

# Задача построения стерео HDR изображений по изображениям с разными выдержками, полученным с двух камер стереомикроскопа

Наталья Пащенко, Кристина Зипа, Алексей Игнатенко  
Факультет вычислительной математики и кибернетики

Московский государственный университет имени М. В. Ломоносова, Москва, Россия  
npashchenko@yandex.ru, kzipa@graphics.cs.msu.ru, ignatenko@graphics.cs.msu.ru

## Аннотация

В работе рассматривается подход к решению задачи, позволяющий значительно сократить время съемки сцены оператором и повышающий качество съемки движущихся объектов.

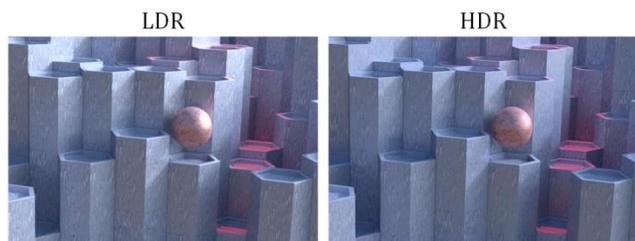
Существующие методы решения задачи предполагают ее декомпозицию на следующие подзадачи: поиск стереосоответствий между изображениями, полученными с двух ракурсов с разными выдержками, и собственно построение HDR изображений.

Авторами собрана база стереоизображений, снятых с разными выдержками с микроскопа. Для решения задачи выбран алгоритм [1], произведена оценка качества его работы на собранной базе по метрике HDR-VDP-2.2 ([14]) с использованием алгоритма поиска стереосоответствий, предложенного в [1], и другими методами, а также представлены дальнейшие планы по улучшению получаемых результатов.

**Ключевые слова:** HDR, стереоизображение, поиск стереосоответствий, диспаритет.

## 1. ВВЕДЕНИЕ

Окружающий мир обладает столь большим разнообразием и огромным диапазоном цветов и яркостей, что порой не получается запечатлеть его с желаемой точностью цветопередачи и высоким уровнем детализации как в темных, так и в светлых областях. В связи с этим возникает задача наиболее корректного отображения реального мира на получаемых при съемке изображениях.

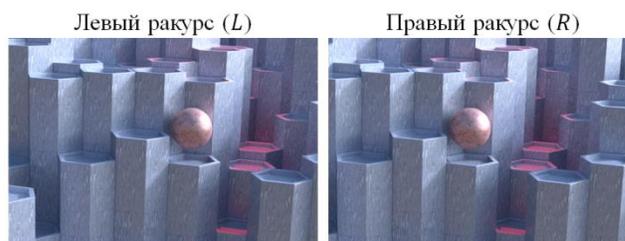


**Рис. 1:** Пример LDR изображения и HDR изображения с примененной к нему тональной компрессией [15] ([1]).

Наиболее распространены изображения, использующие целочисленное 8-битное представление значений яркости, такие изображения принято называть изображениями узкого динамического диапазона (LDR или SDR – low/standard dynamic range images), они отлично подходят для отображения на большинстве современных дисплеев. Также существуют целочисленные форматы, использующие большее число битов, например, RAW формат.

Человеческий глаз способен адаптироваться к условиям освещения в пределах десяти порядков изменения яркости с возможностью одновременно охватить около пяти порядков, в то время как на 8-битном LDR изображении может быть отражено лишь около двух порядков ([4], [16]). Поэтому закономерно появилась необходимость в создании изображений, способных представить произвольный диапазон значений яркости, – изображений широкого динамического диапазона (HDR – high dynamic range images) (Рис. 1), такие изображения могут использовать представление яркости в виде числа с плавающей запятой.

