

Оптический журнал

Иконика
Iconics

DOI: 10.17586/1023-5086-2023-90-08-44-54

УДК 004.932.4, 004.832.32

Метод повышения резкости совмещённых стереоснимков при наличии оптических дисторсий

Роман Олегович Малашин¹✉, Мария Анатольевна Михалькова²^{1, 2}Институт физиологии им. И.П. Павлова Российской академии наук, Санкт-Петербург, Россия¹Государственный университет аэрокосмического приборостроения, Санкт-Петербург, Россия¹malashinroman@infran.ru<https://orcid.org/0000-0002-2493-839X>²mikhalkova_maria@infran.ru<https://orcid.org/0000-0002-7186-5744>

Аннотация

Предмет исследования. Предметом исследования являются методы совмещения изображений, снятых с разных ракурсов при наличии дисторсий объективов. Цель работы — разработка способа снижения размытия изображения при переводе одного изображения в систему координат другого изображения. **Метод.** Рассмотрена задача получения единого преобразования для перевода изображения из системы координат исходных двух снимков. Получено выражение, представляющее собой суперпозицию нескольких геометрических преобразований, связанных с искажениями, вызванными дисторсиями двух объективов и индивидуальным сдвигом всех пикселов, определяемых методами стереозрения или оптического потока. **Результаты.** Особенностью задачи является то, что часть преобразований описывается аналитически (дисторсии), а часть — описывается матрицей сдвигов пикселов. Показано, что более удобными для использования в данной задаче являются преобразования, описываемые не оптическим потоком, а матрицей отображения. Выведены выражения, которые связывают системы координат двух снимков непосредственно. Предложенный подход инкапсуляции уравнений дисторсии в матрицу, описывающую оптический поток ректифицированных снимков, позволяет повысить качество изображения, переведённого в систему координат второго кадра. Это достигается за счёт отказа от последовательного перевода в несколько промежуточных систем координат, которые связаны с применением билинейной интерполяции и, следовательно, избыточным сглаживанием. Приведена экспериментальная проверка разработанного подхода, подтверждающая лучшие характеристики результирующего изображения по сравнению с традиционным подходом. **Практическая значимость.** Разработанный метод может быть применён в оптико-электронных системах с несколькими объективами в задачах, требующих, например, улучшения изображения широколучевой камеры с помощью изображения узкопольного объектива.

Ключевые слова: стереозрение, оптический поток, дисторсии объектива, отождествление пикселов двух снимков, билинейная интерполяция

Благодарность: работа выполнена при поддержке Госпрограммы 47 ГП «Научно-технологическое развитие Российской Федерации» (2019–2030), тема 0134-2019-0006.

Ссылка для цитирования: Малашин Р.О., Михалькова М.А. Метод повышения резкости совмещённых стереоснимков при наличии оптических дисторсий // Оптический журнал. 2023. Т. 90. № 8. С. 44–54. <http://doi.org/10.17586/1023-5086-2023-90-08-44-54>

Коды OCIS: 070.2025, 110.2960, 080.1753, 080.2720, 100.2980, 100.3010, 100.4994.

Method of sharpening of combined stereo images in presence of optical distortions

ROMAN MALASHIN¹✉, MARIA MIKHALKOVA²

^{1, 2}Pavlov Institute of Physiology, Russian Academy of Sciences, Saint Petersburg, Russia

¹Saint Petersburg State University of Aerospace Instrumentation, Saint Petersburg, Russia

¹malashinroman@infran.ru <https://orcid.org/0000-0002-2493-839X>

²mikhalkova_maria@infran.ru <https://orcid.org/0000-0002-7186-5744>

Abstract

Subject of the research. The methods of combining images taken from different angles in the presence of lens distortions are the subject of the research. **Purpose of the work.** Development of a way to beautifully blur an image when transferring one image into another image's coordinate system. **Method.** The problem of obtaining a single transformation for transferring of the image from the coordinate system of the original two images is considered. An expression representing a superposition of several geometric transformations associated with distortions caused by the distortions of two lenses and the individual shift of all pixels, determined by stereovision or optical flow methods is obtained. **Results.** A feature of the problem is that some of the transformations are described analytically (distortions), and some are described by a matrix of pixel shifts. It is shown that transformations described not by an optical flow, but by a mapping matrix, are more convenient for use in this problem. Expressions that connect the coordinate systems of two images directly are derived. The proposed approach of encapsulating the distortion equations into a matrix describing the optical flow of rectified images makes it possible to improve the quality of the image converted to the coordinate system of the second frame. This is achieved by means of avoiding the sequential transfer of the image into several intermediate coordinate systems, which are associated with the use of bilinear interpolation and, consequently, excessive smoothing. An experimental verification of the developed approach confirming the better characteristics of the resulting image compared to the traditional approach was carried out. **Practical significance.** The developed method can be applied in optoelectronic systems with multiple lenses in tasks that require, for example, improving the image of a wide-angle camera using the image of a narrow-field lens.

Keywords: stereo vision, optical flow, lens distortions, identification of pixels in two images, bilinear interpolation

Acknowledgment: the research leading to these results was supported by State Program 47 of the State Enterprise "Scientific and Technological Development of the Russian Federation" (2019–2030), topic 0134-2019-0006.

For citation: Malashin R.O., Mikhalkova M.A. The method of sharpening combined stereo images in the presence of optical distortions [In Russian] // Opticheskii Zhurnal. 2023. V. 90. № 8. P. 44–54. <http://doi.org/10.17586/1023-5086-2023-90-08-44-54>

OCIS codes: 070.2025, 110.2960, 080.1753, 080.2720, 100.2980, 100.3010, 100.4994.

ВВЕДЕНИЕ

Современные оптические системы часто оснащаются не одним, а сразу несколькими объективами. Это нужно для вычисления дальности до объектов сцен по дисперсиям пикселов двух (или более) изображений или достижения эффекта сверхразрешения. Часто подобные оптические системы оснащаются короткофокусными и длиннофокусными объективами. Это делается для возможности съёмки и общих планов, и портретов (в мо-

бильных телефонах) или для построения эффективных систем обнаружения и распознавания (когда предварительно цель обнаруживается с помощью широкоугольной, а потом распознается с помощью снимков узкопольной камеры). При наличии двух снимков, снятых с разных ракурсов, возможно их объединение в единственный более качественный (в частности, за счёт уточнения значений яркости по двум измерениям). Для решения этой задачи, как правило, требуется перевести

два кадра в единую систему координат, после чего их можно объединить. Для этого требуется скомпенсировать смещение пикселов на двух снимках, вызванное следующими факторами:

1. Разным взаимным положением двух камер.

