

Восстановление трехмерных сцен с помощью методов факторизации: принцип работы и оценка погрешностей.

Н.В. Свешникова*, Д.В. Юрин**.

* Московский физико-технический институт

** ФГУП НПП ОПТЭКС

sveshnikova_n@list.ru, yurin_d@inbox.ru

Аннотация

К настоящему времени разработано большое количество разнообразных алгоритмов восстановления трехмерных сцен по набору цифровых изображений. В настоящей работе представлен обзор группы алгоритмов, основанных на факторизации матриц. Предложена методика оценки погрешностей: а) апостериорная, основанная на данных измерений, требуемых алгоритмами факторизации, никакой дополнительной информации не требуется; б) априорная, использующая предварительные оценки интегральных характеристик формы сцены и движения камер относительно нее.

Введение

К настоящему времени разработано большое количество различных алгоритмов и подходов к задаче восстановления трехмерных сцен по набору цифровых изображений [1-10]. Эти подходы базируются на различных физических принципах и математических методах и имеют различные диапазоны применимости. Кроме того, алгоритмы в различной степени требуют наличия тех или иных априорных данных о восстанавливаемой сцене или о параметрах камер. К примеру, восстановление трехмерной сцены по стереопаре способно дать качественный результат и восстановить трехмерный образ практически для всех точек исходного изображения, однако требует высокой точности калибровки камер стереопары.

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. На данный момент для алгоритмов восстановления по стереопаре создана база тестовых изображений [11], позволяющая производить восстановление одних и тех же сцен различными алгоритмами и сравнивать полученные результаты с правильным ответом (ground truth), полученным непосредственным обмером трехмерной сцены. Это позволяет выявить основные достоинства и недостатки тестируемых алгоритмов, но, в силу ограниченности набора тестовых изображений, не дает ответа на вопросы: на каких сценах обеспечивается сходимость алгоритма и указанная точность восстановления, сходится ли алгоритм и какова точность восстановления конкретной обрабатываемой сцены. 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 отличие от алгоритмов восстановления трехмерных сцен по стереопаре, алгоритмы, основанные на факторизации матриц, не требуют априорных данных ни о камерах, ни о форме сцены, но нуждаются в достаточно большой (не менее ~20) последовательности кадров и в выделении на всей этой последовательности набора так называемых характеристических отметок, что является отдельной задачей, изучаемой и решаемой в области компьютерного зрения [16,17]. Алгоритмы, основанные на факторизации матриц, восстанавливают трехмерные координаты характеристических отметок и камер, ориентацию камер в пространстве, а также позволяют получить точность результата [13,14]. Восстановленный объект представляется в виде меша, построенного на базе восстановленных трехмерных точек. Как правило, количество этих точек недостаточно, чтобы построенный меш адекватно отображал форму объекта.

Таким образом, наиболее разумным представляется комбинирование различных подходов. Алгоритмы, основанные на факторизации матриц способны предоставить необходимую информацию, такую как трехмерные координаты камер, их ориентацию в пространстве и точность, с которой эти величины известны, а также координаты и точность некоторых точек трехмерной сцены для работы алгоритмов, дающих плотное восстановление поверхности сцены (для каждого пиксела изображения), к примеру, для восстановления по стереопаре [4,7,10].

В настоящей работе рассматриваются два метода, основанных на факторизации матриц: приближение масштабированной ортографической проекции (МОП) [1] и итерационный метод в перспективной проекции (ИПП) [2]. 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.

Основной принцип, на котором базируется построение оценок погрешностей, является избыточность исходных данных для методов факторизации, и аналогичен, по сути, оценке статистической погрешности измерений при аппроксимации экспериментальных данных теоретической кривой методом наименьших квадратов. 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.

1. Структура решения задачи восстановления трехмерной сцены.

Объектом, или трехмерной сценой, будем называть предмет, который не деформируется во время получения последовательности его изображений, или группу таких предметов, неподвижных относительно друг друга.

Восстановление трехмерной сцены по последовательности цифровых изображений можно разбить на несколько этапов (рис. 1). Пусть $f = 1..F$ кадров получено в результате съемки трехмерной сцены при различных взаимных положениях камер(ы) и сцены.

