

Оценка погрешностей восстановления трехмерных сцен алгоритмами факторизации¹

Н.В. Свешникова*, Д.В. Юрин**.

* Московский физико-технический институт, ЦОС и ВТ

** ФГУП НИИ ОПТЭКС

oduduka@cos.ru, yurin@cos.ru

Abstract

The accuracy of 3D recovery factorization based algorithms was investigated. The method of reliability and recovery errors estimation simultaneously with shape and motion restoration was developed. The analytical expression for 3D scene recovery accuracy was derived for thin object. The formula joins the expected accuracy with the scene and motion average statistical parameters. The verification of the results obtained was performed by numerical data modeling and demonstrates high precision of the errors estimation.

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

Аннотация

Исследована точность восстановления трехмерных сцен и движения камеры алгоритмами факторизации. Разработана методика, позволяющая одновременно с восстановлением сцены и движением камеры оценить достоверность и погрешность результатов. Для случая тонкого объекта получена аналитическая формула, связывающая ожидаемую погрешность восстановления со среднестатистическими характеристиками объекта и движения камеры. Проверка оценок проводилась методами численного моделирования и показала высокую точность.

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

1. ВВЕДЕНИЕ

Многообразие методов решения задачи восстановления трехмерных сцен по набору цифровых изображений [1-6] в настоящее время сопровождается достаточно фрагментарным анализом областей применимости и погрешностей каждого подхода. Как правило, он представляет собой тестирование алгоритма на нескольких известных сценах, чего явно не достаточно, для адекватного определения области применимости алгоритма и оценки достоверности результатов его работы на конкретном наборе исходных данных.

В предлагаемой работе получены аналитические формулы для оценки погрешности восстановления формы сцены и ориентаций камер для двух алгоритмов, основанных на факторизации матриц – линейного алгоритма в приближении масштабированной ортографической проекции (МОП) [1] и итерационного метода в перспективной проекции (ИПП) [2]. Полученные оценки могут быть использованы на стадии планирования условий съемки сцены, для которой известны

среднестатистические характеристики. Разработана методика оценки достоверности и точности результатов в процессе выполнения восстановления трехмерной сцены. В последнем случае не требуется никакой дополнительной информации сверх обычных исходных данных для алгоритмов факторизации.

Основной принцип, на котором базируется построение оценок, 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 объекта.

Проверка полученных оценок осуществлялась на модельных данных и показала высокую точность.

2. ВОССТАНОВЛЕНИЕ ТРЕХМЕРНОЙ СЦЕНЫ МЕТОДОМ ФАКТОРИЗАЦИИ МАТРИЦ

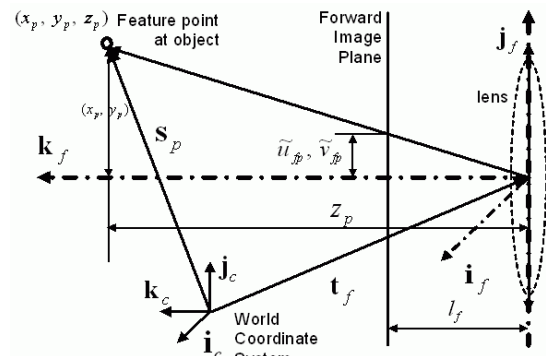


Рисунок 1: Постановка задачи восстановления трехмерной сцены.

Кратко рассмотрим принципы, на которых базируются 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 координа-

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

тами \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:

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

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

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 $l_f, \tilde{u}_{fp}, \tilde{v}_{fp} \cdot \tilde{u}_f^e$ - размер одного пикселя в единицах длины на фоточувствительной матрице; N - ширина изображения в пикселях, $\beta_{\max f}$ - максимальный угол зрения камеры. Вводя обозначение $z'_f = -\mathbf{k}_f \mathbf{t}_f / g_f$, и полагая $\mathbf{k}_f \mathbf{s} / z_f \ll 1$ Получим формулы для приближения МОП:

$$\begin{cases} u_{fp} = \mathbf{m}_f \mathbf{s}_p + t_{xf} \\ v_{fp} = \mathbf{n}_f \mathbf{s}_p + t_{yf} \end{cases}, \quad (2)$$

$$\text{где } \begin{cases} \mathbf{m}_f = \mathbf{i}_f / z'_f, & \begin{cases} t_{xf} = -\mathbf{i}_f \mathbf{t}_f / z'_f \\ t_{yf} = -\mathbf{j}_f \mathbf{t}_f / z'_f \end{cases} \\ \mathbf{n}_f = \mathbf{j}_f / z'_f \end{cases},$$

