

Алгоритмы факторизации: достоверность результата и применение для восстановления эпиполярной геометрии

Н.В. Свешникова*, Д.В. Юрин**

*Московский физико-технический институт, Москва, Россия
sveshnikova_n@list.ru

** ФГУП НПП ОПТЭКС, Зеленоград, Россия
yurin_d@inbox.ru

Аннотация

Для алгоритмов восстановления трехмерных сцен, основанных на факторизации матриц, предлагается вывод априорных и апостериорных оценок погрешностей восстановления трехмерной сцены, ориентаций камер и расстояния до них, не использующий ограничений на форму сцены или способ движения камеры вокруг объекта. Высокая степень достоверности результата выбранной группы алгоритмов и наличие вычисляемых во время работы оценок апостериорных ошибок, позволяют использовать их выходные данные для достоверного вычисления фундаментальной матрицы и эпиполярных линий, широко используемых в стерео алгоритмах.

Ключевые слова: Восстановление трехмерных сцен, метод факторизации, погрешности восстановления, эпиполярная геометрия.

1. ВВЕДЕНИЕ

В условиях огромного разнообразия существующих решений для задачи восстановления трехмерных сцен ограниченная применимость любого подхода является посылкой к выводу о необходимости комбинирования различных алгоритмов в единую систему. Однако структурирование подобной системы невозможно без четко определенных критериев применимости интегрируемых алгоритмов и без оценок точностей их результата. При этом оценки точностей требуются как априорные, использующие предварительные оценки интегральных характеристик формы сцены и движения камер относительно нее, так и апостериорные, основанные на данных измерений, поступающих на вход алгоритма. В последнем случае может и не быть информации о трехмерной сцене и способе съемки.

Этой проблеме была посвящена работа [1]. В ней были рассмотрены два метода, основанные на факторизации матриц: приближение масштабированной ортографической проекции (МОП) [2] и итерационный метод в перспективной проекции (ИПП) [3]. Результатом проведенного в [1] анализа стала методика оценки априорной и апостериорной погрешностей восстановления глубины сцены и ориентации камер. Однако в [1] вывод оценок строился на гипотезе, связанной с приближением сцены малой глубины. Проверка полученных зависимостей была также проведена на модели с малой глубиной.

В представленной работе для алгоритмов МОП и ИПП предлагается аналитический вывод априорных и апостериорных оценок погрешностей восстановления трехмерной сцены, ориентаций камер, а также погрешностей расстояния до камер, не использующий ограничений.

Основной принцип, на котором базируется построение оценок погрешностей, как и в [1], аналогичен, по сути, оценке статистической погрешности измерений. В данном случае, роль ограничений, накладываемых теоретической зависимостью, выполняют ограничения, накладываемые законами проективной геометрии на возможные изображения жесткого трехмерного объекта.

Проверка полученных оценок осуществлялась на модельных данных и показала высокую точность. Используемая в данной работе модель представляет собой объект, характерные размеры которого могут выбираться произвольно. Способ движения камер вокруг сцены также можно менять.

Проведенный анализ погрешностей и высокая степень достоверности результата выбранной группы алгоритмов, позволяет использовать их выходные данные в качестве входных для других алгоритмов. Особый интерес представляет вычисленные фундаментальной матрицы и эпиполярных линий [4, 5, 11], широко используемых в стерео алгоритмах, которые способны улучшить результат восстановления сцены алгоритмами факторизации, т.к. восстанавливают трехмерные координаты практически всех пикселей изображения.

2. ОБЗОР ЗАДАЧИ ВОССТАНОВЛЕНИЯ СЦЕНЫ

Кратко изложим структуру задачи восстановления трехмерной сцены, под которой будем подразумевать предмет, который не деформируется во время получения последовательности его изображений, или группу таких предметов, неподвижных относительно друг друга. Случай движущихся объектов рассмотрен в [8].

Восстановление трехмерной сцены по последовательности цифровых изображений состоит в простейшем случае из трех последовательных шагов.

1. Получение последовательности F изображений в результате съемки трехмерной сцены при различных взаимных положениях камер(ы) и сцены.
2. Обработка последовательности F изображений, программой обнаружения и прослеживания характеристических точек (например, подходы [6, 7]).
3. Восстановления трехмерной сцены алгоритмами факторизации [1, 2, 3, 9]

Результатом работы программы обнаружения и прослеживания характеристических отметок является набор списков двумерных координат для P характеристических точек на всех кадрах, где они были обнаружены. Его удобно представлять в виде матрицы:

$$\mathbf{W} = \begin{pmatrix} u_{11} & u_{12} & \dots & u_{1P} \\ v_{11} & v_{12} & \dots & v_{1P} \\ \vdots & \vdots & \dots & \vdots \\ u_{F1} & u_{F2} & \dots & u_{FP} \\ v_{F1} & v_{F2} & \dots & v_{FP} \end{pmatrix} \quad (1)$$

Матрица \mathbf{W} поступает на вход одного из алгоритмов, основанного на факторизации. В результате его работы в некоторой системе координат восстанавливаются трехмерные координаты всех камер, их пространственные ориентации и трехмерные координаты характеристических точек. Кроме того, алгоритм предоставляет информацию о достоверности полученного результата.

Определим используемые далее обозначения, не останавливаясь на подробном изложении принципов и структуры алгоритмов, основанных на факторизации матриц [1].

