

# Упорядочивание изображений для построения по ним 3d моделей

Д.Б. Волегов\*, Д.В. Юрин\*\*

\*Московский физико-технический институт (государственный университет)

\*\*Московский государственный университет им. М.В.Ломоносова

dvolegov@rambler.ru yurin\_d@inbox.ru

## Аннотация

В работе рассматривается задача последовательного построения 3d модели по фотографиям. Сначала обрабатывается небольшое число фотографий и строится первичная модель, затем обрабатываются следующие фотографии, модель уточняется и дополняется. Во-первых, рассматривается задача выбора последовательности изображений. Во-вторых, вводится понятие определенности фундаментальной матрицы, которое позволяет судить о точности реконструкции по паре кадров без выполнения таковой. Приводятся результаты применения предлагаемого метода к реальным и синтезированным последовательностям изображений.

**Keywords:** упорядочивание изображений, построение 3d моделей, фундаментальная матрица.

## 1. ВВЕДЕНИЕ

На текущий момент разработано множество методов автоматического построения 3d моделей объектов по изображениям. Методы реконструкции могут быть разделены на два типа: пакетная реконструкция и последовательная реконструкция. Методы первой группы обрабатывают все изображения сразу. К ним относятся методы на основе факторизации матриц и его вариации. Методы второй группы основаны на последовательной обработке изображений - модель сначала строится по небольшому числу кадров и постепенно дополняется и уточняется с подключением остальных кадров. К таким методам относятся [1], [2], [3], [4].

В настоящей работе рассматривается задача упорядочивания изображений для методов последовательной реконструкции. В [3], [4] порядок изображений известен, так как изображения образуют последовательность. В [2] проблема построения последовательности решается явно. А именно подход [2] состоит сначала в выборе пары изображений с максимальным количеством соответствий. А следующим обрабатывается изображение, для которого число парных соответствий с предыдущим максимально. В [1] изображения упорядочиваются таким образом, чтобы число уже реконструированных точек, видимых на добавляемом изображении, было максимальным.

Названные подходы к упорядочиванию изображений могут таить в себе опасность. Например, пусть сцена снимается из трех точек, первые две из которых расположены друг к другу гораздо ближе, чем к третьей. Скорее всего, для пары изображений (1,2) найдется больше соответствий, чем для пар (1,3) или (2,3). Но, в силу малого расстояния между точками 1,2 взаимное положение камер в этих точках плохо

определено, и подход, используемый в [2] может привести к ошибочному результату. Это же относится и к подходу, используемому в [1], так как большое количество общих точек между обработанными и новым кадром не гарантирует большой базы съемки.

Часто для реконструкции координат точек и камер используются методы [5] на основе нелинейной минимизации некоторой ошибки. В этом случае чем ближе начальная структура сцены соответствует минимуму, тем меньше вероятность попасть в ложный локальный минимум. Следовательно, несмотря на малое число соответствий, более удачным является вычисление положений камер в точках 1,3 (или 2,3) и последующем вычислении положения в точке 2 (или 1).

Удачный выбор последовательности реконструкции важен по следующим причинам. Во-первых, если для реконструкции используются методы на основе нелинейной минимизации [5], то уменьшается вероятность попадания в ложные локальные минимумы и нахождения ложного решения, т.е. увеличивается стабильность алгоритмов реконструкции. Во-вторых, может уменьшиться время работы алгоритмов минимизации, так как уменьшается число итераций.

В настоящей работе построение 3d модели (облака точек) состоит из следующих основных этапов.

- На всех изображениях ищутся особенности. В настоящей работе применяется многомасштабный детектор особенностей типа blob, в качестве функции отклика детектора используется нормированный Лапласиан [6].
- Вычисляются дескрипторы особенностей, найденных на всех изображениях. В настоящей работе используется дескриптор SIFT [7].
- Строится индекс дескрипторов, который позволяет для заданного дескриптора быстро найти несколько ближайших дескрипторов для особенностей этого же и других изображений. Время поиска составляет  $O(N^\alpha)$ , где  $N$  - общее число особенностей на всех изображениях,  $\alpha \in [0, 1]$ . Параметр  $\alpha$  контролирует степень компромисса между качеством и скоростью поиска. Основная идея в построении индекса состоит в том, чтобы сократить размерность дескриптора со 128 до примерно 30-40 методом анализа главных компонент (principal component analysis), а затем использовать деревья для ускорения поиска. В ходе экспериментов на реальных данных было выяснено, что приемлемым значением является  $\alpha \in [0.3, 0.4]$ . Детальное описание процесса построения индекса выходит за рамки настоящей работы и будет опубликовано в ближайшее время
- Строится таблица парных соответствий. Для каждого дескриптора  $D$ , используя индекс

дескрипторов, ищется ближайший дескриптор  $\mathcal{D}_1$  (изображения, соответствующие  $\mathcal{D}$  и  $\mathcal{D}_1$  различны). Для  $\mathcal{D}_1$  также ищется ближайший дескриптор  $\mathcal{D}_2$ . Если дескрипторы  $\mathcal{D}$  и  $\mathcal{D}_2$  совпадают, то соответствие  $\{\mathcal{D}, \mathcal{D}_1\}$  заносится в таблицу. Вместо поиска одного ближайшего дескриптора возможен поиск взаимных пар среди  $k$  ближайших соседей. Это увеличивает время работы, но и число парных соответствий.

