

Априорный и апостериорный расчет погрешностей восстановления трехмерных сцен алгоритмами факторизации¹.

Н.В. Свешникова*, Д.В. Юрин**.

* Московский физико-технический институт, ЦОС и ВТ

** ФГУП НПП ОПТЭКС

oduduka@cos.ru, yurin@cos.ru

Аннотация

К настоящему времени разработано огромное количество различных алгоритмов восстановления трехмерных сцен по набору цифровых изображений. В зависимости от сцены и условий съемки некоторые алгоритмы могут дать хороший результат, а некоторые неприемлемый. В настоящей работе для группы алгоритмов, основанных на факторизации матриц, получены критерии, позволяющие:

а) зная среднестатистические характеристики сцены и условий съемки, предсказать, применимы ли эти алгоритмы, и какую точность результатов следует ожидать.

б) при восстановлении неизвестной сцены вместе с результатами вычислить их точность, т.е. достоверность.

Предложена модификация алгоритма с адаптивным выбором приближения. Проверка полученных критериев и оценок путем численного моделирования показала их высокую точность и надежность.

Введение

Разнообразие сфер применения систем восстановления трехмерных сцен и сложность проблемы привели к тому, что к настоящему времени разработано большое количество различных алгоритмов и подходов к задаче восстановления трехмерных сцен по набору цифровых изображений [1-6]. Эти подходы базируются на различных физических принципах и математических методах и имеют различные диапазоны применимости. Однако вопрос об условиях применимости и точности алгоритмов восстановления трехмерных сцен обычно исследуется фрагментарно. Типичным является подход, когда разработанный алгоритм тестируется на нескольких заранее известных сценах и дается сравнение результатов восстановления этим алгоритмом и истинной формы сцены, полученной независимым измерением. Остается открытым два вопроса. Как зависит качество восстановления от особенностей сцены, применим ли этот алгоритм или, если это возможно, может быть следовало снимать набор изображений по другому (другие ракурсы, разрешение, количество кадров)? Какова точность полученных результатов при восстановлении данной неизвестной сцены? Последний вопрос наиболее актуален. Ясно, что для любого алгоритма существуют плохие случаи (сцены), когда результаты восстановления неприемлемы. В этом случае хотелось бы, чтобы алгоритм был способен, как минимум, проинформировать, что восстановить сцену он не может. Более того, предположим, имеется два разных алгоритма и оба используются для восстановления одной и той же трехмерной сцены, но дают различные результаты – какому из них верить?

В настоящей работе рассматриваются два метода, основанных на факторизации матриц: приближение масштабированной ортографической проекции (МОП) и итерационный метод в перспективной проекции (ИПП). Первый применим в том случае, когда перспективные искажения не велики и сводится к прямому решению задачи линейной алгебры. Второй – имеет существенно большую вычислительную сложность, но, при наличии перспективных искажений, позволяет устранить неоднозначность направления выпуклости объекта [1], восстановить фокусное расстояние камеры и дает высокую точность восстановления формы и движения.

В настоящей работе на основе аналитического исследования и численного моделирования получен критерий для адаптивного выбора одного из этих двух алгоритмов как на основе априорных знаний о сцене и условиях съемки, так и в процессе обработки данных. Предложена модификация алгоритмов, позволяющая одновременно с результатами восстановления получать оценку точности этих результатов. Получены аналитические формулы для оценки погрешности восстановления формы сцены и ориентаций камер в результате работы исследуемых алгоритмов по средне статистическим характеристикам сцены и расположения камер. Проведено детальное исследование возможности применения этих подходов к задаче трехмерной съемки Земной поверхности с искусственных спутников Земли. Основная идея предлагаемого подхода основана на избыточности исходных данных для методов факторизации и аналогична, по сути, статистической оценке погрешности измерений при аппроксимации экспериментальных данных теоретической кривой методом наименьших квадратов. В качестве «теоретической кривой» в данном случае выступают ограничения, накладываемые законами проективной геометрии на возможные изображения жесткого объекта.

¹ Работа выполнена при частичной поддержке РФФИ, грант 02-01-01063-а.

1. Принципы восстановления данных методом факторизации матриц.

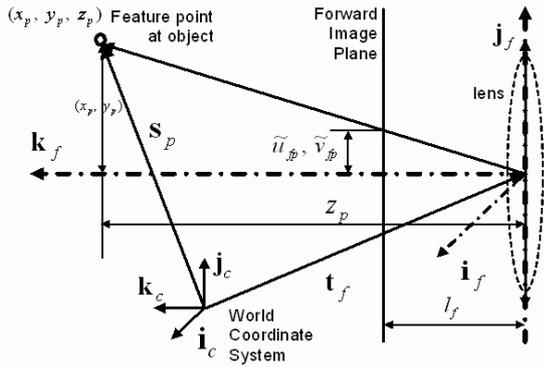


Рис. 1. Постановка задачи восстановления трехмерной сцены.

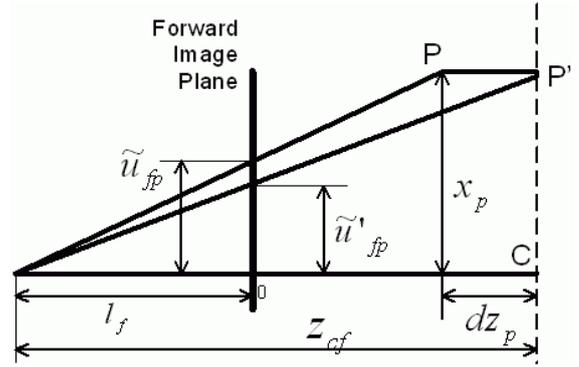


Рис. 2. Постановка задачи построения критерия применимости МОП.

Кратко рассмотрим принципы, на которых базируется подходы [1-3].

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

$$u_{fp} = g_f \frac{\mathbf{i}_f(\mathbf{s}_p - \mathbf{t}_f)}{\mathbf{k}_f(\mathbf{s}_p - \mathbf{t}_f)}, \quad v_{fp} = g_f \frac{\mathbf{j}_f(\mathbf{s}_p - \mathbf{t}_f)}{\mathbf{k}_f(\mathbf{s}_p - \mathbf{t}_f)}, \quad (1)$$

$$\text{где } g_f = \frac{l_f}{\tilde{u}_f^e N} = \frac{1}{2 \operatorname{tg}(\beta_{\max f} / 2)}, \quad u_{fp} = \frac{\tilde{u}_{fp}}{\tilde{u}_f^e N}, \quad v_{fp} = \frac{\tilde{v}_{fp}}{\tilde{u}_f^e N},$$

Здесь $\mathbf{i}_f, \mathbf{j}_f, \mathbf{k}_f$ - правая тройка ортонормированных векторов, задающих систему координат, связанную с f - й камерой, вектор \mathbf{k}_f направлен на объект, а $\mathbf{i}_f, \mathbf{j}_f$ лежат в плоскости изображения; l_f - фокусное расстояние камеры в единицах длины. Безразмерные величины $g_f, u_{fp}, v_{fp} \in [-1/2, 1/2]$ соответствуют величинам в единицах длины $l_f, \tilde{u}_{fp}, \tilde{v}_{fp}$. \tilde{u}_f^e - размер одного пикселя на фоточувствительной матрице; N - ширина изображения в пикселях, $\beta_{\max f}$ - максимальный угол зрения камеры. Вводя обозначение $z_f = -\mathbf{k}_f \mathbf{t}_f$, $z'_f = z_f / g_f$, и полагая $\mathbf{k}_f \mathbf{s} / z_f \ll 1$ Получим формулы для приближения МОП:

$$\begin{cases} u_{fp} = \mathbf{i}_f(\mathbf{s}_p - \mathbf{t}_f) / z'_f \\ v_{fp} = \mathbf{j}_f(\mathbf{s}_p - \mathbf{t}_f) / z'_f \end{cases} \text{ или } \begin{cases} u_{fp} = \mathbf{m}_f \mathbf{s}_p + t_{xf} \\ v_{fp} = \mathbf{n}_f \mathbf{s}_p + t_{yf} \end{cases}, \text{ где } \begin{cases} \mathbf{m}_f = \mathbf{i}_f / z'_f \\ \mathbf{n}_f = \mathbf{j}_f / z'_f \end{cases}, \begin{cases} t_{xf} = -\mathbf{i}_f \mathbf{t}_f / z'_f \\ t_{yf} = -\mathbf{j}_f \mathbf{t}_f / z'_f \end{cases}, \quad (2)$$

Уравнения (2) можно представить в матричной форме:

$$\begin{aligned} \mathbf{W} &= \mathbf{M}\mathbf{S} + \mathbf{T} \\ \mathbf{W}' &= \mathbf{W} - \mathbf{T} \end{aligned} \quad (2')$$

где

$$\mathbf{W} = \begin{pmatrix} u_{11} & u_{12} & \dots & u_{1P} \\ v_{11} & v_{12} & \dots & v_{1P} \\ \vdots & \vdots & \dots & \vdots \\ u_{F1} & u_{F2} & \dots & u_{FP} \\ v_{F1} & v_{F2} & \dots & v_{FP} \end{pmatrix} \quad (3)$$

$$\mathbf{T} = (t_{x1} \quad t_{y1} \quad \dots \quad t_{xF} \quad t_{yF})^T, \quad \mathbf{S} = (\mathbf{s}_1 \quad \dots \quad \mathbf{s}_P), \quad \mathbf{M} = (\mathbf{m}_1 \quad \mathbf{n}_1 \quad \dots \quad \mathbf{m}_F \quad \mathbf{n}_F)^T,$$

Матрица W' представима в виде произведения матриц, ранг которых не превосходит величины ζ , которая равна трем, если объект объемный, двум, если плоский и единице, если одномерный [3]. Поэтому и ранг матрицы W' не превышает указанных значений. Производится сингулярное разложение матрицы W' :

$$W' = U\Sigma V^T = (U\sqrt{\Sigma})(\sqrt{\Sigma}V^T) = \widehat{M}\widehat{S}, \quad rank(W') \leq \zeta \quad (4)$$

здесь U, V - ортогональные матрицы, а $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_n)$, $n = \min(2F, P)$, $\sigma_1, \dots, \sigma_n$ - сингулярные числа матрицы W . Не ограничивая общности, будем считать, что сингулярные числа упорядочены в порядке убывания, тогда выберем из матрицы Σ верхнюю подматрицу размера $\zeta \times \zeta$. Выбирая из матриц U и V подматрицы из строк, соответствующих ζ наибольшим сингулярным числам, получаем аппроксимацию к матрице W . Понятно, что такое разложение не единственно, поскольку между матрицами \widehat{M} и \widehat{S} всегда можно вставить произведение прямой и обратной матриц Q ранга 3, тогда можно записать:

$$W' = \widehat{M}\widehat{S} = \widehat{M}(QQ^{-1})\widehat{S} = (\widehat{M}Q)(Q^{-1}\widehat{S}) = MS, \quad (5)$$

Здесь M и S обозначают истинные матрицы движения и формы.

