

# Поиск карты смещений по пирамиде детальности

Д.Б. Волегов\*, Д.В. Юрин\*\*

\* Московский физико-технический институт (государственный университет)

\*\* ФГУП НПП ОПТЭКС,

[dvolegov@dgap.mipt.ru](mailto:dvolegov@dgap.mipt.ru) [yurin\\_d@inbox.ru](mailto:yurin_d@inbox.ru)

## Аннотация

В работе предложен алгоритм поиска карты смещений, используя пирамиду детальности исходных изображений. Приведены результаты работы алгоритма. Для увеличения стабильности алгоритма используется информация о текстуре и уровне шума на изображениях.

*Keywords: совмещение изображений, пирамида детальности, карта смещений, восстановление 3D*

## 1. ВВЕДЕНИЕ

Рассматривается задача восстановления трехмерной структуры сцены по паре снимков. Картой смещений (disparity map) называется двумерный массив, элементами которого являются двумерные векторы. Каждый вектор задает сдвиг от точки на первом изображении до соответствующей точки на втором изображении. В дальнейшем для карты смещений используется сокращение (КС).

Обычно поиск КС является заключительной стадией процесса восстановления сцены (плотное восстановление) [1,2,3]. К этому моменту уже найдена фундаментальная матрица, изображение может быть ректифицировано, и следует искать сдвиг только по горизонтали. В настоящей работе поиск КС рассматривается как предварительный этап (когда фундаментальная матрица еще не найдена), целью которого является получение как можно большего числа соответствующих точек. Поэтому сдвиг ищется как по горизонтали, так и по вертикали, и может достигать до 20-30% от размеров изображения. Предлагаемый метод является развитием подхода [4] с использованием пирамиды разрешений, что частично преодолевает требование малости отклонений в [4], с другой стороны предполагается до этого произвести грубое совмещение изображений алгоритмом [5], что также помогает расширить диапазон применимости. Далее предполагается поиск особенностей на первом изображении, соответствия для которых могут быть получены из карты смещений (с минимальной проверкой по разностному кадру). Таким образом, предлагаемый алгоритм следует рассматривать не как алгоритм плотного восстановления, а как новый подход к поиску соответствующих точек на изображении. Предлагаемая связка двух алгоритмов направлена на преодоление возможных проблем подходов типа [6] для сцен, содержащих регулярные структуры типа окон в современных зданиях.

Преимуществом предлагаемого подхода является использование пирамиды детальности (ПД). Сначала КС ищется для сильно уменьшенных копий изображений. Найденные значения являются начальными для КС для более детальных копий и т.д. Таким образом, на каждом уровне детальности требуется лишь обновить КС, что значительно

сокращает время вычисления и вероятность нахождения ложных значений. Вместе с тем, алгоритм предполагает наличие на сцене достаточно крупных объектов, т.е. кусочно-гладкую КС.

## 2. АЛГОРИТМ

Имеется пара изображений, для которой требуется построить КС. Строятся две ПД - для каждого изображения. Пирамида строится путем сглаживания изображения на предыдущем уровне и выбором более одного пикселя при помощи билинейной интерполяции. Полуширина  $\sigma$  функции Гаусса связана с отношением  $k$  ( $k > 1$ ) размеров изображений пирамиды на соседних уровнях:

$$\sigma = \frac{2k}{\pi}$$

Такой выбор, с одной стороны, оставляет в сглаженном изображении только те частоты, которые будут содержать уменьшенное изображение, а с другой стороны, не приводит к потере деталей.

Также строится ПД для КС. Поиск КС происходит постепенно, начиная с вершины ПД. Обработка изображений вблизи вершины пирамиды и вблизи основания различна.

Для изображений вблизи вершин пирамид (сильно уменьшенные изображения) ищется геометрическое преобразование (аффинное, проективное) изображения в первой ПД, совмещающее его в целом с изображением на том же уровне второй ПД. Поиск преобразования представляет собой модификацию [4] и описан в подразделе 2.1. Найденное преобразование между изображениями на заданном уровне ПД позволяет вычислить КС на том же уровне. При переходе к следующему уровню более детального разрешения, изображение в первой пирамиде преобразуется согласно КС для текущего уровня, таким образом, на следующем уровне ищется только уточнение КС.