- Строится список пар изображений, между которыми есть достаточное ( $\sim 30$ ) число соответствий, используя таблицу парных соответствий. Для каждой такой пары вычисляется фундаментальная матрица робастным (устойчивым к ложным соответствиям) методом [8] (р. 290) на основе RANSAC [9].
- Строится граф, вершины которого соответствуют изображениям. Пару вершин соединяет ребро, если между соответствующими изображениями вычислена фундаментальная матрица. Граф разбивается на связные компоненты, каждая из которых обрабатывается независимо.
- Ребрам графа назначаются веса. В настоящей работе вводится мера определенности фундаментальной матрицы (см. 2.1), которая является весом ребра. Введенная мера обосновывается в разделе 2.2. Используя алгоритм раздела 3., строится последовательность, в соответствии с которой обрабатываются изображения. Сначала вычисляются матрицы первых двух камер и структура сцены, затем вычисляется матрица следующей камеры, структура уточняется и дополняется и т.д.

Для вычисления определенности фундаментальной матрицы необходимы лишь координаты соответствующих точек. В [10] рассматривается вычисление матрицы ковариации элементов фундаментальной матрицы и демонстрируются несколько применений: оценка ширины полосы вдоль эпиполярной линии, ошибок координат реконструированных точек и ошибки параметров калибровки камер. Целью предлагаемой ОФМ является простой способ оценки достижимой точности реконструкции на раннем этапе обработки и выбор подходящей последовательности реконструкции.

## 2. ОПРЕДЕЛЕННОСТЬ ФУНДАМЕНТАЛЬНОЙ МАТРИЦЫ

### 2.1 Определение

Для однородных координат соответствующих точек для пары изображений выполняется соотношение [8]:

$$\vec{y}_i^T \mathbf{F} \vec{x}_i = 0, \forall i = 1 \dots N \quad (1)$$

Через  $\vec{x}_i$  обозначены векторы однородных координат точки  $i$  на первом изображении, через  $\vec{y}_i$  - соответствующей ей точки на втором изображении,  $\mathbf{F}$  - фундаментальная матрица между парой изображений,  $\vec{x}_{i,3} = \vec{y}_{i,3} = 1$ ,  $N$  - число соответствий между парой кадров.

Используя обозначения,

$$\vec{f} = (F_{11} F_{21} F_{31} F_{12} F_{22} F_{32} F_{13} F_{23} F_{33})^T \quad (2)$$

$$Z_{ij} = x_{i,k} y_{i,l}, \quad j = 3(k-1) + l \quad (3)$$

$$k, l = 1 \dots 3, \quad j = 1 \dots 9$$

соотношения (1) переписываются в матричной форме:

$$\mathbf{Z} \vec{f} = 0 \quad (4)$$

Сингулярное разложение матрицы  $\mathbf{Z}$  записывается в виде:

$$\mathbf{Z} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T \quad (5)$$

Вводится мера определенности фундаментальной матрицы:

$$q(\mathbf{F}) = \frac{\sigma_2 - \sigma_1}{\sqrt{\langle \delta \sigma_1^2 \rangle + \langle \delta \sigma_2^2 \rangle}} \quad (6)$$

В (6)  $\sigma_1, \sigma_2$  - пара наименьших сингулярных чисел матрицы  $\mathbf{Z}$ ,  $\sigma_1 < \sigma_2$ , а  $\langle \delta \sigma_1^2 \rangle, \langle \delta \sigma_2^2 \rangle$  - средние квадраты отклонений сингулярных чисел  $\sigma_1, \sigma_2$ , связанные с погрешностью координат соответствующих точек на изображениях и вычисляются по (27), (29).

Определенность фундаментальной матрицы обладает следующими свойствами. Во-первых, если одно изображение может быть получено из другого проективным преобразованием, то определенность фундаментальной матрицы равна нулю. Названная ситуация имеет место, например, когда центры камер совпадают или сцена плоская и реконструкция невозможна. Во-вторых, в среднем определенность фундаментальной матрицы растет с увеличением базы съемки. Таким образом, введенная мера является полуквантитативной оценкой того, насколько точно можно вычислить структуру сцены по паре кадров без непосредственного ее вычисления.

### 2.2 Обоснование

Фундаментальная матрица связана с геометрией съемки и матрицами внутренней калибровки камер следующим образом [8] (р. 246):

$$\mathbf{F} = \mathbf{K}_2^{-T} [\vec{t}]_{\times} \mathbf{R} \mathbf{K}_1^{-1} \quad (7)$$

В (7),  $\mathbf{K}_1, \mathbf{K}_2$  - матрицы  $3 \times 3$  внутренних параметров камер,  $\mathbf{R}$  - ориентация второй камеры относительно первой,  $\vec{t}$  - сдвиг от фокуса первой камеры к фокусу второй. Матрица  $[\vec{t}]_{\times}$  имеет вид:

$$[\vec{t}]_{\times} = \begin{pmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{pmatrix} \quad (8)$$

**Теорема 1** 1. Если координаты точек связаны проективным преобразованием, то размерность нуль-пространства матрицы  $\mathbf{Z}$  не менее трех. 2. Если центры камер совпадают, то элементам нуль-пространства соответствует множество фундаментальных матриц, у которых вектор  $\vec{t}$  в факторизации (7) может быть произвольным.

Согласно условию теоремы координаты соответствующих точек связаны проективным преобразованием:

$$\vec{y}_i = \mathbf{H} \vec{x}_i \quad (9)$$

Подставив (9) в (1):

$$\vec{x}_i^T \mathbf{H}^T \mathbf{F} \vec{x}_i = 0, \forall i = 1 \dots N \quad (10)$$

Очевидно, что если

$$\mathbf{H}^T \mathbf{F} = \mathbf{A} \quad (11)$$

где  $\mathbf{A}$  произвольная антисимметричная матрица, то (10) выполняется для произвольных  $\vec{x}_i$ .