Так, как $MM^T = MQQ^T M^T$, то, в силу ортонормированности базисных векторов, задающих КСК, получим следующие системы уравнений относительно матрицы Q :

$$\mathbf{m}_f Q Q^T \mathbf{m}_f^T - \mathbf{n}_f Q Q^T \mathbf{n}_f^T = 0, \quad \mathbf{m}_f Q Q^T \mathbf{n}_f^T = 0, \quad \mathbf{m}_1 Q Q^T \mathbf{m}_1^T = 1. \quad (6)$$

Последнее условие необходимо для того, чтобы система (6) имела нетривиальное решение. В данной работе система (6) решалась методом наименьших квадратов (МНК) с применением SVD [7].

Неоднозначность восстановления формы сцены связана с тем, что решением системы (6) может быть как матрица Q , так и матрица вида $Q \text{diag}(\pm 1 \pm 1 \pm 1)$.

Выбрав направление осей таким же, как у системы, связанной с одной из камер, неоднозначность в первых двух знаках можно устранить [1]:

$$\check{M} = MR_0, \quad \check{S} = R_0^T S, \quad R_0 = [\mathbf{i}_f, \mathbf{j}_f, \mathbf{k}_f], \quad \text{где } \mathbf{k}_f = \mathbf{i}_f \times \mathbf{j}_f, \quad f = 1..F \quad (7)$$

Неоднозначность в третьем знаке связана с тем [1], что в рассмотренном линейном приближении (МОП), глубиной объекта по сравнению с расстоянием до него пренебрегают. Как следствие, это приводит к неоднозначности восстановления формы сцены S (выпуклая поверхность может заменяться на вогнутую и наоборот, единым для всей сцены образом) и движения камер M .

Наличие на изображениях перспективных искажений является дополнительным источником информации, позволяющим устранить перечисленные недостатки.

В перспективной проекции уравнения имеют вид (1). Полагая, что фокусное расстояние не зависит от f , т.е. $g_f = g$, перепишем (1) в следующем виде, используя уже встречавшуюся замену z_f (2) на z_f' и обозначая проекции вектора \mathbf{t}_f на тройку векторов $(\mathbf{i}_f, \mathbf{j}_f, \mathbf{k}_f)$, определяющих ориентацию камер, (t_{xf}, t_{yf}, t_{zf}) соответственно:

$$\left(1 + \frac{1}{g} \frac{(\mathbf{k}_f \mathbf{s}_p)}{z_f'}\right) u_{fp} = \frac{1}{z_f'} ((\mathbf{i}_f, \mathbf{s}_p) + t_{xf}), \quad \left(1 + \frac{1}{g} \frac{(\mathbf{k}_f \mathbf{s}_p)}{z_f'}\right) v_{fp} = \frac{1}{z_f'} ((\mathbf{j}_f, \mathbf{s}_p) + t_{yf}). \quad (8)$$

Таким образом, правая часть полученных уравнений соответствует правой части уравнений для МОП (2), а левая содержит добавку. В матричной форме получим:

$$W' = W_1 + \xi W_2, \quad \text{где } \xi = 1/g, \quad W_1 = W \text{ из(3), а}$$

$$W_2 = \begin{pmatrix} \frac{\mathbf{k}_1 \mathbf{s}_1}{z_1'} u_{11} & \frac{\mathbf{k}_1 \mathbf{s}_2}{z_1'} u_{12} & \dots & \frac{\mathbf{k}_1 \mathbf{s}_P}{z_1'} u_{1P} \\ \frac{\mathbf{k}_1 \mathbf{s}_1}{z_1'} v_{11} & \frac{\mathbf{k}_1 \mathbf{s}_2}{z_1'} v_{12} & \dots & \frac{\mathbf{k}_1 \mathbf{s}_P}{z_1'} v_{1P} \\ \vdots & \vdots & \dots & \vdots \\ \frac{\mathbf{k}_F \mathbf{s}_1}{z_F'} u_{F1} & \frac{\mathbf{k}_F \mathbf{s}_2}{z_F'} u_{F2} & \dots & \frac{\mathbf{k}_F \mathbf{s}_P}{z_F'} u_{FP} \\ \frac{\mathbf{k}_F \mathbf{s}_1}{z_F'} v_{F1} & \frac{\mathbf{k}_F \mathbf{s}_2}{z_F'} v_{F2} & \dots & \frac{\mathbf{k}_F \mathbf{s}_P}{z_F'} v_{FP} \end{pmatrix}. \quad (9)$$

Т.е. матрица W представлена в виде двух слагаемых, первое из которых, W_1 , соответствует введённой ранее матрице измерений, а второе, W_2 , зависит от $W_1, \mathbf{s}_p, \mathbf{k}_f$ и имеет смысл поправки на перспективные искажения:

W_2 предлагается [2] вычислять, используя значения $\mathbf{s}_p, \mathbf{k}_f, z_f'$, полученные на предыдущем шаге итераций, а параметр ξ подбирать таким образом, чтобы для реальных зашумленных данных

минимизировать отношение $\sigma_{\zeta+1} / \sigma_{\zeta}$, где ζ равен (4) трем или двум для объемного и плоского объекта соответственно [3]. То есть система уравнений (8) может быть решена итерациями, путём уточнения фокусного расстояния.

Фактически это приводит к двум уровням итераций. Внешним уровнем итерации является вычисление матрицы W_2 по формуле (9) на основе величин $\mathbf{s}_p, \mathbf{k}_f, z'_f$, полученных на предыдущем шаге итераций или в результате работы алгоритма в линейном приближении. После нахождения ξ и соответствующего ему сингулярного разложения W' , решается система (6) и находятся новые величины $\mathbf{s}_p, \mathbf{k}_f, z'_f$. Внутренним уровнем итераций является решение задачи нахождения $\xi = \arg \min_{\xi} (\sigma_{\zeta+1} / \sigma_{\zeta})$, где $\sigma_{\zeta+1}, \sigma_{\zeta}$ - сингулярные числа матрицы W' из (9). Этот шаг был реализован с помощью процедуры FMIN [7], которая носит итерационный характер.

Блок-схема итерационного алгоритма приведена на рис. 3. Штриховыми и точечными контурами обозначены блоки, которые следует внести в алгоритм по результатам анализа, описываемого в данной работе. Их описание будет приведено далее по ходу изложения.

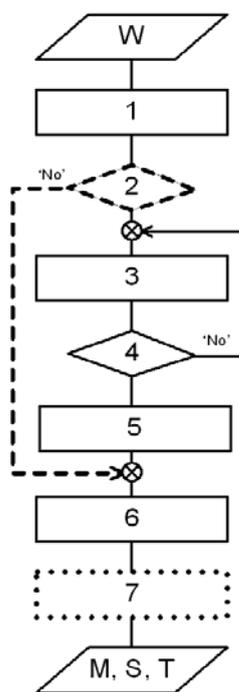


Рис. 3 Блок-схема итерационного алгоритма.

1. Восстановление 3D в линейном приближении (Масштабируемая ортографическая проекция или параперспективная проекция) [1]. Последняя предпочтительнее, т.к. дает существенно лучшую точность, если камера сориентирована не на центр масс объекта. В результате работы блока (формулы (2'), (4), (6)) получаются величины $\mathbf{s}_p, \mathbf{k}_f, z'_f$. Положить $q=0, \xi^{(0)} = 0$.
2. Определение целесообразности выполнения итераций на основе критерия (14).
3. Одна итерация перспективного алгоритма.
 - 3.1. Вычислить W_2 , по формуле (9), используя предыдущие значения $\mathbf{s}_p, \mathbf{k}_f, z'_f$
 - 3.2. Найти $\xi^{(q)} = \arg \min_{\xi} (\sigma_{\zeta+1} / \sigma_{\zeta})$.
 - 3.3. Вычислить W' , по формуле (9).
 - 3.4. Решить задачу в МОП для W' (4)-(6).
4. Определение необходимости дальнейших итераций: $|\xi^{(q)} - \xi^{(q-1)}| < \varepsilon$, где q – номер итерации, ε – требуемая погрешность определения параметра ξ .
5. Устранение неоднозначности по z-компоненте объекта (направление выпуклости объекта). Правильному направлению выпуклости соответствует $\xi > 0$. Если это не выполнено, то следует поменять знак элементов третьей строки матрицы S и третьего столбца матрицы M .
6. Устранение неоднозначности по x и y компонентам объекта (выравнивание по f-той камере, (7)).
7. Вычисление погрешностей восстановления формы объекта и ориентаций камер (28), (35).

Различие между приближением МОП и перспективной проекцией может быть проинтерпретировано геометрически. В модели перспективной проекции (ПП) лучи идут именно так, как это происходит в реальности, то есть по прямой. На рис. 2 реальной точке объекта P в результате перспективного проецирования на ППИ соответствует точка \tilde{u}'_{fp} .

В приближении масштабируемой ортографической проекции все точки объекта сначала проецируются параллельно оптической оси камеры на воображаемую плоскость, проходящую через центр масс точек объекта параллельно ППИ и лишь затем по прямой на ППИ [1]. На рис. 2 камера направлена на центр масс объекта, точка P' является проекцией реальной точки объекта P на описанную выше воображаемую плоскость, а точка \tilde{u}'_{fp} – проекцией, в свою очередь, точки P' на ППИ, соответствующей приближению МОП.

Описанные алгоритмы реализованы в виде программы на C++. На вход программы подается матрица W (для МОП также подается фокусное расстояние g_f). В результате работы программы получаются матрицы трехмерных координат точек объекта и положения камер, а также тройки векторов ориентаций камер. Также формируется VRML-файл для представления результата в трехмерном виде [8].

Далее будут приведены результаты анализа описанных выше алгоритмов. Проверка полученных критериев и оценок погрешностей проводилась на модельных данных. Работа строилась по следующей схеме:

1. Построение трехмерной модели в Matlab в виде сетки характеристических точек.
2. Вычисление изображения модели (точек) с привнесением шумов и погрешностей оцифровки и построение матрицы W .
3. Восстановление трехмерного образа модели алгоритмами перспективной проекции и в приближении МОП программой C++.
4. Совмещение исходной модели и ее восстановленного образа и вычисление погрешностей восстановления формы объекта и ориентации камер по формулам:

$$\varepsilon_{Shape} = \frac{1}{a} \sqrt{\frac{1}{P} \sum_{p=1}^P (\mathbf{s}_p - \mathbf{s}_p^0)^2}; \quad S = [\mathbf{s}_1^T \quad \mathbf{s}_2^T \quad \dots \quad \mathbf{s}_P^T]; \quad S^0 = [\mathbf{s}_1^{0T} \quad \mathbf{s}_2^{0T} \quad \dots \quad \mathbf{s}_P^{0T}]. \quad (10)$$

$$\varepsilon_{Rotation} = \sqrt{\frac{1}{3F} \sum_{f=1}^{3F} (\mathbf{r}_f - \mathbf{r}_f^0)^2}; \quad R = [\mathbf{r}_1^T \quad \mathbf{r}_2^T \quad \dots \quad \mathbf{r}_{3F}^T]; \quad R^0 = [\mathbf{r}_1^{0T} \quad \mathbf{r}_2^{0T} \quad \dots \quad \mathbf{r}_{3F}^{0T}]. \quad (11)$$

Здесь обозначения с нулем обозначают эталонные данные, а без нуля – восстановленные; a – характерный размер объекта.