Начиная с некоторого уровня в ПД более нельзя найти простое геометрическое преобразование, совмещающее изображение. Изображение разбивается на квадраты предопределенного размера (со стороной от 8 до 16 пикселей), и для каждого ищется геометрическое преобразование (обычно сдвиг) и КС тем же способом, что и ранее для всего изображения. Очевидно, что поиск преобразования уместен только в случае, если один квадрат не захватывает объекты на разном расстоянии от камеры, в противном случае разные части квадрата должны испытывать разный сдвиг. С целью уменьшения скачков смещений на границах квадратов, квадраты перекрываются, а по общим частям квадратов выполняется усреднение смещений. Чем больше степень перекрытия квадратов, тем более сглаженной

получается карта смещений, но вместе с тем, возрастает время работы.

## 2.1 Совмещение изображений

Рассматривается пара изображений  $I_1(\vec{x})$ ,  $I_2(\vec{x})$  и КС  $\vec{d}(\vec{x}, \vec{\omega})$  пикселей первого изображения. Вектор  $\vec{\omega}$  есть вектор параметров, задающий модель КС. Ниже рассматриваются две модели КС: сдвиг и преобразование подобия.

**Сдвиг.**

$$\vec{\omega} = (t_x, t_y) \quad (1)$$

$$\vec{d}(\vec{x}, \vec{\omega}) = l\vec{t}$$

**Преобразование подобия.**

$$\vec{\omega} = (t_x, t_y, \lambda, \varphi)$$

$$\vec{d}(\vec{x}, \vec{\omega}) = (1 + \lambda)\mathbf{R}_\varphi \vec{x} + l\vec{t} - \vec{x} \quad (2)$$

$$\mathbf{R}_\varphi = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}$$

Параметр  $l$  вводится для обезразмеривания компонент  $\vec{\omega}$  и приведения их значений к одному порядку величины так, что каждая компонента не превосходит нескольких десятых.

Его численное значение полагается равным размеру диагонали изображения:

$$l = \sqrt{w^2 + h^2}, \quad (3)$$

где  $w$  и  $h$  ширина и высота изображения соответственно.

В обоих случаях:

$$\vec{d}(\vec{x}, \vec{0}) = 0 \quad (4)$$

Поиск КС сводится к поиску  $\vec{\omega}$ . Далее выкладки текущего подраздела аналогичны [4]. Пусть  $\mathcal{E}(\vec{\omega})$  обозначает интеграл от квадрата разности второго и преобразованного первого изображений:

$$\mathcal{E}(\vec{\omega}) = \int (I_1(\vec{x} + \vec{d}(\vec{x}, \vec{\omega})) - I_2(\vec{x}))^2 d\vec{x} \quad (5)$$

Ниже используется обозначение:

$$\vec{y} = \vec{x} + \vec{d}(\vec{x}, \vec{\omega}) \quad (6)$$

Требуется найти  $\vec{\omega}$ , минимизирующий  $\mathcal{E}(\vec{\omega})$ :

$$\frac{\partial \mathcal{E}}{\partial \vec{\omega}} = \int (I_1(\vec{y}) - I_2(\vec{x})) \frac{\partial I_1(\vec{y})}{\partial \vec{y}} \frac{\partial \vec{y}}{\partial \vec{\omega}} d\vec{x} = 0 \quad (7)$$

Ниже используются обозначения:

$$\vec{g}^T \equiv \frac{\partial I_1(\vec{y})}{\partial \vec{y}}, \quad \mathbf{Y} \equiv \frac{\partial \vec{y}}{\partial \vec{\omega}} = \frac{\partial \vec{d}(\vec{x}, \vec{\omega})}{\partial \vec{\omega}} \quad (8)$$