Но даже когда проблема с передачей цвета на изображении решена, может возникнуть другая проблема, заключающаяся в сложности восприятия человеком объекта, находящегося в плоскости экрана (2D), в силу потери такой характеристики, как глубина пространства. Для ее решения используется съемка объекта с двух (или более) камер, результатом которой является стереопара (Рис. 2), позволяющая с помощью специальных очков увидеть стереоизображение (3D).



**Рис. 2:** Стереопара (стереоизображение) ([1]).

Возникновение технологии построения стерео HDR изображений (SHDR – stereoscopic high dynamic range) позволило решить обе указанные проблемы. Данная технология, объединив преимущества HDR изображений и стереоизображений, позволила увеличить реализм снимаемых сцен.

Помимо съемки классических сюжетов она может быть использована для качественной съемки хирургических операций, а также для съемки объектов, рассматриваемых под стереомикроскопом (для облегчения восприятия глубины), при которой часто можно наблюдать резкие перепады интенсивности света, например, при наличии бликов на объекте. В работе рассмотрено последнее применение данной технологии.

## 2. ПОСТАНОВКА ЗАДАЧИ

Стандартная схема построения стерео HDR изображений заключается в независимом построении HDR изображений для левого и правого ракурсов путем съемки нескольких изображений с разной экспозицией (Рис. 3а). В работе

предлагается использовать схему, изображенную на Рис. 3б, в которой изображения с различными экспозициями распределены между двумя ракурсами, такая схема позволяет существенно сократить время съемки. При использовании стандартной схемы возможны ошибки, связанные с движением объектов между кадрами. Предложенная схема позволяет уменьшить влияние движения на результат, особенно при использовании двух значений экспозиции.

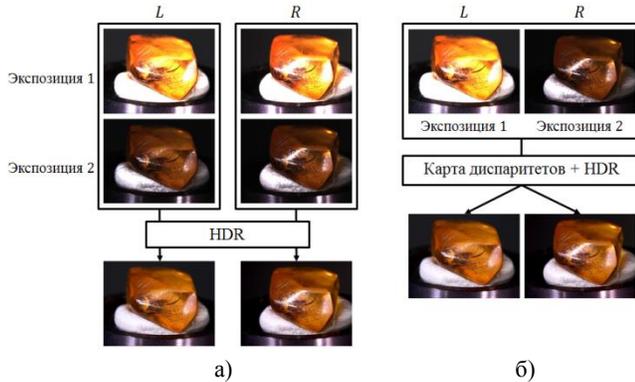


Рис. 3: Стандартная (а) и предлагаемая (б) схемы в случае двух экспозиций.

Таким образом, задача заключается в построении стереопары HDR изображений по изображениям, снятым с двух камер стереомикроскопа с разными выдержками, распределенными между ракурсами.

На вход алгоритму подаются

- изображений ( $N = 2$ )  $\{Img_i\}_{i=1}^N$ , снятых с двух ракурсов с разными выдержками;
- $N$  значений выдержек  $\{\Delta t_i\}_{i=1}^N$ .

Выходом алгоритма является пара построенных HDR изображений для левого и правого ракурсов:

$$\{HDR_L, HDR_R\} = SHDR(\{Img_i\}_{i=1}^N, \{\Delta t_i\}_{i=1}^N). \quad (1)$$

### 3. ОБЗОР СУЩЕСТВУЮЩИХ МЕТОДОВ

С использованием видеоизмененной схемы с Рис. 3б возникает необходимость проведения дополнительного этапа – построения карты диспаратетов. Он нужен для установления соответствия между изображениями, снятыми с разных ракурсов, чтобы иметь возможность использовать информацию о цвете с каждого из них (Рис. 4).

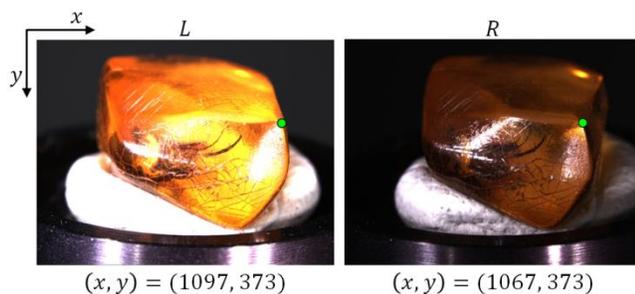


Рис. 4: Соответствующие точки на левом и правом изображениях и их координаты.