5. Сравнение полученных погрешностей и их зависимостей от параметров с теоретическими оценками, вывод которых дан в последующих разделах.

2. Модельные данные.

Модель «Куб», представляющая собой три грани прямоугольного куба со стороной a (см. рис. 3, 4), на каждой из которых характеристические точки образуют равномерную прямоугольную сетку. Система координат задается так, что ее начало совпадает с общей для всех граней точкой, а общие ребра граней совпадают с ее положительными полуосями (см. рис. 4). Камеры располагаются случайно равномерно в слое толщины dr сектора α окружности с началом в точке O и радиусом R . Все камеры ориентированы на центр масс объекта C .

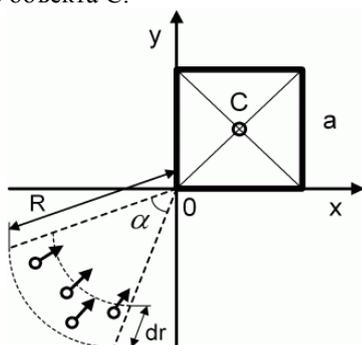


Рис. 4 Схема построения модели «Куб».

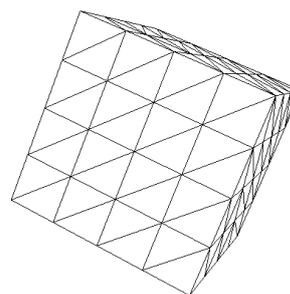


Рис. 5 Восстановленная модель «Куб» в представлении VRML в виде сетки характеристических точек.

Модель «Ступенька» (рис. 6.)- модель спутника, летящего по окружности радиуса $R+H$, где R – радиус Земли, H – высота орбиты над поверхностью, $H \sim 1000$ км, над участком плоскости размера $a \times a$, имеющим небольшую приподнятую часть малой высоты h .

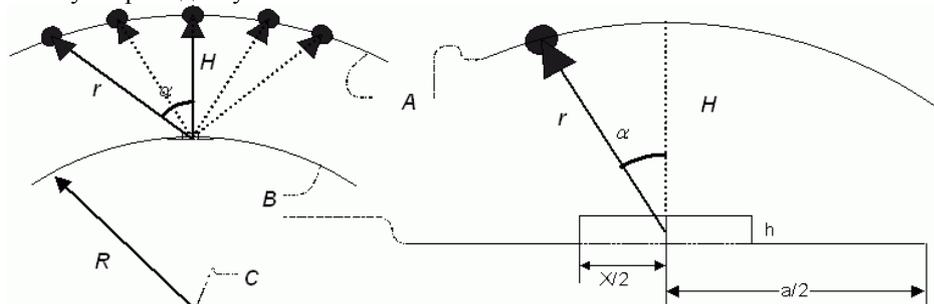


Рис. 6. Схема построения модели съемки спутником плоскости с прямоугольным выступом (сцена малой глубины). Обозначения: А – орбита спутника, В – поверхность Земли, С – центр Земли.

Плоскость со ступенькой помещается в плоскость осей XU и ее центр совпадает с началом координат. Спутник движется по окружности, центр которой находится в точке $(0,0,-R)$. Окружность лежит в плоскости YZ Съемка F кадров производится через равный шаг по углу в диапазоне углов к надиру $|\alpha| \leq 90^\circ$.

Параметрами, определяющими модель являются высота ступеньки h , площадь, занимаемая ею на снимаемой поверхности b^2/a^2 , разрешение в метрах при наблюдении в надир, определяющее видимый размер объекта в пикселях, допустимые углы съемки α и количество кадров F . На модели задается прямоугольная сетка точек.

Пример восстановленной модели приведен на рис. 7 в виде сетки характеристических точек.

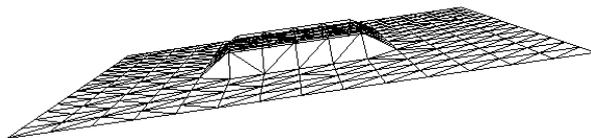


Рис. 7. Восстановленная модель «Спутник» в представлении VRML в виде сетки характеристических точек.

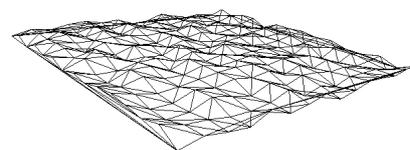


Рис. 8. Восстановленная модель «Спутник над возмущенной плоскостью» в представлении VRML в виде сетки характеристических точек.

Модель «Рельеф» аналогична модели «Ступенька», но восстанавливаемый объект представляет собой плоскость, каждая точка которой возмущена случайным образом относительно плоскости $z=0$ так, что среднее квадратичное отклонение равно h . На рис. 8 приведен пример восстановленной модели такой возмущенной плоскости.

3. Критерий применимости приближения МОП.

Напомним, что на рис. 2 реальной точке объекта P соответствует ее перспективная проекция на ППИ – точка \tilde{u} , а ее образу P' в приближении МОП соответствует проекция \tilde{u}' на ППИ. Т.к. цифровое изображение имеет конечное пиксельное разрешение, то разница между проекциями точек P и P' на ППИ

$$G = |\tilde{u}_{fp} - \tilde{u}'_{fp}| / \tilde{u}_f^e = N |u_{fp} - u'_{fp}|; \quad (12)$$

не будет заметна, если она не превышает погрешности определения координат в пикселях μ . Здесь \tilde{u}_f^e – размер пикселя, N – количество пикселей в строке изображения. Таким образом, результат восстановления в приближении МОП не будет отличен от результата перспективного метода, если

$$G < \mu \quad (13)$$

Если величина μ определяется только шумом оцифровки, то

$$\mu = \sqrt{\int_{-1/2}^{1/2} x^2 dx} = \sqrt{\frac{1}{12}} \approx 0.3,$$

Пусть $(x_p, y_p, z_{cf} + dz_p)$ – координаты некоторой характеристической точки объекта в СК, связанной с f -той камерой, z_{cf} – расстояние от ЦМ объекта до f -той камеры. Тогда из подобия треугольников выразим \tilde{u} и \tilde{u}' , перейдем к безразмерной величине для g_f из (1) и получим (см. рис. 2) выражение:

$$G = Ng_f \frac{x_p \cdot dz_p}{z_{cf}(z_{cf} + dz_p)},$$

Наибольшее значение G_{\max} величина G принимает для точек, наиболее удаленных от ЦМ. Обозначим соответствующие координаты наиболее удаленных точек $x_{\max}, y_{\max}, dz_{\max}$. Учитывая также, что безразмерный фокус из (1) выражается через максимальный угол зрения камеры, для оценки G_{\max} можно записать:

$$G_{\max} = \eta \chi_x \chi_z N \operatorname{ctg} \frac{\beta_{\max}}{2}, \quad \text{где} \quad \chi_x = \frac{x_{\max}}{(z_{cf} - dz_{\max})} \approx \frac{x_{\max}}{z_{cf}}, \quad \chi_z = \frac{dz_{\max}}{z_{cf}} \quad (14)$$

Выражения (13) и (14) дают критерий применимости МОП. Если это условие не выполняется, то перспективные искажения существенны, и необходимо использовать итерационный алгоритм (рис. 3). Если условие выполнено, то результаты, полученные в приближении масштабируемой ортографической проекции и из итерационного алгоритма, будут неотличимы. Т.к. съемка различных кадров производится с различных расстояний и углов и распределение точек на поверхности объекта может быть различным, а совместная обработка всех данных методом факторизации приводит к статистическому усреднению, это выражение справедливо с точностью до коэффициента η :

Путем численного моделирования на основе моделей «Куб» и «Рельеф» проводился поиск величины коэффициента η и проверка критерия (13),(14). Была разработана тестовая программа для сравнения величины G , определенной в (12), получаемой из опыта и G_{\max} , вычисляемой по формуле (14).

Для получения величины из опыта, модель куба восстанавливалась методом масштабируемой ортографической проекции. Трехмерный образ проецировался на ППИ без шума оцифровки. Исходная модель также проецировалась на ППИ без шума оцифровки. В качестве G выбиралось максимальное по f и g значение выражения (12). Величина G_{\max} рассчитывалась по формуле (14), где значения $\chi_x, \chi_z, N, \beta_{\max}$ определялись непосредственно из модели. Коэффициент η вычислялся из условия $G = G_{\max}$.

Результаты численного моделирования представлены на рис. 9-10 в зависимости от χ_z . Параметры модели «Куб» брались следующие: угол обзора 60 градусов, камеры расположены на одинаковом расстоянии до объекта, распределение камер по углу равномерное. Длина ребра куба принята равной 1. Фокус выбирается таким, что объект занимает примерно половину линейного размера изображения. Камера ориентирована на центр масс. Параметры модели «Рельеф»: размер площадки 2 x 2 км, разрешение 2000 пикселей при точности детектора 0.1 пиксель. По оси ординат отложено количество пикселей, на которое сместился на ППИ образ восстановленного объекта относительно исходной модели. Маркерами отмечены величины G , полученные (12) описанным выше способом. Каждому маркеру на рисунках соответствует случайное распределение камер для модели «Куб» и точек рельефа для модели «Рельеф». Через них проводится аппроксимирующая прямая («4» – штрих-пунктир). Сплошная линия «2» – та же величина, вычисленная подстановкой параметров в (14), пунктирная линия «3» отмечает смещение в 1 пиксель. Зоне применимости приближения МОП соответствует область, где пунктирная линия выше сплошной.

Заметим, что поскольку данные для задачи восстановления искажены шумом оцифровки, зависимость величины G от χ_x, χ_z, N носит нелинейный характер в отличие от теоретической зависимости G_{\max} . Действительно, в области малых перспективных искажений формула (12) дает оценку погрешности детектора μ , а не перспективных искажений. Поэтому данные, представленные на рисунках маркерами разбивались на две области, в области наличия перспективных искажений они аппроксимировались прямой, проходящей через начало координат, а в области их отсутствия – горизонтальной прямой, соответствующей погрешностям оцифровки. Граница раздела выбиралась из условия минимума суммы дисперсий данных относительно этих прямых.