Уравнения (2) можно представить в матричной форме:

$$\mathbf{W} = \mathbf{M}\mathbf{S} + \mathbf{T}; \quad \mathbf{W}' = \mathbf{W} - \mathbf{T}, \quad (2')$$

где \mathbf{W} - матрица, размера $2F \times P$, пары строк которой соответствуют координатам точек на изображении, а

$$\begin{aligned} \mathbf{T} &= (t_{x1} \quad t_{y1} \quad \dots \quad t_{xF} \quad t_{yF})^T, \\ \mathbf{S} &= (\mathbf{s}_1 \quad \dots \quad \mathbf{s}_P), \\ \mathbf{M} &= (\mathbf{m}_1 \quad \mathbf{n}_1 \quad \dots \quad \mathbf{m}_F \quad \mathbf{n}_F)^T, \end{aligned} \quad (3)$$

Матрица \mathbf{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 \mathbf{W}' не превышает указанных значений.

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

$$\mathbf{W}' = \mathbf{W}^0 + \omega \quad (4)$$

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

$$\langle \omega^2 \rangle = \rho^2; \quad \langle \omega \rangle = 0. \quad (5)$$

Матрица \mathbf{W}^0 - неискаженная матрица измерений ранга ζ . Величина шума определяется как безразмерная, что следует из (1)-(5), 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 алгоритма, основанного на факторизации матриц, выполняется сингулярное разложение матрицы \mathbf{W}' . Затем выбираются ζ наибольших сингулярных числа, а все остальные полагаются шумовыми и отбрасываются.

3. РАЗРЕШИМОСТЬ МАЛОЙ ГЛУБИНЫ СЦЕНЫ В ПРИБЛИЖЕНИИ МОП

Рассмотрим модельную задачу. Пусть восстанавливаемый объект - две параллельные плоскости, находящиеся на малом расстоянии h друг от друга:

$$\begin{aligned} L_1: \quad & Ax + By + z + D - h = 0; \\ L_2: \quad & Ax + By + z + D + h = 0; \end{aligned} \quad (6)$$

Набор точек, лежащих на этих плоскостях, 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.

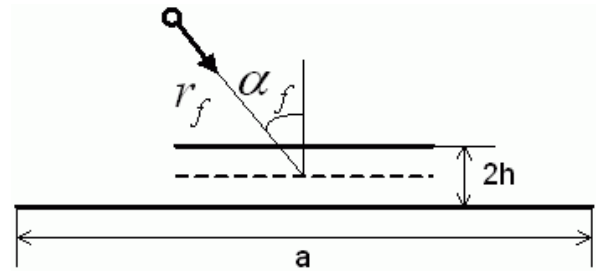


Рисунок 2: Постановка задачи разрешимости сцены малой глубины.

На рис. 2 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.

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 (из (3) и (6) видно, что 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:

$$\sigma_3 > \sigma_n. \quad (7)$$

Рассмотрим идеальную матрицу W^0 в условиях данной задачи. Полагая x и y независимыми координатами, мы можем z выразить через них. Тогда

$$z_p = -(Ax_p + By_p + D) \pm h = z_p^{(1)} \pm h. \quad (8)$$

Вводя обозначения

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

$$S = S_1 + h \cdot S_2,$$

$$S_1 = (s_1^{(1)} \quad \dots \quad s_p^{(1)}) \quad , \quad S_2 = (s_1^{(2)} \quad \dots \quad s_p^{(2)})$$

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

$$W^0 = M \cdot S_1 + h \cdot M \cdot S_2 = W_1^0 + W_2^0 \quad (9)$$

Т.к. при $h=0$ задача вырождается в двумерную: $\text{rank } S_1 = 2 \Rightarrow \text{rank } W = 2$, то, разложив матрицу W^0 на две составляющие: не содержащую h и содержащую, мы разделим задачу на плоскую и добавку. Очевидно, что $\text{rank } S_2 = 1$ и $\text{rank } W_2^0 = 1$, а так как 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_2^0 и будет оценкой для σ_3 .

