот (рис. 1.б), где в качестве такой поверхности выступает пол с малыми перепадами высот; сцен с поверхностью со значительными перепадами высот (рис. 1.в), где в качестве поверхности со значительными перепадами высот выступает лестница.

Каждому участку сложной поверхности  $\text{Surf}_i$  на каждом изображении  $I_i$ , входящему в состав тестового набора данных, соответствует некоторый набор значений параметров  $P_i$ . Описание данных параметров представлено ниже:

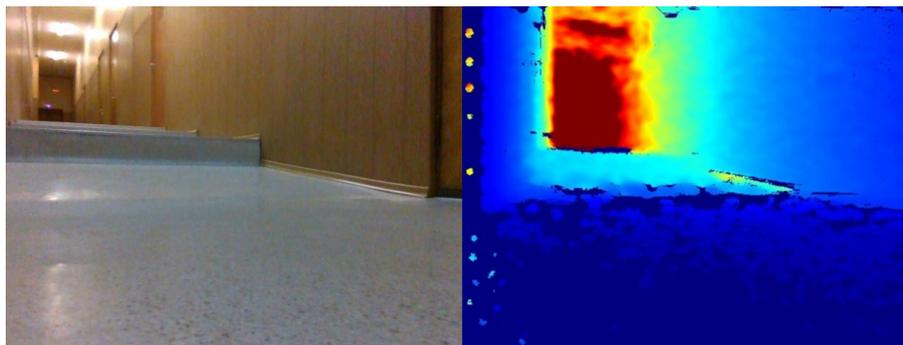
а. Тип поверхности с точки зрения угла наклона  $U$ . Включает в себя 3 различных типа угла наклона поверхности: горизонтальные поверхности; наклонные поверхности; поверхности с переменным наклоном.

Таблица 1. Матрицы ошибок классификации  
 [Table 1. Classification error matrices]

	Поверхность относится к данному типу поверхности	Поверхность не относится к данному типу поверхности
Алгоритм указал на принадлежность к данному типу поверхности	TP	FP
Алгоритм не указал на принадлежность к данному типу поверхности	FN	TN



а)



б)



в)

Рис. 1. Примеры RGB изображений и соответствующих им карт глубины из тестового набора данных для оценки качества разработанного подхода: а) сцены с ровной поверхностью; б) сцены с поверхностью с малыми перепадами высот; в) сцены с поверхностью со значительными перепадами высот

[Fig. 1. Examples of RGB images and their corresponding depth maps from the test dataset to assess the quality of the developed approach: a) scenes with a flat surface; b) scenes with a surface with small height differences; c) scenes with a surface with significant height differences]

б. Тип поверхности с точки зрения изменчивости высот  $V$ : ровная поверхность; поверхность с малыми перепадами высот; поверхность со значительными перепадами высот.

Для оценки качества разработанного подхода были сформированы различные количественные оценки точности работы предложенного решения на тестовом наборе данных. Соответствующие показатели были определены следующим образом:

1. Для тестового набора данных **Surf** на основе анализа результатов применения разработанного подхода к классификации сложных поверхностей были сформированы матрицы ошибок классификации  $\mathbf{M}_{U_i}$  и  $\mathbf{M}_{V_j}$ , где  $i$  ( $i = 1..3$ ) и  $j$  ( $j = 1..2$ ) — число классов поверхностей с точки зрения угла наклона  $U$  и перепадов высот  $V$  соответственно. Данные матрицы формировались независимо для каждого типа поверхностей по каждой из классификационных группировок (табл. 1).

В матрицах  $\mathbf{M}_{U_i}$  ячейки  $\mathbf{TP}_i$  отражают число поверхностей класса  $i$ , которые были корректно классифицированы алгоритмом и отнесены к соответствующему классу  $i$ ; ячейки  $\mathbf{FP}_i$  отображают число экземпляров данных, которые были ложно отнесены алгоритмом к рассматриваемому классу  $i$ ;  $\mathbf{FN}_i$  показывает количество случаев, когда алгоритм ложно показал отсутствие принадлежности исследуемых поверхностей к типу поверхности  $i$ ; ячейки  $\mathbf{TN}_i$  отражают число верно определенных случаев отсутствия принадлежности исследуемых поверхностей к классу  $i$ . Аналогично для матриц  $\mathbf{M}_{V_j}$  и ячеек  $\mathbf{TP}_j$ ,  $\mathbf{FP}_j$ ,  $\mathbf{FN}_j$  и  $\mathbf{TN}_j$  соответственно.