Поскольку расположение камер на определенном расстоянии от объекта задается случайным распределением, результирующая ошибка имеет некоторый разброс и оценка коэффициента η носит статистический характер. На вспомогательных графиках рис. 9-10 отображены результаты оценки коэффициента η , построенные на основании данных основных графиков путем деления значений, принимаемых аппроксимирующей прямой «4», на значения теоретической прямой «2». Излом и разрыв на графиках соответствует точке перехода от области применимости приближения МОП к области целесообразности перспективной проекции и обусловлен тем, что в каждой из областей зависимость G_{\max} от χ_z аппроксимировалась отдельной прямой. В области применимости приближения МОП смещение образа восстановленной модели относительно ее исходного изображения обуславливается только шумом оцифровки, который неизменен. Перспективные искажения линейно зависят от глубины объекта и суммируются с шумом оцифровки.

Результаты численного моделирования показали, что величина коэффициента η лежит в диапазоне 1-4, и имеет слабую зависимость от разрешения изображения N и отношения расстояния до объекта к размеру объекта (χ_x, χ_z). Обращает на себя внимание систематическое занижение теоретической оценки G_{\max} по сравнению с результатами моделирования. Это объясняется тем, что величина G выбиралась как максимум по f и g , что соответствует точке объекта, которая восстановилась хуже всех. Выбор максимальной, а не средней, как обычно, оценки для G и G_{\max} соответствует отданию некоторого предпочтения в переходной области итерационному алгоритму, дающему лучшую точность, чем МОП. Как очевидно из рис. 2, величина перспективных искажений растет с удалением точки изображения от центра кадра, поэтому при любых условиях съемки часть точек не будет содержать перспективных искажений, и в зависимости от распределения найденных детектором характеристических точек по поверхности объекта, эта часть может быть значительна. Поэтому логику указанного выбора можно сформулировать иначе: если есть возможность – дать перспективному алгоритму шанс улучшить точность восстановления, и только если перспективных искажений нет вообще – отказаться от итерационного алгоритма и не строить иллюзий, что удалось восстановить фокусное расстояние и соответственно синхронизировать масштабы в котором восстановлены объект и положения камер.

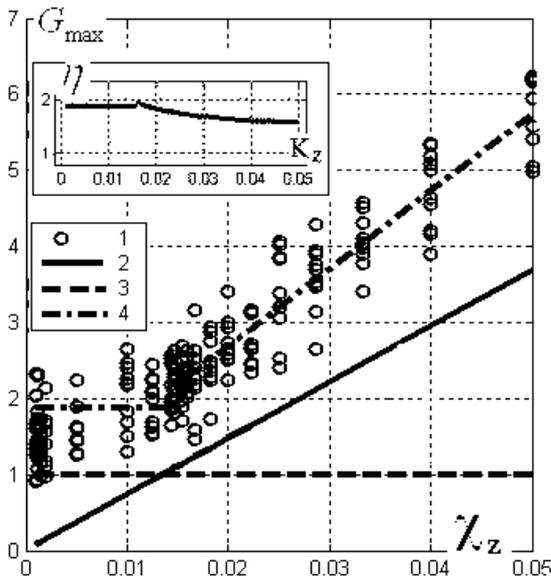


Рис. 9. Зависимость величины G_{\max} от коэффициента χ_z на примере модели «Куб»

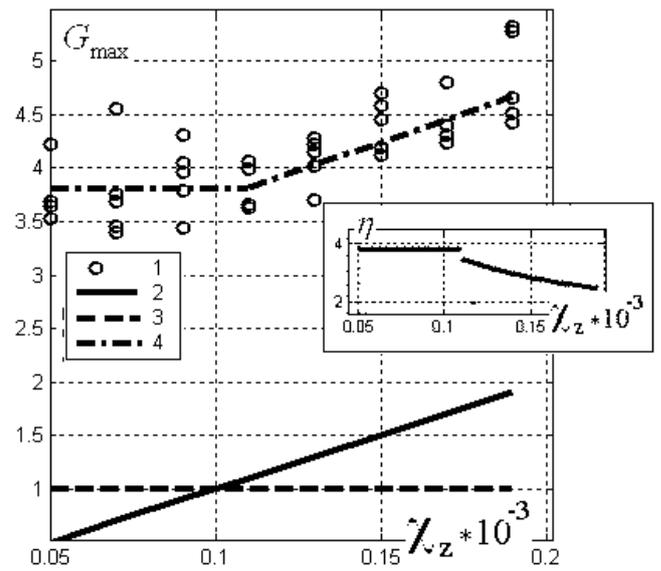


Рис. 10. Зависимость величины G_{\max} от коэффициента χ_z на примере модели «Рельеф»

На основе критерия (13),(14) был реализован адаптивный выбор приближения в алгоритме восстановления рис. 3. В алгоритм добавлен блок «2». На основе результатов восстановления, полученных в МОП (блок «1») вычисляется (14) величина G_{\max} и сравнивается (13) с погрешностью детектора точек μ . По результатам сравнения принимается решение о целесообразности итеративного уточнения результатов.

4. Разрешимость малой глубины сцены в приближении масштабируемой ортографической проекции (МОП).

Рассмотрим модельную задачу. Пусть восстанавливаемый объект – две параллельны плоскости, находящиеся на малом расстоянии друг от друга. Их уравнения:

$$\begin{aligned} L_1: & Ax + By + z + D - h = 0; \\ L_2: & Ax + By + z + D + h = 0; \end{aligned} \quad (15)$$

где h мало.

Набор точек плоскостей снимается камерой с конечным пиксельным разрешением и обрабатывается методом факторизации матриц для восстановления трехмерного образа. Требуется оценить, при каких величинах h метод факторизации матриц сможет различить две плоскости.

На рис. 6 изображена постановка задачи восстановления земного рельефа по последовательности кадров, полученных с искусственного спутника Земли и модель – сцена малой глубины, представляющая собой две параллельные плоскости – основу и выступ малой высоты. Характерные параметры модели: a – ширина плоскости основы, x – ширина выступа, α – угол отклонения камеры от вертикали, g – расстояние от камеры до центра масс объекта, H – высота орбиты над поверхностью земли.

Метод факторизации матриц базируется на том, что для трехмерного объекта ранг идеальной матрицы W' , обозначенный выше ζ равен 3, для плоскости – 2, для прямой – 1:

$$W' = MS, \quad (16)$$

где $f = 1..F$ – номер изображения, $p = 1..P$ – номер точки.

В реальности матрица W' всегда зашумлена и ранг ее в общем случае равен $\min(2F, P)$. Реальную матрицу W' можно представить:

$$W' = W^0 + \omega, \quad (17)$$

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

$$\langle \omega^2 \rangle = \frac{1}{2FP} \sum_{i=1}^{2F} \sum_{j=1}^P \omega_{ij}^2 = \rho^2; \quad \langle \omega \rangle = \frac{1}{2FP} \sum_{i=1}^{2F} \sum_{j=1}^P \omega_{ij} = 0. \quad (18)$$

Величина шума определяется как безразмерная, что следует из (17), и зависит от погрешности детектора и дискретизации изображения, но не от снимаемой сцены. Матрица ω имеет, в общем случае, ранг $\min(2F, P)$ и произвольные сингулярные числа, наибольшее из которых обозначим σ_n .

При работе метода факторизации матриц выполняется сингулярное разложение матрицы W' . Затем выбираются три (для трехмерного объекта) наибольших сингулярных числа. Все остальные полагаются шумовыми и отбрасываются.

Если h мало и рассматриваемый объект близок к двумерному, сингулярное число, отвечающее за третье измерение может стать сопоставимым с шумовым сингулярным числом или меньше него (т.к. из (15) для плоского объекта z будет линейно зависеть от x и y , и S будет иметь ранг 2). Тогда собственный вектор-столбец в матрице V соответствующий сингулярному числу ответственный за шум будет ошибочно принят за составляющую формы [3]. Чтобы этого избежать, необходима априорная оценка шумового сингулярного числа σ_n , и. условием разрешимости будет:

$$\sigma_3 > \sigma_n \quad (19)$$

Рассмотрим идеальную матрицу W' в условиях данной задачи. Полагая x и y независимыми координатами, мы можем z выразить через них. Тогда

$$z_p = -(Ax_p + By_p + D \mp h) = -(Ax_p + By_p + D) \pm h = z_p^{(1)} \pm h. \quad (20)$$

Тогда, вводя обозначения

$$s_p^{(1)} = (x_p, y_p, z_p^{(1)})^T, \quad s_p^{(2)} = (0, 0, \pm 1)^T$$

$$S = S_1 + h \cdot S_2, \quad S_1 = (s_1^{(1)} \quad \dots \quad s_p^{(1)}) \quad , \quad S_2 = (s_1^{(2)} \quad \dots \quad s_p^{(2)}).$$

Представим матрицу W' при данных (x_p, y_p, z_p) :

$$W' = MS_1 + h \cdot MS_2 = W_1' + W_2' \quad (21)$$

Т.к. при $h=0$ задача вырождается в двумерную: $\text{rank } S_1 = 2 \Rightarrow \text{rank } W = 2$, то, разложив матрицу W' на две составляющие: не содержащую h и содержащую, мы разделим задачу на плоскую и добавку. Очевидно, что $\text{rank } S_2 = 1$ и $\text{rank } W_2' = 1$, а т.к. по построению именно эта компонента отвечает за возникновение третьего ненулевого сингулярного числа в разложении исходной матрицы W' , то в приближении малого h , оценка единственного ненулевого сингулярного числа матрицы W_2' и будет оценкой для σ_3 .

Для оценки сингулярных чисел воспользуемся понятием нормы Фробениуса [9] (в случае матрицы ранга 1 она совпадает с Эрмитовой и равна этому единственному сингулярному числу). В случае матрицы произвольного ранга (шумовая добавка) норма Фробениуса дает завышенную оценку наибольшего сингулярного числа. Тогда норма матрицы выразится следующим образом:

$$\|A\|^2 = \|A\|_F^2 = \sum \sum a_{ij}^2 = \sum \sigma_i^2;$$

Подставляя величины, характеризующие шумовую матрицу (18) получим:

$$\sigma_n^2 \leq \|\omega\|^2 = \sum_{f=1}^{2F} \sum_{p=1}^P w_{fp}^2 = 2FP\rho^2; \quad (22)$$

Перейдем к вычислению нормы матрицы W_2' . Т.к. $\text{rank } W_2' = 1$, то норма Фробениуса будет совпадать со значением единственного сингулярного числа, которое в приближении малого h дает оценку σ_3 для W' .

$$\sigma_3^2 = \|W_2'\|^2 = FPh^2 J, \quad \text{где } J = \frac{1}{F} \sum_{f=1}^F \left(\left(\frac{g_f}{z_{cf}} \right)^2 (i_{zf}^2 + j_{zf}^2) \right) \quad (23)$$

Вычисления по формуле (23) могут быть произведены, если для всех кадров известны ориентации и положения камер. На практике это применимо, когда общие статистические характеристики сцены известны *a priori*. Однако вычисление конечных сумм представляется достаточно неудобным для оценивания поэтому, полагая, что число кадров F велико, перейдем к интегралу. Преобразуем формулу (23) следующим образом. Пусть \vec{r} - радиус-вектор, определяющий положение камеры. Будем считать, что камера всегда направлена на центр масс объекта, тогда, зная положение камеры, можно вычислить ее ориентацию, $i_z = i_z(\vec{r})$, $j_z = j_z(\vec{r})$, $g = g(\vec{r})$, $z = |\vec{r}|$. Введем закон случайного распределения камер $\nu = \nu(\vec{r})$. Тогда значение суммы (23), можно приблизить интегралом:

$$J = \int_{-\infty}^{+\infty} \left(\frac{g(\vec{r})}{|\vec{r}|} \right)^2 (i_z^2(\vec{r}) + j_z^2(\vec{r})) \nu(\vec{r}) d\vec{r} \quad (24)$$

В условиях конкретной модели в результате вычисления интеграла можно получить аналитическую формулу для априорной оценки разрешимости задачи восстановления почти плоского объекта. Ниже будет приведен пример построения такой формулы в условиях модели спутниковой съемки Земной поверхности.