В предположении сглаженности  $I_1(\vec{x})$ , малого отличия  $\vec{x}$  от  $\vec{y}$  и с учетом (4) с точностью до членов первого порядка получается:

$$\vec{g}(\vec{y}) = \vec{g}(\vec{x})$$

$$I_1(\vec{y}) = I_1(\vec{x}) + (\vec{g}, \vec{y} - \vec{x}) \quad (9)$$

$$\vec{y} - \vec{x} = \mathbf{Y}\vec{\omega} \quad (10)$$

Используя (9), (10) уравнение (7) переписывается в виде:

$$\int (I_1(\vec{x}) - I_2(\vec{x}) + \vec{g}^T \mathbf{Y} \vec{\omega}) \vec{g}^T \mathbf{Y} d\vec{x} = 0 \quad (11)$$

Из (11) получается система относительно  $\vec{\omega}$ :

$$\mathbf{A}\vec{\omega} = \vec{b} \quad (12)$$

$$\mathbf{A} = \int \mathbf{Y}^T \vec{g} \vec{g}^T \mathbf{Y} d\vec{x}, \quad (13)$$

$$\vec{b} = \int (I_2(\vec{x}) - I_1(\vec{x})) \mathbf{Y}^T \vec{g} d\vec{x} \quad (14)$$

Ниже приводятся  $\mathbf{Y}$  конкретных моделей КС.

**Сдвиг.**

$$\mathbf{Y} = \begin{pmatrix} l & 0 \\ 0 & l \end{pmatrix} \quad (15)$$

**Преобразование подобия.**

$$\mathbf{Y} = \begin{pmatrix} l & 0 & x_1 & -x_2 \\ 0 & l & x_2 & x_1 \end{pmatrix} \quad (16)$$

**Алгоритм 1 (Совмещение изображений).**

1. Вычислить производные первого изображения путем свертки с производными Гаусса. Полуширина функции Гаусса должна быть порядка ожидаемых значений в КС.
2. Составить матрицу  $\mathbf{A}$  системы (13).
3. Составить правую часть  $\vec{b}$  (14) и решить систему (12).
4. Преобразовать второе изображение согласно обратному преобразованию, построенному из  $\vec{\omega}$ . Это исключает необходимость пересчета  $\mathbf{A}$  на каждой итерации.
5. Если  $\vec{\omega}$  такой, что каждый пиксель сдвигается не более чем на заданную величину, то выйти.
6. Перейти к п.3.

## 2.2 Решение системы

Поиск решения системы (12) требует аккуратности, так как изображение (или его фрагмент) может не обладать достаточной текстурой для достоверного определения  $\vec{\omega}$ .

Идея состоит в том, чтобы оценить погрешность матрицы  $\mathbf{A}$  по априорному значению шума  $\mathcal{N}$  на изображении и радиусу сглаживания  $\sigma$  исходных изображений. Далее система (12) решается с применением SVD-разложения, и сингулярные числа меньше порогового обнуляются [7].

Для оценки порогового сингулярного числа матрицы  $\mathbf{A}$  рассмотрим случайное изображение с дисперсией интенсивности пикселей  $\langle n^2 \rangle$  и нулевым средним значением:  $\langle n \rangle = 0$ . Для  $\sigma$  не сильно меньших единицы, когда сглаживание имеет смысл, выполняется оценка для дисперсии  $\mu^2$  производных рассматриваемого изображения:

$$\mu^2 \equiv \frac{\langle |\vec{g}|^2 \rangle}{2} = \frac{\langle n^2 \rangle}{8\pi\sigma^4} \quad (17)$$

Оценка (17) может быть легко получена путем перехода в спектральную область. Спектр рассматриваемого изображения является однородным и случайным [8], дисперсия которого может быть вычислена из теоремы о равенстве интеграла от квадрата модуля сигнала интегралу от квадрата его спектра [9]. Дифференцирование изображения сводится к умножению его спектра на соответствующую производную функции Гаусса, что позволяет оценить  $\mu^2$ .