Для оценки сингулярных чисел воспользуемся понятием нормы Фробениуса [7]. В случае матрицы произвольного ранга матрицы (шумовая добавка), норма Фробениуса дает завышенную оценку наибольшего сингулярного числа.

Подставляя величины, характеризующие шумовую матрицу (5) получим:

$$\sigma_n^2 \leq \| \omega \|^2 = \sum_{f=1}^{2F} \sum_{p=1}^P \omega_{fp}^2 = 2FP\rho^2; \quad (10)$$

Перейдем к вычислению нормы матрицы W_2^0 . Так как $\text{rank } W_2^0 = 1$, то норма Фробениуса будет совпадать со зна-

чением единственного сингулярного числа, которое в приближении малого h дает оценку σ_3 для W' .

$$\sigma_3^2 = \| W_2^0 \|^2 = FPh^2 J, \quad (11)$$

$$где J = \frac{1}{F} \sum_{f=1}^F \left(\left(\frac{g_f}{z_{cf}} \right)^2 (i_{zf}^2 + j_{zf}^2) \right)$$

Вычисления по формуле (11) могут быть произведены, если для всех кадров известны ориентации и положения камер. На практике это применимо, когда общие статистические характеристики сцены известны а priori. Однако вычисление конечных сумм представляется достаточно неудобным для оценивания, поэтому при большом числе кадров F удобно перейти к интегралу. Преобразуем формулу (11) 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 (11), можно приблизить интегралом:

$$J = \int_{-\infty}^{+\infty} \left(\frac{g(\vec{r})}{|\vec{r}|} \right)^2 (i_z^2(\vec{r}) + j_z^2(\vec{r})) \nu(\vec{r}) d\vec{r} \quad (12)$$

В условиях конкретной съемки можно задать модель движения камеры и, в результате вычисления интеграла (12), получить аналитическую формулу для априорной оценки разрешимости задачи восстановления тонкого объекта (его рельефа).

4. ОЦЕНКА ПОГРЕШНОСТЕЙ ВОССТАНОВЛЕНИЯ ФОРМЫ СЦЕНЫ И ОРИЕНТАЦИИ КАМЕР

Предположим, что минимальная разрешимая высота при данных условиях съемки равна абсолютной погрешности восстановления формы любого объекта. Подтверждение данной гипотезы, основанное на результатах численного моделирования, будет приведено ниже.

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:

$$\frac{\delta h}{h} = \frac{h_{\min}}{h} = \frac{\sigma_n}{\sigma_3} \quad (13)$$

Где δh - абсолютная погрешность восстановления формы объекта, h - его средняя высота.

Сингулярное разложение зашумленной матрицы W' с наложением ограничений на ранг матриц M и S по сути является решением задачи методом наименьших квадратов с регуляризацией. Решение находится в смысле

$$\min_{M, S} \left\| \mathbf{W}'_s - \mathbf{M} \cdot \mathbf{S} \right\| \leq \|\omega\| \quad (14)$$

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

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

$$\begin{aligned} \mathbf{W}' &= (\mathbf{M} + \delta\mathbf{M}) \cdot (\mathbf{S} + \delta\mathbf{S}) \approx \\ &\approx \mathbf{M} \cdot \mathbf{S} + \delta\mathbf{M} \cdot \mathbf{S} + \mathbf{M} \cdot \delta\mathbf{S} \end{aligned} \quad (15)$$

Сравнивая (13) и (14), получаем, что при условии $\|\delta\mathbf{M} \cdot \mathbf{S}\| \approx \|\mathbf{M} \cdot \delta\mathbf{S}\|$ и противоположных знаках $\delta\mathbf{M}$ и $\delta\mathbf{S}$ 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:

$$\|\delta\mathbf{M}\| \approx \|\mathbf{M}\| \cdot \frac{\|\delta\mathbf{S}\|}{\|\mathbf{S}\|} \quad (16)$$

Метод факторизации требует полностью заполненной матрицы \mathbf{W}' , поэтому для трехмерных объектов диапазон позиций наблюдения ограничен, остроугольным конусом, так как в противном случае некоторые отметки будут закрываться объектом. Погрешность восстановления формы при таких условиях съемки наиболее велика в генеральном направлении наблюдения. 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 $\delta\mathbf{S} = [\mathbf{0}, \mathbf{0}, \delta\mathbf{S}_z]^T$, а первые две положим равными нулю. Это позволяет применять оценку минимальной разрешимой высоты, полученную в предыдущем разделе и заменить в формуле (16) $\|\mathbf{M}\|$ на $\|\mathbf{m}_z\|$ - норму третьего вектор-столбца матрицы \mathbf{M} , а норму матрицы формы $\|\mathbf{S}\|$ на $h \cdot \|\mathbf{S}_2\|$, определяющую величину третьего сингулярного числа:

Построим матрицу $\delta\mathbf{M}$ как приращение матрицы \mathbf{M} , полагая, что варьируется только ориентация камер, т.е. если

$$\delta\mathbf{M} = \begin{pmatrix} \frac{g_1}{z_1} \delta\mathbf{a}_1 & \frac{g_1}{z_1} \delta\mathbf{j}_1 & \dots & \frac{g_F}{z_F} \delta\mathbf{a}_F & \frac{g_F}{z_F} \delta\mathbf{a}_F \end{pmatrix}^T \quad (17)$$

Считая, из соображений, описанных выше, что отличен от нуля только третий столбец матрицы $\delta\mathbf{M}$, получим выражение для нормы:

$$\|\delta\mathbf{M}\|^2 = \sum_{f=1}^F \left(\frac{g_f}{z_f} \right)^2 (\delta i_{zf}^2 + \delta j_{zf}^2) \quad (18)$$

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

$$\|\delta\mathbf{M}\|^2 = \left(\langle \delta i_z^2 \rangle + \langle \delta j_z^2 \rangle \right) \sum_{f=1}^F \left(\frac{g_f}{z_f} \right)^2 \quad (19)$$

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 (19) принимает вид:

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

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, учитывая что

$$\|\mathbf{M}\| = \sqrt{2 \sum_{f=1}^F \left(\left(\frac{g_f}{z_{cf}} \right)^2 \right)} \quad (20)$$

$$\|\delta\mathbf{M}\| \approx \mathcal{G}_j \|\mathbf{M}\| \quad (21)$$

Подставляя (13), (16) в выражение (21), получим:

$$\|\delta\mathbf{M}\| \approx \mathcal{G}_j \|\mathbf{M}\| \approx \|\mathbf{m}_z\| \frac{\|\delta\mathbf{S}\|}{h \|\mathbf{S}_2\|} = \|\mathbf{m}_z\| \cdot \frac{\sigma_n}{\sigma_3} \quad (22)$$

Выражая \mathcal{G}_j из (22), получим окончательно:

$$\mathcal{G}_j = \frac{\|\mathbf{m}_z\|}{\|\mathbf{M}\|} \cdot \frac{\sigma_n}{\sigma_3} \quad (23)$$

5. РЕЗУЛЬТАТЫ

Проверка полученных оценок проводилась методами численного моделирования.

Модель «Рельеф» - 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 $a \times a$, 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 . Съёмка 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. 3 в виде сетки характеристических точек.

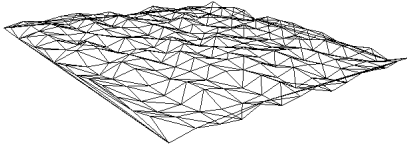


Рисунок 3: Модель «Рельеф».

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:

$$\mathcal{E}_{Shape} = \frac{1}{a} \sqrt{\frac{1}{P} \sum_{p=1}^P (\mathbf{s}_p - \mathbf{s}_p^0)^2}; \quad (24)$$

$$S = [\mathbf{s}_1^T \quad \mathbf{s}_2^T \quad \dots \quad \mathbf{s}_P^T];$$

$$S^0 = [\mathbf{s}_1^{0T} \quad \mathbf{s}_2^{0T} \quad \dots \quad \mathbf{s}_P^{0T}].$$

$$\mathcal{E}_{Rotation} = \sqrt{\frac{1}{3F} \sum_{f=1}^{3F} (\mathbf{r}_f - \mathbf{r}_f^0)^2}; \quad (25)$$

$$R = [\mathbf{r}_1^T \quad \mathbf{r}_2^T \quad \dots \quad \mathbf{r}_{3F}^T];$$

$$R^0 = [\mathbf{r}_1^{0T} \quad \mathbf{r}_2^{0T} \quad \dots \quad \mathbf{r}_{3F}^{0T}].$$

5. Сравнение полученных погрешностей и их зависимостей от параметров с теоретическими оценками (13), (23).

