

С.И. Протасов, аспирант,  
Воронежский государственный университет  
А.А. Крыловецкий, канд. физ.-мат. наук, доц.,  
Воронежский государственный университет  
С.Д. Кургалин, д-р физ.-мат. наук, доц.,  
Воронежский государственный университет

## **Об одном подходе к решению задачи ректификации стереоизображений по сцене без калибровки камер**

В статье рассматривается метод предобработки изображений в системах стереозрения, основанный на модификации алгоритма стабилизации видео [1]. Метод описывает ректификацию изображений как набор последовательных преобразований, каждое из которых находится как решение оптимизационной задачи. В статье описана математическая модель, соответствующая основным положениям подхода к стабилизации видео, приводятся оптимизационные методы нахождения параметров преобразований и описываются статистические алгоритмы уточнения решения. Предлагаемый подход не требует предварительной калибровки камер и использует только изображения сцены. Практической областью применения данного решения являются персональные системы стереозрения.

*стереозрение, машинное зрение, стереоизображения, калибровка, камеры, стабилизация видео.*

Ректификация изображений (image rectification) является одной из важных задач при создании системы стереозрения. Под ректификацией обычно понимают приведение набора изображений к общей координатной системе так, чтобы горизонтальные линии на изображениях соответствовали

одной плоскости. Как правило, данной задаче уделяется мало внимания из-за того, что она решается однократно для конкретной системы камер. Тем не менее, системы стереокамер становятся все более востребованными в различных областях [2]. В персональных системах стереозрения источником изображений могут быть камеры, установленные в мобильных телефонах, ноутбуках и др. В данной статье мы хотим предложить методику ректификации изображений без необходимости калибровки и взаимной калибровки камер. Предлагаемый подход позволит конечному пользователю получить ректифицированный стерео-поток, например, подключив дополнительную веб-камеру к ноутбуку.

### ***Постановка задачи***

Предлагаемая в статье модель решения является адаптацией задачи стабилизации видеопотока к ректификации изображений. Основой задачи стабилизации видеопотока является определение параметров перемещения камеры между последовательными кадрами или эквивалентная ей задача определения перемещения изображения. Под перемещением изображения понимается набор необходимых трансформаций для превращения кадра в последующий. Как правило, рассматривается проекционная модель камеры или, в случае пренебрежения дисторсиями, вносимыми оптикой, модель камеры-обскуры. В общей постановке задачи учитываются все шесть степеней свободы камеры: перемещения вдоль трёх осей декартовых координат, а также повороты вокруг этих осей [3]. Нахождение всех шести параметров перемещения является нетривиальной задачей, поэтому в модели применяются упрощения. Производится ограничение вариаций углов до малых значений (это позволяет применять следствия первого замечательного предела и приводить задачу к линейной), а также исключаются некоторые степени свободы. В данной статье рассматривается решение аналогичной задачи с учётом специфики стереозрения. В предлагаемой нами модели плоскости изображения

принимаются параллельными, то есть оптические оси камер считаются коллинеарными. Такое упрощение правомерно в рамках практических применений: например, закрепление дополнительной камеры на крышке ноутбука. Следствием этого упрощения является исключение из рассмотрения двух поворотных степеней свободы. Ещё одно допущение делается в отношении угла поворота камеры вокруг оптической оси: будем считать его малым.

### *Декомпозиция и анализ задачи*

Рассмотрим математическое описание процесса трансформации изображения одной камеры в изображение другой с учётом принятых ограничений. В рамках модели четыре степени свободы можно трактовать как три преобразования вида  $\mathbb{R}^2 \Rightarrow \mathbb{R}^2$  над точкой изображения  $p \in \mathbb{R}^2$ . Перемещение  $T(p)$  вдоль осей, лежащих в плоскости изображения, можно описать единым вектором перемещения  $\vec{v}_0 = (d_x, d_y)^T$ :

$$T(p) = p + (d_x, d_y)^T.$$

Вращение  $R(p)$  камеры вокруг оптической оси эквивалентно повороту изображения на определённый угол  $\Theta$  вокруг точки  $p_c$ :

$$R(p) = p_c + \begin{pmatrix} \cos \Theta & \sin \Theta \\ -\sin \Theta & \cos \Theta \end{pmatrix} \cdot (p - p_c).$$

Движение вдоль оптической оси  $Z(p)$  можно описать через деформацию растяжения относительно точки  $p_c$ , учитывая при этом возможное изменение соотношения сторон:

$$Z(p) = p_c + \begin{pmatrix} kx & 0 \\ 0 & ky \end{pmatrix} \cdot (p - p_c).$$

Результатом решения задачи ректификации является набор величин  $p_c$ ,  $\vec{v}_0$ ,  $\Theta$ ,  $kx$ ,  $ky$ . Все три преобразования являются линейными по отношению к точке, поэтому с точностью до коэффициентов могут

производиться к любому порядку. В предлагаемом нами решении поиск параметров производится в последовательности  $T \rightarrow R \rightarrow Z$ .

