

**Ю. В. Войтешонок**

## **ПОСТРОЕНИЕ КАРТЫ ГЛУБИНЫ ИЗОБРАЖЕНИЯ ДЛЯ ПОРТАТИВНЫХ УСТРОЙСТВ**

Балтийский федеральный университет им. И. Канта, Калининград, Россия

Поступила в редакцию 24.02.2022 г.

Принята к публикации 10.03.2022 г.

14

**Для цитирования:** *Войтешонок Ю. В.* Построение карты глубины изображения для портативных устройств // Вестник Балтийского федерального университета им. И. Канта. Сер. Физико-математические и технические науки. 2022. №1. С. 14–20.

*Приводится описание и сравнение алгоритмов построения стереоизображений с использованием одной или двух видеокамер для применения в портативном устройстве в режиме реального времени. На основе замеров производительности на наиболее распространенных процессорах одноплатных компьютеров выбран наиболее подходящий алгоритм.*

**Ключевые слова:** стереоизображение, полуглобальное сопоставление блоков, сопоставление блоков, одноплатный компьютер, алгоритмы, сопоставление блоков

На данный момент разработано много способов получения информации об удаленности. В статье рассмотрены алгоритмы, пригодные для использования в портативных устройствах при построении двурмерного стереоизображения в реальном времени.

Наиболее сбалансированной технологией получения стереоизображений является метод, в котором используются разнесенные по одной оси видеокамеры, с применением алгоритмов сопоставления блоков. Данный метод при достаточной юстировке дает наиболее детальную и гладкую карту глубины. Существуют две наиболее популярные реализации алгоритма сопоставления блоков: алгоритм, использующий окно SAD для нахождения точек сопоставления между двумя изображениями, и метод полуглобального сопоставления блоков (SGBM).

Основной механизм алгоритма сопоставления блоков заключается в коррекции и выравнивании двух снятых изображений таким образом, чтобы сравнения проводились только вдоль отдельных строк. Далее алгоритм выполняет поиск строк в двух изображениях для сопоставления групп пикселей. Благодаря своей простоте и предварительной юстировке камер этот алгоритм достаточно быстр.



Для получения стереоизображения большинство алгоритмов обычно следуют четырем шагам [1]: вычисление стоимости сопоставления, агрегация стоимости, вычисление диспаратности, обработка диспаратности. Алгоритмы вычисления диспаратности делятся на локальные и глобальные методы. Локальные методы оценивают один пиксель за один проход, используя для вычислений только соседние пиксели. Это объясняет ошибки данных методов при обнаружении внезапных изменений глубины и окклюзий. Глобальные методы рассматривают информацию, доступную на всем изображении.

Полуглобальное сопоставление является компромиссом между вычислительной сложностью глобального метода и неизбежностью ошибок локального. В отличие от глобального, этот метод рассматривает небольшой блок пикселей. Данный метод также использует информацию от соседних пикселей в нескольких направлениях для вычисления диспаратности пикселя.

Алгоритм SGBM использует блочное полуглобальное сопоставление стоимости, которое сглаживается информацией из нескольких направлений (рис. 1). Этот алгоритм использует взаимную информацию в качестве меры локального сопоставления и повышает согласованность по направлениям, отличным от горизонтальной линии. Однако его вычислительная сложность выше предыдущего алгоритма.

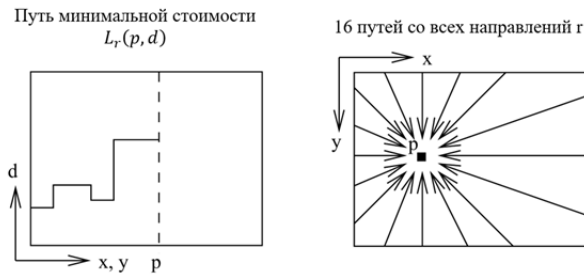


Рис. 1. Агрегация стоимости

На рисунке 2 показано, что алгоритм ищет соответствующий пиксель в левом изображении [2] среди  $D$  пикселей в правом изображении. Полученные значения  $D$  представляют собой уровни диспаратности  $D$  для пикселя в левом изображении. Часть левого изображения не используется, так как для них отсутствуют для сравнения пиксели в правом изображении.