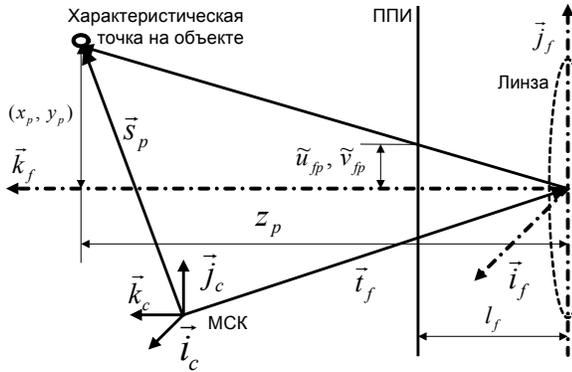


Рис. 1. Постановка задачи восстановления трехмерной сцены. Пусть имеется $f = 1..F$ изображений трехмерной сцены, полученных с различных камер, имеющих неизвестное пространственное положение \vec{t}_f в мировой системе координат (МСК). Будем считать, что на трехмерной сцене выделено $p = 1..P$ характеристических точек с координатами \vec{s}_p в МСК, $\tilde{u}_{fp}, \tilde{v}_{fp}$ - их координаты на передней плоскости изображения (ППИ) в единицах длины. Тогда, из подобия треугольников (рис. 1) можно записать формулы для перспективной проекции:

$$u_{fp} = g_f \frac{\mathbf{i}_f(\mathbf{s}_p - \mathbf{t}_f)}{\mathbf{k}_f(\mathbf{s}_p - \mathbf{t}_f)}, \quad v_{fp} = g_f \frac{\mathbf{j}_f(\mathbf{s}_p - \mathbf{t}_f)}{\mathbf{k}_f(\mathbf{s}_p - \mathbf{t}_f)}, \quad (2)$$

где $g_f = \frac{l_f}{\tilde{u}_f^e N}$, $u_{fp} = \frac{\tilde{u}_{fp}}{\tilde{u}_f^e N}$, $v_{fp} = \frac{\tilde{v}_{fp}}{\tilde{u}_f^e N}$,

Здесь $\vec{i}_f, \vec{j}_f, \vec{k}_f$ - правая тройка ортонормированных векторов, задающих систему координат, связанную с f -й камерой (КСК), вектор \vec{k}_f направлен на объект, а \vec{i}_f, \vec{j}_f лежат в плоскости изображения; l_f - фокусное расстояние камеры в

единицах длины. \tilde{u}_f^e - размер пикселя на фоточувствительной матрице; N - ширина изображения в пикселях. Безразмерные величины $g_f, u_{fp}, v_{fp} \in [-1/2, 1/2]$ соответствуют величинам в единицах длины $l_f, \tilde{u}_{fp}, \tilde{v}_{fp}$. Вводя обозначение $z_f = -(\vec{k}_f \vec{t}_f)$, $z'_f = z_f / g_f$, и полагая $(\vec{k}_f \vec{s}_p) / z_f \ll 1$, получим формулы для приближения МОП [1]:

$$\begin{cases} u_{fp} = \vec{i}_f \vec{s}_p / z'_f + t_{xf} \\ v_{fp} = \vec{j}_f \vec{s}_p / z'_f + t_{yf} \end{cases}, \quad \text{где} \quad \begin{cases} t_{xf} = -(\vec{i}_f \vec{t}_f) / z'_f \\ t_{yf} = -(\vec{j}_f \vec{t}_f) / z'_f \end{cases}, \quad (3)$$

Уравнения (3) можно представить в матричной форме:

$$\mathbf{W} = \mathbf{M}\mathbf{S} + \mathbf{T} = \mathbf{W}' + \mathbf{T}, \quad \text{где} \quad \mathbf{S} = (\vec{s}_1 \dots \vec{s}_P)^T, \quad (4)$$

$$\mathbf{M} = \begin{pmatrix} \vec{i}_1 & \vec{j}_1 & \dots & \vec{i}_F & \vec{j}_F \\ z'_1 & z'_1 & \dots & z'_1 & z'_1 \end{pmatrix}^T$$

Матрица \mathbf{W} определяется в (1), а матрица \mathbf{T} состоит из столбцов $\vec{T}_p = (t_{x1} \ t_{y1} \ \dots \ t_{xF} \ t_{yF})^T$, $p = 1..P$. Алгоритм ИПП [3] заключается в итерационном решении Уравнений (3),(4) но на каждом шаге в матрицу измерений \mathbf{W} вносится предсказание по результатам предыдущей итерации. Поэтому приводимый ниже анализ применим к ИПП.

Матрицу \mathbf{W}' будем называть идеальной, т.к. в ней не учитывается шум, который присутствует всегда в исходных данных. Зашумленную матрицу $\tilde{\mathbf{W}}'$ можно представить в виде:

$$\tilde{\mathbf{W}}' = \mathbf{W}' + \boldsymbol{\omega} \quad (5)$$

где $\boldsymbol{\omega} = [\omega_{ij}]$ - матрица шума размера $2F \times P$ такая, что математическое ожидание ее компонент равно нулю, величину среднеквадратичного отклонения обозначим ρ :

$$\langle \omega^2 \rangle = \rho^2; \quad \langle \omega \rangle = 0. \quad (6)$$

Прежде чем приступить к анализу алгоритмов и построению оценок погрешностей на основе исследования матриц \mathbf{W}' и $\tilde{\mathbf{W}}'$, определим дополнительно несколько матриц, которые будут использоваться в дальнейшем.

Рассмотрим матрицу моментов объекта $\mathbf{\Lambda}$ и матрицу моментов движения $\mathbf{\Omega}$ размера 3×3 [8]:

$$\mathbf{\Lambda} = \frac{1}{P} \mathbf{S}\mathbf{S}^T, \quad \mathbf{\Omega} = \frac{1}{F} \mathbf{M}^T \mathbf{M} \quad (7)$$

3. НОРМЫ МАТРИЦ

Будем строить оценки, используя определение матричной нормы Фробениуса [10]. Для произвольной матрицы \mathbf{A} размера $m \times n$, квадрат ее нормы выражается как сумма квадратов ее компонент или как сумма квадратов ее сингулярных чисел $\hat{\sigma}_i^2$. Стоит заметить, что величины $\hat{\sigma}_i^2$ являются собственными значениями матрицы $\mathbf{A}\mathbf{A}^T$. Вместе с тем след

квадратной матрицы есть сумма ее собственных значений, таким образом, получаем:

$$\| \mathbf{A} \|^2 = Sp \mathbf{A} \mathbf{A}^T \quad (8)$$

Для столбцов и строк матриц будем пользоваться нормой, определенной аналогично, как квадратный корень из суммы квадратов компонент.