1. Построим три взаимно ортогональных вектора  $\vec{f}_k$ ,  $k = 1, 2, 3$ , лежащих в нуль-пространстве матрицы  $\mathbf{Z}$ . Для этого рассмотрим матрицу  $\mathbf{H}^{-1} \mathbf{H}^{-T}$ . Так как она симметрична, то может быть представлена в виде:

$$\mathbf{H}^{-1} \mathbf{H}^{-T} = \mathbf{Q} \mathbf{S} \mathbf{Q}^T \quad (12)$$

где собственные векторы матрицы  $\mathbf{H}^{-1} \mathbf{H}^{-T}$  являются столбцами матрицы  $\mathbf{Q}$  и обозначаются через  $\vec{q}_k$ , а ее собственные значения сосредоточены на диагонали  $\mathbf{S}$ . Определим матрицы  $\mathbf{A}_k$ ,  $\mathbf{F}_k$ :

$$\mathbf{A}_k = [\vec{q}_k]_{\times} \quad (13)$$

$$\mathbf{F}_k = \mathbf{H}^{-T} \mathbf{A}_k \quad (14)$$

Положим векторы  $\vec{f}_k$  соответствующими матрицам  $\mathbf{F}_k$  из (14) согласно (2). Векторы  $\vec{f}_k$  лежат в нуль-пространстве матрицы  $\mathbf{Z}$ , так как (10) выполняется для  $\mathbf{F}$ , равных  $\mathbf{F}_k$  из (14). Покажем, что  $\vec{f}_k$  взаимно ортогональны.

Из (2), используя (14), следует:

$$\begin{aligned} (\vec{f}_k, \vec{f}_l) &= \text{tr} \left( \mathbf{F}_k^T \mathbf{F}_l \right) = -\text{tr} \left( \mathbf{A}_k \mathbf{H}^{-1} \mathbf{H}^{-T} \mathbf{A}_l \right) = \\ &= -\text{tr} \left( \mathbf{A}_l \mathbf{A}_k \mathbf{H}^{-1} \mathbf{H}^{-T} \right), \quad k, l = 1, 2, 3 \end{aligned} \quad (15)$$

Подставив (13), (12) в (15), и, используя ортогональность векторов  $\vec{q}_k$ , получаем, например, для  $k = 1, l = 2$

$$\begin{aligned} (\vec{f}_1, \vec{f}_2) &= -\text{tr} \left( [\vec{q}_2]_{\times} [\vec{q}_1]_{\times} \mathbf{Q} \mathbf{S} \mathbf{Q}^T \right) = \\ &= -\text{tr} \left( [\vec{q}_2]_{\times} (\vec{0} |\vec{q}_3| - \vec{q}_2) \mathbf{S} \mathbf{Q}^T \right) = \\ &= -\text{tr} \left( (\vec{0} |\vec{q}_1| \vec{0}) \mathbf{S} \mathbf{Q}^T \right) = -\text{tr} \left( \mathbf{Q}^T (\vec{0} |\vec{q}_1| \vec{0}) \mathbf{S} \right) = \\ &= -\text{tr} \left( \vec{0} \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} \mathbf{S} \right) = 0 \end{aligned} \quad (16)$$

Аналогично, ортогональны и остальные пары  $\vec{f}_k$ ,  $\vec{f}_l$  для несовпадающих  $k, l$ .

Таким образом, построены три взаимно ортогональных вектора  $\vec{f}_k$ , в нуль-пространстве матрицы  $\mathbf{Z}$  и соответствующие им фундаментальные матрицы  $\mathbf{F}_k$ . Следовательно, размерность нуль-пространства  $\mathbf{Z}$  не менее трех.

2. Пусть центры камер совпадают. Обозначим через  $\mathcal{N}$  трехмерное пространство, натянутое на векторы

$\vec{f}_k, k = 1, 2, 3$ . Легко показать, что матрица  $\mathbf{H}$  имеет вид:

$$\mathbf{H} = \mathbf{K}_2 \mathbf{R} \mathbf{K}_1^{-1} \quad (17)$$

Согласно (7):

$$[\vec{t}]_{\times} = \mathbf{K}_2^T \mathbf{F} \mathbf{K}_1 \mathbf{R}^T \quad (18)$$

Выражая  $\mathbf{F}$  из (11) и подставляя в (18), получим:

$$\begin{aligned} [\vec{t}]_{\times} &= \mathbf{K}_2^T \mathbf{H}^{-T} \mathbf{A} \mathbf{K}_1 \mathbf{R}^T = \\ &= \mathbf{K}_2^T \mathbf{K}_2^{-T} \mathbf{R} \mathbf{K}_1^T \mathbf{A} \mathbf{K}_1 \mathbf{R}^T = \mathbf{R} \mathbf{K}_1^T \mathbf{A} \mathbf{K}_1 \mathbf{R}^T \end{aligned} \quad (19)$$

В (19) матрица  $\mathbf{A}$  является произвольной антисимметричной матрицей, и это же относится к матрице  $[\vec{t}]_{\times}$ , так как для произвольной  $[\vec{t}]_{\times}$  найдется матрица  $\mathbf{A} = \mathbf{K}_1^{-T} \mathbf{R}^T [\vec{t}]_{\times} \mathbf{R} \mathbf{K}_1^{-1}$ , согласно (19). Таким образом, вектор  $\vec{t}$  в факторизации (7) фундаментальной матрицы может быть любым ■

**Результат 1** Если центры камер совпадают или точки съемки расположены на плоскости, то определенность фундаментальной матрицы равна нулю.