### ***Решение задачи***

Решение задачи состоит из четырёх основных этапов, базируется на операциях с векторным полем перемещения точек и является задачей оптимизации. Линейность каждого из преобразований позволила нам разделить решение задачи на независимые этапы, каждый из которых также является задачей оптимизации. На первом этапе производится поиск и фильтрация точечных особенностей на изображении первой камеры. На втором этапе строится векторное поле перемещения точек и находятся параметры перемещения  $T(p)$ . На третьем этапе вычисляются параметры трансформации  $R(p)$  с учётом предварительной оценки координат точки  $p_c$ . На последнем шаге вычисляются параметры преобразования  $Z(p)$  и находятся поправки для координат точки  $p_c$ .

### ***Поиск точечных особенностей***

В задаче стабилизации видеопотока применяется расчёт векторного поля перемещения точек изображения. Опорные точки поля находятся в узлах регулярной сетки. Для уменьшения вероятности некорректного определения перемещений мы предлагаем использовать в качестве опорных точки с максимальной дисперсией окрестности. Под дисперсией окрестности понимается дисперсия яркости пикселей в малой квадратной симметричной окрестности рассматриваемой точки. Для того чтобы опорные точки были распределены относительно равномерно, применяем фильтрацию через регулярную сетку. Таким образом, на каждую ячейку регулярной сетки будет приходиться не более одной точечной особенности, как и в оригинальном методе, но каждая из них будет обладать локально максимальной дисперсией. Ячейки, внутри которых дисперсия окрестности точек не превышает заданный порог, будут «пустыми». Кроме того, для снижения вероятности ложных срабатываний

поиск точечных особенностей производится с заданным отступом от края изображения, так как края изображения с первой камеры с высокой вероятностью будут отсутствовать на изображении со второй камеры и наоборот.

### ***Нахождение векторного поля перемещения***

Для каждой точки на изображении с первой камеры, отобранной на предыдущем этапе, найдём соответствие на изображении со второй камеры. В общем случае камеры не являются идентичными, имеют различную оптику и чувствительностью матрицы, а значит изображения будут иметь различную яркость и цветовой баланс. Поэтому до поиска соответствий проводим цветовую калибровку изображений [4]. Поскольку в нашей модели заложена возможность поворота изображений на малый угол, то нам необходимо учитывать влияние поворота на функцию соответствия. Поэтому в качестве критерия подобия мы используем взвешенное средне-квадратичные отклонение (weighted MSE) [1]. Весовой коэффициент мы положили равным  $(1 + \sqrt{|\Delta x| + |\Delta y|})^{-1}$ . В результате для каждой изначально отобранной точки получили индивидуальный вектор её перемещения  $T_p$  (рис. 1) и значение минимума  $wMSE$ , являющееся критерием достоверности результата. Для ускорения работы алгоритма мы применили одну из модификаций алгоритма поиска соответствий с адаптивным окном [5].

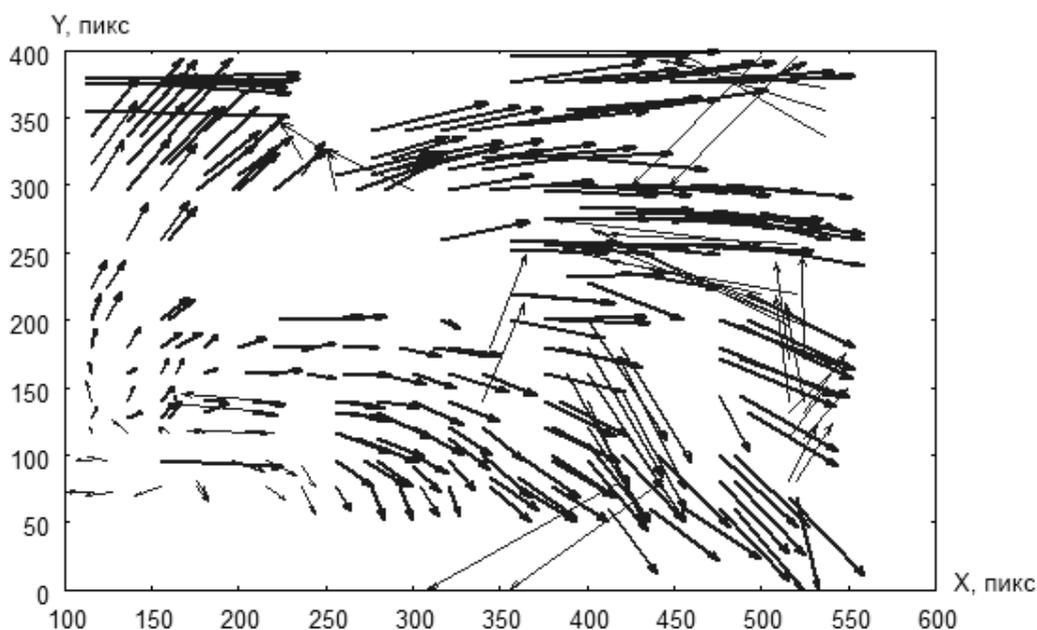


Рис. 1. Векторное поле перемещения точек с указанием отфильтрованных векторов.