Карта диспаратетов представляет собой числовую матрицу такого же размера, как и входные изображения, каждый

элемент которой соответствует пикселю с теми же координатами. Элементы такой матрицы имеют значения смещений точек сцены на правом изображении относительно левого и наоборот.

Общая идея существующих алгоритмов заключается в поиске функции отклика камеры, построении карты диспаратетов и генерации HDR изображений с ее использованием ([19], [1], [20]). Функция отклика камеры используется для преобразования исходных изображений с различными экспозициями в пространство энергетической яркости (radiance space), в котором каждая точка объекта на изображениях будет иметь примерно одинаковую яркость. Это позволяет получать более качественные карты диспаратетов для изображений с различными экспозициями. Таким образом, алгоритмы могут отличаться способами вычисления функции отклика камеры ([3], [13], [17], [5], [21], [8]), методами построения карты диспаратетов, классификация и описание которых представлены в [18], а также алгоритмами генерации HDR изображения ([9], [17], [3]).

В [19] предлагается вычислить начальные карты диспаратетов для левого и правого изображений с гамма-коррекцией, по ним выбрать достоверные значения и, используя только цвета пикселей с достоверными диспаратетами, вычислить полиномиальную функцию отклика камеры. Затем на основе полученной функции провести преобразование входных изображений в пространство энергетической яркости, значения пикселей в котором позволяют откорректировать и дополнить начальную карту диспаратетов. Карты диспаратетов строятся с использованием алгоритма минимизации энергии [2]. Полученная карта используется далее для построения HDR изображения ([13]).

В отличие от алгоритма [19] в [1] вычисление функции отклика камеры выполняется один раз, и производится оно с использованием изображений, снятых с одного ракурса, что обеспечивает большую надежность вычислений, однако требует статичности снимаемой сцены. Затем карты диспаратетов могут быть получены с использованием изображений в пространстве энергетической яркости, и необходимость во втором этапе стереосопоставления исчезает. Для вычисления карты диспаратетов в качестве функции стоимости соответствия пикселей используется нормированная взаимнокорреляционная функция, которая, как показано в [7] и [20], является одной из наиболее подходящих для изображений с различными экспозициями. Полученные значения стоимостей усредняются по крестообразным областям, построенным вокруг каждого пикселя ([11]). Далее каждому пикселю присваивается значение диспаратета с минимальной стоимостью, производится уточнение карты диспаратетов и генерация HDR изображения, основанная на алгоритме [17].

Наличие фиксированных камер в стереомикроскопе позволяет провести вычисление функции отклика камеры один раз, как предложено в [1]. Также в [1] перед генерацией HDR изображения предлагается по изображению с левого ракурса строить битовую маску, в которой значение 1 имеют недоэкспонированные или переэкспонированные пиксели. При построении HDR изображения пиксели, вошедшие в маску, корректируются с использованием соответствующих пикселей изображения с правого ракурса, значения остальных пикселей берутся без изменений. Подобное использование маски при построении HDR изображения на

изображениях со стереомикроскопа может позволить отсечь фон, не интересующий оператора и не нуждающийся в расширении динамического диапазона. Таким образом, вследствие указанных особенностей в качестве исследуемого алгоритма был выбран [1].