Из условий теоремы следует, что точки связаны проективным преобразованием, тогда согласно теореме 1 размерность нуль-пространства матрицы  $\mathbf{Z}$  не менее трех:

$$\sigma_1 = \sigma_2 = \sigma_3 = 0 \quad (20)$$

и согласно (6) определенность фундаментальной матрицы равна нулю ■

**Теорема 2** Погрешность фундаментальной матрицы обратно пропорциональна разности  $\sigma_2 - \sigma_1$ .

Используя результат [11] о том, что возмущение угла поворота сингулярного вектора обратно пропорционально разности между соответствующим сингулярным числом и ближайшим к нему сингулярным числом (sinθ-теорема), получаем:

$$\frac{\|\delta \mathbf{F}\|}{\|\mathbf{F}\|} \sim \frac{\|\delta \mathbf{Z}\|}{\sigma_2 - \sigma_1} \quad (21)$$

где  $\|\delta \mathbf{F}\|$  - возмущение фундаментальной матрицы, связанное с возмущением компонент матрицы  $\mathbf{Z}$ , т.е. погрешностями координат соответствующих точек на изображениях. Полагая норму фундаментальной матрицы равной единице, получаем:

$$\frac{\|\delta \mathbf{F}\|}{\|\mathbf{F}\|} \sim \|\delta \mathbf{Z}\| \quad (22)$$

Используя (6) получаем, что погрешность фундаментальной матрицы обратно пропорциональна введенной мере ее определенности ■

**Результат 2** Погрешность реконструкции обратно пропорциональна определенности фундаментальной матрицы с коэффициентом, зависящим от сцены и положения камер:

$$\epsilon \sim \frac{C}{q(\mathbf{F})} \quad (23)$$

В (23)  $\epsilon$  - относительная ошибка реконструкции на паре кадров,  $q(\mathbf{F})$  - определенность фундаментальной

матрицы, константа  $C$  зависит от сцены и расположения камер.

Используя результаты (26), (10) работы [10] заключаем, что погрешность реконструированных точек прямо пропорциональна погрешности фундаментальной матрицы. И, используя, результат теоремы 2 заключаем, что погрешность реконструированных по паре кадров точек обратно пропорциональна степени определенности фундаментальной матрицы матрицы между этой парой кадров. Численные значения коэффициента  $C$  обсуждаются в конце подраздела 2.3 ■

Остаток настоящего раздела посвящен выводу формул для вычисления  $\langle \delta\sigma_n^2 \rangle$ , используемых в (6).

Из-за конечной точности детектора особенностей их координаты подвержены шуму. Тогда элементы матрицы  $\mathbf{Z}$  также подвержены возмущениям  $\delta Z_{ij}$ . Отклонение  $\delta\sigma_n$  сингулярного числа  $\sigma_n$  записывается в виде:

$$\delta\sigma_n = \sum_{i,j=(1,1)}^{i,j=(N,9)} \frac{\partial\sigma_n}{\partial Z_{ij}} \delta Z_{ij} \quad (24)$$

Средний квадрат отклонения сингулярных чисел  $\langle \delta\sigma_n^2 \rangle$  записывается в виде:

$$\begin{aligned} \langle \delta\sigma_n^2 \rangle &= \left\langle \sum_{i,j=1,1}^{i,j=N,9} \frac{\partial\sigma_n}{\partial Z_{ij}} \delta Z_{ij} \sum_{k,l=1,1}^{N,9} \frac{\partial\sigma_n}{\partial Z_{kl}} \delta Z_{kl} \right\rangle = \\ &= \sum_{i,j,k,l=1,1,1,1}^{N,9,N,9} \frac{\partial\sigma_n}{\partial Z_{ij}} \frac{\partial\sigma_n}{\partial Z_{kl}} \langle \delta Z_{ij} \delta Z_{kl} \rangle \quad (25) \end{aligned}$$

В [12] показывается, что производные сингулярных чисел произвольной матрицы по элементам этой матрицы записываются в виде:

$$\frac{\partial\sigma_k}{\partial Z_{ij}} = U_{ik} V_{jk}, \quad k = 1 \dots 9 \quad (26)$$

где  $U_{ik}$ ,  $V_{jk}$  - элементы матриц  $\mathbf{U}$ ,  $\mathbf{V}$  в сингулярном разложении (5). Используя (26) и независимость отклонения координат различных точек:

$$\begin{aligned} \langle \delta\sigma_n^2 \rangle &= \sum_{i,j,k,l=1,1,1,1}^{N,9,N,9} U_{in} V_{jn} U_{kn} V_{ln} \langle \delta Z_{ij} \delta Z_{kl} \rangle = \\ &= \sum_{i=1}^N U_{in}^2 \vec{v}_n^T \mathbf{W}_i \vec{v}_n \quad (27) \end{aligned}$$

В (27) вектор  $\vec{v}_n$  -  $n$ -ый столбец матрицы  $\mathbf{V}$ . Матрица  $\mathbf{W}_i$  имеет размеры  $9 \times 9$  и определяется моделью ошибки координат  $i$ -ой пары соответствующих точек, а ее элементы имеют вид:

$$W_{i,kl} \equiv \langle \delta Z_{ik} \delta Z_{il} \rangle \quad (28)$$

Если положить, что координаты всех точек подвержены независимому нормальному шуму с дисперсией  $s^2$ , то из (3) матрицы  $\mathbf{W}_i$  симметричны и имеют вид (29)

### 2.3 Численные эксперименты

В настоящем разделе рассматриваются численные эксперименты, проведенные для исследования

определенности фундаментальной матрицы. Эксперименты проводились для различных конфигураций точек в пространстве и траекторий движения камеры.