Проверка критерия разрешимости по высоте (19),(22),(23) проводилась путем численного моделирования на примере модели «Рельеф». На рис 11-14 приведены зависимости третьего и шумового сингулярных чисел в зависимости от таких параметров как глубина сцены и разрешение изображения. Остальные параметры модели выбирались следующими: $h = 0.1$ км, размер площадки 2 X 2 км, разрешение 2000 пикселей при точности детектора 0.1 пиксель [10]. Для каждого маркера проводилась новая генерация случайной поверхности. На основных графиках приведено по три зависимости. Маркер «1» соответствует сингулярному числу, взятому из сингулярного разложения исходной матрицы W' из (3). Маркер «2» отображает значения сингулярных чисел, полученных из сингулярного разложения матрицы W' из (9) после

выполнения итераций. Видно, что маркеры практически совпадают, т.е. в данной задаче в процессе итераций третье сингулярное число изменяется мало. Сплошная линия «3» соответствует величине сингулярного числа, вычисленного по формулам (22, 23) путем прямой подстановки туда параметров, взятых из модели. На вспомогательном графике рис. 11-12 приведена относительная погрешность оценки σ_3 . Видно, что оценка дает результат с точностью порядка 15%. На вспомогательном графике рис. 13-14 приведена относительная погрешность оценки σ_n после выполнения итераций, когда σ_n сильно уменьшается за счет устранения перспективных добавок. Остается только шум дискретизации, оценка которого и была проведена в (22). Точность оценки не хуже 40%.

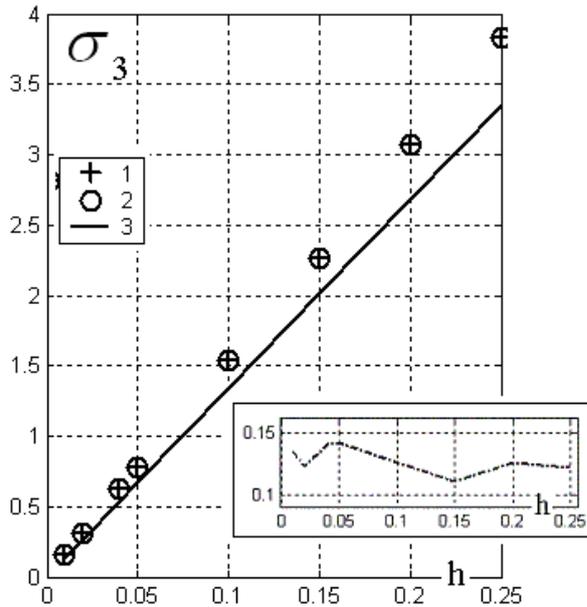


Рис. 11 Зависимость третьего сингулярного числа от глубины сцены.

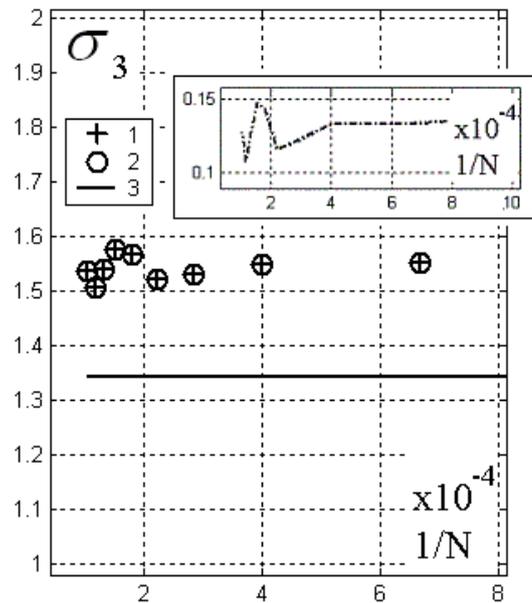


Рис. 12. Зависимость третьего сингулярного числа от разрешения изображения в пикселях

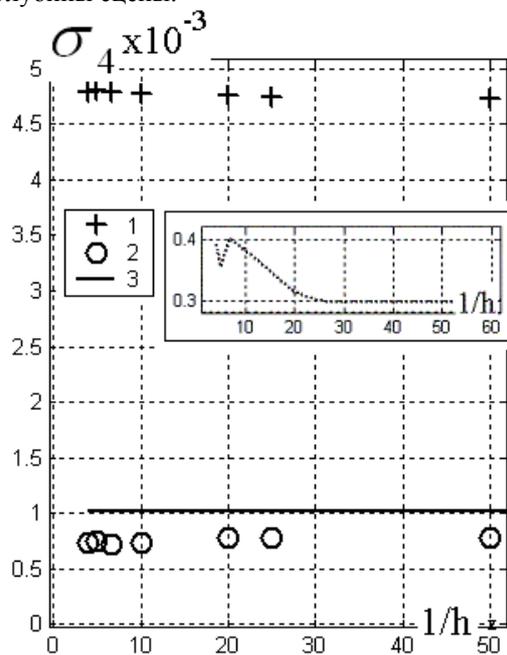


Рис. 13 Зависимость шумового сингулярного числа от глубины сцены.

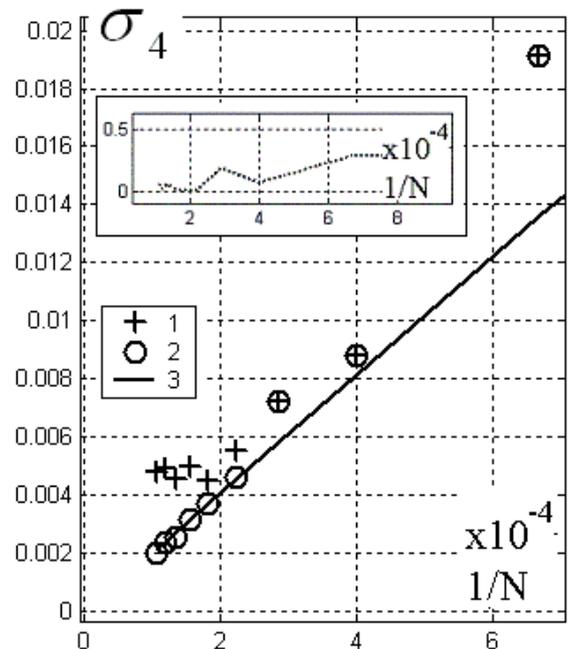


Рис. 14. Зависимость шумового сингулярного числа от разрешения изображения в пикселях

5. Оценка погрешностей восстановления формы сцены и ориентации камер.

В предыдущем разделе была решена задача о минимальной высоте ступеньки на плоской поверхности, которая может быть обнаружена по наблюдаемым данным (W') методом факторизации. Выдвинем гипотезу, что полученная оценка есть одновременно абсолютная погрешность восстановления формы объекта произвольной формы. Далее справедливость этой гипотезы и высокая точность оценки будет показана численным моделированием. Сделаем некоторые предварительные замечания.

Метод факторизации требует полностью заполненной матрицы W' , поэтому для сплошных объектов диапазон углов наблюдения ограничен, по крайней мере, полусферой, или $\pm 90^\circ$ от выбранного направления на объект, так как в противном случае некоторые отметки будут закрываться объектом. На самом деле для объекта отличного от плоскости, диапазон углов еще меньше и определяется высотой и крутизной шероховатостей объекта. Ясно, что погрешность восстановления формы наиболее велика в генеральном направлении наблюдения, особенно если камеры распределены в небольшом диапазоне углов. Поэтому наибольший интерес представляет оценка погрешности формы в этом направлении. Выбирая систему координат, связанную с центром масс отметок объекта с осью z вдоль генерального направления наблюдения, будем оценивать только третью компоненту $\delta S = [\mathbf{0}, \mathbf{0}, \delta S_z]^T$, а первые две положим равными нулю. Это позволяет применять оценку минимальной разрешимой высоты, полученную в предыдущем разделе и заменить в формуле (28) $\|M\|$ на $\|m_z\|$ - норму третьего вектор-столбца матрицы M , а норму матрицы формы $\|S\|$ на $h\|S_2\|$, определяющую величину третьего сингулярного числа:

Как было показано в предыдущем разделе в предположении малой глубины объекта в генеральном направлении наблюдения $\sigma_3 \sim h$, а минимальная разрешимая глубина сцены h_{\min} определяется из условия $\sigma_3 = \sigma_n$, тогда гипотеза может быть переформулирована следующим образом:

$$\frac{\delta h}{h} = \frac{h_{\min}}{h} = \frac{\sigma_n}{\sigma_3} \quad (25)$$

Где δh - абсолютная погрешность восстановления формы объекта.

Напомним, что при восстановлении трехмерных сцен методом факторизации ищется решение математической задачи: (4) где W' - зашумленная масштабированная матрица измерений. Задача решается путем сингулярного разложения матрицы $W' = W^0 + \omega$, ограничения на ранг матриц M и S позволяют отделить регулярную часть W^0 от шумовой ω . По сути - это задача наименьших квадратов с регуляризацией. Решение находится в смысле

$$\min_{M,S} \|W'_s - M \cdot S\| \leq \|\omega\| \quad (26)$$

Заметим, что равенство $W^0 = M \cdot S$ точное.