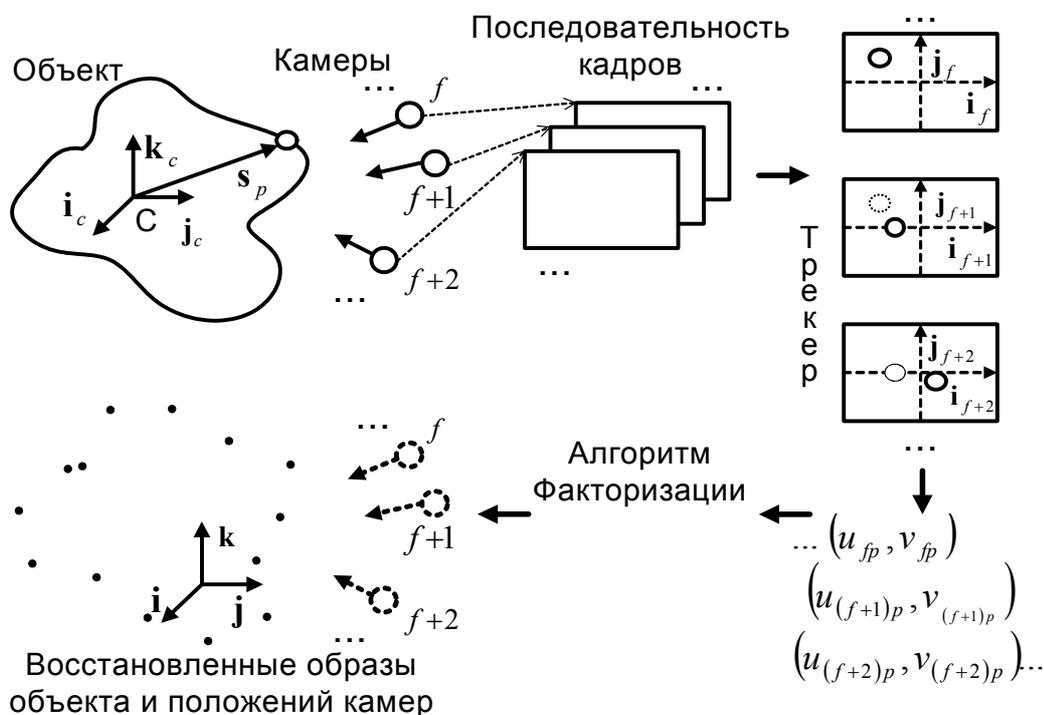


Рис. 1. Этапы восстановления трехмерной сцены.

Последовательность изображений сначала должна быть обработана программой обнаружения и прослеживания характеристических точек. Примерами подходов, служащих этим целям, могут послужить алгоритмы, описанные в [16, 17].

Как изображено в верхнем правом углу рис. 1 программа обнаружения и прослеживания характеристических точек фиксирует некоторую точку с координатами (u_{fp}, v_{fp}) на одном из кадров и ищет ее координаты $(u_{(f+1)p}, v_{(f+1)p})$, $(u_{(f+2)p}, v_{(f+2)p})$ и т.д. на последующих кадрах. Результатом работы программы является список двумерных координат для данной характеристической точки на всех кадрах последовательности, где она может быть обнаружена, (см. рис 1, справа внизу). Как правило, количество характеристических точек, обнаруживаемых на последовательности кадров, имеет порядок десятков.

Эти данные поступают на вход программы восстановления трехмерной сцены алгоритмами факторизации. В результате их работы в некоторой системе координат восстанавливаются трехмерные координаты всех камер, их пространственные ориентации и трехмерные координаты характеристических

точек, выделенных на предыдущем этапе (см. рис.1 слева внизу). Кроме того, алгоритм предоставляет информацию о достоверности полученного результата.

Количество характеристических точек не всегда достаточно для получения узнаваемого образа исходного объекта. Однако, восстановление информации об ориентации и расположении камер, а также точностей этих величин, позволяет использовать эти результаты в качестве дополнительных входных данных для алгоритмов, способных дать более качественный результат. Известные координаты восстановленных точек объекта и их погрешности могут быть использованы для наложения дополнительных ограничений и оценки достоверности результата таких алгоритмов.

Таким образом, алгоритмы факторизации представляются хорошим выбором в качестве начального приближения для восстановления произвольной неизвестной сцены.

2. Принципы восстановления данных методом факторизации матриц.

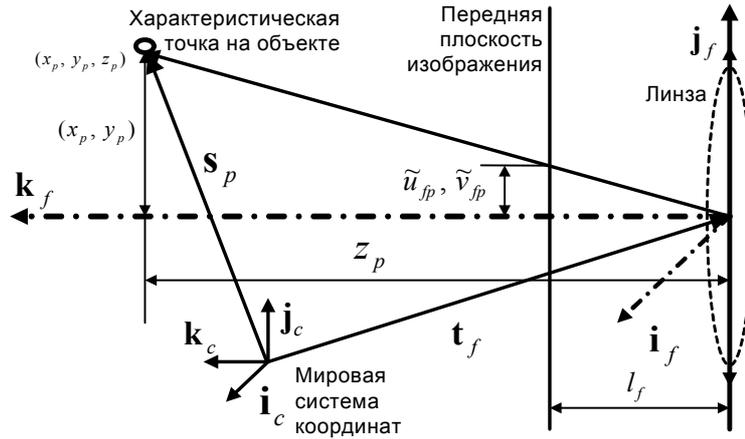


Рис. 2. Постановка задачи восстановления трехмерной сцены.

Рассмотрим принципы, на которых базируются алгоритмы [1-3].