#### Конфигурации точек

1. Точки равномерно распределены внутри сферы с центром в начале координат и радиусом 1.
2. Точки равномерно распределены внутри эллипсоида с центром в начале координат и полуосями (1,1,0.25).
3. Точки равномерно распределены внутри эллипсоида с центром в начале координат и полуосями (0.25,1,1).

**Траектории движения камеры.** Перемещения камеры происходят в плоскости  $xz$ .

1. Боковое движение. Камера движется от точки (0,0,-4) до точки (3,0,-4) вдоль прямой; оптическая ось остается параллельной.
2. Круговое движение. Камера движется от точки (0,0,-4) к точке (4,0,0) вдоль окружности, с центром в начале координат и проходящей через начальную и конечные точки; оптическая ось камеры проходит через начало координат.
3. Движение к сцене. Камера движется от точки (0,0,-4) до точки (0,0,-1) вдоль прямой; оптическая ось остается параллельной.

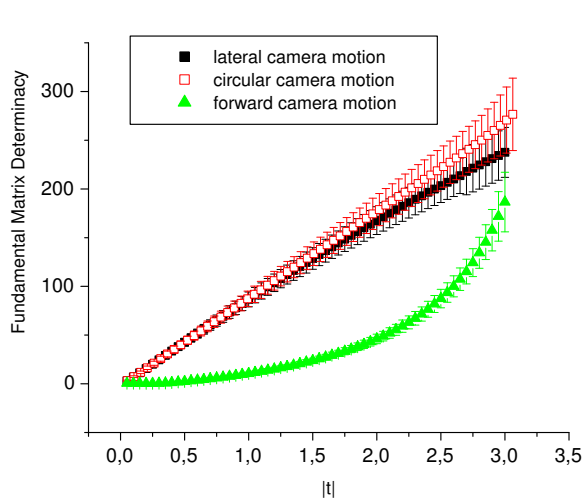
Рассматриваются именно названные траектории движения камер, так как они моделируют часто встречающиеся в реальности последовательности фотографий, получаемых при приближении к объекту или при его круговом обходе.

На Рис.1 представлены зависимости ОФМ для различных траекторий движения камеры и конфигурации точек 1. Для всех представленных результатов число точек равно 100. Результаты экспериментов усреднялись по 100 случайным конфигурациям точек. Пределы ошибки на рисунках соответствуют стандартному отклонению при усреднении. Матрицы внутренних параметров  $\mathbf{K}_1$ ,  $\mathbf{K}_2$  обеих камер полагались единичными. К проекциям точек добавлялся случайный шум со стандартным отклонением  $10^{-3}$ .

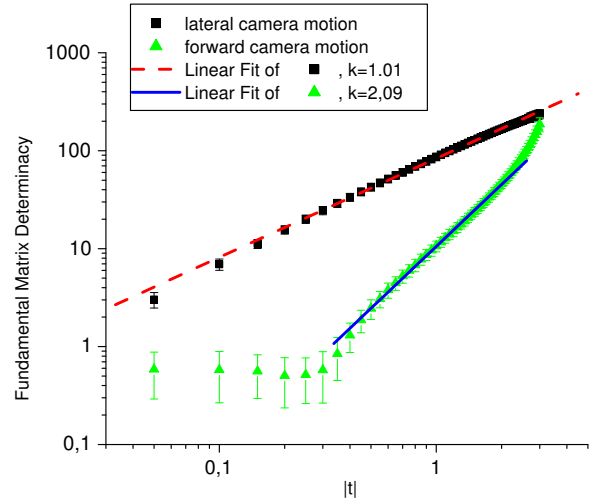
На Рис.1а видно, что зависимости для траекторий 1 и 2 очень похожи, поэтому в дальнейшем рассматриваются лишь траектории 1 и 3. На Рис.1б зависимости представлены в двойном логарифмическом масштабе. Следует отметить, что в широких пределах изменения ОФМ ( $\gtrsim 3$ ) выдерживается степенная зависимость: для траектории 1 - линейная, для траектории 3 - квадратичная.

На Рис.2 представлены зависимости средней относительной погрешности реконструированных точек от ОФМ. Каждая точка соответствует одному эксперименту. Для эксперимента задается положение камер и генерируется облако точек. Вычисляются проекции точек, к ним добавляется шум. Из фундаментальной матрицы вычисляются структура сцены [13] и матрицы камер [14]. Далее ищется преобразование, совмещающее реконструированные

$$W_i = s^2 \begin{pmatrix} x_{i,1}^2 + y_{i,1}^2 & y_{i,1}y_{i,2} & y_{i,1}y_{i,3} & x_{i,1}x_{i,2} & 0 & 0 & x_{i,1}x_{i,3} & 0 & 0 \\ - & x_{i,1}^2 + y_{i,2}^2 & y_{i,2}y_{i,3} & 0 & x_{i,1}x_{i,2} & 0 & 0 & x_{i,1}x_{i,3} & 0 \\ - & - & y_{i,3}^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ - & - & - & x_{i,2}^2 + y_{i,1}^2 & y_{i,1}y_{i,2} & y_{i,1}y_{i,3} & x_{i,2}x_{i,3} & 0 & 0 \\ - & - & - & - & x_{i,2}^2 + y_{i,2}^2 & y_{i,2}y_{i,3} & 0 & x_{i,2}x_{i,3} & 0 \\ - & - & - & - & - & y_{i,3}^2 & 0 & 0 & 0 \\ - & - & - & - & - & - & x_{i,3}^2 & 0 & 0 \\ - & - & - & - & - & - & - & x_{i,3}^2 & 0 \\ - & - & - & - & - & - & - & - & 0 \end{pmatrix} \quad (29)$$



(a) Линейный масштаб



(b) Двойной логарифмический масштаб. Для бокового движения зависимость примерно линейная, для движения вперед - примерная квадратичная

**Figure 1:** Зависимость ОФМ от модуля сдвига камеры для различных конфигурация точек и траекторий движения камеры

точки с исходными, и вычисляется средняя относительная погрешность реконструированных точек.