Полученное поле по-прежнему содержит в себе векторы, соответствующие ложным срабатываниям. Корректное определение векторного поля перемещений точек изображения является существенным для точности работы всех дальнейших этапов. Именно поэтому на этом этапе необходимо обеспечить минимизацию количества ошибочно найденных векторов. Нами предложен способ фильтрации векторов: изображение разбивается на достаточно крупные блоки, в рамках которых, тем не менее, угол отклонения каждого вектора перемещения от среднего вектора по блоку невелик. Задав порог максимального отклонения от среднего, мы можем избавиться от заведомо некорректных векторов. В нашем эксперименте мы использовали шесть блоков (3x2) и допуск угла отклонения  $\pm \frac{\pi}{6}$ . Величина угла отклонения получена в рамках проведённых численных экспериментов. На рис. 1 можно видеть векторы, которые были отсеяны в результате фильтрации (они изображены тонкими линиями).

Для сцен, все точки которых являются равноудалёнными от камер, средний вектор перемещения  $T_{avg}$  позволяет произвести предварительную

оценку точки  $p'_c$  перемещением центра изображения  $p_{im} = \left( \frac{width}{2}, \frac{height}{2} \right)^T$  с первой камеры:

$$p'_c = p_{im} + T_{avg} \cdot$$

### **Нахождение угла поворота**

Найденная оценка  $p'_c$  позволяет гарантировать, что после прибавления ко всем векторам  $\vec{v}_p$  вектора  $-T_{avg}$  искомая точка  $p_c$  будет находиться в пределах видимого изображения (рис. 2).

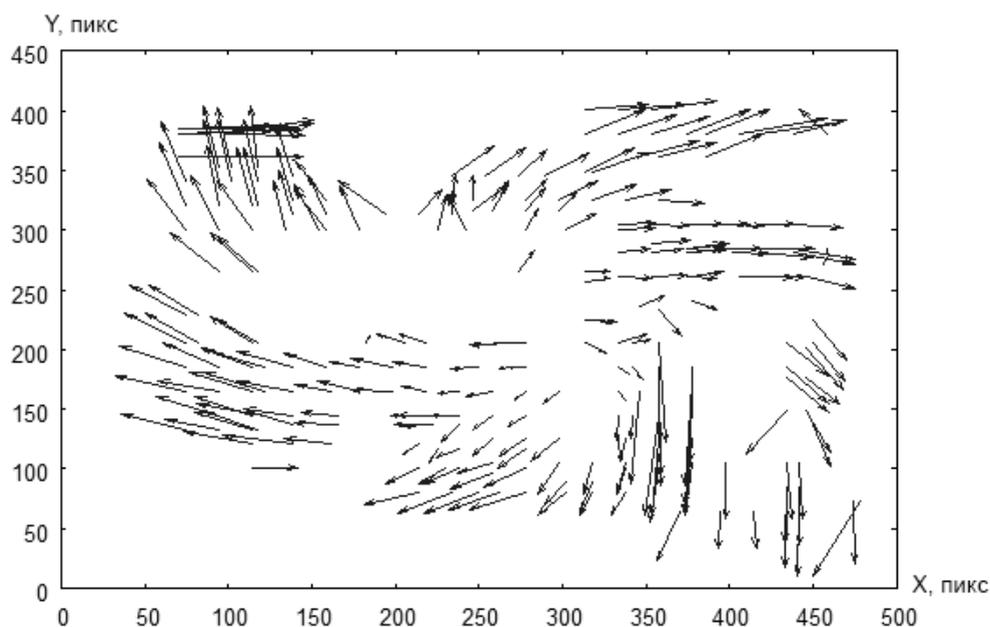


Рис. 2. Векторное поле перемещения точек после смещения

**Лемма.** Исходная область векторного поля перемещений точек, смещённая на вектор  $-T_{avg}$ , будет содержать в себе точку, относительно которой осуществляется операция деформации растяжения.

**Доказательство** (от противного).

Предположим, что точка  $p_c$  находится за пределами смещённой области по некоторой координате  $Q$ . Тогда средний вектор перемещения:

1. устремлён к точке  $p_c$  вдоль оси  $Q$  в случае, если коэффициент деформации по координате  $Q$  меньше 1,

2. устремлён от точки  $p_c$  вдоль оси  $Q$  в случае, если коэффициент деформации по координате  $Q$  больше 1,

3. в проекции на ось  $Q$  будет равен 0, если коэффициент деформации по координате  $Q$  равен 1.