2. Перемещением объектов в поле зрения двух камер за время, которое прошло между двумя снимками.

Необходимость перевода возникает, как в традиционных алгоритмах получения сверхразрешения или устранения шума за счёт итеративной оптимизации [1], так и в алгоритмах глубокого обучения, поскольку в них используется слой измерения и компенсации движения. При этом, если камеры синхронизированы по времени, смещения, вызванные движением объектов, нивелируются. В этом случае задачу сопоставления пикселов можно свести к задаче стереозрения: поиск соответствующих пикселов можно производить вдоль эпиполярных линий. Если пренебречь движением объектов нельзя, то задача сопоставления является более сложной и в общем случае отождествления нужно искать на всём кадре. Алгоритмы восстановления этих движений называются алгоритмами оптического потока.

На практике часто оказывается, что полноценно выровнять эпиполярные линии двух снимков не удается из-за сложностей калибровки камер, связанных, в частности, с тем, что аналитические уравнения дисторсий не могут идеально точно описать геометрические aberrации объективов. Восстановление более точных параметров дисторсий (например, локальных по площади изображений) не всегда возможно. Заметим, что в силу нелинейности оптических дисторсий после корректировки основной их составляющей, описываемых аналитически, дополнительных ограничений на область поиска, кроме ограничения на максимальный сдвиг, наложить нельзя. Для решения этой задачи используют методы оптического потока с ограничениями. Именно поэтому мы не ограничиваем совокупность методов, к которым применим предлагаемый подход (подразумевающий предварительную стереокалибровку камер) только методами поиска диспаратностей.

Большинство методов с использованием матрицы корреляционной энергии подразумева-

ет, что, если эпиполярное ограничение выполнено, то можно ограничить область смещения окна корреляции горизонтальным смещением. Это значит, что к изображениям предварительно применено корректирующее проективное преобразование.

Стоит отметить, что эпиполярная геометрия применима только к снимкам, когда устранены дисторсии объектива, поэтому предварительно требуется их устранить, преобразовав соответствующим образом изображения. Процедуру устранения дисторсий, а также (если необходимо) применение корректирующего проективного преобразования для выравнивания эпиполярных линий называют ректификацией.

Операцию ректификации применяют в классических методах решения задачи стереозрения [2–7]. Классическое стереосопоставление обычно включает извлечение локальных соответствий и полуглобальную регуляризацию [2]. Методы оптического потока [3] одновременно анализируют смещения, окклюзию и симметрию движения [4, 5] для решения менее ограниченных и сложных проблем двумерного сопоставления. Несмотря на явные различия между размерами пространства поиска, стереосопоставление и оптический поток имеют схожие допущения, такие как постоянство яркости и непрерывность сохранения границ [6, 7]. В современных работах [8–13] также применяют ректификацию для решения задачи стереосопоставления.

Для оценки оптического потока, начиная с Flownet [14], был предложен ряд сквозных моделей глубокого обучения [14–20]. В [20] авторы представляют новый метод решения неоднозначности зашумлённых оценок сопоставления — агрегация затрат с трансформаторами (CATs++). Предложенные методы превосходят предыдущие современные методы оценки оптического потока с большим отрывом.

Другим важным направлением исследований является метод оценки оптического потока, обученный без учителя [21–23]. В [24] авторы исследуют несколько способов применения эпиполярного ограничения при оценке потока. Все упомянутые выше методы так или иначе могут получить преимущество от работы с ректифицированными изображениями. В частности, в методах, использующих матри-

цу корреляционной энергии, её можно рассчитывать только с учётом допустимых сдвигов.

Любые пространственные преобразования изображения, связанные с субпиксельными смещениями, понижают разрешение, т.к. для вычисления яркости пикселя в новой системе координат требуется усреднение яркостей нескольких пикселов исходного изображения. Если стоит задача совместить два кадра в исходной системе координат, то в общем случае требуется произвести следующие преобразования:

1. Ректифицировать (устранить дисторсии) первое изображение.

2. Ректифицировать (устранить дисторсии) второе изображение.

3. Решить задачу стереосопоставления, восстановив карту смещений.

4. Перевести первое ректифицированное изображение в систему координат второго ректифицированного изображения.

5. Применить дисторсии второго изображения к изображению, полученному в п. 4.

Требуется значительное количество операций, каждая из которых будет снижать резкость полученного изображения.

Таким образом, целью данной работы было разработать способ снижения размытия изображения при переводе его в систему координат другого изображения. В результате был разработан метод, который позволяет заменить преобразования изображений преобразованиями над картой смещений, полученной в п. 3. В этом случае:

1. Преобразования будут снижать точность карты смещений, что, однако, не приведёт к потере разрешения изображений, поскольку изображение будет преобразовано лишь один раз из исходной системы координат первого изображения в систему координат второго изображения.

2. Часто дисторсии могут быть описаны аналитически. В этом случае преобразование векторов движений (поскольку они описываются вещественными числами) может быть выполнено более точно, чем преобразование изображений.

Чтобы не накладывать дополнительные ограничения, в данной работе все выражения выведены с использованием карты произвольных сдвигов, а не карты смещений вдоль эпиполярных линий (диспаратностей), поскольку последнее является частным случаем первого.

1. ОБЩИЕ ОБОЗНАЧЕНИЯ И ТЕРМИНЫ

1.1 Оптический поток

Пусть \mathbf{A} и \mathbf{B} — два изображения одинакового размера $W \times H$, которые определены на координатах пикселов $\Omega = \mathbb{N}^{W \times H} = \{\mathbf{x} \in \mathbb{N}^2\}$.