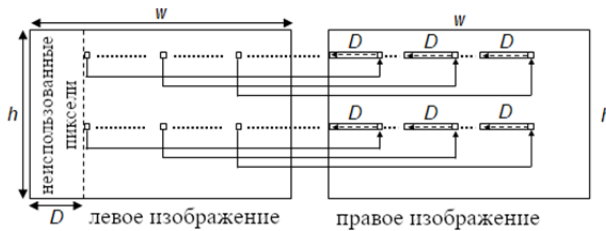
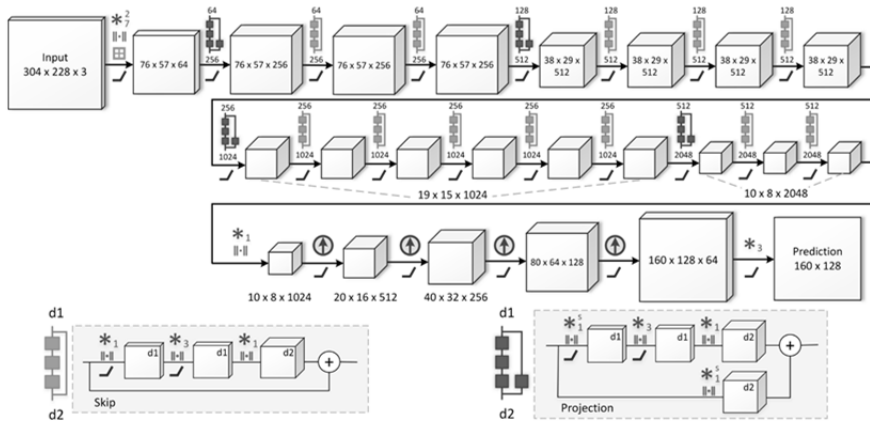


Рис. 2. Работа SGBM для двух входных изображений:  $w$  – ширина изображения,  $h$  – высота изображения



Из одного изображения возможно вычислить карту глубины при помощи нейронной сети архитектуры FCRN, основанной на CNN. Эта архитектура (рис. 3) основана на ResNet-50 с заменой полносвязных слоев на блоки повышения разрешения, дающие на выходе примерно половину входного разрешения [4].



- $*_n^{(s)}$  свертка  $n \times n$  со страйдом  $s$
- $\wedge$  ReLU
- $\text{max}$  слой объединения по максимуму (max-pooling)  $3 \times 3$  со страйдом 2
- $\text{||} \cdot \text{||}$  пакетная нормализация
- $\uparrow$  up-convolution

Рис. 3. Архитектура FCRN

Для сравнительных тестов используется изображение, которое показано на рисунке 4 [5].

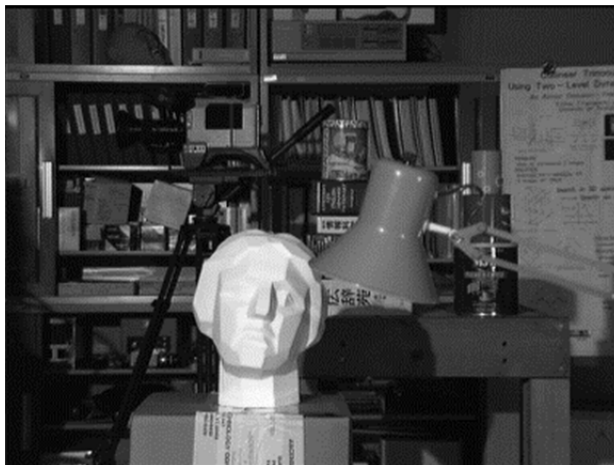
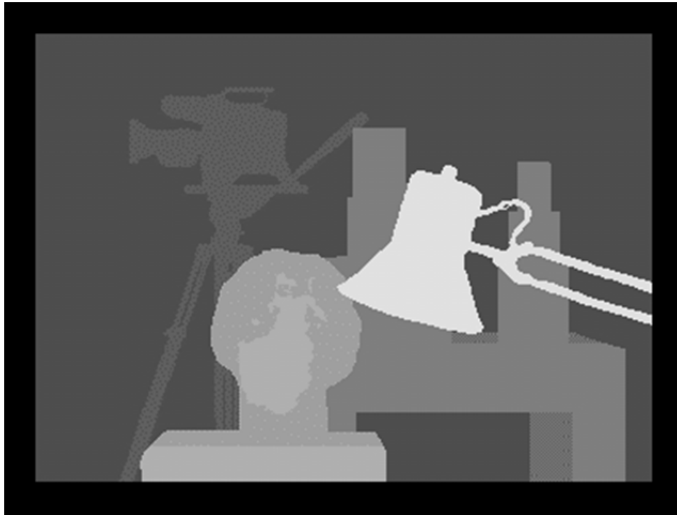


Рис. 4. Фото, снятое с левой камеры



На рисунке 5 изображена истинная карта глубины.