Пусть имеется $f = 1..F$ 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 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. 2) 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 (КСК), 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. \tilde{u}_f^e - размер одного пикселя на фоточувствительной матрице; N - ширина изображения в пикселях, $\beta_{\max f}$ - максимальный угол зрения камеры. 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 $z_f = -\mathbf{k}_f \mathbf{t}_f$, $z'_f = z_f / g_f$, и полагая $\mathbf{k}_f \mathbf{s} / z_f \ll 1$, получим формулы для приближения МОП [1]:

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

Уравнения (2) можно представить в матричной форме:

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

где

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

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

Матрица \mathbf{W}' представима в виде произведения матриц, ранг которых не превосходит величины ζ . Если не все камеры ориентированы одинаково, то $\text{rank } \mathbf{M} = 3$ и ζ равна трем, если объект объемный, двум, если плоский и единице, если одномерный [3].

$$\zeta = \text{rank } \mathbf{W} = \text{rank } \mathbf{S} = \begin{cases} 1 & - \text{ для } 1D \text{ объекта} \\ 2 & - \text{ для } 2D \text{ объекта} \\ 3 & - \text{ для } 3D \text{ объекта} \end{cases} \quad (4)$$

Поэтому и ранг матрицы \mathbf{W}' не превышает указанных значений. Производится сингулярное разложение матрицы \mathbf{W}' :

$$\mathbf{W}' = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T = (\mathbf{U}\sqrt{\mathbf{\Sigma}})(\sqrt{\mathbf{\Sigma}}\mathbf{V}^T) = \widehat{\mathbf{M}}\widehat{\mathbf{S}}, \quad \text{rank}(\mathbf{W}') \leq \zeta \quad (5)$$

здесь \mathbf{U} , \mathbf{V} - ортогональные матрицы, а $\mathbf{\Sigma} = \text{diag}(\sigma_1, \dots, \sigma_n)$, $n = \min(2F, P)$, $\sigma_1, \dots, \sigma_n$ - сингулярные числа матрицы \mathbf{W} . Не ограничивая общности, будем считать, что сингулярные числа упорядочены в порядке убывания, тогда выберем из матрицы $\mathbf{\Sigma}$ верхнюю подматрицу размера $\zeta \times \zeta$. Выбирая из матриц \mathbf{U} и \mathbf{V} подматрицы из строк, соответствующих ζ наибольшим сингулярным числам, получаем аппроксимацию к матрице \mathbf{W} . Понятно, что такое разложение не единственно, поскольку между матрицами $\widehat{\mathbf{M}}$ и $\widehat{\mathbf{S}}$ всегда можно вставить произведение прямой и обратной матриц \mathbf{Q} ранга 3, тогда можно записать [1]:

$$\mathbf{W}' = \widehat{\mathbf{M}}\widehat{\mathbf{S}} = \widehat{\mathbf{M}}(\mathbf{Q}\mathbf{Q}^{-1})\widehat{\mathbf{S}} = (\widehat{\mathbf{M}}\mathbf{Q})(\mathbf{Q}^{-1}\widehat{\mathbf{S}}) = \mathbf{M}\mathbf{S},$$

Здесь \mathbf{M} и \mathbf{S} обозначают истинные матрицы движения и формы.

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

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

Последнее условие необходимо для того, чтобы система (6) была неоднородной и имела нетривиальное решение. Система является линейной относительно матрицы $\widetilde{\mathbf{Q}} = \mathbf{Q}\mathbf{Q}^T$. В данной работе система (6) решалась методом наименьших квадратов (МНК) с применением SVD [12]. После нахождения матрицы $\widetilde{\mathbf{Q}}$ искомая матрица \mathbf{Q} извлекалась посредством факторизации.

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

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

$$\tilde{M} = MR_0, \quad \tilde{S} = R_0^T S, \quad R_0 = [\mathbf{i}_{f_0}, \mathbf{j}_{f_0}, \mathbf{k}_{f_0}], \quad \text{где } \mathbf{k}_{f_0} = \mathbf{i}_{f_0} \times \mathbf{j}_{f_0}, \quad (7)$$

где f_0 выбирается из диапазона $1..F$.

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 уточнить восстанавливаемый результат [2].

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 определяется из (5). 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 $\sigma_{\zeta+1}$, σ_{ζ} - сингулярные числа матрицы W' из (9). Этот шаг был реализован с помощью процедуры FMIN [12], которая носит итерационный характер.

Блок-схема итерационного алгоритма приведена на рис. 3. Точечным контуром обозначен блок, который следует внести в алгоритм по результатам анализа, описываемого в данной работе. Его описание будет приведено далее по ходу изложения.

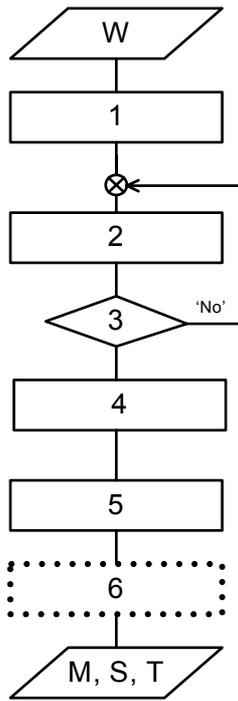


Рис. 3 Блок-схема итеративного алгоритма.