Матрицы M и S тесно взаимосвязаны. Так изображения похожих объектов могут выглядеть с различных ракурсов одинаково, то есть различия или схожесть изображений (матрица измерений) объектов (матрица формы S) могут быть объяснены изменением точки и направления зрения (матрица M). Наличие шума $\|\omega\| \neq 0$ вносит неоднозначность в восстановление M и S . Пусть форма объекта восстановлена с погрешностью δS , тогда возникнет погрешность в восстановлении матрицы движения δM .

$$W' = (M + \delta M) \cdot (S + \delta S) = M \cdot S + \delta M \cdot S + M \cdot \delta S + \delta M \cdot \delta S \approx M \cdot S + \delta M \cdot S + M \cdot \delta S \quad (27)$$

Сравнивая (25) и (26), получаем, что при условии $\|\delta M \cdot S\| \approx \|M \cdot \delta S\|$ и противоположных знаках δM и δS погрешности формы и движения компенсируют друг друга; приближенное равенство следует понимать «с точностью до $\|\omega\|$ ». Таким образом, можно оценить максимальную погрешность в матрице движения, которая может вызываться погрешностью в восстановлении формы:

$$\|\delta M\| \approx \|M\| \cdot \frac{\|\delta S\|}{\|S\|} \quad (28)$$

Перейдем к рассмотрению погрешности ориентации камер. Построим матрицу δM как приращение матрицы M , полагая, что варьируется только ориентация камер, т.е. если

$$M = \begin{pmatrix} g_1 \mathbf{i}_1 & g_1 \mathbf{j}_1 & \dots & g_F \mathbf{i}_F & g_F \mathbf{j}_F \\ z_1 & z_1 & \dots & z_F & z_F \end{pmatrix}^T, \text{ то } \delta M = \begin{pmatrix} g_1 \delta \mathbf{i}_1 & g_1 \delta \mathbf{j}_1 & \dots & g_F \delta \mathbf{i}_F & g_F \delta \mathbf{j}_F \\ z_1 & z_1 & \dots & z_F & z_F \end{pmatrix}^T \quad (29)$$

Считая, из соображений, описанных выше, что отличен от нуля только третий столбец матрицы δM , получим выражение для нормы:

$$\|\delta M\|^2 = \sum_{f=1}^F \left(\frac{g_f}{z_f} \right)^2 (\delta i_{zf}^2 + \delta j_{zf}^2) \quad (30)$$

Так как $\delta i_{zf}^2 \in [0, \delta i_0^2]$ и $\delta j_{zf}^2 \in [0, \delta j_0^2]$, то есть ограничено, а $(g_f / z_f)^2$ знакопостоянно, то, по теореме о среднем, существуют $\langle \delta i_z^2 \rangle$ и $\langle \delta j_z^2 \rangle$ такие, что

$$\| \delta M \|^2 = \left(\langle \delta i_z^2 \rangle + \langle \delta j_z^2 \rangle \right) \sum_{f=1}^F \left(\frac{g_f}{z_f} \right)^2 \quad (31)$$

Погрешность детектора характеристических точек и/или оцифровки приводит к тому, что направление на каждую точку объекта определяется с ошибкой, что, в конечном итоге и приводит к погрешности в восстановлении ориентации камер. Будем считать, что пиксели квадратные а погрешность детектора равновероятна и статистически независима вдоль строк и столбцов изображения. Тогда $\langle \delta i_z^2 \rangle \approx \langle \delta j_z^2 \rangle$ и (31) принимает вид:

$$\| \delta M \|^2 = 2 \langle \delta j_z^2 \rangle \sum_{f=1}^F \left(\frac{g_f}{z_f} \right)^2$$

Величина $\langle \delta j_z^2 \rangle$ является дисперсией погрешности восстановления ориентаций камер и, при малых $\langle \delta j_z^2 \rangle$, среднеквадратичная угловая ошибка определения ориентации камер $\mathcal{G}_j = \sqrt{\langle \delta j_z^2 \rangle}$. Тогда

$$\| M \|^2 = \sum_{f=1}^F \left(\left(\frac{g_f}{z_{cf}} \right)^2 (\mathbf{i}_f^2 + \mathbf{j}_f^2) \right) = \sum_{f=1}^F \left(\left(\frac{g_f}{z_{cf}} \right)^2 \right) \quad (32)$$

$$\| \delta M \|^2 \approx \mathcal{G}_j^2 \| M \|^2 \quad (33)$$

Подставляя (32, 33) в выражение (25), получим:

$$\| \delta M \|^2 \approx \mathcal{G}_j^2 \| M \|^2 \approx \| \mathbf{m}_z \|^2 \frac{\| \delta S \|^2}{h \| S_2 \|^2} = \| \mathbf{m}_z \|^2 \frac{\delta h}{h} = \| \mathbf{m}_z \|^2 \cdot \frac{\sigma_n}{\sigma_3} \quad (34)$$

Выражая \mathcal{G}_j из (34), получим окончательно:

$$\mathcal{G}_j = \frac{\| \mathbf{m}_z \|^2}{\| M \|^2} \cdot \frac{\sigma_n}{\sigma_3} \quad (35)$$

Результаты численного моделирования представлены на рис. 15-18. Приведенные графики строились на модели «Рельеф» при разрешении 2000 пикселей в строке изображения и погрешности детектора 0.1 пикселя, среднеквадратичная высота перепадов рельефа $h=0.1$ km, $a=2$ km, $P \sim 400$, $F = 41$. Делалось несколько случайных сбросов, результаты которых затем были усреднены. На основании критерия (13),(14), областью применимости приближения МОП для зависимости от h является $h < 0.1$ или для обратной величины $1/h > 10$, а в зависимости от N - $N < 5000$, для обратной величины соответственно $1/N > 2 * 10^{-4}$. На рис. 15-18 маркерам «1» и «2» соответствуют величины погрешностей, полученные из (10)-(11) для приближения МОП и перспективной проекции соответственно. Пунктирной «4» и сплошной линии «3» соответствуют оценки погрешностей (25), (35), где сингулярные числа брались из разложения матрицы измерения W . Для сплошной линии «3» использовалось разложение исходной матрицы W (2), что соответствует оценке погрешности для приближения МОП. Для пунктирной линии «4» использовалась матрица, полученная в ходе итераций (9), что соответствует оценке погрешности для перспективной проекции. Штрихпунктирная линия «5» получена путем подстановки в (25, 35) аналитических выражений для сингулярных чисел (22) и (23). На представленных графиках видно, что результаты вычислений по формулам оценок (25), (35) и (22),(23) хорошо согласуются с фактическими значениями погрешностей перспективной проекции. Таким образом, сформулированная гипотеза верна.

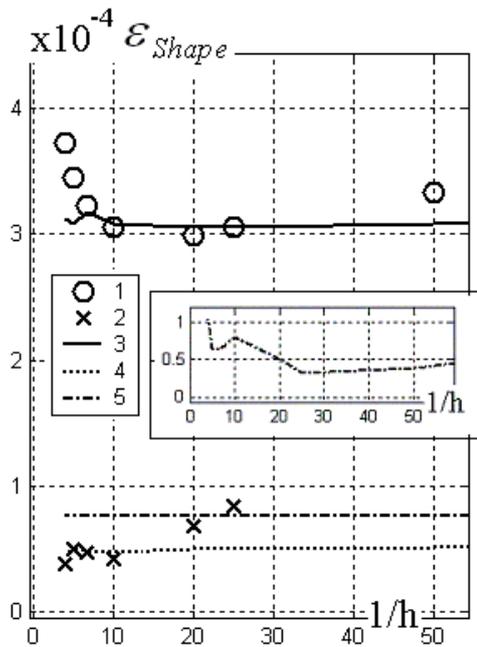


Рис. 15. Зависимость абсолютной погрешности восстановления формы от величины $1/h$.

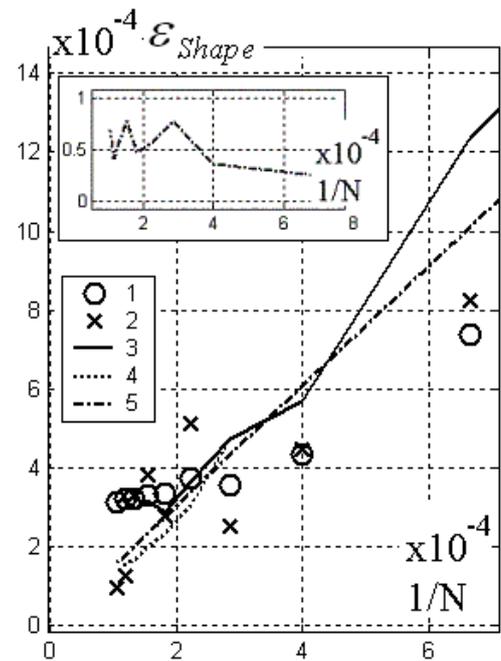


Рис. 16. Зависимость абсолютной погрешности восстановления формы от $1/N$.

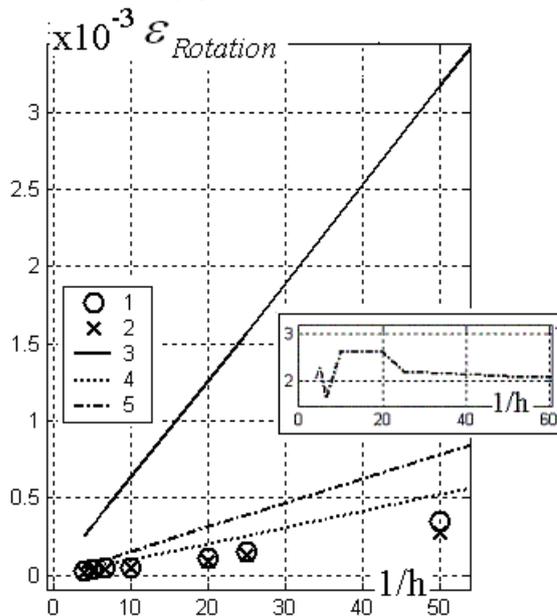


Рис. 17. Зависимость абсолютной погрешности восстановления ориентации камер от $1/h$.

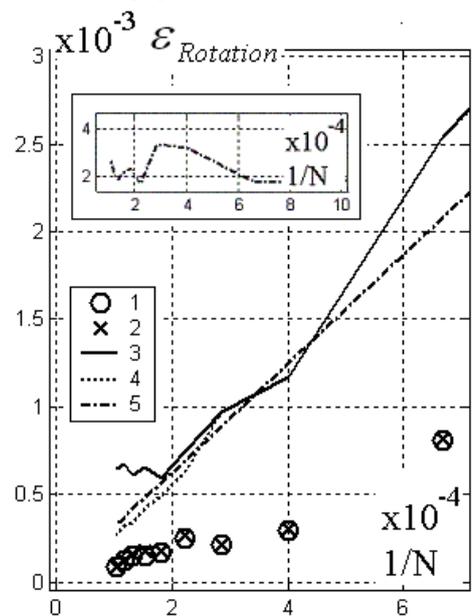


Рис. 18. Зависимость абсолютной погрешности восстановления ориентации камер от $1/N$.

6. Пример: применение результатов анализа алгоритмов к задаче восстановления рельефа по спутниковым изображениям.

Критерий выбора приближения и критерий разрешимости задачи малой глубины и оценки погрешностей проверялись также на модели «Ступенька». Данная модель, как и модель «Рельеф» имеют своей целью проиллюстрировать принципиальную возможность предлагаемого подхода восстановить данные, полученные с искусственного спутника Земли с удовлетворительной точностью.