Пусть $\mathbf{OF}_{\mathbf{A} \rightarrow \mathbf{B}}$ — поле оптического потока, отображающее координаты каждого пикселя изображения \mathbf{A} в вектор смещения \mathbf{z}

$$\mathbf{OF}_{\mathbf{A} \rightarrow \mathbf{B}}(\mathbf{x}) = \mathbf{z}(\mathbf{x}), \mathbf{z}(\mathbf{x}) \in \mathbb{R}^2, \forall \mathbf{x} \in \Omega. \quad (1)$$

такое что

$$\mathbf{A}[\mathbf{x}] \approx \mathbf{B}[\mathbf{x} + \mathbf{OF}_{\mathbf{A} \rightarrow \mathbf{B}}(\mathbf{x})], \forall \mathbf{x} \in \Omega, \quad (2)$$

где символ \approx означает, что для 3D-точки бегутся наиболее близкие к ней пиксели изображений \mathbf{A} и \mathbf{B} . Заметим, что из-за окклюзии не каждый пикセル в \mathbf{B} имеет соответствие в \mathbf{A} , но мы опускаем этот случай, считая, что в области окклюзии оптическим потоком выбирается наиболее подходящее «ложное» соответствие.

1.2. Отображение

Оптический поток $\mathbf{OF}_{\mathbf{A} \rightarrow \mathbf{B}}$ может быть представлен в виде отображения $M_{\mathbf{A} \rightarrow \mathbf{B}}$

$$M_{\mathbf{A} \rightarrow \mathbf{B}}(\mathbf{x}) = \mathbf{x} + \mathbf{OF}_{\mathbf{A} \rightarrow \mathbf{B}}(\mathbf{x}), \forall \mathbf{x} \in \Omega. \quad (3)$$

В нижнем индексе M символ слева от стрелки соответствует изображению, система координат которого находится в области определения. Справа от стрелки находится изображение, система координат которого задаёт область значений. Отображение $M_{\mathbf{A} \rightarrow \mathbf{B}}(\mathbf{x})$ можно интуитивно описать так: нам нужно значение пикселя с координатой \mathbf{x} изображения \mathbf{A} , но у нас есть только изображение \mathbf{B} , тогда нам нужно преобразовать координату \mathbf{x} из системы координат \mathbf{A} в систему координат \mathbf{B} , используя отображение M . Таким образом, можем сказать, что

$$M_{\mathbf{A} \rightarrow \mathbf{B}} : X_{\mathbf{A}} \rightarrow X_{\mathbf{B}}, \quad (4)$$

где $X_{\mathbf{A}}$ обозначает пространство координат изображения \mathbf{A} , а $X_{\mathbf{B}}$ — пространство координат изображения \mathbf{B} .

По определению $\mathbf{OF}(\mathbf{x})$ должен указывать на один из пикселов изображения, поэтому

мы могли бы (ошибочно) заключить, что $M: \Omega \rightarrow \Omega$, что справедливо, если бы векторы оптического потока были целыми числами, но они действительные. Тем не менее, основное отличие состоит лишь в том, что при доступе к элементу изображения с вещественным индексом должна применяться интерполяция.

Использование отображения более удобно, чем использование оптического потока, потому что оно естественным образом выражает

$$\mathbf{A}[\mathbf{x}] \approx \mathbf{B}[M_{\mathbf{A} \rightarrow \mathbf{B}}(\mathbf{x})]. \quad (5)$$

Оно также удобнее для представления суммирования нескольких оптических потоков (см. ниже). Причина, по которой в программах часто используется оптический поток, заключается в том, что значение каждого элемента отображения содержит фиксированное значение $\mathbf{x} \in \Omega$, равное индексу пикселя по строке и столбцу в изображении. Это значение велико для некоторых пикселов в то время, как без наложения ограничений общности матожидание значений элементов оптического потока равно нулю. Это делает отображение гораздо менее компактным для хранения в памяти.

1.3. Преобразование систем координат

При использовании отображения $M_{\mathbf{B} \rightarrow \mathbf{A}}$ для каждой координаты $\mathbf{x} \in \Omega$ получается изображение $\mathbf{B}_\mathbf{A}$ — изображение \mathbf{B} , преобразованное в систему координат \mathbf{A}

$$\mathbf{B}_\mathbf{A}(\mathbf{x}) = \mathbf{B}_\mathbf{B}[M_{\mathbf{A} \rightarrow \mathbf{B}}(\mathbf{x})], \forall \mathbf{x} \in \Omega, \quad (6)$$

где нижние индексы в $\mathbf{B}_\mathbf{A}$ и $\mathbf{B}_\mathbf{B}$ определяют систему координат кадра. Изначально имеем изображение $\mathbf{B}_\mathbf{B}$.

Для компактности обозначим функцию преобразования системы координат через

$$\mathbf{B}_\mathbf{A} = \gamma(\mathbf{B}_\mathbf{B}, M_{\mathbf{A} \rightarrow \mathbf{B}}). \quad (7)$$

1.4. Интерполяция

Основная проблема в том, что изображение, оптический поток и отображение $M_{\mathbf{B} \rightarrow \mathbf{A}}$ определяются на \mathbb{N}^2 , но оптический поток вычисляется с субпиксельной точностью, поэтому область значений $M_{\mathbf{B} \rightarrow \mathbf{A}}$ определяется на \mathbb{R}^2 . В связи с этим, яркость пикселов изображения $\mathbf{B}_\mathbf{A}$ при использовании формулы (6) нужно

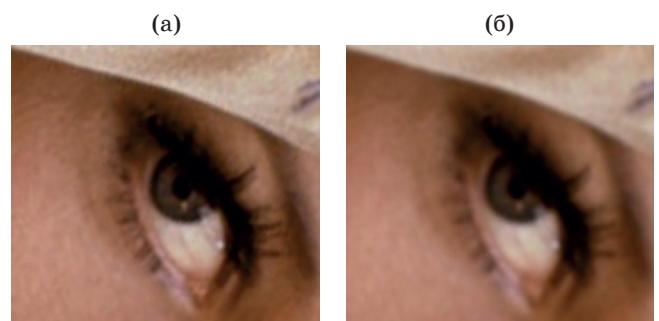


Рис. 1. Эффект размытия на изображении, повернутом с применением билинейной интерполяции. (а) Изображение, повернутое 1 раз на 60° , (б) изображение, повернутое 6 раз на 10°

Fig. 1. The blur effect on an image rotated using bilinear interpolation. (a) Image rotated 1 time by 60° , (b) image rotated 6 times by 10°

вычислять с использованием интерполяции. Наиболее популярно использование билинейной интерполяции, когда яркость пикселя будет определяться с учётом яркости и величины перекрытия четырёх соседних пикселов.