Вычислим нормы матриц \mathbf{S} и ее строк \bar{s}_i и \mathbf{M} и ее столбцов \bar{m}_i , $i = \{x, y, z\}$, а затем нормы матриц \mathbf{W}' и $\tilde{\mathbf{W}}'$.

$$\| \mathbf{S} \|^2 = \sum_{i=\{x, y, z\}} \|\bar{s}_i\|^2, \quad \text{где} \quad \|\bar{s}_i\|^2 = \sum_{p=1}^P s_{ip}^2 \quad (9)$$

$$\| \mathbf{M} \|^2 = \sum_{f=1}^F \frac{2}{z_f'^2}, \quad \|\bar{m}_i\|^2 = \sum_{f=1}^F \frac{i_{if}^2 + j_{if}^2}{z_f'^2} = \sum_{f=1}^F \frac{1 - k_{if}^2}{z_f'^2}$$

В силу (8) выражения (9) также можно записать через диагональные компоненты $\mathbf{\Lambda}$ и $\mathbf{\Omega}$:

$$\| \mathbf{S} \|^2 = P Sp \mathbf{\Lambda}, \quad \|\bar{s}_i\|^2 = P \Lambda_{ii}, \quad \| \mathbf{M} \|^2 = 2F Sp \mathbf{\Omega}, \quad (10)$$

$$\|\bar{m}_i\|^2 = 2F \Omega_{ii}, \quad \text{где} \quad i = \{x, y, z\}$$

Для матрицы \mathbf{W}' квадрат ее нормы выражается с одной стороны как сумма квадратов всех ее компонент, а с другой как след матрицы $2FP(\mathbf{\Omega}\mathbf{\Lambda})$, т.к. собственные значения этой матрицы есть квадраты сингулярных чисел матрицы \mathbf{W}' [8]:

$$\| \mathbf{W}' \|^2 = Sp(\mathbf{W}' \mathbf{W}'^T) = 2FP Sp(\mathbf{\Omega}\mathbf{\Lambda}) \quad (11)$$

В общем случае, если выражение (11) расписывать покомпонентно, получаются довольно громоздкие и сложные для анализа формулы. Без потери общности будем рассматривать задачу системе координат (СК), оси которой совпадают с осями симметрии эллипсоида, задаваемого матрицей $\mathbf{\Lambda}$, а начало СК – с его центром, т.е. с центром масс восстанавливаемого объекта, представленного в виде набора точек. Т.е. выбираем такую СК, в которой матрица $\mathbf{\Lambda}$ диагональная:

$$\mathbf{\Lambda} = diag(\Lambda_x, \Lambda_y, \Lambda_z), \quad \Lambda_i = \frac{1}{P} \|\bar{s}_i\|^2, \quad (12)$$

$$i = \{x, y, z\}$$

Кроме того, в силу того, что начало СК совпадает с центром масс, построчное суммирование компонент матрицы \mathbf{W}' также будет равно нулю:

$$\frac{1}{P} \sum_{p=1}^P w_{fp} = 0, \quad f = 1..F \quad (13)$$

В такой системе координат матрица $(\mathbf{\Omega}\mathbf{\Lambda})$ предстает в менее громоздком виде, а след такой матрицы есть сумма произведений диагональных компонент матриц $\mathbf{\Lambda}$ и $\mathbf{\Omega}$. Используя равенства (10), можно записать:

$$2FP Sp(\mathbf{\Omega}\mathbf{\Lambda}) = \sum_{i=\{x, y, z\}} \|\bar{m}_i\|^2 \|\bar{s}_i\|^2 \quad (14)$$

Т.е. если выбрать СК такой что матрица $\mathbf{\Lambda}$ диагональная, то норма матрицы \mathbf{W}' , выражается через нормы строк матрицы

формы и столбцов матрицы движения. Т.е. из (11) и (14) получаем окончательно:

$$\| \mathbf{W}' \|^2 = \sum_{i=\{x, y, z\}} \|\bar{m}_i\|^2 \|\bar{s}_i\|^2 \quad (15)$$

Для зашумленной матрицы $\tilde{\mathbf{W}}'$, получим следующее выражение для нормы:

$$\| \tilde{\mathbf{W}}' \|^2 = Sp \mathbf{W}' \mathbf{W}'^T + Sp(\mathbf{W}' \mathbf{\omega}^T + \mathbf{\omega} \mathbf{W}'^T) + Sp \mathbf{\omega} \mathbf{\omega}^T \quad (16)$$

Величина последнего слагаемого следует из (6):

$$Sp \mathbf{\omega} \mathbf{\omega}^T = \| \mathbf{\omega} \|^2 = \sum_{f=1}^{2F} \sum_{p=1}^P w_{fp}^2 = 2FP \rho^2 \quad (17)$$

Или, используя гипотезу из [1], которая была проверена численным моделированием и показала хорошую точность, можно приблизить величину нормы шумовой матрицы величиной наибольшего из шумовых сингулярных чисел:

$$\| \mathbf{\omega} \|^2 \approx \sigma_n^2 \quad (18)$$

Формула (17) предназначена для априорных оценок и требует информации о величине шума. Формула (18) подходит для апостериорного анализа и использует шумовое сингулярное число, получаемые в ходе работы алгоритма, основанного на факторизации.

Рассмотрим приращение функционала $\Phi = \| \mathbf{W}' \|^2$ за счет шумовой добавки как усредненную разницу квадратов норм зашумленной и идеальной матриц \mathbf{W}' и $\tilde{\mathbf{W}}'$:

$$\Delta \Phi = \langle Sp \tilde{\mathbf{W}}' \tilde{\mathbf{W}}'^T - Sp \mathbf{W}' \mathbf{W}'^T \rangle \quad (19)$$

Учитывая, что $\mathbf{\omega}$ - случайная матрица, а \mathbf{W}' - регулярная, и усредненное произведение шумовой и идеальной матрицы можно представить как $\langle Sp(\mathbf{W}' \mathbf{\omega}^T) \rangle = \langle \mathbf{\omega} \rangle \sum_{f,p} w_{fp}$, что в силу (6) и (13) равно нулю. Таким образом, приращение функционала есть квадрат нормы шумовой матрицы:

$$\Delta \Phi = \langle Sp \mathbf{\omega} \mathbf{\omega}^T \rangle \quad (20)$$

Используя равенство (20), будем выражать погрешности восстановления различных параметров системы с помощью частных производных функционала Φ по параметрам системы и приближая приращение функционала дифференциалом по интересующим нас переменным.