Далее для каждой матрицы вида  $\mathbf{M}$  были определены показатели качества классификации accuracy  $A$ , recall  $R$  и precision  $Pr$  в соответствии со следующими выражениями:

$$A = \frac{TP + TN}{TP + FP + FN + TN};$$

$$R_{ij} = \frac{TP}{TP + FN};$$

$$Pr_{ij} = \frac{TP}{TP + FP}.$$

На рис. 2 представлены диаграммы полученных значений показателей  $A$ ,  $R$ ,  $Pr$  для

каждого варианта классификации сложных поверхностей: по типу угла наклона  $U$  поверхности и по типу перепадов высот  $V$ .

Основываясь на представленных выше результатах, можно заключить, что тип поверхности влияет на качество классификации сложных поверхностей, в частности, рассчитанные показатели  $A$ ,  $R$ ,  $Pr$  имеют более низкие значения для следующих типов поверхностей: поверхности с переменным наклоном; поверхности с малыми перепадами высот. Предполагается, что подобные результаты могут быть связаны с возникающими погрешностями при построении карт глубины, а также с возможными погрешностями при сегментации сложных поверхностей по изображениям. Наиболее высокие значения рассчитанных показателей имеют горизонтальные поверхности, а также ровные поверхности. Усредненные значения показателей  $A$ ,  $R$ ,  $Pr$  при классификации поверхностей по типу угла наклона составили  $\{0.74, 0.68, 0.67\}$ , а значения данных показателей в отношении классификации поверхностей по типу перепадов высот составили  $\{0.76, 0.68, 0.74\}$ .

На основании полученных данных были определены результирующие оценки качества классификации поверхностей из тестового набора данных (табл. 2). Классификация поверхности считалась успешной в случае, если поверхность была корректно классифицирована в рамках обеих классификационных групп.

Как видно из представленных выше данных, результирующая доля корректно классифицированных поверхностей из тестового набора данных составила 62,6 %; поверхности были верно классифицированы по углу наклона, но неверно классифицированы по типу перепадов высот в 9,4 % случаев; доля поверхностей верно классифицированных по типу перепадов высот, но неверно классифицированных по углу наклона составила 12,8 %; поверхности были неверно классифицированы по обоим вариантам классификации в 15,2 % случаев. Исходя из полученных результатов можно заключить, что предложенное решение демонстрирует довольно высокое качество классификации сложных поверхностей