Результаты численных экспериментов показывают, что:

1. Ошибка реконструкции не бывает маленькой при маленькой ОФМ, и большое значение ОФМ не гарантирует маленькой ошибки. Это значит, что большое значение ОФМ является необходимым, но недостаточным условием малой ошибки.
2. Выдерживается грубая обратная зависимость между ОФМ и относительной ошибкой реконструкции в соответствии с результатом 2.
3. Для большинства сцен коэффициент пропорциональности  $C$  лежит в диапазоне  $C \in [0.1, 1]$

### 3. УПОРЯДОЧИВАНИЕ ИЗОБРАЖЕНИЙ

Упорядочивание изображений состоит в определении последовательности, согласно которой вычисляются структура сцены и матрицы камер. Для этого строится взвешенный граф  $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ , вершины

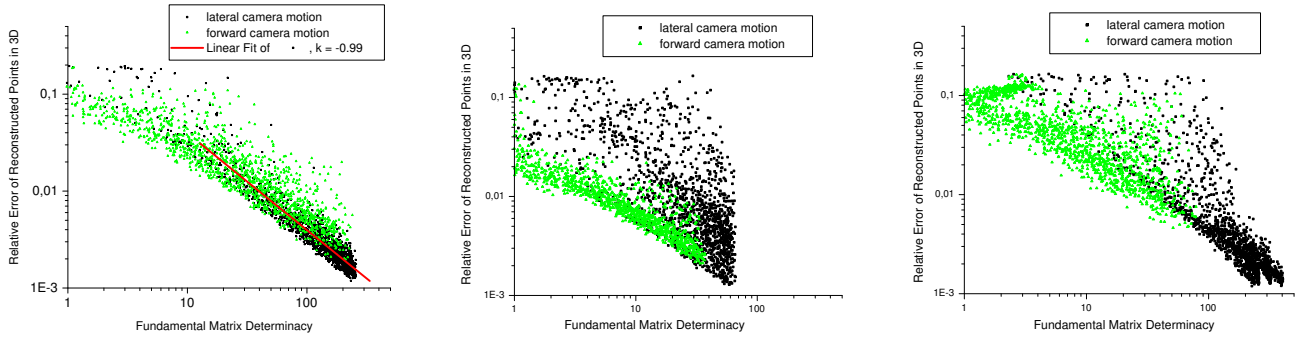
$\{v_i\}$  которого соответствуют изображениям. Пару вершин соединяет ребро, если между соответствующими вершинам изображениями вычислена фундаментальная матрица. Вес ребра  $w(v_i, v_j)$  между парой вершин  $v_i, v_j$  полагается равным определению соответствующей фундаментальной матрицы. Сначала обрабатывается пара изображений, соответствующих вершинам  $v_1, v_2$ , потом добавляется  $v_3$  и т.д. Назначение номеров вершинам выполняется следующим образом.

#### Алгоритм упорядочивания изображений

- Выбрать пару вершин, между которыми вес ребра максимален. Назначить этим вершинам номера 1 и 2.
- Следующий номер назначается той вершине, для которой сумма весов ребер из этой вершины к уже выбранным максимална среди пронумерованных вершин:

$$\sum_{k=1}^{i-1} w(v_k, v_i) \geq \sum_{k=1}^{i-1} w(v_k, v), \quad \forall v \in \{\mathcal{V} \setminus \{v_1, \dots, v_i\}\}$$

Отметим, что в [2] выбирается изображение, для которого число соответствий между ним и предыдущим



(a) Конфигурация точек 1. Наклон прямой, проведенной через точки в двойном логарифмическом масштабе, означает примерно обратную зависимость точности реконструкции от ОФМ

(b) Конфигурация точек 2

(c) Конфигурация точек 3

**Figure 2:** Зависимость относительной ошибки реконструкции координат точек в пространстве от ОФМ для различных конфигураций точек и траекторий движения камеры

максимально. Предлагаемый алгоритм упорядочивания похож на [2], но отличается способом назначения весов ребрам графа.

## 4. РЕЗУЛЬТАТЫ

### 4.1 Синтезированные данные

Генерируется облако из  $N = 100$  точек внутри единичной сферы с центром в начале координат. Генерируются положения 30 камер внутри сферы радиусом 4 и вне единичной сферы с центрами в начале координат. Оптические оси камер проходят через начало координат. Количество соответствий между парой кадров  $i, j$

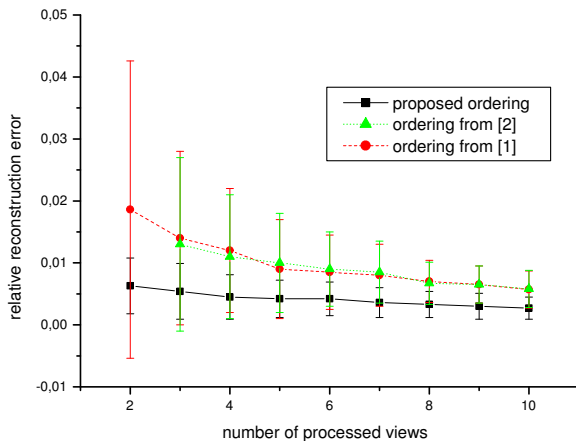
задается формулой:

$$N_{ij} = N \left(1 - \frac{\alpha_{ij}}{\alpha_0}\right) \left(1 - \frac{|d_i - d_j|}{d_i + d_j}\right) \quad (30)$$

В (30)  $\alpha_{ij}$  - угол между оптическими осями камер  $i$  и  $j$ , а  $d_i, d_j$  - расстояние от фокусов камер до начала координат,  $\alpha_0 = \frac{\pi}{3}$ . Зависимость (30) числа соответствий моделирует уменьшение их числа при увеличении угла съемки и масштаба между кадрами. К проекциям точек добавляется шум и, используя (30), вычисляются фундаментальные матрицы между кадрами, их меры определенности и строится последовательность реконструкции.