Спецификой традиционных методов восстановления земного рельефа по последовательности цифровых изображений, полученных с искусственного спутника Земли, является высокое требование к точности восстановления ориентаций камер, определяемое как угловое разрешение одного пикселя (до 10^{-5} – 10^{-6} радиан).

Модель «Ступенька» (рис. 6) рассматривалась при следующих параметрах: $H = 1000$ км; $N = 2000 \dots 10000$; $a = 2$ км; $F = 10 \dots 100$, $P = 20 \dots 1000$. $\alpha = 30 \dots 80$. Камера всегда ориентирована на центр масс объекта.

Отсюда параметры для вычисления критерия разрешимости задачи:

$$g_f = H/a = g; \quad i_{zf} = 0; \quad j_{zf} = -\sin\alpha_f;$$

$$z_{cf} = R\left(\sqrt{\cos^2 \alpha_f + b^2} - \cos \alpha_f\right), \quad \text{где } b^2 = 2\frac{H}{R} + \frac{H^2}{R^2}; \quad (36)$$

$$\rho = \frac{1}{N} \mu = \frac{1}{N} \sqrt{\int_{-1/2}^{1/2} x^2 dx} = \frac{1}{N} \sqrt{\frac{1}{12}} \approx 0.3 \frac{1}{N},$$

Где μ – погрешность детектора в определении координаты пикселя, уже вводившаяся выше. Здесь точность детектора полагалась равной одному пикселю. Тогда критерий приобретает вид:

$$\sigma_3 > \sigma_n,$$

$$\text{где } \sigma_3 = \frac{h}{2} \sqrt{2FPJ}, \quad \text{где } J = \left(\frac{H}{aR}\right)^2 \frac{1}{2F} \sum_{f=1}^F \left(\frac{\sin \alpha_f}{\left(\sqrt{\cos^2 \alpha_f + b^2} - \cos \alpha_f\right)} \right)^2; \quad (37)$$

$$\sigma_4 = \frac{0.3}{N} \sqrt{2FP}.$$

В данной модели положение камер определяется единственным параметром – углом α . Тогда введем функцию случайного распределения камер $\nu = \nu(\alpha)$. Тогда получим выражение для J из (24):

$$J = \left(\frac{H}{R}\right)^2 \int_{-\infty}^{+\infty} \left(\frac{\sin \alpha}{\left(\sqrt{\cos^2 \alpha + b^2} - \cos \alpha\right)} \right)^2 \nu(\alpha) d\alpha. \quad (38)$$

Подстановка параметров задачи в полученный критерий показала, что минимальная разрешимая высота h соизмерима с разрешением изображения. К примеру при разрешении 1м в 1 пикселе минимальная разрешимая высота составляет 2 м.

Рис 19-22 по содержанию и обозначениям аналогичны рис 15-18. Параметры модели «Ступенька» брались следующими: разрешение 2000 пикселей в строке изображения и погрешности детектора 0.1 пикселя, высота ступеньки 0.1 km, $a=2$ km, $P \sim 400$, $F = 41$. На основании критерия (13), (14) областью применимости приближения МОП для зависимости от h является $h < 0.2$ или для обратной величины $1/h > 5$, а для зависимости от N - $N < 10000$, для обратной величины соответственно $1/N > 10^{-4}$.

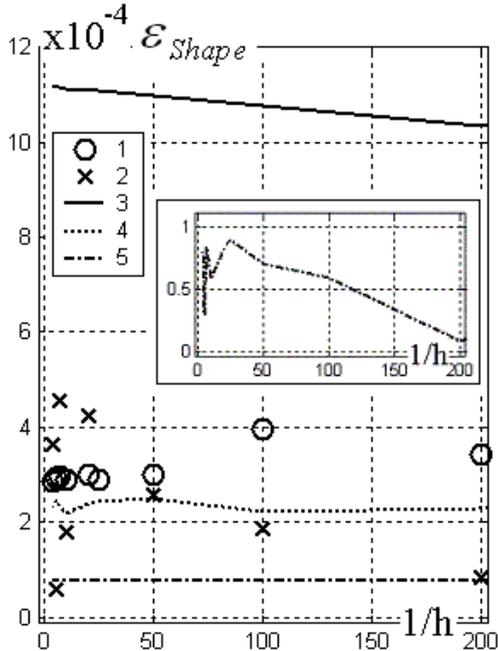


Рис. 19. Зависимость абсолютной погрешности восстановления формы от высоты h .

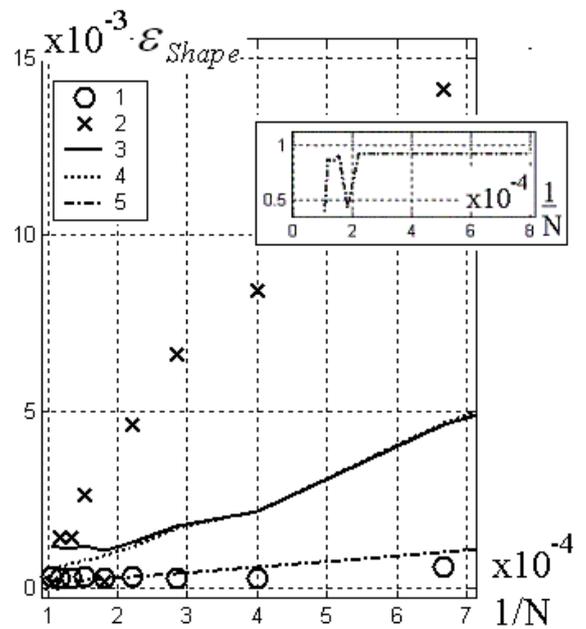


Рис. 20. Зависимость абсолютной погрешности восстановления формы от разрешения изображения в пикселях.

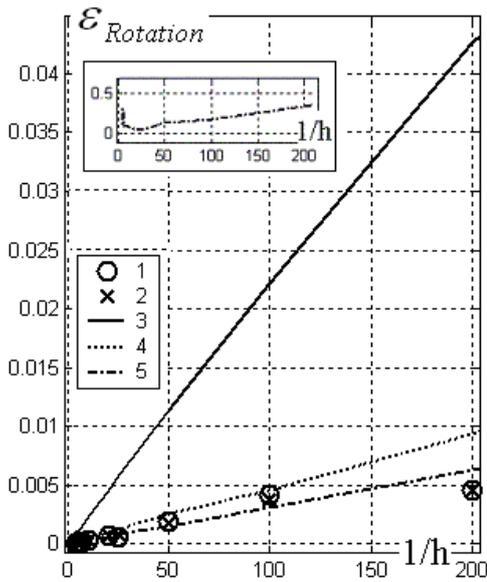


Рис. 21. Зависимость абсолютной погрешности восстановления ориентации камер от высоты ступеньки.

Рассмотрим частный случай формулы (38). Пусть камеры распределены равномерно, что соответствует ситуации, когда спутник, двигаясь по орбите производит съемку через равные углы, тогда

$$\nu(\alpha) = \begin{cases} \frac{1}{2\tilde{\alpha}}, & |\tilde{\alpha}| \leq \alpha, \\ 0, & |\tilde{\alpha}| \geq \alpha. \end{cases}$$

а выражение (37) для σ_3 приобретает вид:

$$\sigma_3^2 = \left(\frac{h}{2}\right)^2 FPJ = \frac{1}{2\alpha} \left(\frac{h}{2a} \frac{H}{R}\right)^2 FPI, \text{ где } I = \int_{-\alpha}^{\alpha} \frac{1 - \cos^2 \tilde{\alpha}}{\left(\sqrt{\cos^2 \tilde{\alpha} + b^2} - \cos \tilde{\alpha}\right)^2} d\tilde{\alpha} \quad (39)$$

Посредством замены $x = \cos \tilde{\alpha}$; $d\tilde{\alpha} = -\frac{dx}{\sqrt{1-x^2}}$ и умножения числителя и знаменателя интеграла

(39) на выражение $\left(\sqrt{x^2 + b^2} + x\right)^2$, интеграл может быть разложен на сумму табличных, или сводящихся к табличным интегралов. После вычисления и упрощения результат для σ_3^2 может быть представлен в виде:

$$\sigma_3^2 = \frac{1}{2\alpha} \left(\frac{h}{2a} \frac{H}{R}\right)^2 FP \frac{1}{b^4} \left(-\frac{1}{4} \sin 2\alpha (\cos 2\alpha + 2b^2) + \frac{1}{2} (1 + 2b^2) \alpha - \right. \\ \left. - \frac{1}{2} (\cos 2\alpha + b^2) \sin \alpha \sqrt{\cos^2 \alpha + b^2} + \left(\frac{1+b^2}{2}\right)^2 \arccos \left(\frac{\cos 2\alpha + b^2}{1+b^2}\right) \right) \quad (40)$$

Формула (40) не удобна для инженерных оценок. Поэтому целесообразно построить функцию, аппроксимирующую данное выражение в диапазоне параметров, представляющих практический интерес: $\alpha \in [10..70]$ градусов и $H \in [200..3000]$ км.

Анализ выражения (40) и поиск аппроксимирующей функции проводились в системе компьютерной математики Maple. Оценивались относительные погрешности аппроксимации – среднеквадратичная и наихудшая по выбранному диапазону параметров. Получено две аппроксимации различной сложности:

$$\sigma_3^2 \cong \frac{1}{4} \frac{h^2 H^2 P (0.9 + 0.8 \cdot b^{1.83}) \sin^e \alpha}{a^2 R^2 \alpha \cdot b^4} \quad (41)$$

Среднеквадратичная относительная погрешность 10%, наихудшая 50%. Более точная аппроксимация имеет погрешности 6% и 9% соответственно и определяется выражением:

$$\sigma_3^2 \cong \frac{1}{4} \frac{h^2 H^2 P (0.7024 + 1.051 \cdot b^{1.9}) \sin^e ((1.197 - 0.2182 \cdot b) \cdot \alpha - 0.021)}{a^2 R^2 \alpha \cdot b^4} \quad (41')$$

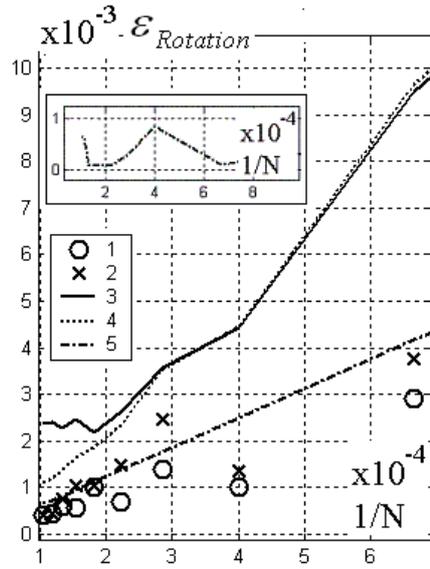


Рис. 22. Зависимость абсолютной погрешности восстановления ориентации камер от разрешения изображения в пикселях.