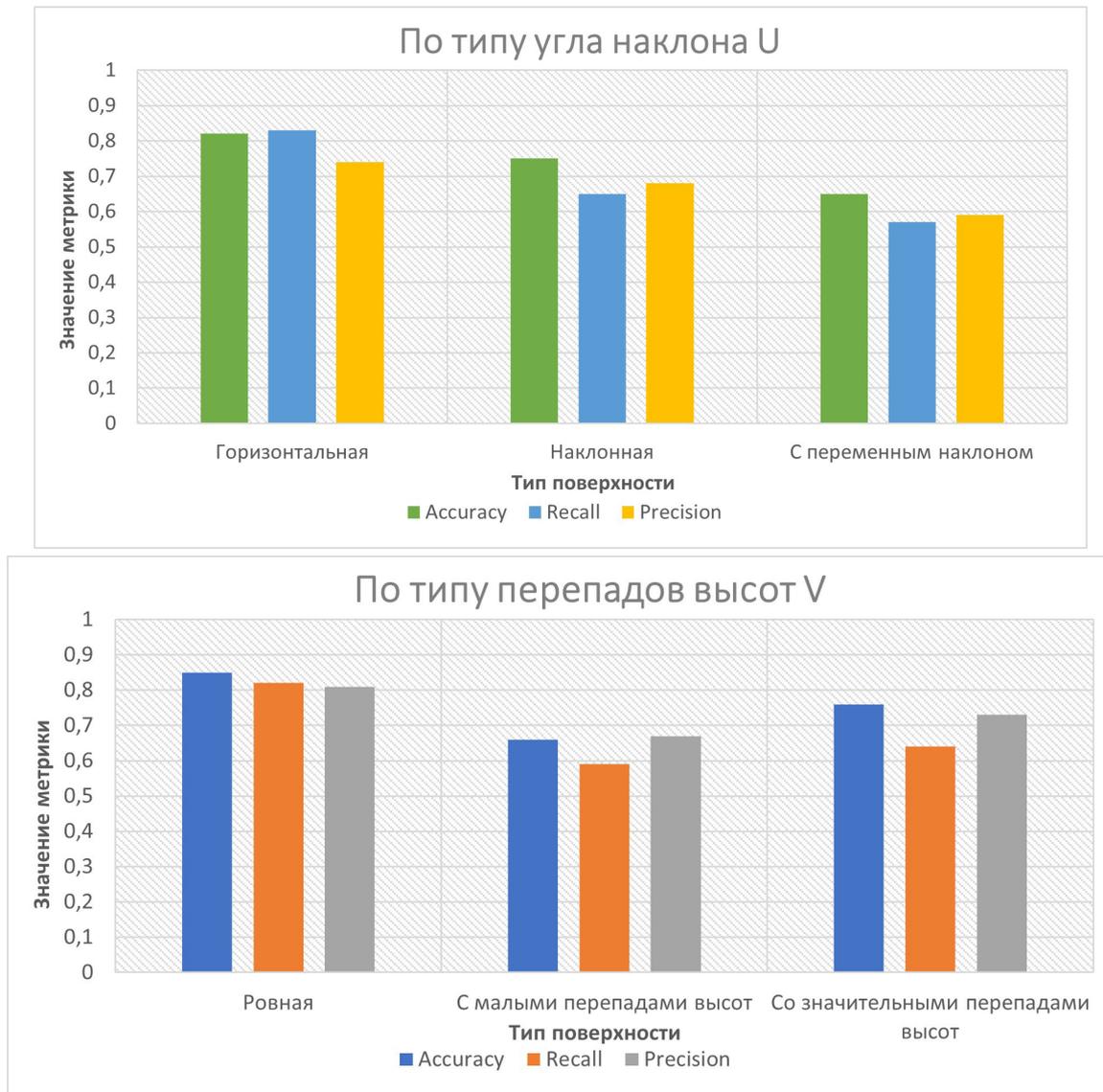


Рис. 2. Диаграммы полученных значений показателей A, R, Pr для каждого варианта классификации сложных поверхностей: по типу угла наклона U и по типу перепадов высот V [Fig. 2. Diagrams of the obtained values of indicators A, R, Pr for each variant of the classification of complex surfaces: by the type of angle of inclination U and by the type of height differences V]

Таблица 2. Результирующие оценки качества классификации поверхностей из тестового набора данных

[Table 2. Resultant quality assessments of surface classification from the test dataset]

	Количество поверхностей	Средняя доля
Поверхность верно классифицирована как по типу угла наклона, так и по типу перепадов высот	3755	62,6 %
Поверхность верно классифицирована по углу наклона, но неверно классифицирована по типу перепадов высот	561	9,4 %
Поверхность верно классифицирована по типу перепадов высот, но неверно классифицирована по углу наклона	768	12,8 %
Класс поверхности определен неверно в обоих случаях	916	15,2 %

по изображениям. Стоит также отметить, что данное решение позволяет не только определить классы поверхности, но и определить значения численных характеристик данных поверхностей, включая такие параметры как: величина угла наклона поверхности к горизонтальной плоскости, ориентация аппроксимирующей плоскости, положение краевых точек. Таким образом, предложенное решение позволяет успешно классифицировать различные поверхности по типу угла наклона и по типу перепадов высот, а также осуществлять оценку параметров данных поверхностей по RGB-D изображениям.

### ЗАКЛЮЧЕНИЕ

По результатам апробации предложенного подхода к классификации и оценке параметров сложных поверхностей на тестовом наборе данных из 4500 изображений, предложенное решение показало высокое качество классификации сложных поверхностей. Усредненные показатели точности (accuracy, recall, precision) классификации по углу наклона и по типу перепадов высот для соответствующего набора данных составили {0.74, 0.68, 0.67} и {0.76, 0.68, 0.74} соответственно. Результирующая доля корректно классифицированных сложных поверхностей составила 62,6 %. Таким образом, предложенное решение позволяет успешно классифицировать различные поверхности по типу угла наклона и по типу перепадов высот, а также осуществлять оценку геометрических параметров сложных поверхностей по RGB-D изображениям.