Зная модель искомого преобразования (15), (16) и оценку (17) можно оценить минимальное сингулярное число матрицы  $\mathbf{A}$ :

$$s_{\min} = \mu^2 l^2 wh \quad (18)$$

Сингулярные числа матрицы  $\mathbf{A}$ , меньшие  $s_{\min}$ , должны быть обнулены при вычислении псевдообратной матрицы. В случае оценки минимального сингулярного значения при работе с фрагментом  $\mathcal{W}$ ,  $h$ ,  $l$  в (18) – это размеры фрагмента.

## 3. РЕЗУЛЬТАТЫ

На рис. 1 приведены результаты работы алгоритма для снимков, полученных фотоаппаратом со стереоприставкой фирмы Logeo ([www.loreo.com](http://www.loreo.com)). На рис. 2,3 представлены результаты работы алгоритма для пары снимков, полученных обычным фотоаппаратом.

Можно видеть, что для кусочно-гладких поверхностей алгоритм работает удовлетворительно. Неверные значения получаются вблизи границ объектов.

Время работы алгоритма на ноутбуке Athlon-XP-M-2200 для пары цветных снимков 480x640 составляет около 7 секунд и пропорционально числу пикселей на изображениях.



Рис 1. Слева – левое изображение, в центре – правое изображение, справа – карта смещений.

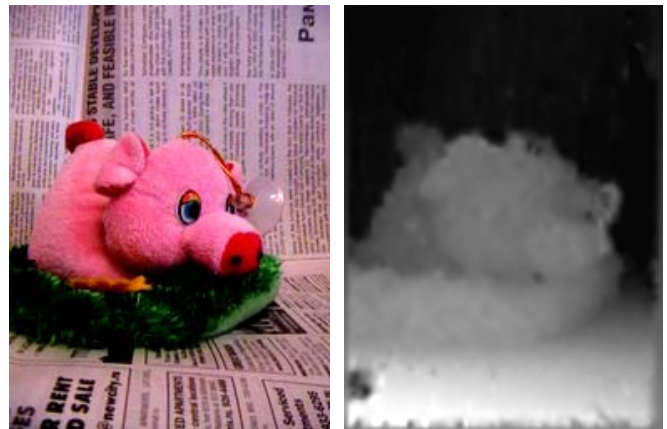


Рис 2. Слева – левое изображение, справа – карта смещений.



Рис 3. Легенда та же, что и на рис. 2

## 4. БУДУЩАЯ РАБОТА

На текущий момент алгоритм имеет два слабых места: во-первых, на границах объектов находятся неправильные значения смещений, во-вторых, предлагаемый подход в принципе неприменим в случае наличия мелких, расположенных близко к камере объектов. Это связано с тем, что на грубых разрешениях объект пропадает, а на высоком разрешении сдвиг объекта слишком большой, чтобы быть найденным предлагаемым методом. Однако анализ разностного кадра может выявить подобные проблемные

объекты, и для них возможно применение других методов в качестве пост-обработки.

Основное направление будущей деятельности – это улучшение работы алгоритма на краях объектов, где КС терпит разрыв. Рассматриваются две возможности. Во-первых, можно улучшать уже полученную КС, используя результат применения детектора краев, а именно, экстраполировать достоверную часть КС внутри области до ее границ, найденных детектором краев, например Канни. Во-вторых, возможно на этапе вычисления КС использовать детектор краев, и, в случае работы вблизи границы, искать КС по каждую сторону от границы отдельно.

Следует отметить также, что в результате работы алгоритма получены взаимно-однозначные соответствия между парами пикселей первого и второго изображения, верные для значительного количества точек изображения (исключая области затенения, occlusions). Это создает возможность использовать робастные методы (RANSAC) для нахождения фундаментальной матрицы и эпиполярной геометрии. Таким образом возможен второй проход алгоритма, выполняемый при наложенных эпиполярных ограничениях.

## 5. ЗАКЛЮЧЕНИЕ

В работе предложен алгоритм поиска КС. Достоинством алгоритма является постепенный поиск КС, начиная с сильно уменьшенных изображений. С увеличением детализации изображения выполняется лишь обновление КС относительно предыдущего уровня, что приводит к уменьшению времени работы алгоритма и вероятности нахождения ложных значений. При вычислении КС используется априорное значение уровня шума на изображении, что увеличивает стабильность алгоритма. Приведенные результаты свидетельствуют об удовлетворительной работе алгоритма для сцен, состоящих из кусочно-гладких поверхностей.