#### 4. БАЗА СТЕРЕОИЗОБРАЖЕНИЙ С МИКРОСКОПА

Для тестирования алгоритмов была собрана база изображений, снятых со стереомикроскопа с разными экспозициями. На экспозицию, представляющую количество света, попадающего на светочувствительный элемент, оказывают влияние три основных параметра: светочувствительность, диафрагма и выдержка. В качестве параметра, изменяющего значение экспозиции, была выбрана выдержка. Данный подход позволяет получить широкий динамический диапазон и не оказывает сильного влияния на качество результатов этапа стереосопоставления. Также съемка серии изображений с различными значениями выдержки в автоматическом режиме поддерживается многими камерами.

Установка для получения стереоизображений включала стереомикроскоп Leica M205 A с двумя высокочувствительными CCD камерами Prosilica GC1380 и системой захвата цифрового стерео, подсоединенный к компьютеру, с которого производилось управление 5-осевым моторизованным держателем объекта, освещением сцены и параметрами съемки. Авторами были выделены следующие классы рассматриваемых объектов:

- контрастные;
- слабоконтрастные;
- высокодетализированные;
- низкодетализированные;
- плоские;
- выпуклые;
- прозрачные, преломляющие;
- непрозрачные.

Было отснято 27 объектов и получено 76 различных сцен; размер изображений 1332×1011 пикселей. Примеры полученных изображений представлены на Рис. 5.



Рис. 5: Примеры изображений из собранной базы.

К специфике рассматриваемой предметной области можно отнести малое расстояние от камер до объекта и, как следствие этого, большой диапазон возможных значений диспаратета (включая отрицательные значения), наличие прозрачных объектов и бликов, а также присутствие на некоторых изображениях фона, не интересующего оператора. Указанные особенности повышают сложность этапа поиска стереосоответствий, а также демонстрируют актуальность применения технологии HDR к изображениям с выдержками,

распределенными между ракурсами, например, в случае блика, присутствующего на всех изображениях с одного ракурса, но отсутствующего на изображениях с другого ракурса.

#### 5. ОЦЕНКА КАЧЕСТВА РЕЗУЛЬТАТОВ

Качество получаемых HDR изображений главным образом определяется используемой схемой поиска стереосоответствий. Существует множество алгоритмов, дающих приемлемые результаты для изображений с одинаковым освещением. Но большинство из них не работают на изображениях с радиометрическими различиями, возникающими в результате изменения экспозиции, освещения или положения источника света, наличия неламбертовских поверхностей, для которых количество отраженного света зависит от угла наблюдения, и т.п.

Согласно отмеченному выше при тестировании алгоритмов, решающих поставленную задачу, целесообразно сравнивать не только результирующие HDR изображения, но и построенные карты диспаратетов. Однако в данном случае необходимо иметь эталонные карты диспаратетов, получение которых для реальных данных является отдельной нетривиальной задачей. Поэтому выбор алгоритма поиска стереосоответствий базируется на исследовании [7], посвященном решению данной задачи для изображений с радиометрическими различиями.

Для оценки результатов алгоритма в целом используется метрика HDR-VDP-2.2 ([14]), определяющая качество получаемого изображения по отношению к контрольному изображению, выраженное средней экспертной оценкой (MOS – mean opinion score), варьирующейся от 0 (наихудшее качество) до 100 (наилучшее качество), и предсказывающая вероятность того, насколько среднестатистическому наблюдателю будет заметна разница между изображениями. На вход ей подаются проверяемое изображение и контрольное, в качестве которого берется HDR изображение, построенное по одному ракурсу.

#### 6. ОПИСАНИЕ Тестируемых алгоритмов

Базовым алгоритмом решения задачи был выбран [1], общая схема работы которого представлена на Рис. 6.

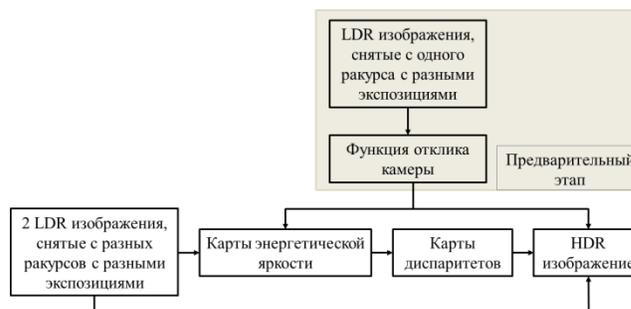


Рис. 6: Схема алгоритма [1].