17

Рис. 5. Истинная карта глубины

Программная реализация алгоритма SGBM взята из библиотеки OpenCV версии 4.5.1 [3]. Приложение вычисления карты глубины написано на Python. Для заданного изображения найдены значения `numDisparities` и `blockSize`. Их выбор осуществлялся на основе рекомендаций документации и сравнения результата с истинной картой глубины. Таким же образом использовался алгоритм BM.

Результаты работы SGBM приводятся на рисунке 6, для BM – на рисунке 7.

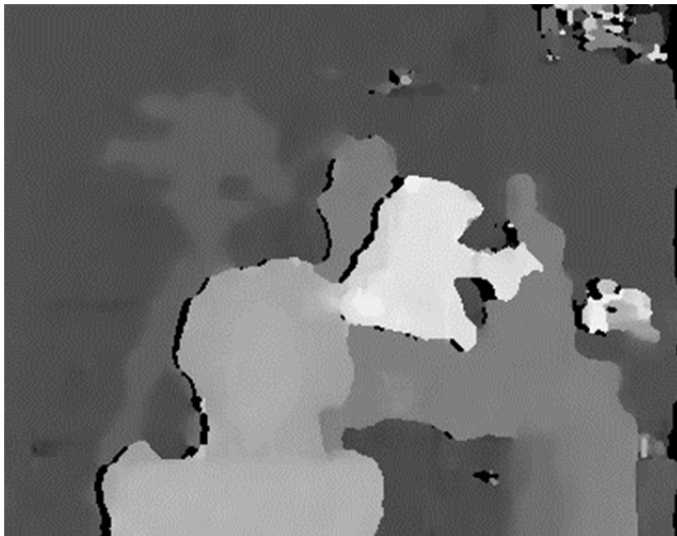


Рис. 6. Карта глубины, полученная при использовании SGBM

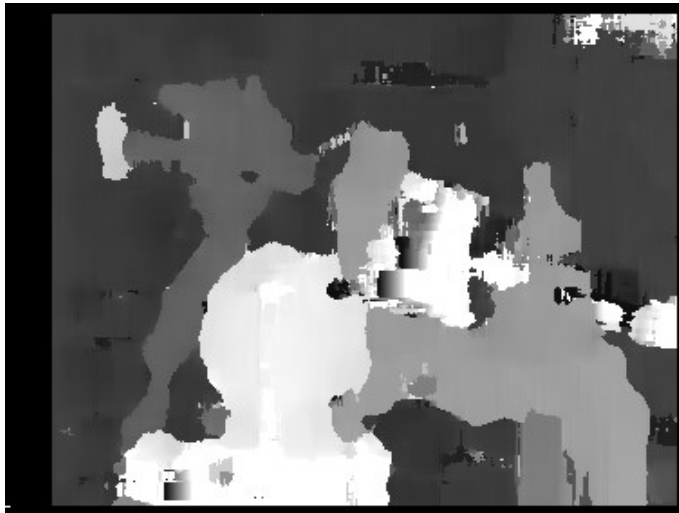


Рис. 7. Карта глубины, полученная при использовании ВМ

Для построения карты глубины на основе одного изображения с помощью FCRN использована его реализация для библиотеки TensorFlow [6]. Данная модель обучена для вычисления карты глубины из единичного входного RGB изображения из наборов данных NYU Depth v2 и Make3D для сцен в помещении и на улице соответственно [7]. Для простоты дальнейшего портирования для работы с моделью задействован фреймворк TensorFlow. Полученная карта глубины для левого изображения показана на рисунке 8.

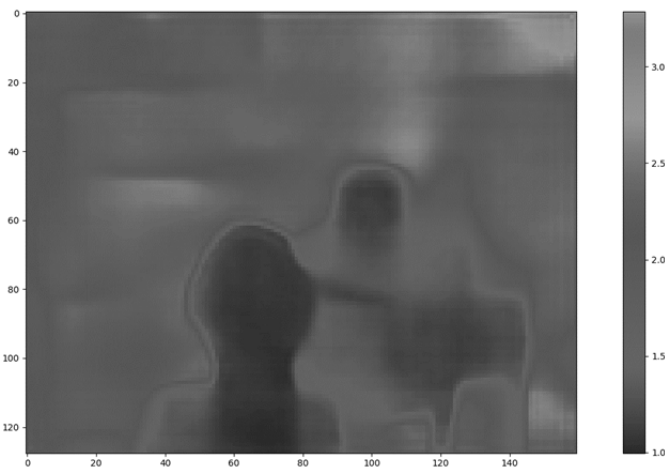


Рис. 8. Результат работы FCRN для левого изображения (рис. 4)