### БЛАГОДАРНОСТИ

Исследование выполнено при финансовой поддержке Российского научного фонда: РФФИ 20-08-01109\_А.

### КОНФЛИКТ ИНТЕРЕСОВ

Авторы декларируют отсутствие явных потенциальных конфликтов интересов, связанных с публикацией настоящей статьи.

### СПИСОК ЛИТЕРАТУРЫ

1. Hashmi, M. Z. What Lies Beneath One's Feet? Terrain Classification Using Inertial Data of Human Walk / M. Z. U. H. Hashmi, Q. Riaz, M. Husain, M. Shahzad // Applied Sciences. – 2019. – V. 9, No. 15. – P. 3099, doi: 10.3390/app9153099.
2. Ebadi, F. Road Terrain detection and Classification algorithm based on the Color Feature extraction / F. Ebadi, M. Norouzi // 2017 Artificial Intelligence and Robotics (IRAN-OPEN). – IEEE, 2017. – P. 139–146, doi: 10.1109/RIOS.2017.7956457.
3. Shill, J. J. Tactile surface classification for limbed robots using a pressure sensitive robot skin / J. J. Shill, Jr, E. G., Collins, E. Coyle, J. Clark // Bioinspiration & biomimetics. – 2015. – V. 10, No 1. – P. 016012.
4. Kertész, C. Rigidity-based surface recognition for a domestic legged robot // IEEE Robotics and Automation Letters. – 2016. – V. 1, No 1. – P. 309–315, doi: 10.1109/LRA.2016.2519949.
5. Mrva, J. Feature Extraction for Terrain Classification with Crawling Robots / J. Mrva, J. Faigl // ITAT. – 2015. – P. 179–185.
6. Jamiruddin, R. [et al.] RGB-depth SLAM review //arXiv preprint arXiv:1805.07696. – 2018.
7. Tao, A. Hierarchical Multi-Scale Attention for Semantic Segmentation / A. Tao, K. Sapra, B. Catanzaro //arXiv preprint arXiv:2005.10821. – 2020.
8. Yu, C. [et al.] Context Prior for Scene Segmentation // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. – 2020. – P. 12416–12425.
9. Rubtsova, J. Approach to Image-Based Segmentation of Complex Surfaces Using Machine Learning Tools During Motion of Mobile Robots // Electromechanics and Robotics. – Springer, Singapore, 2022. – P. 191–200, doi:10.1007/978-981-16-2814-6\_17.
10. Grady, L. Random walks for image segmentation // IEEE transactions on pattern analysis and machine intelligence. – 2006. – V. 28, No 11. – P. 1768–1783, doi:1768-1783. 10.1109/TPAMI.2006.233.
11. Ohlander, R. Picture segmentation using a recursive region splitting method // Computer

- Graphics and Image Processing. – 1980. – V. 13. – P. 222–241, doi: 10.1016/0146-664X(78)90060-6.
12. Wu, Z. An optimal graph theoretic approach to data clustering: Theory and its application to image segmentation / Z. Wu, R. Leahy // IEEE transactions on pattern analysis and machine intelligence. – 1993. – V. 15, No 11. – P. 1101–1113, doi:10.1109/34.244673.
13. He, K. [et al.] Mask r-cnn // Proceedings of the IEEE international conference on computer vision. – 2017. – P. 2961–2969.
14. Zhang, Z. Microsoft kinect sensor and its effect // IEEE multimedia. – 2012. – V. 19, No 2. – P. 4–10. doi:10.1109/MMUL.2012.24.
15. Fernald, F. G. Analysis of atmospheric lidar observations: some comments // Applied optics. – 1984. – V. 23, No 5. – P. 652–653, doi: 10.1364/AO.23.000652.
16. Keselman, L. Intel realsense stereoscopic depth cameras / L. Keselman, J. Iselin Woodfill, A. Grunnet-Jepsen, A. Bhowmik // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. – 2017. – P. 1–10.
17. Karsch, K. Depth extraction from video using non-parametric sampling / K. Karsch, C. Liu, S. B. Kang // In Proc. Europ. Conf. Computer Vision (ECCV). – 2012. – P. 775–788. doi:10.1007/978-3-642-33715-4\_56.
18. Eigen, D. Predicting depth, surface normals and semantic labels with a common multiscale convolutional architecture / D. Eigen R. Fergus // In Proc. Int. Conf. Computer Vision (ICCV), 2015.
19. Laina, I. Deeper depth prediction with fully convolutional residual networks / I. Laina, C. Rupprecht, V. Belagiannis, F. Tombari, N. Navab // In: 2016 Fourth international conference on 3D vision (3DV). – 2016. – P. 239–248. doi:10.1109/3DV.2016.32
20. Devernay, F. Automatic calibration and removal of distortion from scenes of structured environments / F. Devernay, O.D. Faugeras // Investigative and Trial Image Processing. – International Society for Optics and Photonics, 1995. – T. 2567. – P. 62–72. doi: 10.1117/12.218487.
21. Измаилов, А. Ф. Численные методы оптимизации / А. Ф. Измаилов, М. В. Солодов. – 2003.
22. Гладков, Л. А. Генетические алгоритмы / Л. А. Гладков, В. В. Курейчик, В. М. Курейчик. – 2010.