При каждом использовании преобразования с применением билинейной интерполяции изображение теряет в резкости, т.к. яркость каждого пикселя в конечном счёте является функцией от всего большего количества пикселов исходного изображения. Такой эффект виден на рис. 1, где изображение, повернутое 6 раз на 10° , более размытое, чем изображение, полученное при повороте сразу на 60° .

2. СУПЕРПОЗИЦИЯ ОТОБРАЖЕНИЙ

Пусть у нас есть третье изображение \mathbf{C} и отображение $M_{\mathbf{C} \rightarrow \mathbf{B}}$. Таким образом, мы имеем

$$\mathbf{B}_\mathbf{A}(\mathbf{x}) = \mathbf{B}_\mathbf{B}[M_{\mathbf{A} \rightarrow \mathbf{B}}(\mathbf{x})], \forall \mathbf{x} \in \Omega, \quad (8)$$

$$\mathbf{C}_\mathbf{B}(\mathbf{x}) = \mathbf{C}_\mathbf{C}[M_{\mathbf{B} \rightarrow \mathbf{C}}(\mathbf{x})], \forall \mathbf{x} \in \Omega. \quad (9)$$

Поскольку $\mathbf{C}_\mathbf{B}$ находится в системе координат \mathbf{B} , мы можем заменить им $\mathbf{B}_\mathbf{B}$ в (8), чтобы получить $\mathbf{C}_\mathbf{A}$

$$\begin{aligned} \mathbf{C}_\mathbf{A} &= \mathbf{C}_\mathbf{B}[M_{\mathbf{A} \rightarrow \mathbf{B}}(\mathbf{x})] = \\ &= \mathbf{C}_\mathbf{C}[M_{\mathbf{B} \rightarrow \mathbf{C}}[M_{\mathbf{A} \rightarrow \mathbf{B}}(\mathbf{x})]] = \\ &= \mathbf{C}_\mathbf{C}[M_{\mathbf{A} \rightarrow \mathbf{C}}(\mathbf{x})], \forall \mathbf{x} \in \Omega, \end{aligned} \quad (10)$$

где

$$M_{A \rightarrow C}(x) = M_{B \rightarrow C}[M_{A \rightarrow B}(x)]. \quad (11)$$

Используя сокращение из (7)

$$M_{A \rightarrow C} = \gamma(M_{B \rightarrow C}, M_{A \rightarrow B}). \quad (12)$$

Мы можем создать цепочку отображений любой длины, но, как уже говорилось выше, из-за интерполяции каждое преобразование координат уничтожает информацию сглаживающим эффектом билинейной интерполяции. Если используется метод ближайшего соседа, то следует усиление выраженности граничных артефактов.

Заметим, что операция (12) не является коммутативной:

$$M_{A \rightarrow C} \neq \gamma(M_{A \rightarrow B}, M_{B \rightarrow C}), \quad (13)$$

Можно представить (10) как суперпозицию отображений:

$$M_{A \rightarrow C} = M_{A \rightarrow B} \circ M_{B \rightarrow C}. \quad (14)$$

Используя полученные равенства (11), (12) и (14) в третьем разделе статьи мы сводим процедуру представления искомого отображения к манипулированию символами области определения и области значения функций (символами до и после стрелки).

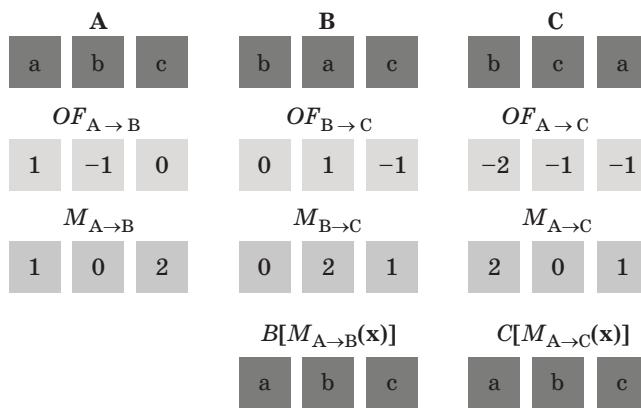


Рис.2. Иллюстрация преобразования изображения с помощью оптического потока. А, В, С — изображения, оптический поток рассчитывается только в направлении оси X

Fig. 2. An illustration of image conversion using optical flow. A, B, C — images, optical flow is calculated only in the direction of the X axis

На рис. 2 представлена иллюстрация применения формул к простому одномерному примеру.

$M_{A \rightarrow C}$ получена по формуле (11). При этом такой же компактной записи (аналога выражения (14)) для получения $OF_{A \rightarrow C}$ из $OF_{A \rightarrow B}$ и $OF_{B \rightarrow C}$ не существует, что показывает преимущество использования отображений перед использованием оптического потока для композиции оптических потоков. Также заметим, что, если поменять местами порядок применения отображений в (11), то получим отображение $M_0 = M_{A \rightarrow B}[M_{B \rightarrow C}] = (1, 2, 0)$, что приведёт к преобразованному изображению (c, a, b)), которое отличается от искомого изображения С. Это демонстрирует некоммутативность операции.

3. ПРЕОБРАЗОВАНИЕ СИСТЕМЫ КООРДИНАТ ДЛЯ РАСЧЁТА ОПТИЧЕСКОГО ПОТОКА С ИСПОЛЬЗОВАНИЕМ УРАВНЕНИЙ ОПТИЧЕСКИХ ДИСТОРСИЙ

Пусть в системе с двумя жёстко закреплёнными друг относительно друга камерами (или двумя объективами) восстановлены внешние и внутренние параметры камер. К внешним параметрам относятся положение и ориентация камер относительно друг друга, ко внутренним параметрам относятся фокусные расстояния, координата точки пересечения оптической оси с матрицей, а также параметры дисторсий.

Рекалификацией называется процесс выравнивания эпиполярных линий за счёт использования внутренних и внешних параметров камер: дисторсии корректируются, а с использованием внешних параметров рассчитываются матрицы проективного преобразования, которые применяются к двум изображениям. Иллюстрация приведена на рис. 3. Внутренние и внешние параметры могут быть восстановлены в процессе калибровки.