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'), (5), (6)) получаются величины $\mathbf{s}_p, \mathbf{k}_f, z'_f$. Положить $q=0, \xi^{(0)}=0$.
2. Одна итерация перспективного алгоритма.
 - 2.1. Вычислить W_2 , по формуле (9), используя предыдущие значения $\mathbf{s}_p, \mathbf{k}_f, z'_f$
 - 2.2. Найти $\xi^{(q)} = \arg \min_{\xi} (\sigma_{\xi+1} / \sigma_{\xi})$.
 - 2.3. Вычислить W' , по формуле (9).
 - 2.4. Решить задачу в МОП для W' (5)-(6).
3. Определение необходимости дальнейших итераций: $|\xi^{(q)} - \xi^{(q-1)}| < \varepsilon$, где q - номер итерации, ε - требуемая погрешность определения параметра ξ .
4. 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 .
5. Устранение неоднозначности по x и y компонентам объекта (выравнивание по f -той камере, (7)).
6. Вычисление погрешностей восстановления формы объекта и ориентаций камер (20), (30).

Различие между приближением МОП и перспективной проекцией может быть проинтерпретировано геометрически. В модели перспективной проекции (ППИ) лучи идут именно так, как это происходит в реальности, то есть по прямой. На рис. 4 реальной точке объекта P в результате перспективного проецирования на ППИ соответствует точка $(\tilde{u}'_{fp}, \tilde{v}'_{fp})$.

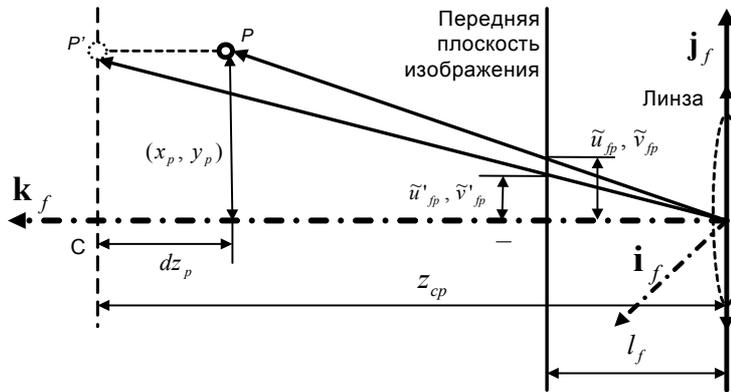


Рис. 4. Геометрическая интерпретация различия между приближением МОП и перспективной проекцией.

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. 4 камера направлена на центр масс объекта, точка P' является проекцией реальной точки объекта P на описанную выше воображаемую плоскость, а точка $(\tilde{u}'_{fp}, \tilde{v}'_{fp})$ - 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 [14].

3. 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:

$$\begin{aligned} L_1: & Ax + By + z + D - h = 0; \\ L_2: & Ax + By + z + D + h = 0; \end{aligned} \quad (10)$$

Расстояние h полагается малым по сравнению с поперечными размерами видимых участков плоскостей.

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.

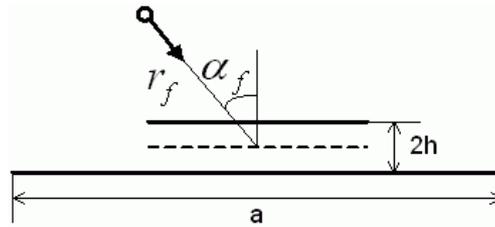


Рис. 5. Постановка задачи разрешимости сцены малой глубины.

На рис. 5 изображена постановка задачи в условиях предложенной модели. Характерными параметрами являются: a – ширина плоскости, α_f – угол отклонения f -той камеры от вертикали, r_f – расстояние от f -той камеры до центра масс объекта.

Метод факторизации матриц базируется на том, что для трехмерного объекта ранг идеальной матрицы W' , обозначенный выше ζ определяется из (4):

$$W' = MS, \quad (11)$$

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:

$$W = W^0 + \omega, \quad (12)$$

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

$$\langle \omega^2 \rangle = \frac{1}{2FP} \sum_{i=1}^{2F} \sum_{j=1}^P \omega_{ij}^2 = \rho^2; \quad \langle \omega \rangle = \frac{1}{2FP} \sum_{i=1}^{2F} \sum_{j=1}^P \omega_{ij} = 0. \quad (13)$$

Величина шума определяется как безразмерная, что следует из (12), 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 [1].