В первом и втором случае мы получаем противоречие, так как новый средний вектор перемещения  $\frac{1}{n} \sum_n (\vec{v}_p - T_{avg}) = \frac{1}{n} \sum_n \vec{v}_p - T_{avg} = T_{avg} - T_{avg} = 0$ . В третьем случае рассматривается тривиальная операция деформации, которая может осуществляться относительно любой точки, что также противоречит изначальному утверждению. **Лемма доказана.**

Найдём очередное приближение точки  $p_c''$ . Воспользуемся гипотезой о малости угла поворота. В данном случае закономерно предположение, что две прямые, параллельные осям координат и проходящие через искомую точку, разбивают плоскость изображения на четыре квадранта знакопостоянства проекций векторов перемещения на оси координат. Воспользуемся численным интегрированием с различным шагом вдоль осей для нахождения точки  $p_c''$ . Глобальные максимумы на графике (рис. 3) соответствуют координатам искомого приближения.

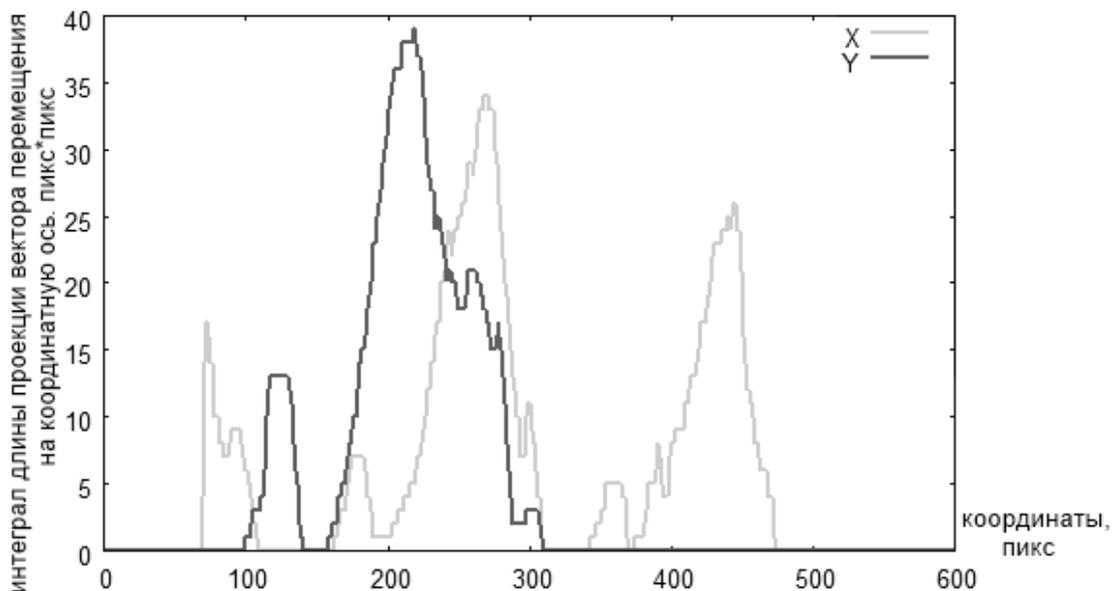


Рис. 3. Поиск координат центральной точки для операции поворота

Поиск угла поворота мы предлагаем осуществлять по принципу максимизации суммы модулей скалярных произведений единичных векторов перемещений  $\vec{nv}_p = \frac{\vec{v}_p}{|\vec{v}_p|}$  на единичные векторы «направления на точку»  $\vec{t}_p = \frac{P - P_c}{|P - P_c|}$ . Максимизация суммы модулей скалярных произведений нормированных векторов гарантирует минимизацию среднего угла отклонения для малых значений угла отклонения вне зависимости от размеров изображений:

$$\forall \left( \Theta : |\Theta| < \frac{\pi}{2} \right) : \sum_n \left| \left( \vec{nv}_p \cdot \vec{t}_p \right) \right| = \sum_n |\cos \Theta_n|.$$

при  $\sum_n |\cos \Theta_n| \rightarrow n, \quad \sum_n \Theta_n \rightarrow 0$

Практически нам удалось показать, что данный метод является эффективным для определения угла поворота вокруг оптической оси (см. рис. 4) в случае, если оценка  $p_c''$  является достаточно точной.