4. ПОГРЕШНОСТИ ВОССТАНОВЛЕНИЯ

Т.к. нас интересуют среднеквадратичные погрешности, то запишем частные производные по s_{ip}^2 , k_{if}^2 и $z_f'^2$, $i = \{x, y, z\}$:

$$\frac{\partial \Phi}{\partial (s_{ip}^2)} = \|\bar{m}_i\|^2, \quad \frac{\partial \Phi}{\partial (k_{if}^2)} = -P \frac{1}{z_f'^2} \Lambda_i, \quad (21)$$

$$\frac{\partial \Phi}{\partial (z_f'^2)} = -P \frac{1}{z_f'^4} \left(\sum_{i=\{x, y, z\}} (1 - k_{if}^2) \Lambda_i \right)$$

Введем обозначения для проекций относительных средне-квадратичных погрешностей формы и ориентаций камер вдоль осей СК:

$$\varepsilon_{i_shape} = \sqrt{\left\langle \frac{d(\bar{s}_i^2)}{\bar{s}_i^2} \right\rangle}, \quad \varepsilon_{i_rotation} = \sqrt{\left\langle d(k_i^2) \right\rangle}, \quad (22)$$

Считая погрешности (22) независимыми друг от друга, будем приближать дифференциал $d\Phi$ по отдельности вдоль каждой из осей:

$$d\Phi = \varepsilon_{i_shape}^2 \|\bar{m}_i\|^2 \|\bar{s}_i\|^2$$

$$d\Phi = \varepsilon_{i_rotation}^2 \frac{\|\mathbf{M}\|^2}{2} \|\bar{s}_i\|^2$$

Общую погрешность по форме, ориентациям камер и расстоянию до них будем считать следующим образом:

$$\varepsilon_{shape} = \sqrt{\varepsilon_{x_shape}^2 + \varepsilon_{y_shape}^2 + \varepsilon_{z_shape}^2}$$

$$\varepsilon_{rotation} = \sqrt{\varepsilon_{x_rotation}^2 + \varepsilon_{y_rotation}^2 + \varepsilon_{z_rotation}^2} \quad (23)$$

$$\varepsilon_{camera_z} = \sqrt{\left\langle \left| \frac{d(z'^2)}{z'^2} \right| \right\rangle}$$

Заменяя дифференциал функционала на квадрат нормы шумовой добавки (20), получим окончательные формулы оценки погрешности восстановления формы, ориентации камер и расстояния до них:

$$\varepsilon_{shape} = \sigma_n \sqrt{\sum_{i=\{x, y, z\}} \frac{1}{\|\bar{m}_i\|^2 \|\bar{s}_i\|^2}}$$

$$\varepsilon_{rotation} = \frac{\sqrt{2}\sigma_n}{\|\mathbf{M}\|} \sqrt{\sum_{i=\{x, y, z\}} \frac{1}{\|\bar{s}_i\|^2}} \quad (24)$$

$$\varepsilon_{camera_z} = \frac{\sigma_n}{\sqrt{\sum_{i=\{x, y, z\}} \sigma_i^2}}$$

Такие формулы удобны для апостериорной оценки, когда уже выполнена факторизация исходной матрицы \mathbf{W} и найдены матрицы \mathbf{M} и \mathbf{S} . Поэтому вместо точного выражения (17) для нормы шумовой матрицы $\|\mathbf{W}\|$ в ней использовалось оценка (18). Для априорных оценок в случае, когда известны только предварительные оценки интегральных характеристик формы сцены и движения камер относительно нее, необходимо (24) преобразовать. Воспользуемся формулами (10), (12) и подставим их в (24):

$$\varepsilon_{shape} = \rho \sqrt{\frac{1}{\Omega_{xx}\Lambda_x} + \frac{1}{\Omega_{yy}\Lambda_y} + \frac{1}{\Omega_{zz}\Lambda_z}}$$

$$\varepsilon_{rotation} = \sqrt{2}\rho \sqrt{\frac{1}{Sp\mathbf{\Omega}} \sum_{i=\{x, y, z\}} \frac{1}{\Lambda_i}} \quad (25)$$

$$\varepsilon_{camera_z} = \frac{\rho}{\sqrt{\sum_{i=\{x, y, z\}} \Omega_{ii}\Lambda_i}}$$

В выбранной нами СК выражения для диагональных компонент матрицы $\mathbf{\Lambda}$ есть дисперсии вдоль осей координат относительно центра масс системы точек. Рассмотрим диагональные компоненты матрицы $\mathbf{\Omega}$:

$$\Omega_{ii} = \frac{1}{2F} \sum_{f=1}^F \frac{1-k_{if}^2}{z'_f{}^2}, \quad \text{где } z'_f = -\frac{(\bar{k}_f, \bar{t}_f)}{g}, \quad i = \{x, y, z\} \quad (26)$$

Пусть известно, что камеры распределены случайным образом по закону $\tau(\bar{t})$ на некотором множестве G (например, в объеме V , на поверхности S или вдоль кривой L). Тогда можно в формуле (26) перейти от суммы к интегралу.

Ориентации камер \bar{k}_f могут быть как зависимы от положения камеры \bar{t}_f (в случае, когда камера всегда направлена на центр масс объекта или на какую-то фиксированную точку), т.е. $\bar{k}_f = \bar{k}_f(\bar{t}_f)$, так и независимы от \bar{t}_f (в случае, когда все камеры направлены случайным образом). Второй случай реализуется редко, т.к. если камеры направлены случайным образом, характеристические отметки будут теряться от кадра к кадру. В первом случае функционалы Ω_{ii} , где $i = \{x, y, z\}$ зависят от одной случайной переменной, а именно от положения камеры \bar{t} . В общем виде, считая функцию $\tau(\bar{t})$ нормированной, получим для функционалов Ω_{ii} , выражение:

$$\Omega_{ii} = g^2 \int_G \frac{1-k_i^2(\bar{t})}{(\bar{k}(\bar{t}), \bar{t})^2} \tau(\bar{t}) dG, \quad i = \{x, y, z\} \quad (27)$$

В условиях конкретной планируемой съемки можно задать модель движения камеры вычислить интеграл (27), подставить результат в формулы (25) и получить аналитическую формулу для априорных оценок погрешностей. Пример такого расчета приведен в следующем разделе.