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 (10) для плоского объекта z будет линейно зависеть от x и y , и S будет иметь ранг 2). Тогда собственный вектор-столбец в матрице V (см. формулу (5)), соответствующий сингулярному числу ответственному за шум будет ошибочно принят за составляющую формы [3]. Чтобы этого избежать, необходима априорная оценка шумового сингулярного числа σ_n , и условием разрешимости будет:

$$\sigma_3 > \sigma_n. \quad (14)$$

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

$$z_p = -(Ax_p + By_p + D \mp h) = -(Ax_p + By_p + D) \pm h = z_p^{(1)} \pm h. \quad (15)$$

Тогда, вводя обозначения

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

$$\mathbf{S} = \mathbf{S}_1 + h \cdot \mathbf{S}_2, \quad \mathbf{S}_1 = (s_1^{(1)} \quad \dots \quad s_p^{(1)}) \quad , \quad \mathbf{S}_2 = (s_1^{(2)} \quad \dots \quad s_p^{(2)}).$$

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

$$W' = \mathbf{M}\mathbf{S}_1 + h \cdot \mathbf{M}\mathbf{S}_2 = W_1' + W_2' \quad (16)$$

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 [15] (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 (13) получим:

$$\sigma_n^2 \leq \|\omega\|^2 = \sum_{f=1}^{2F} \sum_{p=1}^P w_{fp}^2 = 2FP\rho^2; \quad (17)$$

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' .

$$\sigma_3^2 = \|W_2'\|^2 = FPh^2 J, \quad \text{где } J = \frac{1}{F} \sum_{f=1}^F \left(\left(\frac{g_f}{z_{cf}} \right)^2 (i_{zf}^2 + j_{zf}^2) \right) \quad (18)$$

Вычисления по формуле (18) 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 (18) 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 (18), можно приблизить интегралом:

$$J = \int_{-\infty}^{+\infty} \left(\frac{g(\vec{r})}{|\vec{r}|} \right)^2 (i_z^2(\vec{r}) + j_z^2(\vec{r})) \nu(\vec{r}) d\vec{r} \quad (19)$$

В условиях конкретной модели в результате вычисления интеграла можно получить аналитическую формулу для априорной оценки разрешимости задачи восстановления почти плоского объекта.

4. 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 h_{\min} есть одновременно абсолютная погрешность восстановления формы объекта произвольной формы δh , то есть $h_{\min} \sim \delta h$. Далее справедливость этой

гипотезы и высокая точность оценки будет показана численным моделированием. Сделаем некоторые предварительные замечания.

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

$$\frac{\delta h}{h} = \frac{\sigma_n}{\sigma_3} \quad (20)$$

Напомним, что при восстановлении трехмерных сцен методом факторизации ищется решение математической задачи: (5) 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

$$\min_{M, S} \|W'_s - M \cdot S\| \leq \|\omega\| \quad (21)$$

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

Матрицы M и S тесно взаимосвязаны. Так изображения похожих объектов могут выглядеть с некоторых 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 .

$$W' = (M + \delta M) \cdot (S + \delta S) = M \cdot S + \delta M \cdot S + M \cdot \delta S + \delta M \cdot \delta S \approx M \cdot S + \delta M \cdot S + M \cdot \delta S \quad (22)$$

Сравнивая (20) и (21), 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:

$$\|\delta M\| \approx \|M\| \cdot \frac{\|\delta S\|}{\|S\|} \quad (23)$$

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 (23) $\|M\|$ на $\|\mathbf{m}_z\|$ - норму третьего вектор-столбца матрицы M :

$$\|\delta M\| \approx \|\mathbf{m}_z\| \cdot \frac{\|\delta S\|}{\|S\|}$$

Т.к. $\frac{\|\delta S\|}{\|S\|}$ - относительная погрешность восстановления формы сцены, то в силу гипотезы (20):

$$\frac{\|\delta S\|}{\|S\|} = \frac{\sigma_n}{\sigma_3}$$

$$\|\delta M\| \approx \|\mathbf{m}_z\| \cdot \frac{\sigma_n}{\sigma_3} \quad (23')$$

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

$$M = \begin{pmatrix} \frac{g_1}{z_1} \mathbf{i}_1 & \frac{g_1}{z_1} \mathbf{j}_1 & \dots & \frac{g_F}{z_F} \mathbf{i}_F & \frac{g_F}{z_F} \mathbf{j}_F \end{pmatrix}^T, \text{ то } \delta M = \begin{pmatrix} \frac{g_1}{z_1} \delta \mathbf{i}_1 & \frac{g_1}{z_1} \delta \mathbf{j}_1 & \dots & \frac{g_F}{z_F} \delta \mathbf{i}_F & \frac{g_F}{z_F} \delta \mathbf{j}_F \end{pmatrix}^T \quad (24)$$

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

$$\|\delta M\|^2 = \sum_{f=1}^F \left(\frac{g_f}{z_f} \right)^2 (\delta i_{zf}^2 + \delta j_{zf}^2) \quad (25)$$

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

$$\|\delta M\|^2 = (\langle \delta i_z^2 \rangle + \langle \delta j_z^2 \rangle) \sum_{f=1}^F \left(\frac{g_f}{z_f} \right)^2 \quad (26)$$

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 (26) 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

$$\|M\| = \sqrt{\sum_{f=1}^F \left(\left(\frac{g_f}{z_{cf}} \right)^2 (\mathbf{i}_f^2 + \mathbf{j}_f^2) \right)} = \sqrt{2 \sum_{f=1}^F \left(\left(\frac{g_f}{z_{cf}} \right)^2 \right)} \quad (27)$$

$$\|\delta M\| \approx \mathcal{G}_j \|M\| \quad (28)$$

Подставляя (28) в (23'), получим:

$$\mathcal{G}_j \|M\| = \|\mathbf{m}_z\| \cdot \frac{\sigma_n}{\sigma_3} \quad (29)$$

Выражая \mathcal{G}_j из (29), получим окончательно:

$$\mathcal{G}_j = \frac{\|\mathbf{m}_z\|}{\|M\|} \cdot \frac{\sigma_n}{\sigma_3} \quad (30)$$

5. Результаты.