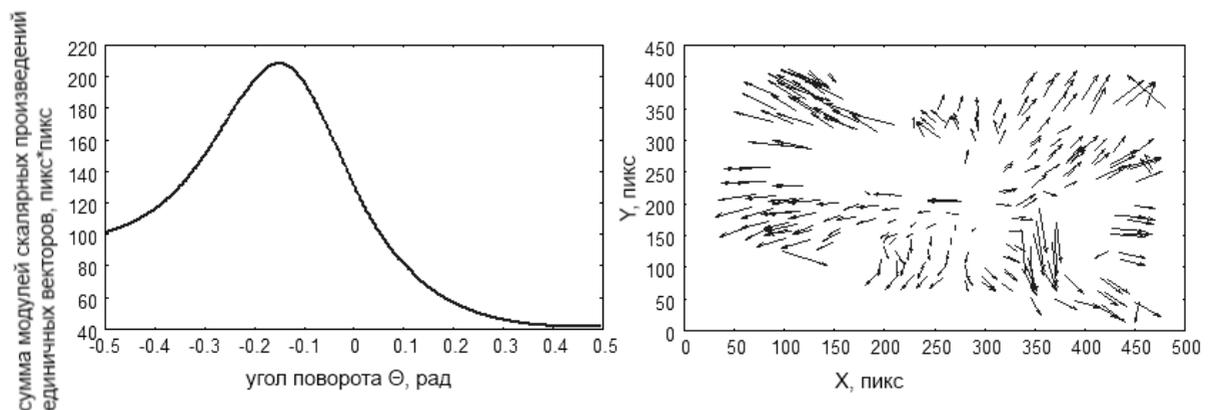


Рис. 4. Зависимость суммы модулей скалярных произведений от угла поворота и результирующее поле

Предположение о малости угла  $\Theta$  позволяет избежать ситуации, когда поворот изображения на угол  $\pi$  будет также максимизировать сумму. Найденный угол  $\Theta$  вместе с оценкой  $p_c''$  определяет преобразование поворота  $R(p)$ , одновременно уточняя преобразование перемещения  $T(p)$ .

### ***Изменение размера изображения***

В общем случае используемые камеры являются неидентичными, имеют различный угол обзора, а их плоскости изображений являются параллельными, но не совпадают. Преобразование  $Z(p)$  учитывает все эти факторы и допускает изменение отношения сторон изображения. Мы предлагаем для нахождения коэффициентов  $k_x$  и  $k_y$  минимизировать сумму модулей векторов перемещения, получаемых после трансформации изображения (1).

$$(k_x, k_y) = \arg \min_{(k_x, k_y)} \sum_n |Z(R(T(p_n))) - p_n^*|, \quad (1)$$

где  $p_n^*$  — координаты соответствующей точки изображения со второй камеры. На рис. 5 приводится график функции (1), на котором хорошо заметен глобальный минимум, соответствующий оптимальному преобразованию  $Z(p)$ .

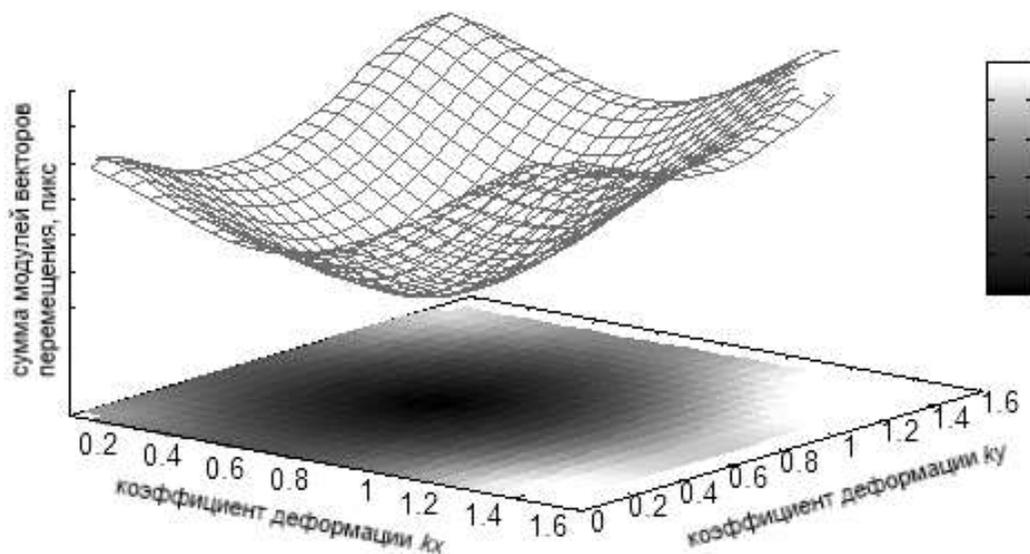


Рис. 5. Минимизация значения суммы модулей векторов в координатах  $(k_x, k_y)$ .

На рис. 6 изображена векторная карта перемещений, полученная после всех преобразований.