На Рис.3 приведена зависимость относительной ошибки реконструкции от числа обработанных кадров для предлагаемого способа построения последовательности и способа из [2]. Погрешность соответствует стандартному отклонению результата при усреднении. По вертикальной оси отложена погрешность реконструкции. По горизонтальной оси отложено число кадров, которое обработано согласно соответствующей последовательности. Повторим, что общее число кадров одинаково для всех экспериментов и всех способов упорядочивания, изменяется лишь число кадров, после обработки которых вычисляется погрешность реконструкции.

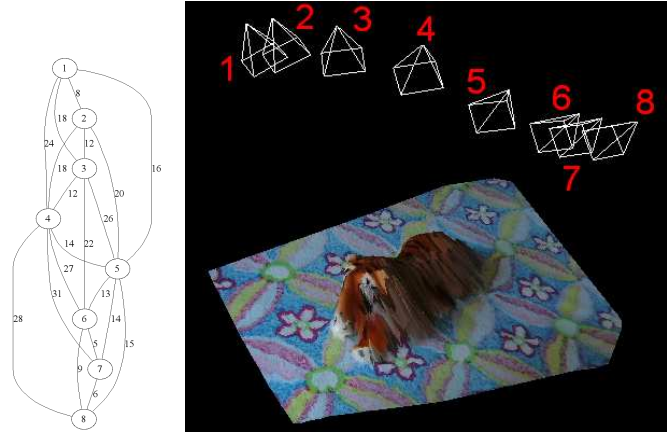
Для предлагаемого способа упорядочивания изображений относительная ошибка и ее разброс меньше, чем в случае упорядочивания, предложенного в [2], [1] особенно в начале процесса реконструкции, когда обработано мало (2-4) кадров. Это связано с тем, что эти алгоритмы выполняют первичную реконструкцию по кадрам, полученным для малой базы. Предлагаемый алгоритм сначала обрабатывает изображения, соответствующие более удаленным камерам, для которых число соответствий меньше. И так как погрешность реконструкции уменьшается с увеличением базы, то предлагаемый подход к построению последовательности реконструкции приводит к более точному результату.



**Figure 3:** Зависимость относительной ошибки реконструкции координат точек от числа обработанных кадров



(a) Пример нескольких изображений из набора восьми снимков



(b) Граф, (c) восстановленная модель и положения соответствующий камер. Числа около камер обозначают порядок, последовательности которым изображения были получены. на Рис.4а Найденная последовательность реконструкции:  $S = \{4, 7, 8, 5, 6, 3, 1, 2\}$ .

**Figure 4:** Результаты применения предлагаемого алгоритма упорядочивания изображений

#### 4.2 Реальные данные

Пример последовательности фотографий, использованных для реконструкции приведен на Рис.4а. На Рис.4б приведен граф изображений. Вершины графа соответствуют изображениям, вес ребра равен определенности фундаментальной матрицы между изображениями. Сначала степень определенности фундаментальной матрицы между изображениями растет с увеличением расстояния между камерами, а потом уменьшается из-за уменьшения числа соответствий.

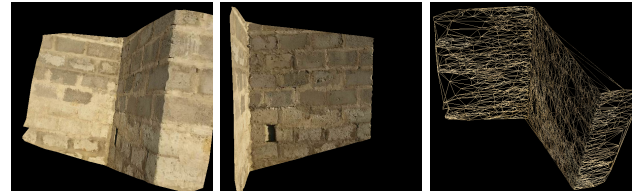
Сцена сначала вычисляется по паре кадров из середины (4) и конца (7) последовательности. Далее добавляется еще изображение из конца последовательности (8), так как, несмотря на близость камер в положениях 7 и 8, велика база между кадрами между 4 и 8. Далее продолжают добавляться изображения из второй половины последовательности (5,6) и лишь потом обрабатывается начало последовательности (3,2,1).

На Рис.4с представлена построенная модель с наложенной текстурой.

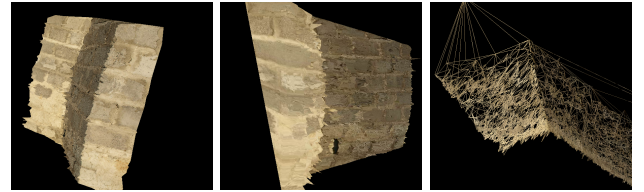
На Рис.5а приведена последовательность из трех фотографий, для которой последовательность реконструкции сильно влияет на погрешность. А именно, база съемки между первыми двумя кадрами значительно меньше, чем между первым (вторым) изображением и третьим. На Рис.5б представлена модель, полученная для предлагаемой последовательности реконструкции. На Рис.5с представлена модель, полученная для последовательности реконструкции, используемой в [2]. Можно видеть, что предлагаемый алгоритм выбора последовательности приводит к более точному



(a) Набор изображений



(b) Трехмерная модель (сеточная и текстурированная) с разных ракурсов, полученная в соответствии с предлагаемым алгоритмом выбора последовательности реконструкции.



(c) Трехмерная модель (сеточная и текстурированная) с разных ракурсов, полученная в соответствии алгоритмом [2] выбора последовательности реконструкции.