Проверка полученных оценок проводилась методами численного моделирования.

Модель «Рельеф» имитирует набор характеристических точек, выделенных на последовательности цифровых изображений, полученной со спутника, летящего по окружности радиуса $R+H$, где R – радиус Земли, H – высота орбиты над поверхностью, $H \sim 1000$ км, над участком плоскости размера $a \times a$, каждая точка которой возмущена случайным образом относительно плоскости $z = 0$ так, что среднеквадратичное отклонение равно h .

Плоскость помещается в плоскость осей XU и ее центр совпадает с началом МСК. 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 $|\alpha| \leq 90^\circ$. В идеальную модель вносился шум оцифровки, определяемый количеством пикселей N в строке изображения.

Пример восстановленной модели приведен на рис. 6 в виде сетки характеристических точек.

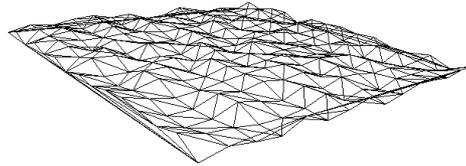


Рис. 6. Модель «Рельеф».

Проверка полученных оценок погрешностей строилась по следующей схеме:

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

$$\varepsilon_{Shape} = \frac{1}{a} \sqrt{\frac{1}{P} \sum_{p=1}^P (\mathbf{s}_p - \mathbf{s}_p^0)^2}; S = [\mathbf{s}_1^T \quad \mathbf{s}_2^T \quad \dots \quad \mathbf{s}_P^T]; S^0 = [\mathbf{s}_1^{0T} \quad \mathbf{s}_2^{0T} \quad \dots \quad \mathbf{s}_P^{0T}]. \quad (31)$$

$$\varepsilon_{Rotation} = \sqrt{\frac{1}{3F} \sum_{f=1}^F (\mathbf{i}_f - \mathbf{i}_f^0)^2 + (\mathbf{j}_f - \mathbf{j}_f^0)^2 + (\mathbf{k}_f - \mathbf{k}_f^0)^2}; \quad (32)$$

$$R = [\mathbf{i}_1 \quad \mathbf{j}_1 \quad \mathbf{k}_1 \quad \dots \quad \mathbf{i}_F \quad \mathbf{j}_F \quad \mathbf{k}_F]; \quad R^0 = [\mathbf{i}_1^0 \quad \mathbf{j}_1^0 \quad \mathbf{k}_1^0 \quad \dots \quad \mathbf{i}_F^0 \quad \mathbf{j}_F^0 \quad \mathbf{k}_F^0].$$

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 (20), (30).

На рис 7-10 приведены зависимости третьего и шумового сингулярных чисел в зависимости от таких параметров как глубина сцены и разрешение изображения. Остальные параметры модели выбирались следующими: $\alpha \leq 60^\circ$, $h = 0.05$ км, размер площадки 2×2 км, разрешение 1000 пикселей при точности детектора 0.1 пиксель [16]. Для каждого маркера проводилось несколько новых генераций случайной поверхности, погрешности восстановления которых затем усреднялись. На основных графиках приведено по три зависимости. Маркер «1» соответствует сингулярному числу, взятому из сингулярного разложения исходной матрицы W' из (2'). Маркер «2» отображает значения сингулярных чисел, полученных из сингулярного разложения матрицы W' , масштабированной в ходе выполнения итерационного алгоритма [2] в перспективной проекции. 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 (17), (18) путем прямой подстановки туда параметров,

взятых из модели. На вспомогательном графике рис. 7, 9 приведена относительная погрешность оценки σ_3 . Видно, что оценка дает результат с точностью порядка 15%. На вспомогательном графике рис. 8, 10 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 (17). Точность оценки не хуже 50%.

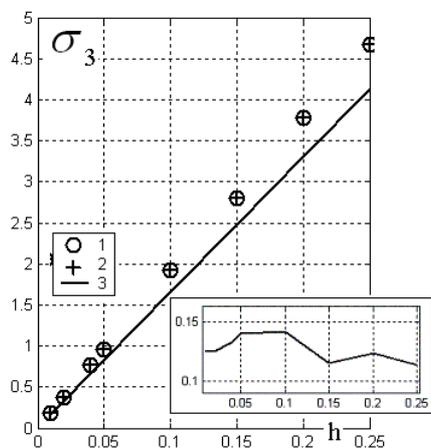


Рис. 7. Зависимость величины третьего сингулярного числа σ_3 матрицы W' от глубины сцены h .