5. РЕЗУЛЬТАТЫ АНАЛИЗА

5.1 Модель «Точки»

Проверка полученных оценок проводилась методами численного моделирования. В среде Matlab была разработана модель «Точки», представляющая собой случайный набор точек, равномерно распределенных внутри эллипсоида (рис. 2):

$$\frac{x^2}{a_x^2} + \frac{y^2}{a_y^2} + \frac{z^2}{a_z^2} \leq 1 \quad (28)$$

Камеры располагаются случайно равномерно в слое толщины h сектора сферы радиуса R , образованного центральным четырехгранным углом $\alpha \times \alpha$. Центр сферы расположен в начале СК. Все камеры ориентированы на центр масс объекта. Положение сектора окружности в пространстве (ориента-

ция четырехгранного угла) определяется генеральным направлением, обозначенном на рис. 2 вектором \vec{B} .

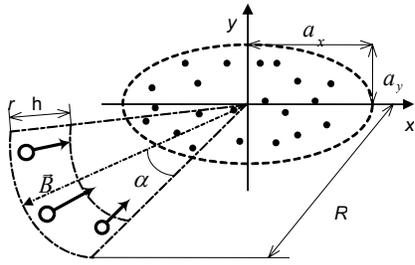


Рис. 2. Модель «Точки»

Проверка полученных оценок строилась по следующей схеме:

1. Построение трехмерной модели в среде MatLab в виде множества характеристических точек.
2. Вычисление изображения модели (характеристических точек) с привнесением шумов и погрешностей оцифровки и построение матрицы \mathbf{W}' .
3. Восстановление трехмерного образа модели алгоритмами перспективной проекции [3] и в приближении МОП [2] программой на C++.
4. Совмещение исходной модели и ее восстановленного образа. Вычисление среднеквадратичных погрешностей восстановления формы объекта и ориентации камер по формулам (23), где компоненты погрешностей для формы и ориентации камер вдоль осей вычисляются следующим образом:

$$\begin{aligned} \varepsilon_{i_shape} &= \frac{1}{a_i} \sqrt{\frac{1}{P} \sum_{p=1}^P (s_{ip} - s_{ip}^0)^2} \\ \varepsilon_{i_rotation} &= \sqrt{\frac{1}{F} \sum_{f=1}^F (k_{if} - k_{if}^0)^2}, \quad i = \{x, y, z\} \\ \varepsilon_{camera_z} &= \frac{1}{\langle |\vec{t}| \rangle} \sqrt{\frac{1}{F} \sum_{f=1}^F (|\vec{t}_f| - |\vec{t}_f^0|)^2}, \quad \langle |\vec{t}| \rangle = \frac{1}{F} \sum_{f=1}^F |\vec{t}_f| \end{aligned} \quad (29)$$

Здесь обозначения с нулем обозначают эталонные данные, а без нуля – восстановленные; a_i – длины полуосей эллипсоида (28).

5. Сравнение полученных погрешностей и их зависимостей от параметров с теоретическими оценками.

5.2 Априорные оценки погрешностей

Апостериорные оценки вычислялись по формулам (24). Для получения априорных оценок на основе формул (25), (27) использовались сферические координаты:

$$\begin{aligned} \vec{k} &= (-\cos\theta\cos\varphi \quad -\cos\theta\sin\varphi \quad -\sin\theta) \\ \vec{t} &= (r\cos\theta\cos\varphi \quad r\cos\theta\sin\varphi \quad r\sin\theta), \\ \text{где } r &\in [R-h, R], \quad \theta \in [\theta_0 + \alpha, \theta_0 - \alpha], \\ \varphi &\in [\varphi_0 + \alpha, \varphi_0 - \alpha] \end{aligned} \quad (30)$$

Тогда интегралы в функционалах (27) легко вычисляются:

$$\begin{aligned} \Omega_{xx} &= \frac{1}{k} \left(1 - A \left(\frac{1}{2} + \cos 2\varphi_0 \frac{\sin 2\alpha}{\alpha} \right) \right) \\ \Omega_{yy} &= \frac{1}{k} \left(1 - A \left(\frac{1}{2} - \cos 2\varphi_0 \frac{\sin 2\alpha}{\alpha} \right) \right) \\ \Omega_{zz} &= \frac{1}{k} A, \quad \text{где } A = \frac{5}{6} - \frac{2}{3} \sin^2 \alpha \cos^2 \theta_0, \\ k &= \frac{R^2}{g^2} \left(1 - \frac{h}{R} + \frac{1}{3} \cdot \frac{h^2}{R^2} \right) \end{aligned} \quad (31)$$

В результате подстановки выражений (31) в (25) получаются формулы для априорной оценки погрешностей. Эти выражения достаточно сложны и громоздки. Поэтому, если необходимо получить простые аналитические выражения для погрешностей, можно аппроксимировать результат более простой зависимостью, как это, к примеру, было сделано в [1]. В данном случае аппроксимация не проводилась. Зависимости, полученные из (31) и (25) сравнивались с практическими погрешностями наряду с апостериорными оценками (24).

5.3 Сравнение теоретических и практических погрешностей

Ниже приведены графики зависимостей погрешностей от отношения τ максимальной из полуосей эллипсоида (28) к расстоянию до него и от разрешения изображения в пикселях N . Остальные параметры модели выбирались следующими: $\alpha = 30^\circ$ (рис. 2), разрешение $N=1000$ пикселей, оси эллипсоида соотносятся как $a_x : a_y : a_z = 1 : 2 : 3$, камеры находятся на расстояниях от начала СК в 4-5 раз превышающие a_z , генеральное направление $\vec{B} = (-1 \quad -1 \quad -1)$. Таким образом, СК, в которых матрицы Λ и Ω диагональные, различны. Для каждого маркера проводилось 50 новых генераций случайной сцены, погрешности восстановления усреднялись, вычислялось их среднеквадратичное отклонение.