На рис 4-7 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 2×2 км, разрешение 2048 пикселей при точности детектора 0.1 пиксель [8]. 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 (10), (11) путем прямой подстановки туда параметров, взятых из модели. На вспомогательном графике рис. 4, 6 приведена относительная погрешность оценки σ_3 . Видно, что оценка дает результат с точностью порядка 15%. На вспомогательном графике рис. 5, 7 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 (10). Точность оценки не хуже 40%.

На рис. 8-11 приведены результаты проверки оценок погрешностей восстановления формы объекта и движения камер (24), (25) на основе той же модели с теми же характеристиками. На рис. 8-11 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 линии «3» соответствуют оценки погрешностей (13), (23), 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 (13), (23) аналитических выражений для сингулярных чисел (10) и (11).

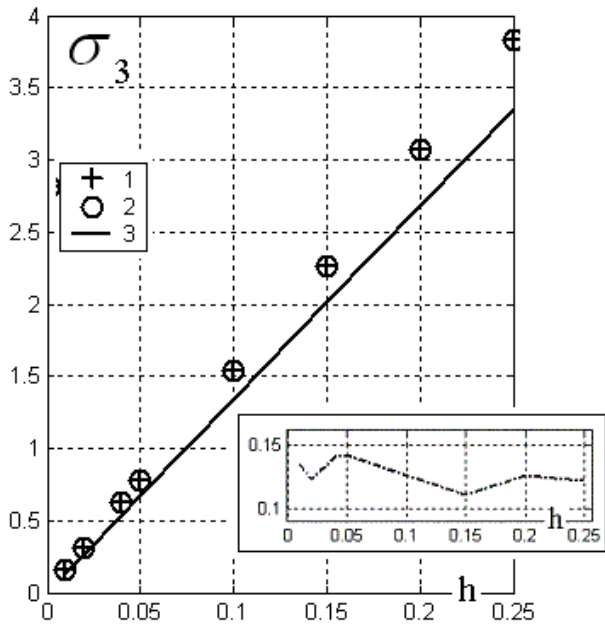


Рисунок 4: Зависимость величины третьего сингулярного числа σ_3 матрицы W' от глубины сцены h .

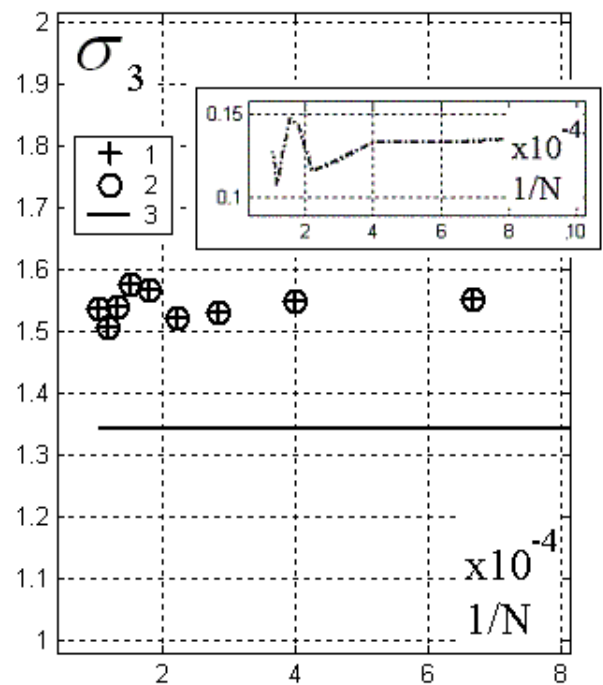


Рисунок 6: Зависимость величины третьего сингулярного числа σ_3 матрицы W' от разрешения изображения N .