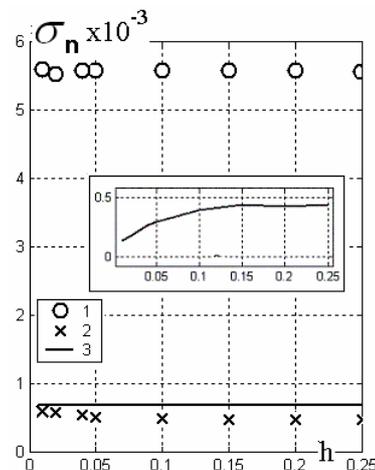


Рис. 8. Зависимость величины шумового сингулярного числа σ_n матрицы W' от глубины сцены h .

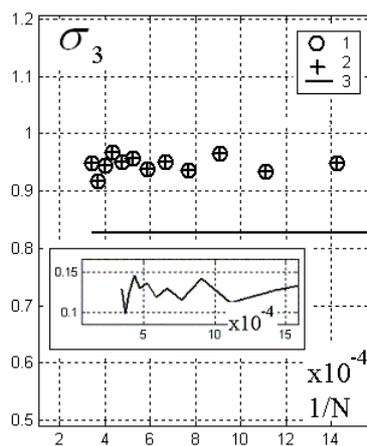


Рис. 9. Зависимость величины третьего сингулярного числа σ_3 матрицы W' от разрешения изображения N .

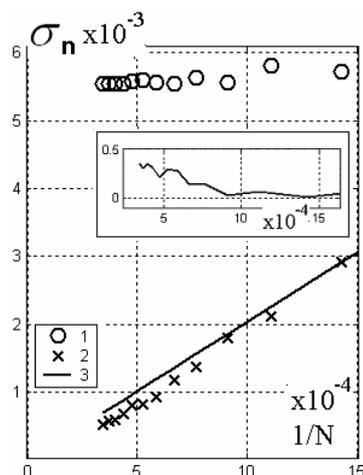


Рис. 10. Зависимость величины шумового сингулярного числа σ_n матрицы W' от разрешения изображения N .

На рис. 11-14 приведены результаты проверки оценок погрешностей восстановления формы объекта и движения камер (31), (32) на основе той же модели с теми же характеристиками. На рис. 11-14 маркерам «1» и «2» соответствуют величины среднеквадратичных погрешностей восстановления формы объекта и движения камер, полученные в результате сравнения исходной модели и ее трехмерного образа, восстановленного в приближении МОП и в перспективной проекции соответственно. Пунктирной «4» и сплошной линией «3» соответствуют оценки погрешностей (20), (30), 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 W' , преобразованная в ходе выполнения итерационного алгоритма [2] в перспективной проекции. Штрихпунктирная линия «5» получена путем подстановки в (20), (30) аналитических выражений для сингулярных чисел (17) и (18).

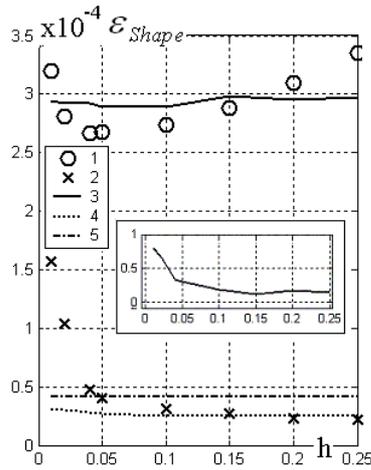


Рис. 11. Зависимость погрешности восстановления формы объекта от глубины сцены h .

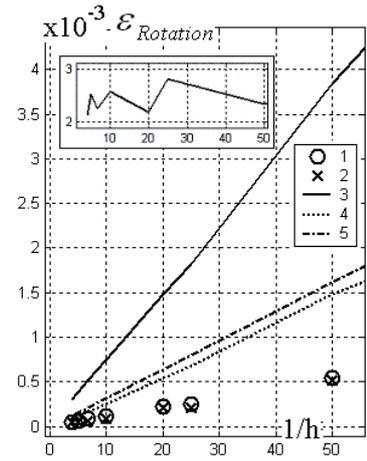


Рис. 12. Зависимость погрешности восстановления ориентаций камер от глубины сцены h .

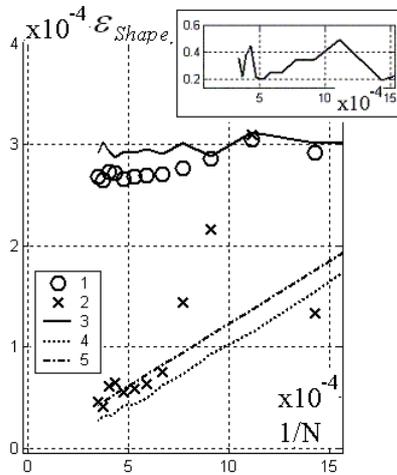


Рис. 13. Зависимость погрешности восстановления формы объекта от разрешения изображения N .