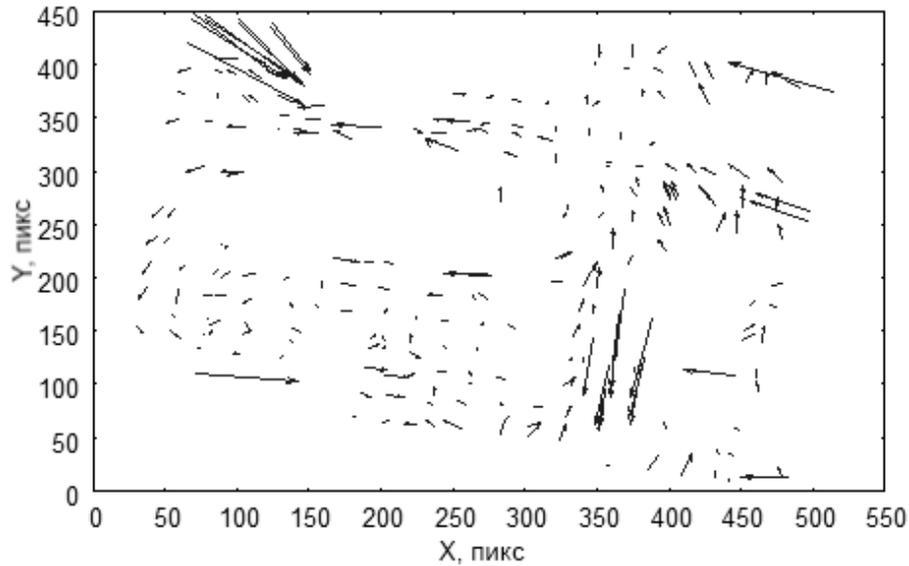


Рисунок 6: Векторное поле перемещений точек после преобразования  $Z(p)$ .

### ***Трансформация изображения и уточнение перемещения***

Нами были найдены параметры всех преобразований  $T(p)$ ,  $R(p)$ ,  $Z(p)$ . На данном этапе изображение с первой камеры может быть преобразовано в координаты второй камеры последовательным применением преобразований. Однако следует отметить, что найденное значение  $p_c''$  является не окончательным решением, а всего лишь оценкой. Для ректификации изображений наиболее важным является размещение эпполярных линий вдоль соответствующих строчек изображений. Ошибка в коэффициенте перемещения  $d_x$  не влияет на расположение строк изображений. Нам необходимо найти корректировку  $\Delta d_y$  для смещения  $d_y$ . Для этого мы предлагаем использовать дискретную функцию суммы яркостей вдоль строчек изображений  $I_s(y) = \sum_x I(x, y)$ , а затем найти смещение, минимизирующее сумму средне-квадратичного отклонения функции (2). В качестве альтернативного подхода мы предлагаем максимизировать взаимную корреляцию.

$$\partial y = \arg \min_{d_y} \sum_y [I_s^*(y) - I_s(y + \partial y)]^2, \quad (2)$$

где  $I(y)$  и  $I^*(y)$  — дискретные функции суммы яркостей изображений первой и второй камер соответственно.

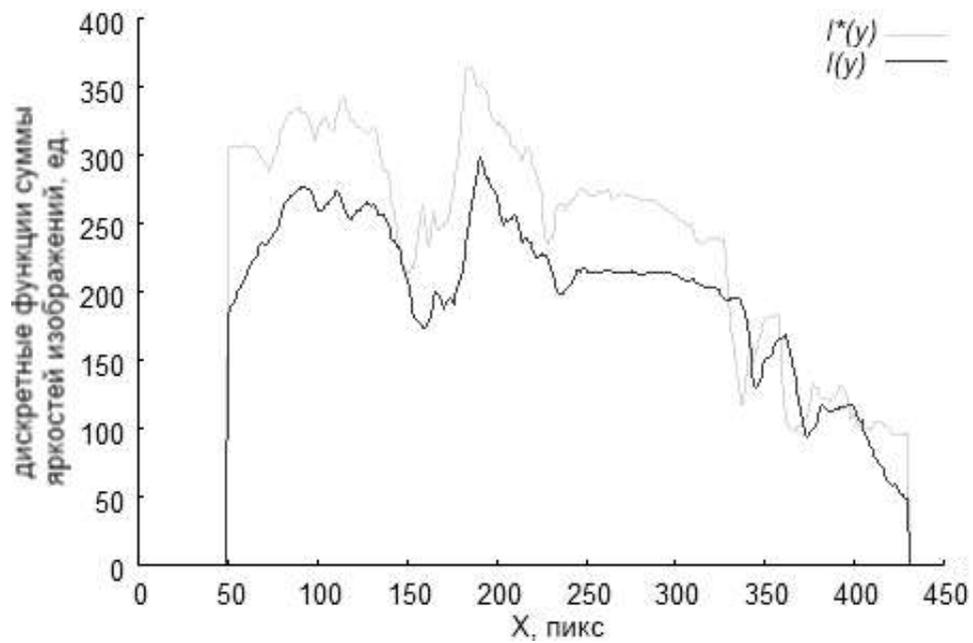


Рис. 7. Дискретные функции суммарной яркости  $I(y)$  и  $I^*(y)$

На рис. 7 можно увидеть «профили» функций  $I_s(y)$  и  $I_s^*(y)$ . Подбор корректировки  $\Delta d_y$  позволяет вычислить окончательные координаты  $p_c = p_c'' + (0, \Delta d_y)^T$ . На рис. 8 представлены результаты работы алгоритма.