**Figure 5:** Пример набора изображений, для которого разные последовательности реконструкции приводят к существенно разной погрешности. База съемки между первыми двумя кадрами значительно меньше, чем между первым (вторым) изображением и третьим.

результату, чем в [2].

#### 5. ЗАКЛЮЧЕНИЕ

В работе рассматривается задача последовательного построения 3d модели по фотографиям. Во-первых, вводится мера определенности фундаментальной матрицы. Полученные результаты показывают, что эта мера может служить оценкой достижимой точности реконструкции по паре кадров. Большое значение определенности фундаментальной матрицы является необходимым условием для маленькой погрешности реконструкции. Во-вторых, введенная мера используется для построения графа, вершины которого соответствуют изображениям, а вес ребра равен определенности фундаментальной матрицы между соответствующей парой изображений. Предлагается алгоритм выбора последовательности реконструкции. Полученные результаты численных экспериментов и обработка реальных изображений показывают, что предлагаемый алгоритм упорядочивания изображений приводит к более точным и стабильным результатам. Предложена априорная оценка (23) точности реконструкции по двум кадрам.

#### 6. БЛАГОДАРНОСТИ

Работа выполнена при поддержке грантов РФФИ 09-01-92470-МНКС\_а, 09-07-92000-НС\_а.

## 7. REFERENCES

- [1] Richard I. Hartley. Euclidean reconstruction from uncalibrated views. In *Proceedings of the Second Joint European - US Workshop on Applications of Invariance in Computer Vision*, pages 237–256, London, UK, 1994. Springer-Verlag.
- [2] Frederik Schaffalitzky and Andrew Zisserman. Multi-view matching for unordered image sets, or "how do i organize my holiday snaps?". In *ECCV '02: Proceedings of the 7th European Conference on Computer Vision-Part I*, pages 414–431, London, UK, 2002. Springer-Verlag.
- [3] P. Beardsley, A. Zisserman, and D. Murray. Sequential updating of projective and affine structure from motion, 1997.
- [4] Andrew W. Fitzgibbon and Andrew Zisserman. Automatic camera recovery for closed or open image sequences. In *ECCV (1)*, pages 311–326, 1998.
- [5] Bill Triggs, Philip McLauchlan, Richard Hartley, and Andrew Fitzgibbon. Bundle adjustment – A modern synthesis. pages 298–375, 2000.
- [6] Tony Lindeberg. Feature detection with automatic scale selection. *International Journal of Computer Vision*, 30(2):77–116, 1998.
- [7] David G. Lowe. Object recognition from local scale-invariant features. In *Proc. of the International Conference on Computer Vision ICCV, Corfu*, pages 1150–1157, 1999.
- [8] Richard Hartley and Andrew Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2003.
- [9] M.A. Fischler and R.C. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, June 1981.
- [10] Gabriella Csurka, Cyril Zeller, Zhengyou Zhang, and Olivier D. Faugeras. Characterizing the uncertainty of the fundamental matrix. *Computer Vision and Image Understanding: CVIU*, 68(1):18–36, 1997.
- [11] Chandler Davis and W. M. Kahan. The rotation of eigenvectors by a perturbation. *SIAM J.*, 7(1):1–46, march 1970.
- [12] Théodore Papadopoulo and Manolis I. A. Lourakis. Estimating the jacobian of the singular value decomposition: Theory and applications. In *ECCV '00: Proceedings of the 6th European Conference on Computer Vision-Part I*, pages 554–570, London, UK, 2000. Springer-Verlag.
- [13] A Zisserman P A Beardsley and D W Murray. Navigation using affine structure from motion. In *Proc 3rd European Conf on Computer Vision, Stockholm*, Lecture Notes in Computer Science, pages 85–96. Springer, 1994.
- [14] Richard I. Hartley. Estimation of relative camera positions for uncalibrated cameras. In *ECCV '92: Proceedings of the Second European Conference on Computer Vision*, pages 579–587, London, UK, 1992. Springer-Verlag.

## ОБ АВТОРАХ

Волегов Дмитрий Борисович - аспирант Московского Физико-Технического Института факультета Общей и Прикладной Физики. [dvolegov@rambler.ru](mailto:dvolegov@rambler.ru).

Юрин Дмитрий Владимирович, к.ф.-м.н., с.н.с. лаборатории Математических методов обработки изображений факультета Вычислительной математики и кибернетики Московского государственного университета им. М.В.Ломоносова. [yurin\\_d@inbox.ru](mailto:yurin_d@inbox.ru)

## ORDERING IMAGE SET TO RECONSTRUCT 3D MODELS

### Abstract

The problem of reconstruction of 3d models from unordered image set is considered. To begin a small number of images is processed and initial model is built, then other images are processed and the model is refined. Firstly the problem of image ordering is considered. Secondly the original measure "determinacy of fundamental matrix" is introduced. This measure allows easily to estimate the precision of reconstruction without actually doing it. The results for real and synthetic sequences are presented.

**Keywords:** *image ordering, 3d reconstruction, fundamental matrix* .

### ABOUT THE AUTHOR

Dmitriy B. Volegov is a Ph.D student at Moscow institute of Physics and Technology, Department of General and Applied Physics. His contact email is [dvolegov@rambler.ru](mailto:dvolegov@rambler.ru)

Dmitry V. Yurin, PhD, is a senior scientist at Laboratory of Mathematical Methods of Image Processing, Chair of Mathematical Physics, Faculty of Computational Mathematics and Cybernetics, Moscow Lomonosov State University. His contact email is [yurin\\_d@inbox.ru](mailto:yurin_d@inbox.ru)