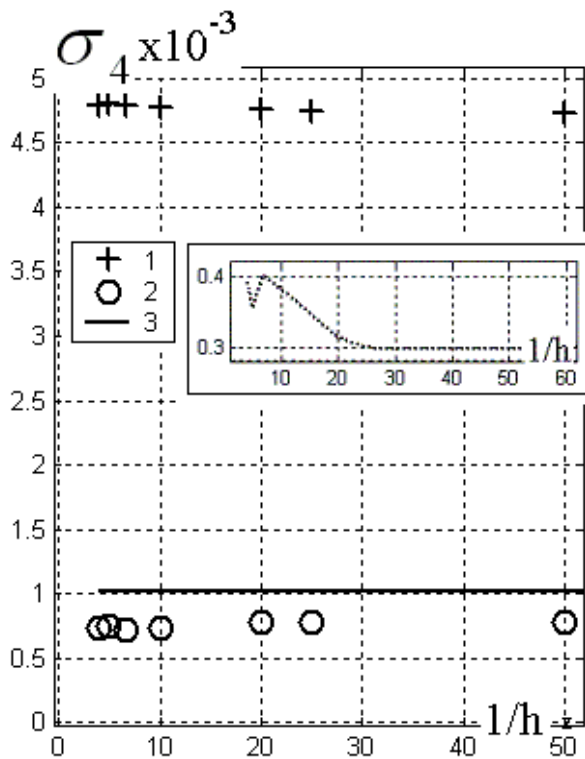


Рисунок 5: Зависимость величины шумового сингулярного числа σ_n матрицы W' от глубины сцены h .

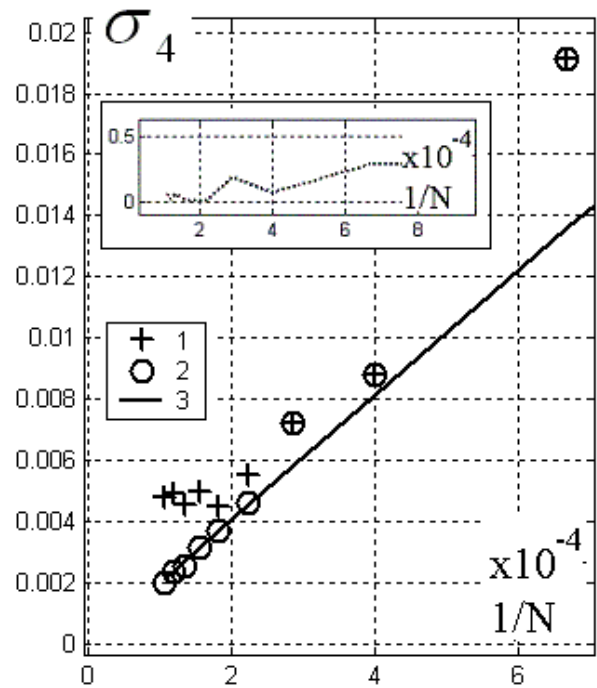


Рисунок 7: Зависимость величины шумового сингулярного числа σ_n матрицы W' от разрешения изображения N .

На представленных графиках видно, что результаты вычислений по формулам оценок (10), (11) и (13), (23) хорошо согласуются с фактическими значениями погрешностей пер-

спективной проекции. Таким образом, сформулированная гипотеза верна.

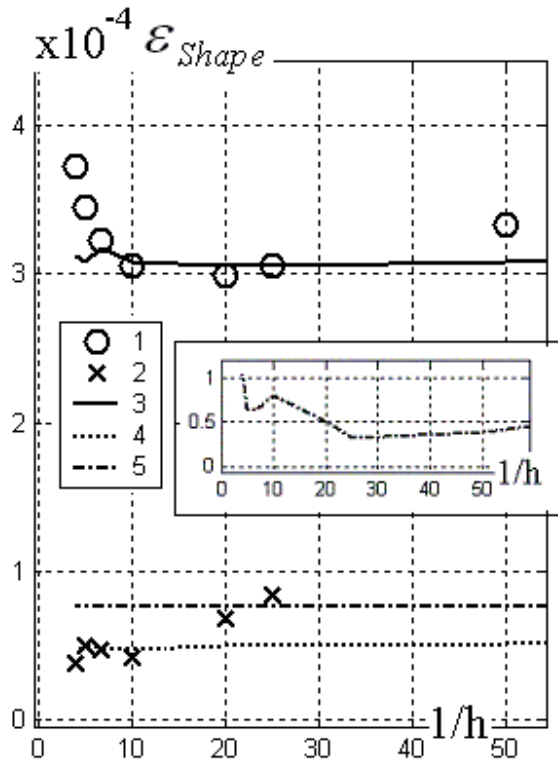


Рисунок 8: Зависимость погрешности восстановления формы объекта от глубины сцены h .