В нем в качестве функции стоимости выбора значения диспаратета  $d$  для пикселя  $p$  используется нормированная взаимнокорреляционная функция  $C(p, d)$  с нулевым средним (ZNCC – zero-mean normalized cross-correlation):

$$C(p, d) = \frac{\sum_{q \in N_p} (I_L(q) - \bar{I}_L(p)) (I_R(q - (d, 0)) - \bar{I}_R(p - (d, 0)))}{\sqrt{\sum_{q \in N_p} (I_L(q) - \bar{I}_L(p))^2 \sum_{q \in N_p} (I_R(q - (d, 0)) - \bar{I}_R(p - (d, 0)))^2}} \quad (2)$$

$p$  – область (окно) с центром в пикселе  $p$  размера  $9 \times 9$ ,  
 $I_L, I_R$  – изображения, снятые с левой и правой камер соответственно,  
 $\bar{I}_L(p), \bar{I}_R(p)$  – средние значения в области  $p$  изображений  $I_L$  и  $I_R$ .

Полученные значения  $C(p, d)$  изменяются в диапазоне от -1 до 1, где -1 означает худшее совпадение, а 1 – лучшее. Поэтому для того чтобы данная функция действительно была стоимостью, следует поменять знаки.

Согласно [7] ZNCC является одной из лучших функций стоимости для работы с изображениями с различными экспозициями среди класса параметрических стоимостей, в то время как лучшей непараметрической признана стоимость, полученная с использованием Census преобразования [22], которое также включено в рассмотрение. При таком преобразовании для каждого пикселя изображения строится битовая строка, в которой все биты соответствуют определенным пикселям в окрестности рассматриваемого пикселя. Бит строки имеет значение 1, если некоторая характеристика соответствующего пикселя (например, яркость или цветовая компонента) меньше характеристики центрального пикселя окрестности. При этом значением стоимости является расстояние Хэмминга между битовыми строками сравниваемых пикселей. Затем производится усреднение карт стоимостей, и диспаратетом пикселя является то значение, при котором стоимость минимальна.

Помимо усреднения стоимостей по крестообразным областям и выбора диспаратетов по стратегии WTA (winner takes all), свойственной локальным алгоритмам построения карт диспаратетов, было решено протестировать полуглобальный алгоритм [6], примененный к стоимостям, полученным с использованием Census преобразования.

Во всех экспериментах для вычисления функции отклика камеры и построения HDR изображений был использован алгоритм [17]. Согласно ему

$$HDR(p) = \frac{\sum_i w(Img_i(p)) \Delta t_i r(Img_i(p))}{\sum_i w(Img_i(p)) \Delta t_i^2}, \quad (3)$$

$HDR$  – результирующее HDR изображение,

$w(\cdot)$  – весовая функция,

$\Delta t_i$  – выдержка изображения  $Img_i$ ,

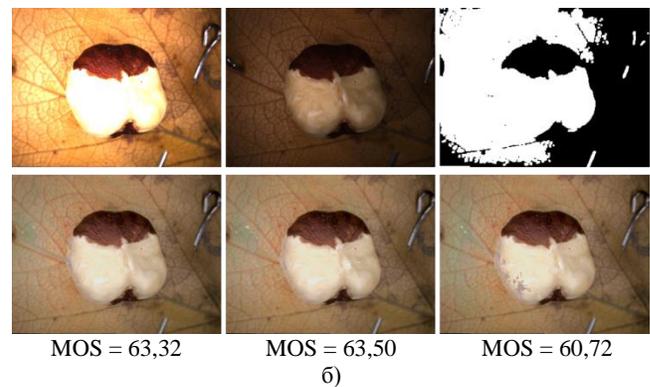
$r(\cdot)$  – отображение, переводящее LDR изображение в пространство энергетической яркости.