**Рубцова Юлия Игоревна** — младший научный сотрудник лаборатории технологий больших данных социкиберфизических систем Санкт-Петербургского института информатики и автоматизации Российской академии наук

E-mail: julia\_rubik@mail.ru

ORCID iD: <https://orcid.org/0000-0002-7335-7559>

**Яковлев Роман Никитич** — младший научный сотрудник лаборатории технологий больших данных социкиберфизических систем Санкт-Петербургского института информатики и автоматизации Российской академии наук

E-mail: iakovlev.r@mail.ru

ORCID iD: <https://orcid.org/0000-0002-6721-9707>

## AN APPROACH TO CLASSIFICATION AND EVALUATION OF GEOMETRIC PARAMETERS OF COMPLEX SURFACES BASED ON RGB-D IMAGES

© 2022 U. I. Rubtsova✉, R. N. Iakovlev

*St. Petersburg Federal Research Center of the Russian Academy of Sciences,  
St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences  
39, 14th Line V.O., 199178 St. Petersburg, Russian Federation*

**Annotation.** When developing control systems for formations of modular robotic systems (RS), the task of classification and evaluation of geometric parameters of complex surfaces is relevant. This approach is based on the use of the Intel RealSense D435 depth camera, which evaluates the distances from the camera lens to the scene sections in metric units, the HRNet pre-trained segmentation neural network model that selects the target area of a complex surface in the image, as well as several analytical algorithms that evaluate key surface parameters and classify complex surfaces representing individual terrain sections based on their geometric characteristics. The approbation and evaluation of the quality of the proposed approach was carried out based on a test data set including 4,500 images. This data set includes images of scenes — rooms containing at least one surface on which the PC is to move. The average accuracy indicators (accuracy, recall, precision) of the classification by angle of inclination and by type of elevation differences for the corresponding data set were {0.74, 0.68, 0.67} and {0.76, 0.68, 0.74}, respectively, while the average proportion of correctly classified surfaces in both classifications was 62.6%. According to the results of the testing, the proposed solution makes it possible to successfully classify various surfaces by the type of angle of inclination and by the type of height differences, as well as to evaluate the geometric parameters of complex surfaces using RGB-D images.

**Keywords:** HRNet, Intel RealSense, surface classification, geometry of complex surfaces, RGB-D images, image analysis.

### CONFLICT OF INTEREST

The authors declare the absence of obvious and potential conflicts of interest related to the publication of this article.

### REFERENCES

1. Hashmi M. Z. U. H., Riaz Q., Hussain M. and Shahzad M. (2019). What lies beneath one's feet? terrain classification using inertial data of human walk. *Applied Sciences*. 9(15). P. 3099. Available from: doi:10.3390/app9153099

2. Ebadi F. and Norouzi M. (2017) Road Terrain detection and Classification algorithm based on the Color Feature extraction. *Artificial Intel-*

*ligence and Robotics*. P. 139–146. Available from: doi:10.1109/RIOS.2017.7956457

3. Shill J. J., Collins Jr E. G., Coyle E. and Clark J. (2015) Tactile surface classification for limbed robots using a pressure sensitive robot skin. *Bio-inspiration & biomimetics*. 10(1). 016012.

4. Kertész C. (2016). Rigidity-based surface recognition for a domestic legged robot. *Robotics and automation letters*. 1(1). P. 309–315. Available from: doi: 10.1109/LRA.2016.2519949

5. Mrva J. and Faigl J. (2015). Feature Extraction for Terrain Classification with Crawling Robots. *In ITAT*. P. 179–185.