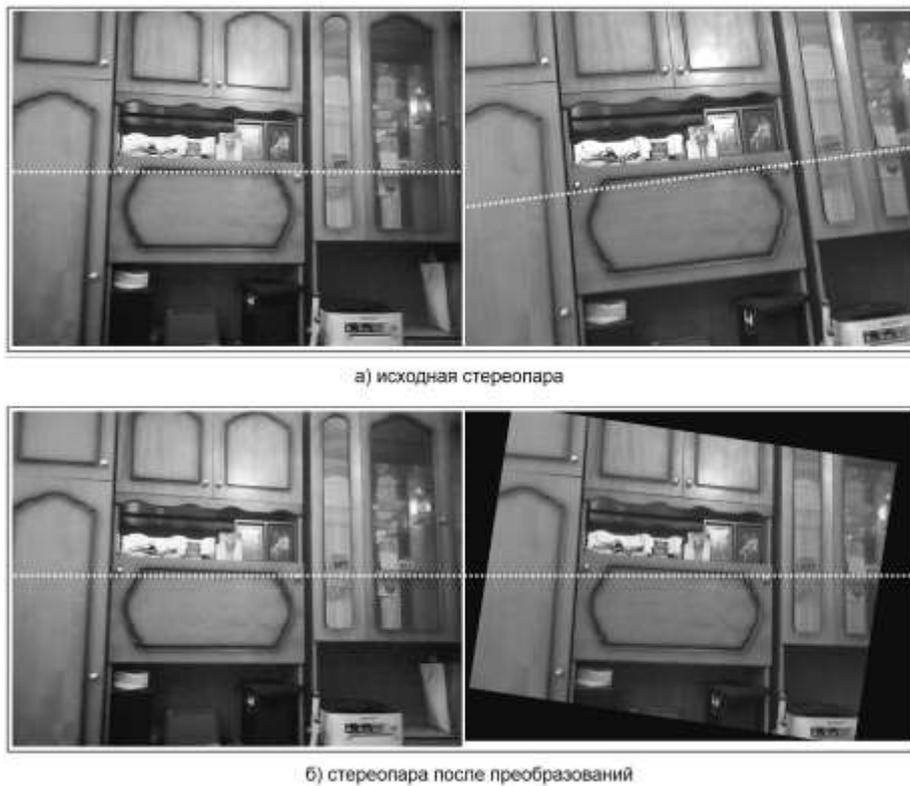


Рис. 8. Входная и результирующая стереопары.

## Статистическая обработка

После нахождения преобразований  $T$ ,  $R$  и  $Z$  в первом приближении мы можем увеличить точность коэффициентов этих преобразований. Так, на рис. 6 можно заметить векторы, существенно выделяющиеся длиной. Данные векторы являются следствием неисключённых ошибок алгоритма поиска точечных соответствий. Поскольку данные векторы легко отделить от прочих, исключим их из выборки и проведём поиск преобразований заново. Эксперимент показал, что в подавляющем большинстве случаев для фильтрации достаточно одной итерации: при повторной фильтрации по длине отбрасывается менее 5% векторов.

Ещё одним способом увеличения точности решения является поиск коэффициентов для нескольких независимых сцен. На рис. 9 в качестве примера показано значение угла поворота  $\Theta$  камеры в радианах, найденное для различных кадров, при одинаковом позиционировании камер. После исключения выбросов можно воспользоваться средним значением коэффициента как окончательным приближением.