На рис. 3-6 приведены результаты сравнения теоретических оценок погрешностей восстановления формы объекта и ориентации камер (24), (25) и практических погрешностей. Маркерам «1» и «2» соответствуют величины среднеквадратичных погрешностей, полученные в результате сравнения исходной модели и ее трехмерного образа, восстановленного в приближении МОП и в перспективной проекции соответственно. Тонкие линии по обе стороны от цепочек маркеров определяют их среднеквадратичное отклонение по всем проведенным генерациям модели. Сплошной толстой «3» и пунктирной толстой линии «4» соответствуют апостериорные оценки погрешностей (24). Для сплошной линии «3» использовались нормы и сингулярные числа, полученные в приближении МОП из исходной матрицы \mathbf{W} . Для пунктирной линии «4» использовались результаты работы ИПП [2]. Штрихпунктирная толстая линия «5» соответствует априорной оценке погрешности и получена путем подстановки аналитических выражений для функционалов Ω_{ii} (31), (25). На графиках видно, что априорные и апостериорные оценки практически совпадают.

На вспомогательных графиках рис. 3-6 приведено отношение теоретических оценок погрешностей к практическим погрешностям для ИПП. Линия «1» соответствует апостери-

орной оценке погрешностей (24), линия «2» - априорной (25), (31). Оценки отличаются от практических величин примерно в два раза. Отношение теоретических оценок погрешностей к практическим погрешностям для линейного алгоритма не приводится, но очевидно, что теоретическая оценка существенно ухудшается с ростом τ . Это связано с тем, что большие τ соответствуют меньшему расстоянию до объекта, а значит большим перспективным искажениям, т.е. алгоритм в приближении МОП становится неприменимым [1].

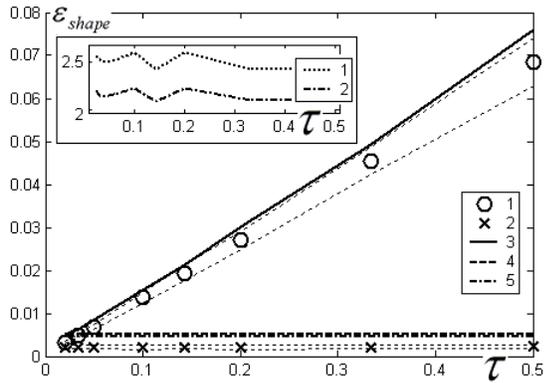


Рис. 3. Зависимость погрешности восстановления формы объекта от параметра τ .

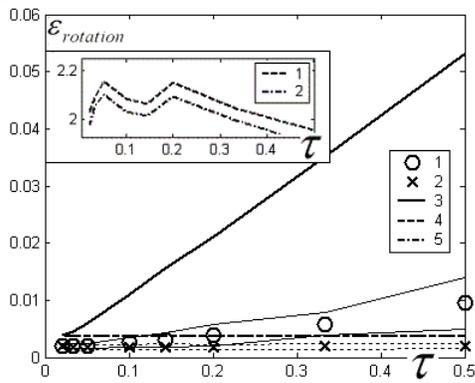


Рис. 4. Зависимость погрешности восстановления ориентаций камер от параметра τ .

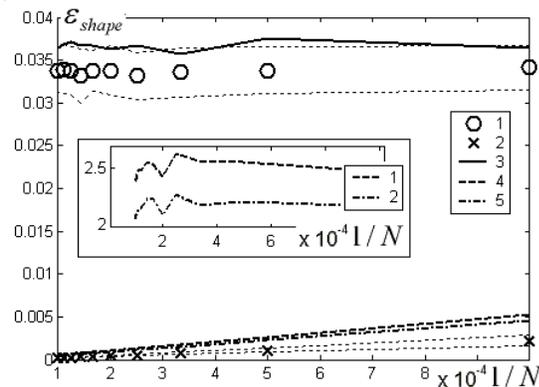


Рис. 5. Зависимость погрешности восстановления формы объекта от разрешения изображения N .

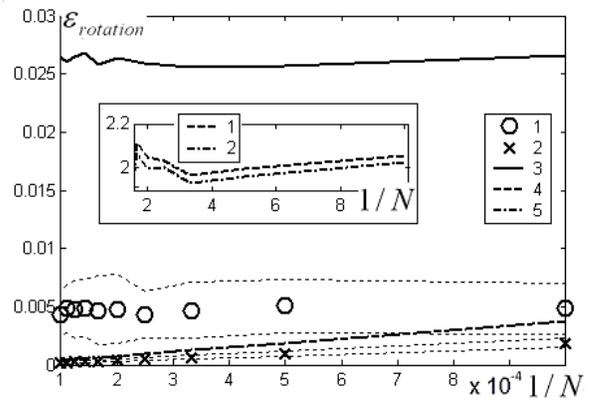


Рис. 6. Зависимость погрешности восстановления ориентаций камер разрешения изображения N .

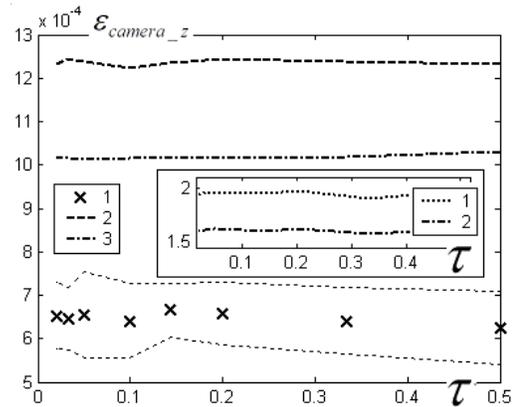


Рис. 7. Зависимость погрешности восстановления расстояния от объекта до камер от параметра τ .