6. Jamiruddin R. [et al.] (2018) RGB-depth SLAM review. *arXiv preprint arXiv:1805.07696*.

7. Tao A., Sapra K. and Catanzaro B. (2020). Hierarchical multi-scale attention for semantic segmentation. *arXiv preprint arXiv:2005.10821*.

8. Yu C. [et al.] (2020) Context Prior for Scene Segmentation. *Proceedings of the IEEE/CVF Con-*

✉ Rubtsova Julia I.  
e-mail: julia\_rubik@mail.ru

*ference on Computer Vision and Pattern Recognition*. P. 12416–12425.

9. Rubtsova J. (2022). Approach to Image-Based Segmentation of Complex Surfaces Using Machine Learning Tools During Motion of Mobile Robots. In *Electromechanics and Robotics*. P. 191–200. Available from: doi:10.1007/978-981-16-2814-6\_17

10. Grady L. (2006). Random walks for image segmentation. *IEEE transactions on pattern analysis and machine intelligence*. 28(11). Available from: doi:1768-1783. 10.1109/TPAMI.2006.233

11. Ohlander R., Price K. and Reddy D. R. (1978). Picture segmentation using a recursive region splitting method. *Computer Graphics and Image Processing*. 8(3). P. 313–333. Available from: doi:10.1016/0146-664X(78)90060-6

12. Wu Z. and Leahy R. (1993). An optimal graph theoretic approach to data clustering: Theory and its application to image segmentation. *IEEE transactions on pattern analysis and machine intelligence*. 15(11). P. 1101–1113. Available from: doi:10.1109/34.244673

13. He K. [et al.] (2017). Mask r-cnn. In *Proceedings of the IEEE Conference on Computer Vision*. P. 2961–2969.

14. Zhang Z. (2012). Microsoft kinect sensor and its effect. *IEEE multimedia*. 19(2). P. 4–10. Available from: doi:10.1109/MMUL.2012.24

15. Fernald F. G. (1984) Analysis of atmospheric lidar observations: some comments. *Applied optics*. V. 23, No 5. P. 652–653. Available from: doi:10.1364/AO.23.000652

16. Keselman L., Iselin Woodfill J., Grunnet-Jepsen A. and Bhowmik A. (2017). Intel realsense stereoscopic depth cameras. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*. P. 1–10.

17. Karsch K., Liu C. and Kang S. B. (2012, October). Depth extraction from video using non-parametric sampling. In *European conference on computer vision*. P. 775–788. Available from: doi:10.1007/978-3-642-33715-4\_56

18. Eigen D. and Fergus R. (2015). Predicting depth, surface normals and semantic labels with a common multi-scale convolutional architecture. In *Proceedings of the IEEE international conference on computer vision*. P. 2650–2658.

19. Laina I., Rupprecht C., Belagiannis V., Tombari F. and Navab N. (2016, October). Deeper depth prediction with fully convolutional residual networks. In *2016 Fourth international conference on 3D vision (3DV)*. P. 239–248. Available from: doi:10.1109/3DV.2016.32

20. Devernay F. and Faugeras O. D. (1995, September). Automatic calibration and removal of distortion from scenes of structured environments. In *Investigative and Trial Image Processing*. (2567). P. 62–72, Available from: doi:10.1117/12.218487.

21. Izmailov A. F. and Solodov M. V. (2003) Chislenny`e metody` optimizacii [Numerical optimization models]. Moscow. (in Russian)

22. Gladkov L. A., Kurejchik V. V. and Kurejchik V. M. (2010) Geneticheskie algoritmy` [Genetic algorithms]. Moscow. (in Russian)

**Rubtsova Julia I.** — Junior Researcher Laboratory of Big Data Technologies of Socio-Cyberphysical Systems of the St. Petersburg Institute of Informatics and Automation of the Russian Academy of Sciences.

E-mail: julia\_rubik@mail.ru

ORCID iD: <https://orcid.org/0000-0002-7335-7559>

**Iakovlev Roman N.** — Junior Researcher Laboratory of Big Data Technologies of Socio-Cyberphysical Systems of the St. Petersburg Institute of Informatics and Automation of the Russian Academy of Sciences.

E-mail: iakovlev.r@mail.ru

ORCID iD: <https://orcid.org/0000-0002-6721-9707>