На рис. 23-24 представлены графики для зависимости (40) и аппроксимации (41') от угла α и от величины параметра b , где $b^2 = 2H/R + (H/R)^2$. Визуально графики практически неразличимы. На вспомогательных графиках приведены относительные погрешности аппроксимации.

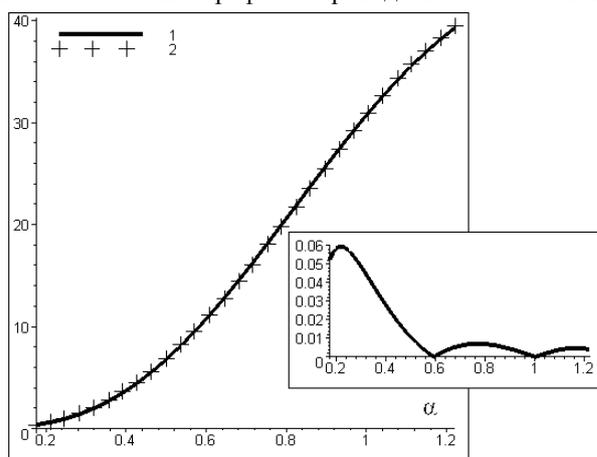


Рис. 23. 2- функция (40) и 1 - ее аппроксимация (41') в зависимости от угла α .

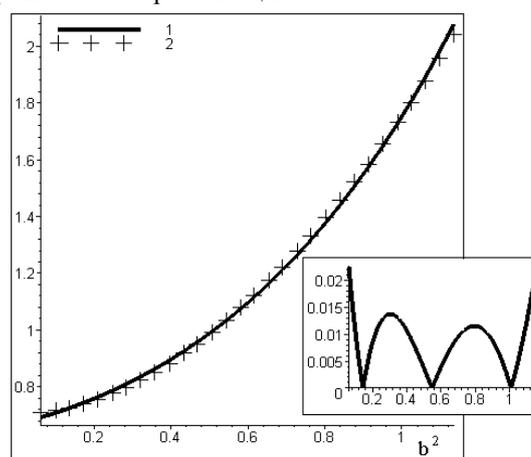


Рис. 24. 2- функция (40) и 1 - ее аппроксимация (41') в зависимости от параметра b .

Видно, что абсолютные величины погрешности не превосходят 6% и 2% для рис.23-24 соответственно.

7. Практическое использование полученных оценок погрешностей.

Полученный критерий выбора приближения и формулы для оценки погрешностей восстановления трехмерных сцен методом факторизации матриц, могут быть использованы в двух принципиально различных задачах:

Априорный анализ качества результатов восстановления сцены может быть произведен в случае, когда известны средние статистические характеристики сцены и условий съемки. Тогда можно с помощью формул (13)-(14) до обработки выбрать алгоритм (МОП или перспектива), который следует применять для восстановления сцены. С помощью формул (25),(35) и (22)-(24) могут быть оценены погрешности, которые следует ожидать от алгоритма восстановления сцены. На стадии конструирования аппаратуры и/или планирования съемки, варьируя параметры, такие как расстояние до объекта, разрешение камеры и диапазон углов, с которых производится съемка можно добиться желаемой точности или показать, что требуемая точность не достижима. Пример методики расчетно-параметрического и аналитического исследования погрешностей применительно к конкретной задаче дан выше (36)-(41).

Апостериорный анализ достоверности восстановленного результата может быть произведен, даже если восстанавливается неизвестная сцена. Перед этим система восстановления трехмерного образа должна быть калибрована (тестирована) с целью определения фактической точности детектора μ . После решения задачи МОП (блок 1 на схеме рис. 3) следует с помощью (14) принять решение о целесообразности применения итерационного алгоритма и, если требуется, выполнить итерационный алгоритм. Затем (блок 7 на схеме рис. 3) убедиться, что после последней итерации σ_3 больше, а σ_4 не превосходит по порядку величины шумовое значение σ_n , полученное на стадии калибровки. Если это условие не выполнено, достоверность полученных результатов не может быть гарантирована при данном подходе. Если выполнено, по формулам (25),(35), используя те же σ_3 и σ_4 следует вычислить погрешности формы и ориентации камер. Таким образом, кроме восстановления трехмерного образа сцены и движения камер вычисляется и погрешность этих величин, то есть четко определена степень достоверности результата восстановления.

Заключение.

Алгоритмы восстановления трехмерных сцен и движения, основанные на факторизации матриц, работают в условиях избыточности информации. Если трехмерные точки образуют жесткую конфигурацию (лежат на поверхности твердого тела), то требования симметрии накладывают жесткие ограничения на то, какими могут быть исходные данные (матрица W). При наличии шумов исходные данные могут слегка нарушать эти закономерности, но тоже закономерным статистическим образом. Это дает уникальную возможность для алгоритмов рассматриваемой группы одновременно с решением задачи восстановления формы и движения оценить погрешность получаемых данных исходя непосредственно из исходных данных задачи (изображений, требуется провести ряд предварительных тестов для оценки средней субпиксельной точности детектора точек). Более того, оказывается, что оценки погрешности можно провести в рамках того же аппарата сингулярного анализа, который используется при решении самой задачи, используя уже полученные в ходе ее решения результаты, а в некоторые оценки получить еще до завершения алгоритма.

В ходе анализа алгоритмов восстановления трехмерных сцен по последовательности цифровых изображений, основанных на факторизации матриц, в приближении масштабируемой ортографической проекции (МОП) и в перспективной проекции были получены следующие результаты.

- Построен критерий выбора приближения (13)-(14). Проверка критерия проводилась на примере различных моделей и показала хорошее совпадение результатов опыта с теоретическими данными в зависимости от различных параметров. При этом величина эмпирического коэффициента η оказалась не превышающей четырех.
- Построенный критерий применен для адаптивного выбора алгоритма, что показано на рис. 3 в виде блока выбора «2». Никакой дополнительной информации, кроме обычных исходных данных для стандартного алгоритма [1] не требуется.
- Предложен критерий априорной оценки разрешимости задачи восстановления сцены малой глубины (19) и для него построены оценка погрешности оцифровки (22) и оценка снизу величины, содержащей информацию об объекте, в приближении сцены малой глубины (23-24). Оценки показали высокую точность при проверке на различных модельных данных. Точность оценки шумов не хуже 60%, точность оценки второй величины не хуже 15%.
- В приближении сцены малой глубины получена оценка погрешностей восстановления формы объекта (25) и ориентаций камер (35). Формулы оценок погрешностей непосредственно связаны с критерием разрешимости. Более высокая точность восстановления будет соответствовать изображению с более слабым шумом (для шума оцифровки это большее разрешение) и/или с объектом большей глубины, по форме близкому к кубическому и занимающему почти все изображение. Результаты численного моделирования показали, что формулы (25, 30) завышают оцениваемые погрешности примерно в 2 раза. Формулы (25, 35) могут использоваться как для априорной оценки погрешностей восстановления сцены, для которой известны среднестатистические характеристики, так и для апостериорной оценки, которая вычисляется исключительно на основе исходной матрицы измерений и результатов восстановления, не требуя какого бы то ни было априорного знания о характеристиках сцены. Подробнее см. раздел 6.
- Предложена модификация алгоритмов восстановления трехмерных сцен, основанных на факторизации матриц таким образом, что он возвращает не только восстановленную трехмерную сцену и движение, но и точность, с которой они были восстановлены (рис. 3, блок «7»). Наличие величин оценок погрешностей позволяет получить объективную информацию о степени достоверности результата.
- Для задачи восстановления рельефа по спутниковым данным, полученные выражения доведены до расчетных аналитических аппроксимирующих формул вычисления ожидаемых погрешностей (37), (41).

Полученные критерии и оценки были проверены на нескольких моделях. Особое внимание уделялось моделям, в которых условия съемки и характеристики объектов имитировали условия съемки земного рельефа искусственным спутником Земли. На примерах этих моделей было установлено, что задача становится разрешимой, т.е. отличной от плоской уже при глубине сцены в 2-3 раза превышающей метрическое разрешение изображения. Однако при столь малых высотах точность восстановления, как формы объекта, так и ориентаций камер, очень низка в силу того, что величина, определяющая глубину сцены близка к уровню шума. (См. критерий (19) и формулы оценки погрешности (25, 35).) Для достижения точностей, которые требуются в задачах спутниковой съемки, необходимо восстанавливать сцену в условиях применимости перспективной проекции. Этому соответствует подбор сцен с большей глубиной и/или повышение разрешения изображения, а также точности детектора. К примеру, для достижения точности восстановления ориентаций камер 10^{-5} радиан, необходимо изображение с разрешением 1 м, детектор с субпиксельной точностью порядка 0.1 пикселя и глубина сцены порядка 100 метров.

Литература.

1. Conrad I. Poelman, Takeo Kanade. A Paraperspective Factorization Method for Shape and Motion Recovery: //Technical Report CMU-CS-93-219 / School of Computer Science, Carnegie Mellon University. — 11 December 1993. http://www.ri.cmu.edu/pubs/pub_1189.html, http://www.ri.cmu.edu/people/person_136_pubs.html
2. 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. http://www.graphicon.ru/2002/pdf/Yanova_Re.pdf
3. 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>
4. Yuri Boykov, Olga Veksler, Ramin Zabih. Fast Approximate Energy Minimization via Graph Cuts: /Cornell University, 1999. <http://www.csd.uwo.ca/faculty/yuri/>, <http://research.microsoft.com/~vnk/>, <http://www.cs.cornell.edu/~rdz/index.htm>.

5. S. K. Nayar and M. Watanabe. Telecentric optics for constant-magnification imaging. Technical report, *Dept. of Computer Science, Columbia University*, New York, USA, September 1995. CUCS-026-95. <ftp://ftp.cs.columbia.edu/pub/CAVE/papers/watanabe/nabe\~nayar\~tele\~iuw\~95.ps.gz>.
6. Ali Azarbayejani Tony Jebara and Alex Pentland. 3d structure from 2d motion. Technical report, *MIT Media Laboratory*, May 1999. Perceptual Computing Tech. Rep. #523. http://cgi.media.mit.edu/vismod/tr_pagemaker.cgi. <http://vismod.media.mit.edu/>.
7. Дж. Форсайт., М. Малькольм., К. Моулер. Машинные методы математических вычислений. Пер с англ. -М.:Мир,1980. 280 стр.
8. Н. В. Свешникова, А. С. Чернышев, Д.В. Юрин. Самокалибрующаяся система трехмерного зрения и анализ ее применимости в робототехнике и спутниковой съемке земного рельефа. – В сб. *Труды конференции. 13-я Международная конференция по Компьютерной графике и Машинному зрению ГрафиКон'2003* – стр. 257-260. Москва, 2003.
9. Дж. Деммель. Вычислительная линейная алгебра. – М.: Мир, 2001.
10. *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