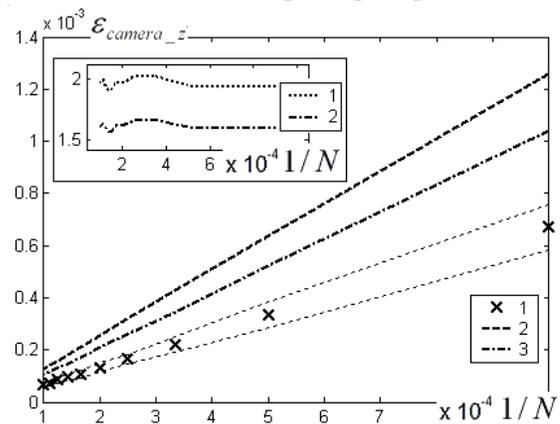


Рис. 8. Зависимость погрешности восстановления расстояния от объекта до камер от разрешения изображения в пикселях N .

На рис. 7-8 приведены аналогичные зависимости для погрешностей восстановления расстояния до камер (24), (25). Т.к. в линейном приближении восстановить расстояние до сцены принципиально невозможно, на графиках представлены зависимости только для алгоритма в перспективной проекции. Маркеру «1» соответствует величина среднеквадратичной погрешности, полученная в результате сравнения исходной модели и ее трехмерного образа. Линиям «2» и «3» соответствуют оценки погрешностей (24) и (25), (31) соответственно. Вспомогательные графики абсолютно аналогичны

описанным выше. Оценки отличаются от практических величин также примерно в два раза.

На представленных графиках видно, что результаты вычислений по формулам оценок (24), (25) достаточно хорошо согласуются с фактическими значениями погрешностей перспективной проекции и линейного приближения МОП.

5.4 Эпиполярная геометрия

В результате работы алгоритма факторизации, как уже упоминалось, в некоторой системе координат восстанавливаются трехмерные координаты всех камер, их пространственные ориентации. Выбирая два наиболее удаленных друг от друга кадра $\mathbf{R}_\chi = (\bar{\mathbf{R}}_1 \ \bar{\mathbf{R}}_2 \ \bar{\mathbf{R}}_3)_\chi^T = (\bar{\mathbf{i}}_\chi \ \bar{\mathbf{j}}_\chi \ \bar{\mathbf{k}}_\chi)^T$, $\bar{\mathbf{i}}_\chi$, \mathbf{g}_χ , $\chi = \{left, right\}$, получаем стерео пару. Используя выражение для фундаментальной матрицы \mathbf{F} , приведенное в [4,11], перепишем его в привычных обозначениях:

$$\mathbf{F}_{bk} = e_{cda} (\bar{\mathbf{i}}_{left} - \bar{\mathbf{i}}_{right})_d (\mathbf{R}_{left})_{ba} (\mathbf{R}_{right})_{kc}, \quad b, k = 1..3 \quad (32)$$

Здесь e_{cda} - единичный антисимметричный тензор третьего ранга. По сути, это значит, что компоненты фундаментальной матрицы выражаются как смешанное произведение вектора трансляции $(\bar{\mathbf{i}}_{left} - \bar{\mathbf{i}}_{right})$ и векторов ориентации левого и правого кадров:

$$\mathbf{F}_{bk} = ((\bar{\mathbf{i}}_{left} - \bar{\mathbf{i}}_{right})_b (\bar{\mathbf{R}}_{left})_k), \quad b, k = 1..3 \quad (33)$$

Подставляя в (33) выходные данные алгоритмов факторизации, получаем фундаментальную матрицу, на основе которой могут быть вычислены эпиполярные линии для любой точки кадра $\bar{\mathbf{x}}_\chi$ вида $(u, v, g)_\chi^T$, $\chi = \{left, right\}$:

$$\bar{\mathbf{x}}_{left}^T \mathbf{F} \bar{\mathbf{x}}_{right} = 0, \quad (34)$$

На рис. 9 приведены изображения нескольких эпиполярных линий для одного из «кадров» модели «Точки». Маркерами «1» обозначены все точки модели, маркеры «2» выделяют точки, для которых строились эпиполярные прямые. Сплошная линия «3» соответствует вычислению по формулам (33), (34) на основе идеальных параметров модели, штрихпунктирная «4» - вычислению на основе результатов работы алгоритма ИПП.

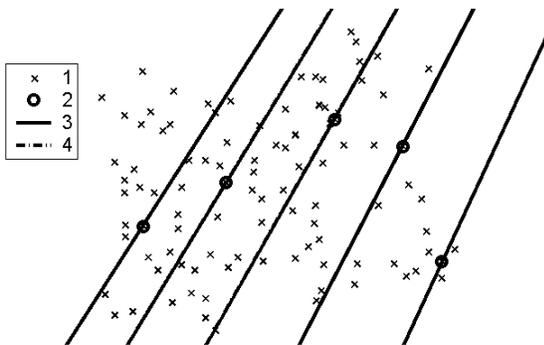


Рис. 9. Эпиполярные линии, построенные на основе параметров съемки, восстановленных алгоритмом ИПП.

Идеальные прямые и прямые, полученные на основе результатов восстановления параметров съемки алгоритмами факторизации, практически полностью совпадают и на рис. 9 неразличимы. Отклонение реальной эпиполярной линии от идеальной по углу порядка 10^{-3} радиана, по расстоянию от начала координат порядка 1 пикселя. При этом были погрешности восстановления: по ориентации камер $2.6 \cdot 10^{-3}$, по расстоянию до камер (относительная погрешность) 10^{-3} , и погрешность восстановления положения камер (относительная) $2.8 \cdot 10^{-3}$.

6. ЗАКЛЮЧЕНИЕ

Данная работа является продолжением и развитием проблемы, поднятой в [1]. Используя избыточность информации, которая является неперенным условием работы группы алгоритмов, основанных на факторизации матриц, и анализируя свойства исходных данных, которые обрабатывают эти алгоритмы, удалось получить апостериорные оценки погрешностей (24), имеющие более общее применение, чем формулы, полученные в [1]. Оценки не накладывают никаких дополнительных требований на форму объекта и способ обхода камеры вокруг него и согласуются с результатами, полученными в [1]. Кроме того, была получена формула для оценки погрешности расстояния до камер (24).