Рассмотрим четыре изображения и три отображения:

1. $S1$ — первое исходное изображение, $S2$ — второе исходное изображение. $R1, R2$ — рекалифицированные изображения ($S1$ и $S2$ после рекалификации).

2. $M_{R1 \rightarrow R2}(x) = x + OF_{R1 \rightarrow R2}$, $x \in \Omega$, где $OF_{R1 \rightarrow R2}$ — оптический поток, вычисляемый по кадрам $R1$ и $R2$.

3. $M_{R2 \rightarrow S2}$ — отображение, соответствующее процедуре рекалификации первого кадра.

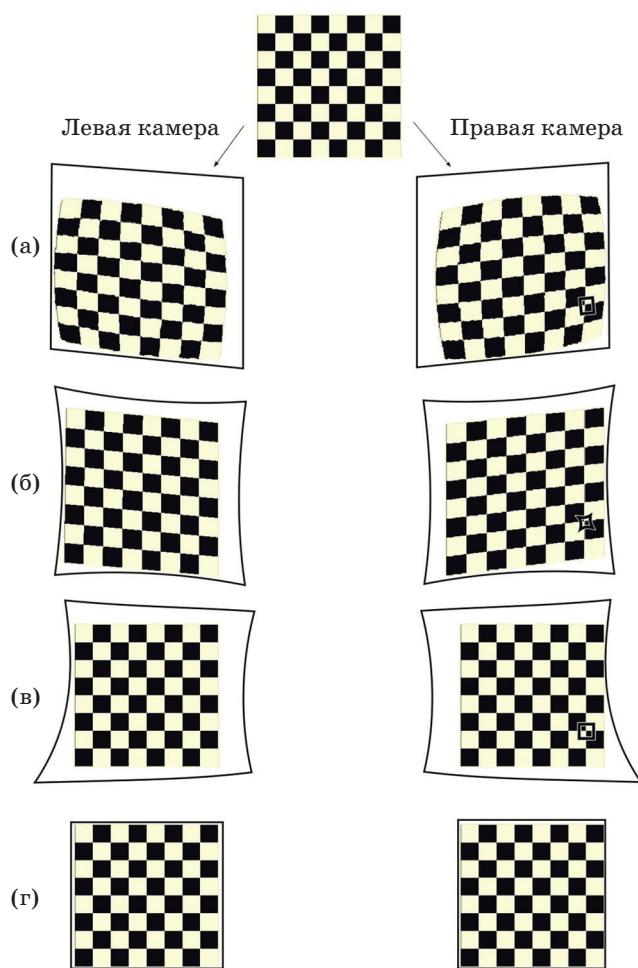


Рис. 3. Стерео-ректификация на примере шахматной доски. (а) Исходные изображения, (б) устранение дисторсии, (в) проективное преобразование для выравнивания эпиполярных линий, (г) результат после обрезания нужного фрагмента

Fig. 3. Stereo-rectification using the example of a chessboard. (a) Original images, (b) elimination of distortion, (c) projective transformation to align epipolar lines, (d) result after cropping the desired fragment

4. $M_{R1 \rightarrow S1}$ — отображение, соответствующее процедуре ректификации второго кадра.

Используя эти отображения, мы хотим вычислить $M_{S1 \rightarrow S2}$, чтобы с помощью него перевести $S2$ в систему координат $S1$.

3.1. Решение

Используя (12), мы можем расширить уравнение, зная окончательную область определения и желаемое отображение:

$$M_{S1 \rightarrow S2}(x) = M_{? \rightarrow S2}[M_{S1 \rightarrow ?}(x)], \quad (15)$$

где символ ‘?’ означает, что должна быть какая-то другая система координат, связывающая две исходные системы координат. У нас есть $M_{R2 \rightarrow S2}$ и мы можем использовать его следующим образом:

$$M_{S1 \rightarrow S2}(x) = M_{R2 \rightarrow S2}[M_{S1 \rightarrow R2}(x)]. \quad (16)$$

Повторим процедуру, заменив $M_{S1 \rightarrow R2}$,

$$\begin{aligned} M_{S1 \rightarrow R1}(x) &= \\ &= M_{R2 \rightarrow S2}[M_{R1 \rightarrow R2}[M_{S1 \rightarrow R1}(x)]]. \end{aligned} \quad (17)$$

У нас нет $M_{S1 \rightarrow R1}$, поскольку восстанавливается преобразование, корректирующее оптические дисторсии, а не обратное к нему. Тем не менее, мы можем получить его из $M_{R1 \rightarrow S1}$ аналитически или приближенно (в зависимости от того, как задаются дисторсии):

$$M_{S1 \rightarrow R1} \approx M_{R1 \rightarrow S1}^{-1}. \quad (18)$$

Тогда окончательное уравнение будет

$$\begin{aligned} M_{S1 \rightarrow R1}(x) &= \\ &= M_{R2 \rightarrow S2}[M_{R1 \rightarrow R2}[M_{R1 \rightarrow S1}^{-1}(x)]]. \end{aligned} \quad (19)$$

Мы можем использовать сокращённую запись (7)

$$\begin{aligned} M_{S1 \rightarrow S2} &= \\ &= \gamma(M_{R2 \rightarrow S2}, \gamma(M_{R1 \rightarrow R2}, M_{R1 \rightarrow S1}^{-1})). \end{aligned} \quad (20)$$

Таким образом, мы получили искомое отображение.

Можно показать, что для уравнения (20) выполняется ассоциативное свойство — преобразования можно производить в другом порядке.

Полученное преобразование $M_{R1 \rightarrow S1}$ применимо непосредственно к исходному изображению, что позволяет избежать промежуточных ректифицированных изображений, которые снижают резкость изображений (рис. 1).

При этом, однако, заметим, что $M_{S1 \rightarrow S2}$ также, как и изображение, описывается матрицей (определяется на \mathbb{R}^2), а, следовательно, для вычисления по формуле (20) также потребуется использование интерполяции. Однако в этом случае:

1. В отличие от изображений значения отображения определяются на \mathbb{R}^2 , а, значит, при

преобразовании может быть сохранено больше информации.

2. Можно использовать аналитические формулы преобразований дисторсий. Например, одним из популярных способов является описание в виде комбинации радиальных и тангенциальных дисторсий [25].

За счёт этого обеспечивается повышение резкости полученного изображения.