## 7. РЕЗУЛЬТАТЫ ТЕСТИРОВАНИЯ

Итак, на собранной со стереомикроскопа базе изображений были протестированы следующие алгоритмы: алгоритм [1] (ZNCC), [1] с использованием [22] (Census) и [1] с использованием [22] и [6] (Census\_SGM). Результаты экспериментов показали превосходство алгоритма ZNCC на большинстве выделенных классов объектов и на всей базе в целом. На оставшихся классах, в том числе на прозрачных объектах, лучшие результаты показал алгоритм Census, у алгоритма Census\_SGM не было выявлено явных преимуществ на данной базе (Таблица 1). Примеры результатов работы алгоритмов приведены на Рис. 7, к полученным HDR изображениям применен оператор тональной компрессии [10].

Класс объектов	ZNCC	Census	Census_SGM
Контрастные	<b>59,76</b>	58,61	58,70
Слабоконтрастные	<b>61,30</b>	61,25	61,01
Высокодетализированные	<b>60,99</b>	60,38	60,01
Низкодетализированные	61,48	<b>61,59</b>	61,33
Плоские	<b>62,34</b>	61,91	61,43
Выпуклые	59,36	<b>59,37</b>	59,27
Прозрачные	59,48	<b>59,58</b>	59,40
Непрозрачные	<b>61,60</b>	61,03	60,81
Все объекты	<b>60,79</b>	60,47	60,21

**Таблица 1:** Значения средней экспертной оценки (MOS), посчитанные по метрике HDR-VDP-2.2, для базы изображений со стереомикроскопа.



**Рис. 7:** Примеры работы алгоритмов на изображениях из собранной базы: а) Flowers\_fabric, б) Nut. Верхний ряд слева направо: входные изображения с левого и правого ракурсов, маска недоэкспонированных и переэкспонированных пикселей; нижний ряд слева направо: результаты алгоритмов ZNCC, Census и Census\_SGM и соответствующие им значения MOS.

## 8. ДАЛЬНЕЙШИЕ ПЛАНЫ

В дальнейшем планируется протестировать работу алгоритма поиска стереосоответствий [23], использующего сверточную нейронную сеть, на который опираются многие алгоритмы, лидирующие в рейтинге Middlebury ([12]). Затем предстоит проанализировать полученные результаты и сделать выбор либо в пользу алгоритма [23] и, возможно, дообучения

нейронной сети на фрагментах из собранной базы, либо в пользу одного из рассмотренных алгоритмов или их комбинации.

Поскольку съемка производится с помощью стереомикроскопа с фиксированными камерами, информация о положении камер и их оптических характеристиках может быть использована для более точного вычисления карты диспаратетов.