Для априорных оценок погрешностей были предложены выражения, использующие интегральные характеристики сцены. На примере модели «Точки» выражения были доведены до расчетных аналитических формул вычисления ожидаемых погрешностей (25), (32).

Теоретические оценки погрешностей показали хорошую точность при проверке на модельных данных. Результаты численного моделирования показали, что формулы (24) и (25), (32) расходятся с практической погрешностью не более чем в два раза. Для задачи оценки погрешности (погрешность погрешности) это достаточно хорошая точность, т.к. при решении подобной задачи важно получить достаточно простую в использовании формулу, не требующую сложных вычислительных операций, и предоставляющую возможность гарантированно оценить порядок величины погрешности.

Результаты восстановления параметров съемки могут быть использованы для дальнейшего уточнения трехмерной модели сцены стерео алгоритмами. Эту возможность демонстрирует вычисление фундаментальной матрицы и эпиполярных линий на примере модельных данных. Результат (рис. 9) представляется достаточно хорошим, однако он требует тщательного анализа с целью построения оценок погрешностей вычислений.

Комбинирование различных подходов, их последовательное выполнение и использованием выходных данных одних алгоритмов в качестве входных для других способно привести к высококачественным результатам восстановления трехмерных сцен. Однако необходимым условием для этого является строгий контроль накопления погрешностей на каждом этапе вычислений.

7. БЛАГОДАРНОСТИ

Работа выполнена при поддержке РФФИ, гранты 05-07-90344 и 05-07-90390 и 06-01-00789. Авторы выражают благодар-

ность зав. кафедрой СИМ МФТИ д.ф.-м.н., проф. Клименко С.В. за проявленное к работе внимание и поддержку.

8. БИБЛИОГРАФИЯ

- [1] Н. В. Свешникова, Д. В. Юрин. *Априорный и апостериорный расчет погрешностей восстановления трехмерных сцен алгоритмами факторизации.* //Программирование 2004, Т.30, № 5, С. 48-68..
- [2] Conrad I. Poelman, Takeo Kanade. *Factorization Method for Shape and Motion Recovery: //Technical Report CMU-CS-93-219 / School of Computer Science, Carnegie Mellon University. — 1993. http://www.ri.cmu.edu/pubs/pub_1189.html, http://www.ri.cmu.edu/people/person_136_pubs.html.*
- [3] Н. В. Янова, Д.В. Юрин. *Итеративный алгоритм восстановления трехмерных сцен, движения и фокусного расстояния камеры в перспективной проекции, основанный на факторизации матриц.* – В сб. Труды конференции. 12-я Международная конференция по Компьютерной графике и Машинному зрению ГрафиКон '2002 – стр. 123-129. Нижний Новгород, 2002.
- [4] H.C. Longuet-Higgins. *A Computer Algorithm for Reconstructing a Scene From Two Projections," Nature, vol. 293, pp. 133-135, Sept 1981.*
- [5] Ali Azarbayejani Tony Jebara and Alex Pentland. *3d structure from 2d motion. Technical report, MIT Media Laboratory, May 1999. Perceptual Computing Tech. Rep. #523.*
- [6] Carlo Tomasi, Takeo Kanade. *Shape and Motion from Image Streams: a Factorization Method, Part 3, Detection and Tracking of Point Features //Technical Report CMU-CS-91-132 / School of Computer Science, Carnegie Mellon University. — April 1991. http://www.ri.cmu.edu/pubs/pub_2543.html*
- [7] Anton Shokurov, Andrey Khropov, Denis Ivanov. *Feature Tracking in Images and Video.* – В сб. Труды конференции. 13-я Международная конференция по Компьютерной графике и Машинному зрению ГрафиКон '2003 – стр. 177-179. Москва, 2003.
- [8] Joao Paulo Salgado Arriscado Costeira. *A multi-body Factorization method for motion analysis: //Tese para obtencao do grau de doutor em Engenharia Electrotecnica e de Computadores. /Universidade Technica de Lisboa Instituto Superior Rechnico. Lisboa, Maio de 1995. <http://omni.isr.ist.utl.pt/~jpc/pubs.html>*
- [9] Peter Sturm, Bill Triggs. *A Factorization Based Algorithm for Multi-Image Projective Structure and Motion: //4th European Conference on Computer Vision, Cambridge, England, April 1996, pp 709-720.*
- [10] Дж. Деммель. *Вычислительная линейная алгебра.* – М.: Мир, 2001.
- [11] Hartley, R. I. and Zisserman, A., *Multiple View Geometry in Computer Vision,* Cambridge University Press, ISBN: 0521623049, 2000.

Об авторах

Свешникова Наталья Владимировна – аспирантка ФФКЭ Московского физико-технического института, кафедра СИМ.
E-mail: sveshnikova_n@list.ru

Юрин Дмитрий Владимирович – кандидат физ.-мат. наук, ФГУП НПП ОПТЭКС (Федеральное Государственное Унитарное Предприятие Научно-Производственное Предприятие "Оптико-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

The Factorization Algorithms: Results Reliability and Application for the Epipolar Geometry Recovery

Abstract

The paper is devoted to the analytic derivation of the a priori and a posteriori shape, cameras position and orientation error estimates for the group of 3D recovery factorization based algorithms. This deviation doesn't impose any constraints on the 3D scene shape or the way of camera movement. The high accuracy of these algorithms and availability of the a posteriori errors estimates makes it possible to use their results to compute the fundamental matrix and epipolar lines reliably, which is widely used in dense stereo.

Keywords: 3D recovery, factorization methods, recovery errors, epipolar geometry.

About the authors

Natalya Sveshnikova is a Ph.D. student at Moscow Institute of Physics and Technology, Department of Physical and Quantum Electronics, sub-faculty of SIM. Her contact email is sveshnikova_n@list.ru.

Dmitry V. Yurin, PhD, is a chief specialist in FGUP 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.