4. ЭКСПЕРИМЕНТАЛЬНАЯ ПРОВЕРКА

Снижение размытия изображения при использовании предлагаемого подхода может быть визуально заметно, если алгоритм оптического потока работает достаточно точно. В связи с этим, для демонстрации преимуществ предлагаемого подхода мы используем синтетические данные, что позволяет нам абстрагироваться от ошибок оптического потока, низкой точности процедуры ректификации и калибровки.

Для простоты будем полагать, что

- стереокамера является идеально калиброванной по положению и ориентации камер (не требуется выполнения проективного преобразования одного из кадров);
- дисторсии объектива 1 ограничиваются радиальными, на объективе R2 дисторсии отсутствуют ($M_{S2 \rightarrow R2}$ является тождественным отображением), значит $R2 = S2$;
- поле движения является равномерным по всему кадру.

Заметим, что ослабление данных ограничений требует использования более сложных преобразований $M_{R2 \rightarrow S2}$, $M_{R1 \rightarrow R2}$ и $M_{R1 \rightarrow S1}$, но не меняет необходимости делать интерполяцию даже большее количество раз, поэтому при полном восстановлении изображения степень размытия традиционного метода будет даже усиlena.

Для синтеза искажённых дисторсиями изображений с помощью метода обратного проецирования нам требуется знать формулы преобразования искажённых координат в восстановленные. Кроме того, согласно выражению (20) для применения разработанного подхода необходимо знать функцию для «декретификации» изображения $M_{R1 \rightarrow S1}^{-1}$ и предпочтительно, чтобы эта функция была аналитической, поэтому предварительно рассмотрим частный случай получения обратного преобразования.

4.1. Инвертирование радиальных дисторсий
Двумя основными видами искажений являются радиальные и тангенциальные искажения. Для их моделирования наиболее часто используется следующая система уравнений для нахождения искажённых точек изображения:

$$\begin{aligned} x_d &= x_c \left(1 + k_1 r^2 + k_2 r^4 \right) + \\ &+ 2p_1 x_c y_1 + p_2 \left(r^2 + 2x_c^2 \right), \\ y_d &= y_c \left(1 + k_1 r^2 + k_2 r^4 \right) + \\ &+ 2p_2 x_c y_1 + p_1 \left(r^2 + 2y_c^2 \right), \end{aligned} \quad (21)$$

где x_d , y_d — координаты положения точки на искажённом изображении, x_c , y_c — координаты для неискажённой точки, k_1 , k_2 — коэффициенты радиальной дисторсии, p_1 , p_2 — коэффициенты тангенциальной дисторсии, r — Евклидово расстояние между точкой искажённого изображения и центром искажения.

Из этой системы уравнений можно выразить коэффициенты x_c , y_c (неискаженного местоположения пикселя) для получения модели обратного преобразования. Для простоты возьмем случай, когда все коэффициенты искажения равны нулю, кроме k_1 , который равен 0,2 (то есть на изображении будет присутствовать только радиальные дисторсии).

Для вывода формул мы воспользовались возможностями символьного вывода Matlab и получили следующие выражения:

$$\begin{aligned} \sigma_1 &= \frac{5y_d^3}{2(x_d^2 + y_d^2)} + \frac{1}{3}\sqrt{\frac{25y_d^6}{4(x_d^2 + y_d^2)^2} + \frac{125y_d^6}{27(x_d^2 + y_d^2)^3}}, \\ \sigma_2 &= \sigma_1 - \frac{5y_d^2}{3(x_d^2 + y_d^2)\sigma_1}, \\ x_c &= \frac{x_d\sigma_2}{y_d}, \\ y_c &= \sigma_2. \end{aligned} \quad (22)$$

Для проверки корректности выведенной формулы на изображение шахматной доски были наложены, а затем скомпенсированы радиальные дисторсии с заданными коэффициентами (0,2, 0, 0, 0).

На рис. 4 представлены изображения с радиальными дисторсиями и изображение с уст-

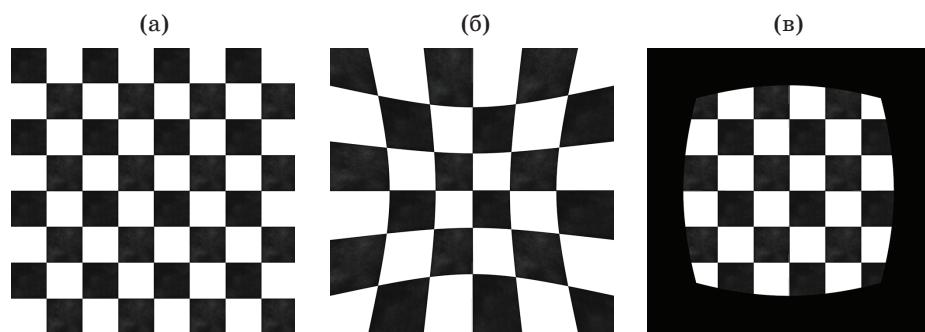


Рис. 4. Использование выведенной формулы для проверки её корректности. (а) Исходное изображение, (б) полученное с помощью формул (22) искажённое изображение, (в) изображение с скомпенсированными искажениями, полученное с помощью формулы (21)

Fig. 4. Using the derived formula to check its correctness. (a) Original image, (b) distorted image obtained using the formulas (22), (v) image with compensated distortions obtained using the formula (21)

ранёнными искажениями с использованием этой формулы и метода обратного проецирования.

Как видно из рис. 4а и 4в, скомпенсированное изображение получилось достаточно точным, что подтверждает корректность выведенных формул.

4.2. Генерация данных

Для демонстрации положительного эффекта от применения предложенного метода были сгенерированы искусственные данные. Для того, чтобы они содержали эталон, с которым можно было бы сравнивать результаты работы двух методов, генерация данных выполнялась в высоком разрешении, а сравниваемые методы работали с пониженным разрешением. Для получения исходных данных $S1$ и $S2$ были проделаны следующие шаги:

1. Изображение шахматной доски размером 3600×3600 (без дисторсий) было принято за исходное изображение высокого разрешения $S2^H$. При этом $S2^H = R2^H$, т.к. мы исходим из того, что на втором изображении дисторсий нет.

2. Оптический поток $\mathbf{OF}_{R1 \rightarrow R2}^H$ был принят за равномерное поле сдвигов (4 пикселя по каждой из осей). С помощью сдвига каждого пикселя изображения $R2^H$ на 4 пикселя было получено изображение $R1^H$.