## 9. ЛИТЕРАТУРА

- [1] M. Bätz, T. Richter, J.-U. Garbas, A. Papst, J. Seiler, and A. Kaup. High dynamic range video reconstruction from a stereo camera setup. *Signal Processing: Image Communication*, 29(2): 191-202, 2014.
- [2] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 23(11): 1222-1239, 2001.
- [3] P. Debevec and J. Malik. Recovering high dynamic range radiance maps from photographs. In *SIGGRAPH*, pages 369-378, 1997.
- [4] J. A. Ferwerda. Elements of early vision for computer graphics. *IEEE Computer Graphics and Applications*, 21(5): 22-33, 2001.
- [5] M. D. Grossberg and S. K. Nayar. Determining the camera response from images: What is knowable? *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 25(11): 1455-1467, 2003.
- [6] H. Hirschmuller. Stereo processing by semiglobal matching and mutual information. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 30(2): 328-341, 2008.
- [7] H. Hirschmuller and D. Scharstein. Evaluation of stereo matching costs on images with radiometric differences. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 31(9): 1582-1599, 2009.
- [8] S. Lin, J. Gu, S. Yamazaki, and H. Y. Shum. Radiometric calibration from a single image. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 2, pages 938-945, 2004.
- [9] S. Mann and R. W. Picard. On being 'undigital' with digital cameras: extending dynamic range by combining differently exposed pictures. In *Proceedings of Information Systems & Technology (IS&T)*, pages 442-448, 1995.
- [10] R. Mantiuk, S. Daly, and L. Kerofsky. Display adaptive tone mapping. *ACM Transactions on Graphics (TOG)*, 27(3): 68, 2008.
- [11] X. Mei, X. Sun, M. Zhou, S. Jiao, H. Wang, and X. Zhang. On building an accurate stereo matching system on graphics hardware. In *IEEE International Conference on Computer Vision Workshops (ICCV Workshops)*, pages 467-474, 2011.
- [12] Middlebury Stereo Vision Page. *Online at <http://vision.middlebury.edu/stereo/eval3>*.
- [13] T. Mitsunaga and S. K. Nayar. Radiometric self calibration. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 374-380, 1999.
- [14] M. Narvaria, R. K. Mantiuk, M. P. Da Silva, and P. Le Callet. HDR-VDP-2.2: a calibrated method for objective quality prediction of high-dynamic range and standard images. *Journal of Electronic Imaging*, 24(1): 010501, 2015.
- [15] E. Reinhard, M. Stark, P. Shirley, and J. A. Ferwerda. Photographic tone reproduction for digital images. *ACM Transactions on Graphics (TOG)*, 21(3): 267-276, 2002.
- [16] E. Reinhard, G. Ward, S. Pattanaik, and P. Debevec. High dynamic range imaging: acquisition, display, and image-based lighting. Morgan Kaufmann, 2010.
- [17] M. A. Robertson, S. Borman, and R. L. Stevenson. Dynamic range improvement through multiple exposures. In *Proceedings of the 1999 International Conference on Image Processing (ICIP)*, volume 3, pages 159-163, 1999.
- [18] D. Scharstein and R. Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal of Computer Vision (IJCV)*, 47(1-3): 7-42, 2002.
- [19] N. Sun, H. Mansour, and R. Ward. HDR image construction from multi-exposed stereo LDR images. In *IEEE International Conference on Image Processing (ICIP)*, pages 2973-2976, 2010.
- [20] A. Troccoli, S. B. Kang, and S. Seitz. Multi-view multi-exposure stereo. In the *Third International Symposium on 3D Data Processing, Visualization, and Transmission*, pages 861-868, 2006.
- [21] Y. Tsing, V. Ramesh, and T. Kanade. Statistical calibration of CCD imaging process. In *Proceedings of the 8<sup>th</sup> IEEE International Conference on Computer Vision (ICCV)*, volume 1, pages 480-487, 2001.
- [22] R. Zabih and J. Woodfill. Non-parametric local transforms for computing visual correspondence. In *European Conference on Computer Vision (ECCV)*, pages 151-158, 1994.
- [23] J. Žbontar and Y. LeCun. Stereo matching by training a convolutional neural network to compare image patches. *Journal of Machine Learning Research (JMLR)*, 17: 1-32, 2016.

## Об авторах

Наталья Пащенко – студент первого курса магистратуры  
ВМК МГУ имени М. В. Ломоносова, prashchenko@yandex.ru.

Кристина Зипа – аспирант ВМК МГУ имени  
М. В. Ломоносова, kzipa@graphics.cs.msu.ru.

Алексей Игнатенко – к.ф.-м.н., с.н.с. лаборатории  
компьютерной графики и мультимедиа ВМК МГУ имени  
М. В. Ломоносова, ignatenko@graphics.cs.msu.ru.