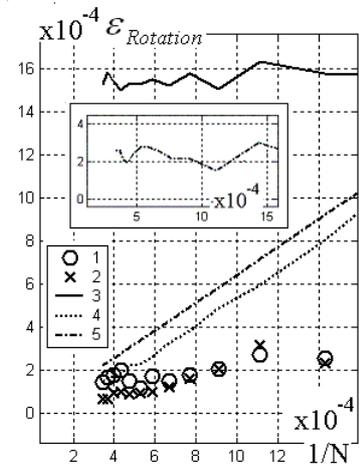


Рис. 14. Зависимость погрешности восстановления ориентаций камер от разрешения изображения N .

На представленных графиках видно, что результаты вычислений по формулам оценок (17), (18) и (20), (30) хорошо согласуются с фактическими значениями погрешностей перспективной проекции.

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 с помощью формул (20),(30) и (17)-(19) 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.

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 восстановления следует убедиться, что σ_3 больше, а σ_4 не превосходит по порядку величины шумовое значения 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 (20),(30), 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.

Погрешность восстановления модельных данных алгоритмами, основанными на факторизации матриц, определялась дискретизацией изображения и составляла менее 0.1 % для формы объекта и менее 1 % для ориентаций камер для алгоритмов МОП и ИПП в ситуациях, когда перспективные искажения были неразличимы на изображении. При увеличении разрешения изображения и/или увеличении глубины объекта погрешность ИПП не меняла порядок величины, в то время как погрешность МОП увеличивалась, достигая величин 0.1 % и 1 % для формы объекта и для ориентаций камер соответственно.

Теоретические оценки погрешностей показали хорошую точность при проверке на модельных данных. Результаты численного моделирования показали, что формула (30) завышает оцениваемые погрешности примерно в 2 раза. Формула (20) дает величину погрешности с точностью 50%. Для задачи оценки погрешности (погрешность погрешности) это очень хорошая точность, т.к. при решении подобной задачи важно было получить достаточно простую в использовании формулу, удобную в расчетах и не требующую сложных вычислительных операций. Формулы (20, 30) 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.

В свою очередь, наличие оценки достоверности восстановленного результата раскрывает широкие возможности для использования алгоритмов, основанных на факторизации матриц, в целях получения начального приближения восстановления трехмерного образа сцены по последовательности цифровых изображений, полученных с некалиброванных камер. Восстанавливаемые сцены при этом могут выбираться с достаточной степенью произвольности, т.к. алгоритм факторизации в состоянии проинформировать о невозможности восстановления сцены.

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. M. Watanabe S. K. Nayar and M. Noguchi. Real-time focus range sensor. *Proc. of Intl. Conf. on Computer Vision*, pages 995--1001, New York, USA, September 1995. <ftp://ftp.cs.columbia.edu/pub/CAVE/papers/nayar/nayar\-nabe\-noguchi\-sensor\ iccv\-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. Vladimir Kolmogorov and Ramin Zabih. Computing Visual Correspondence with Occlusions using Graph Cuts. In: *International Conference on Computer Vision*, July 2001. <http://www.cs.cornell.edu/rdz/Papers/KZ-ICCV01-tr.pdf>
8. Ruo Zhang, Ping-Sing Tsai, James Edwin Cryer, and Mubarak Shah. Analysis of shape from shading techniques. In *IEEE CVPR - 94*, pages 377--384, Seattle, Washington, June 1994. <http://www.cs.ucf.edu/~vision/papers/cvpr943.pdf> См. также University of Central Florida, Orlando,

- Computer Science Department, Technical Report, 1993,
<http://www.cs.ucf.edu/~vision/papers/analysisOfShapeFromShadingTechniques.pdf>
9. Eric Krotkov Fabio Cozman. Depth from scattering. Technical report, *Robotics Institute, Carnegie Mellon University*, Pittsburgh,
http://www.ri.cmu.edu/pub_files/pub2/cozman_fabio_1997_1/cozman_fabio_1997_1.pdf.
 10. Michael H. Lin, Carlo Tomasi: Surfaces with Occlusions from Layered Stereo. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2003.
 11. D. Scharstein and R. Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *IJCV*, 47:7–42, 2002. <http://www.middlebury.edu/stereo/>.
 12. Дж. Форсайт., М. Малькольм., К. Моулер. Машинные методы математических вычислений. Пер с англ. -М.:Мир,1980. 280 стр.
 13. Н. В. Свешникова, Д. В. Юрин. Априорный и апостериорный расчет погрешностей восстановления трехмерных сцен алгоритмами факторизации. //Программирование 2004, Т.30, № 5, С. 48-68.
 14. Н. В. Свешникова, Д. В. Юрин. Оценка погрешностей восстановления трехмерных сцен алгоритмами факторизации. – В сб. *Труды конференции. 14-я Международная конференция по Компьютерной графике и Машинному зрению ГрафиКон'2004* – стр. 200-207. Москва, 2004. [http://www.graphicon.ru/2004/Proceedings/Technical_ru/1\[1\].pdf](http://www.graphicon.ru/2004/Proceedings/Technical_ru/1[1].pdf).
 15. Дж. Деммель. Вычислительная линейная алгебра. – М.: Мир, 2001.
 16. 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
 17. 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.