3. Изображение $S1^H$ было получено наложением дисторсий на $R1^H$.

4. Исходные данные для работы алгоритмов $S1$, $S2$, $\mathbf{OF}_{R1 \rightarrow R2}$ были получены уменьшением изображений $S1^H$, $S2^H$ и поля движений $\mathbf{OF}_{R1 \rightarrow R2}^H$ в 8 раз.

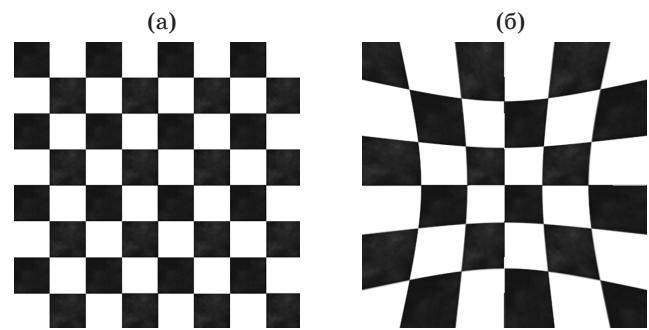


Рис. 5. Примеры сгенерированных данных. (а) Изображение $S2$, (б) изображение $S1$

Fig. 5. Examples of generated data. (a) Image $S2$, (b) image $S1$

Такая процедура позволяет сгенерировать данные, которые получаются при съёмке на стереокамеру (с ограничениями, описанными выше) при, например, съёмке плоского объекта, располагающегося перпендикулярно оптической оси при наличии смещения камеры по двум направлениям одновременно. Примеры сгенерированных эталонных изображений представлены на рис. 5.

4.3. Сравнение предложенного и основного метода

При использовании традиционного метода, чтобы получить первый кадр (изображение $S1$), нужно с помощью оптического потока получить из $S2$ изображение $R1$, а затем восстановить искомый кадр $S2$ наложением дисторсий. При этом, при каждом переводе в несколько промежуточных систем координат

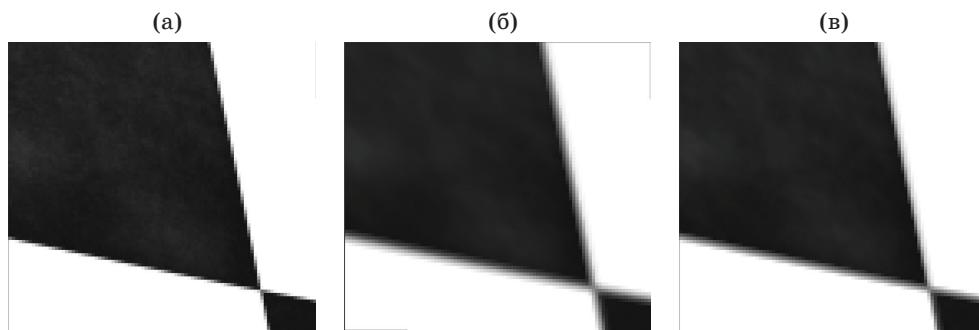


Рис. 6. Левый верхний угол полученного изображения S_1 . (а) Эталонное изображение, (б) полученное изображение с помощью исходного подхода, (в) полученное изображение с помощью предложенного подхода

Fig. 6. The upper left corner of the resulting image S_1 . (a) Reference image, (b) image obtained using the original approach, (v) image obtained using the proposed approach

применяется билинейная интерполяция, тем самым изображение сглаживается 2 раза (в случае наличия дисторсий на изображении S_2 это происходило бы три раза) и становится размытым. Предлагаемое преобразование позволяет избежать создания промежуточных изображений, которые снижают итоговую резкость. Сравнение полученных изображений S_1 с помощью двух подходов представлено на рис. 6.

Из рис. 6 видно, что изображение, полученное с помощью предложенного подхода, более резкое, чем при применении традиционного метода.

ЗАКЛЮЧЕНИЕ

В данной работе была рассмотрена задача получения единого преобразования для совмещения снимков, снятых с разных ракурсов при наличии дисторсии объективов. Для этого получено выражение, представляющее собой суперпозицию нескольких геометрических преобразований, связанных с искажениями, вызванными дисторсиями двух объективов, а также матрицы отождествлений пикселов, полученной по ректифицированным снимкам. Показано, что более удобными для использования в данной задаче являются пре-

образования, описываемые не оптическим потоком, а матрицей отображения. Выведено выражение (20), которое связывает системы координат двух снимков непосредственно.

Проведена экспериментальная проверка разработанного подхода, которая подтверждает визуальное превосходство изображений, полученных с помощью предложенного подхода, по сравнению с изображениями, полученными традиционным методом.

Особенностью предложенного подхода является возможность «встраивания» уравнений дисторсии в матрицу, описывающую оптический поток ректифицированных снимков, что позволяет повысить качество изображения после перевода в систему координат второго кадра. Это достигается за счёт отказа от последовательного перевода в несколько промежуточных систем координат, которые связаны с применением билинейной интерполяции и, следовательно, сглаживанием изображения.

Разработанный метод может быть применён в оптико-электронных системах с несколькими объективами в задачах, требующих, например, улучшения изображения широколучальной камеры с помощью изображения узкопольного объектива.

СПИСОК ИСТОЧНИКОВ

- Chambolle A. An algorithm for total variation minimization and applications // Journal of Mathematical Imaging and Vision. 2004. V. 20. P. 89–97. <https://doi.org/10.1023/B:JMIV.0000011325.36760.1e>
- Hirschmuller H., Scharstein D. Evaluation of stereo matching costs on images with radiometric differen-
ces // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2009. V. 31. № 9. P. 1582–1599. <https://doi.org/10.1109/TPAMI.2008.221>
- Li S.Z. Markov random field models in computer vision // Proceedings of the Third European Conference on Computer Vision. 1994. V. 11. P. 361–370. <https://doi.org/10.1007/BFb0028368>