## 6. БЛАГОДАРНОСТИ

Работа выполнена при поддержке РФФИ, гранты 06-01-00789-а, 05-07-90345-в, 05-07-90390-в. Авторы выражают благодарность зав. кафедрой СИМ МФТИ д.ф.-м.н., проф. Клименко С.В. за проявленное к работе внимание и поддержку.

## 7. ЛИТЕРАТУРА

- [1] Michael H. Lin, Carlo Tomasi: *Surfaces with Occlusions from Layered Stereo*. //IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003.
- [2] D. Scharstein and R. Szeliski. *A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms*. *IJCV* 47(1/2/3):7-42, April-June 2002. <http://cat.middlebury.edu/stereo/>.
- [3] V. Kolmogorov and R. Zabih. *Computing visual correspondence with occlusions using graph cuts*. In *International Conference on Computer Vision, Vancouver, Canada, July 2001*. <http://www.cs.cornell.edu/rdz/Papers/KZ-ICCV01-tr.pdf>.
- [4] Jianbo Shi and Carlo Tomasi, "Good features to track", *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'94) (Seattle), June 1994*.

[5] Д.Б.Волегов, Д.В.Юрин. Грубое совмещение изображений по найденным на них прямым линиям//В сб. Труды конференции. 16-я Международная Конференция по Компьютерной Графике и её приложениям ГрафиКон'2006 - С. 463-466. Новосибирск, 1-5 июля 2006/ <http://www.graphicon.ru/proceedings2006/proceedings.html>.

[6] . Dufournaud, C. Schmid, and R. Horaud. Matching images with different resolutions. In *CVPR*, pages 612-618, 2000. <http://citeseer.ist.psu.edu/dufournaud00matching.html>.

[7] Дж. Форсайт, М. Малькольм, К. Моулер, *Машинные методы математических вычислений, пер. с англ.* –М.: Мир, 1980, 280 с.

[8] Saeed V. Vaseghi, *Advanced Digital Signal Processing and Noise Reduction*, John Willey & Sons Ltd., 2000

[9] Кудрявцев Л. Д., *Курс математического анализа (в двух томах)*. -М.: Высшая школа, 1981, т.II:-584 с., ил.

## Об авторах

Волегов Дмитрий Борисович – студент ФОФФ Московского Физико-Технического Института, кафедра СИМ.

Телефон: +7-(926)-354-21-92 E-mail: [dvolegov@dgap.mipt.ru](mailto:dvolegov@dgap.mipt.ru)

Юрин Дмитрий Владимирович – кандидат физ.-мат. наук, ФГУП НПП ОПТЭКС (Федеральное Государственное Унитарное Предприятие Научно-Производственное Предприятие "Опτικο-электронные комплексы и системы"), главный специалист; с.н.с. кафедры Математической физики ВМиК МГУ им. М.В.Ломоносова

Телефон: +7-(916)-387-7367, +7-(495)-535-4887

E-mail: [yurin\\_d@inbox.ru](mailto:yurin_d@inbox.ru)

## Finding disparity map via image pyramid

### Abstract

Algorithm for finding disparity map using image pyramid is presented. Stability of algorithm is improved by using image texture information and noise level. Results are presented.

**Keywords:** *disparity map, image registration, image pyramid, 3D reconstruction*

### About the authors

Dmitry B. Volegov is a student at Moscow Institute of Physics and Technology, Department of General and Applied Physics, sub-faculty of SIM. His contact email is [dvolegov@dgap.mipt.ru](mailto:dvolegov@dgap.mipt.ru)

Dmitry V. Yurin, PhD, is a chief specialist of FSUE RPE OPTECS (The Federal State Unitary Enterprise Research and Production Enterprise "Optoelectronic Complexes and Systems") and a senior researcher at Moscow State University, Department of Computational Mathematics and Cybernetics. His contact email is [yurin\\_d@inbox.ru](mailto:yurin_d@inbox.ru)