Как можно заметить, алгоритм SGBM вычисляет достаточно равномерную карту глубины, сохраняя большинство деталей. ВМ на рисун-



ке 7 предоставляет схожую с SGBM информацию об удаленности объектов, теряя незначительное количество информации. FCRN на рисунке 8 предоставляет меньше всего информации.

Тестирование алгоритмов проводилось для использования в SBC в переносном устройстве. Для SGBM как более сложного среди алгоритмов сопоставления блоков тесты проводились на BCM2711 (Broadcom) с ядром ARM A72 [8] (ARM v8) и BCM2835 (Broadcom) с микроархитектурой ARM1176JZF-S (ARM v6). Помимо очевидных различий в тактовой частоте, ISA BCM2711 расширяется SIMD по технологии Neon [9] (это отражено при компиляции OpenCV). Для изображения  $384 \times 288$  время построения карты для BCM2711 составило 0,044 с, для BCM2835 – 0,74 с.

Исходя из проведенного сравнения трех алгоритмов построения карты глубины стереоизображения для переносных устройств можно сделать следующий вывод. Алгоритм простого сопоставления блоков имеет более низкую вычислительную сложность. Алгоритм SGBM работает медленнее простого сопоставления, но карта глубины изображения является более плавной и информативной. Алгоритм FCRN позволяет получить карту глубины при использовании одной камеры, но большой размер модели накладывает ограничения на ее использование. Таким образом, наиболее подходящим алгоритмом построения карты глубины изображения для использования в портативных устройствах в режиме реального времени является SGBM.

### Список литературы

1. Spangenberg R., Langner T., Roja R. Computer Analysis of Images and Patterns // Weighted Semi-Global matching and Center-Symmetric Census Transform for Robust Driver Assistance. Berlin, 2013. P. 34–41.
2. Setyawan R., Sunoko R., Choiron M., Mudjirahardjo P. Implementation of Stereo Vision Semi-Global Block Matching Methods for Distance Measurement // Indonesian Journal of Electrical Engineering. 2018. №12 (2). P. 585–591.
3. Stereo datasets with ground truth. 2011 // Stereo datasets : [сайт]. URL: <https://vision.middlebury.edu/stereo/data/scenes2001/> (дата обращения: 04.05.2021).
4. Deeper Depth Prediction with Fully Convolutional Residual Networks. 2017 // GitHub : [сайт]. URL: <https://github.com/iro-cp/FCRN-DepthPrediction> (дата обращения: 02.01.2021).
5. Deeper Depth Prediction with Fully Convolutional Residual Networks. Stanford, 2016.
6. Cortex-A72 // Arm Developer : [сайт]. URL: <https://developer.arm.com/ip-products/processors/cortex-a/cortex-a72> (дата обращения: 23.03.2021).
7. Neon // Arm Developer : [сайт]. URL: <https://developer.arm.com/architectures/instruction-sets/simd-isas/neon> (дата обращения: 09.03.2021).

### Об авторе

Юрий Владимирович Войтешонок – студ., Балтийский федеральный университет им. И. Канта, Калининград, Россия.

E-mail: JVVojteshonok@stud.kantiana.ru



*Yu. V. Voiteshonok*

## BUILDING AN IMAGE DEPTH MAP FOR PORTABLE DEVICES

Immanuel Kant Baltic Federal University, Kaliningrad, Russia

Received 24 February 2022

Accepted 10 March 2022

**To cite this article:** Voiteshonok Yu. V. 2022, Building an image depth map for portable devices, *Vestnik of Immanuel Kant Baltic Federal University. Series: Physical-mathematical and technical sciences*, №1. P. 14–20.

20

*A description and comparison of algorithms for constructing stereo images using one or two video cameras for use in a portable device in real time is given. Based on performance measurements for the most common processors for single-board computers, the most appropriate approach was selected.*

**Keywords:** stereo image, Semi-Global Block Matching, Block Matching, single board computer, algorithms, block matching

### The author

Yury V. Voiteshonok, Student, Immanuel Kant Baltic Federal University, Kaliningrad, Russia.

E-mail: JVVojteshonok@stud.kantiana.ru