4. Hur J., Roth S. Mirrorflow: Exploiting symmetries in joint optical flow and occlusion estimation // 2017 IEEE International Conference on Computer Vision (ICCV). 2017. P. 312–321. <https://doi.org/10.1109/ICCV.2017.42>
5. Yamaguchi K., McAllester D., Urtasun R. Efficient joint segmentation, occlusion labeling, stereo and flow estimation // European Conference on Computer Vision. 2014. P. 756–771. https://doi.org/10.1007/978-3-319-10602-1_49
6. Revaud J., Weinzaepfel P., Harchaoui Z. et al. Epicflow: Edge-preserving interpolation of correspondences for optical flow // 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2015. P. 1164–1172. <https://doi.org/10.1109/CVPR.2015.7298720>
7. Sun D., Yang X., Liu M.-Y. et al. Pwc-net: Cnns for optical flow using pyramid, warping, and cost volume // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2018. P. 8934–8943. <https://doi.org/10.1109/CVPR.2018.00931>
8. Darmon F., Monasse P. The polar epipolar rectification // Image Processing On Line. 2021. P. 56–75. <https://doi.org/10.5201/ipol.2021.328>
9. Raad L., Oliver M., Ballester C. et al. On anisotropic optical flow inpainting algorithms // Image Processing On Line. 2020. P. 78–104. <https://doi.org/10.5201/ipol.2020.281>
10. Dagobert T., Monzón N., Sánchez J. Comparison of optical flow methods under stereomatching with short baselines // Image Processing On Line. 2019. P. 329–359. <https://doi.org/10.5201/ipol.2019.217>
11. Rodríguez M., Facciolo G., Morel J.-M. Robust homography estimation from local affine maps // Image Processing On Line. 2023. P. 65–89. <https://doi.org/10.5201/ipol.2023.356>
12. Garamendi J.F., Lazcano V., Ballester C. Joint TV-L1 optical flow and occlusion estimation // Image Processing On Line. 2019. P. 432–452. <https://doi.org/10.5201/ipol.2019.118>
13. Dagobert T., Grompone von Gioi R., C. de Franchis et al. Cloud detection by luminance and inter-band parallax analysis for pushbroom satellite imagers // Image Processing On Line. 2020. P. 167–190. <https://doi.org/10.5201/ipol.2020.271>
14. Dosovitskiy A., Fischer P., Ilg E., Hausser P. et al. Flownet: Learning optical flow with convolutional networks // 2015 IEEE International Conference on Computer Vision (ICCV). 2015. P. 2758–2766. <https://doi.org/10.1109/ICCV.2015.316>
15. Ren Z., Yan J., Ni B. et al. Unsupervised deep learning for optical flow estimation // Proceedings of the AAAI Conference on Artificial Intelligence. 2017. P. 1495–1501. V. 31. № 1. <https://doi.org/10.1609/aaai.v31i1.10723>
16. Ilg E., Mayer N., Saikia T. et al. Flownet 2.0: Evolution of optical flow estimation with deep networks // in Proceedings of the IEEE conference on computer vision and pattern recognition. 2017. P. 2462–2470. <https://doi.org/10.1109/CVPR.2017.179>
17. Ranjan A., Black M.J. Optical flow estimation using a spatial pyramid network // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2017. P. 2720–2729. <https://doi.org/10.1109/CVPR.2017.291>
18. Hui T.-W., Tang X., Loy C.C. LiteFlowNet: A light-weight convolutional neural network for optical flow estimation // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2018. P. 8981–8989. <https://doi.org/10.1109/CVPR.2018.00936>
19. Hur J., Roth S. Iterative residual refinement for joint optical flow and occlusion estimation // 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2019. P. 5747–5756. <https://doi.org/10.1109/CVPR.2019.00590>
20. Seokju C., Sungwan H., Seungryong K. CATs++: Boosting cost aggregation with convolutions and transformers // arXiv preprint arXiv:2202.06817. 2022
21. Luo K., Wang C., Liu S. et al. Upflow: Upsampling pyramid for unsupervised optical flow learning // The IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2021. P. 1045–1054. <https://doi.org/10.1109/CVPR46437.2021.00110>
22. Meister S., Hur J., Roth S. UnFlow: Unsupervised learning of optical flow with a bidirectional census loss // Thirty-Second AAAI Conference on Artificial Intelligence. 2018. V. 32. № 1. P. 7251–7259. <https://doi.org/10.1609/aaai.v32i1.12276>
23. Jonschkowski R., Stone A., Barron J.T. et al. What matters in unsupervised optical flow // The European Conference on Computer Vision (ECCV). 2020. V. 12347. https://doi.org/10.1007/978-3-030-58536-5_33
24. Zhong Y., Ji P., Wang J. et al. Unsupervised deep epipolar flow for stationary or dynamic scenes // 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2019. P. 12087–12096. <https://doi.org/10.1109/CVPR.2019.01237>
25. Hartley R., Zisserman A. Multiple view geometry in computer vision. N.Y.: Cambridge University Press, 2003. 607 p.

АВТОРЫ

Роман Олегович Малашин — канд. техн. наук, старший научный сотрудник, Институт физиологии им. И.П. Павлова РАН, 199034, Санкт-Петербург, Россия; доцент, Государственный университет аэрокосмического приборостроения, 190000, Санкт-Петербург, Россия; Scopus ID: 55789083900, <https://orcid.org/0000-0002-2493-839X>, malashinroman@mail.ru

Мария Анатольевна Михалькова — магистр, младший научный сотрудник, Институт физиологии им. И.П. Павлова РАН, 199034, Санкт-Петербург, Россия; <https://orcid.org/0000-0002-7186-5744>, marusia_mi@mail.ru

Статья поступила в редакцию 30.09.2022
Одобрена после рецензирования 17.05.2023
Принята к печати 26.06.2023

AUTHORS

Roman O. Malashin — PhD, Senior Researcher, Pavlov Institute of Physiology RAS, 199034, Saint-Petersburg, Russia; Associate Professor, Saint-Petersburg State University of Aerospace Instrumentation, 190000, Saint-Petersburg, Russia; Scopus ID: 55789083900, <https://orcid.org/0000-0002-2493-839X>, malashinroman@mail.ru

Maria A. Mikhalkova — MS, Junior Researcher, Pavlov Institute of Physiology RAS, 199034, Saint-Petersburg, Russia; <https://orcid.org/0000-0002-7186-5744>, marusia_mi@mail.ru

The article was submitted to the editorial office 30.09.2022
Approved after review 17.05.2023
Accepted for publication 26.06.2023