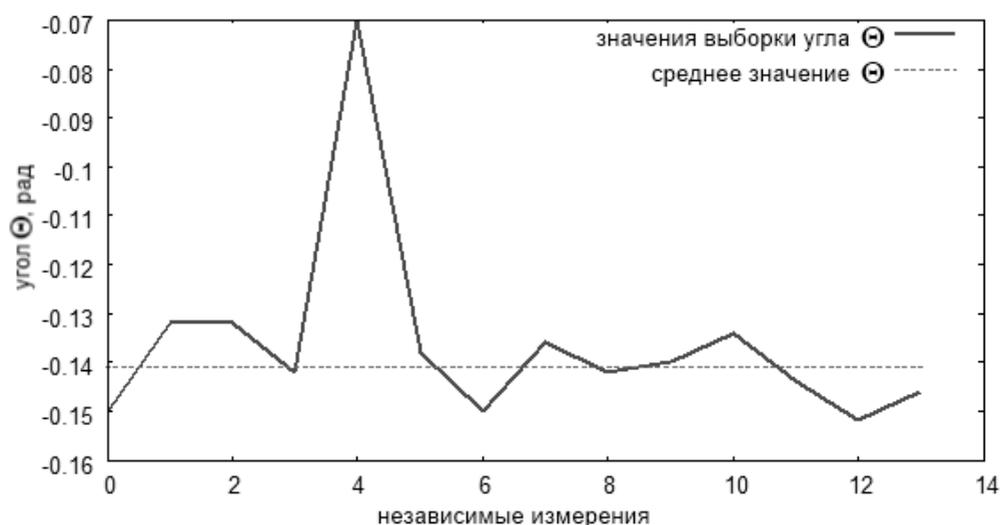


Рис. 9. Значения угла поворота для разных кадров

## Заключение

Нами был разработан метод и алгоритм ректификации стереоизображений по сцене, основанный на модификации подхода к

стабилизации камеры. Получаемые изображения могут применяться для создания стерео-видеопотоков и стереоизображений.

Предложенный метод содержит несколько ограничений. Так, на уровне модели исключены две степени свободы, а поворот изображения вокруг оптической оси камеры рассматривался в диапазоне  $\pm \frac{\pi}{6}$ .

Поскольку в модель заложена возможность поворота камеры, мы использовали для сопоставления критерий взвешенного среднеквадратичного отклонения (weighted MSE) для увеличения достоверности сопоставления. Данный критерий, применяемый локально, может давать некорректные результаты на регулярных паттернах (одинаковые полосы, решётки, обои и т.п.). В качестве альтернативы можно рассматривать применение специальных изображений-паттернов с высокой цветовой дисперсией для нахождения параметров преобразований.

Поскольку найденные преобразования, полученные за один проход по единственной стереопаре, могут содержать неточные значения коэффициентов, мы предложили два статистических улучшения, которые повышают точность найденного решения. Во-первых, на рис. 6 видно, что часть векторов с относительно большими значениями нормы с высокой вероятностью являются результатом ошибочного срабатывания алгоритма wMSE; если отфильтровать такие векторы и повторить решение, то можно повысить точность найденных параметров преобразования. Во-вторых, выборка из нескольких решений, выполненных при одинаковой конфигурации камер, позволяет использовать математическое ожидание коэффициента в качестве окончательного решения.

## Литература

1. Ting Chen, "Video Stabilization Algorithm Using a Block-Based Parametric Motion Model". Winter 2000, Stanford University, pp. 3-4
2. Бондаренко С., Бондаренко М. Создание 3D-изображений: теория и практика [Электронный ресурс] // Сайт проекта 3Domen. 2011, 27 февраля. URL: <http://3domen.com/index.php?newsid=5794> (дата обращения 26.07.2012)
3. Jesse S Jin, Zhigang Zhu, Guangyou Xu. Digital Video Sequence Stabilization Based on 2.5D Motion Estimation and Inertial Motion Filtering. IEEE International Conference on Intelligent Vehicles (2001) Volume 7, Issue 4, pp. 357-365.
4. Протасов С.И. Использование web-камер как источника стереоизображений в реальном времени / С.И. Протасов, А.А. Крыловецкий // Информатика: проблемы, методология и технологии: материалы XI международной научно-методической конференции, Воронеж, 10-11 февраля 2011г.: в 3 т.- Воронеж: издательско-полиграфический центр Воронежского государственного университета, 2011, Т.2, С. 229-232
5. Takeo Kanade, Masatoshi Okutomi. "A stereo matching algorithm with an adaptive window: theory and experiment". IEEE Transactions on Pattern Analysis and Machine Intelligence, Sep 1994, pp. 920-932.
6. Eddy Vermeulen. "Real-time Video Stabilization For Moving Platforms". 21st Bristol UAV Systems Conference - April 2007.

Protasov S.I., Krylovetsy A.A., Kurgalin S.D.

**An approach to solve stereoimage rectification problem without camera calibration.**

This article is devoted to the method of initial image processing to use in stereovision systems. It is based on a modification of video stabilization approach. The method considers image rectification process as a sequence of transformations. Each transformation is found as a solution of optimization problem. The article describes mathematical model that fits main principles of video stabilization method, provides optimization techniques to find transformation parameters and describes statistical approach to solution refinement. The method provided does not require camera calibration and uses only scene images. This approach can be applied to create individual stereovision systems.

**Keywords:** *stereo-vision, computer vision, stereoimages, camera calibration, video stabilization.*