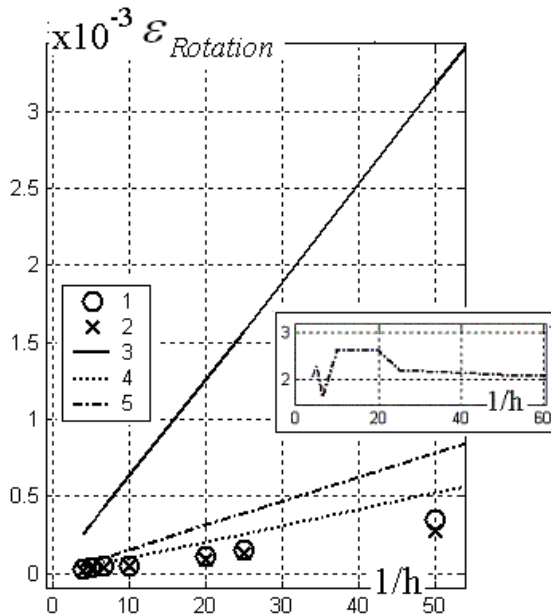


Рисунок 9: Зависимость погрешности восстановления ориентаций камер от глубины сцены h .

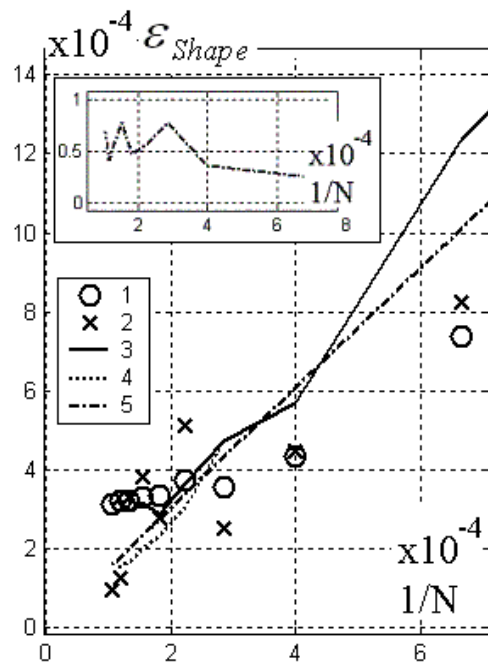


Рисунок 10: Зависимость погрешности восстановления формы объекта от разрешения изображения N .

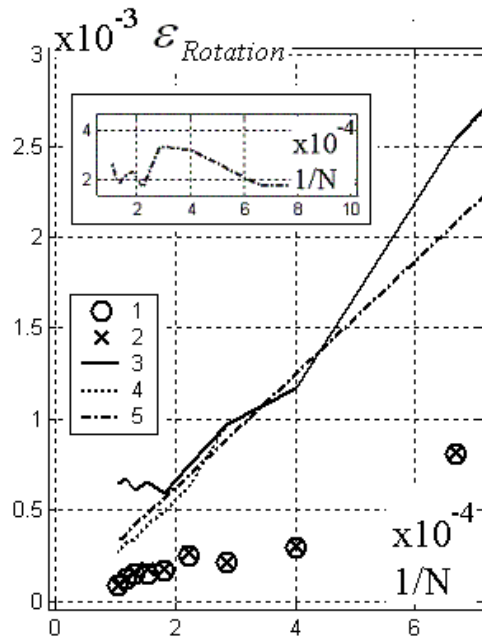


Рисунок 11: Зависимость погрешности восстановления ориентаций камер от разрешения изображения N .

6. ЗАКЛЮЧЕНИЕ

Наличие шумов в исходных данных, обусловленных погрешностью детектора характеристических точек и конечным пиксельным разрешением изображения, вносит в результаты восстановления статистические погрешности. Алгоритмы восстановления трехмерных сцен, основанные на факториза-

ции матриц, используют избыточное количество информации, предоставляемое длинной последовательностью цифровых изображений объекта, выделяя из нее самосогласованную информацию. Это дает возможность получить оценку статистической погрешности восстановления трехмерной сцены и достоверности восстановления. Проведенный анализ дал следующие результаты.

Построены оценка погрешности оцифровки σ_n (10) и оценка величины σ_3 , содержащей информацию об объекте, в приближении сцены малой глубины (11). Точность оценки шумов, полученная методами численного моделирования, не хуже 60%, точность оценки второй величины не хуже 15%.

Предложен критерий оценки разрешимости задачи восстановления сцены малой глубины (7). В условиях априорного анализа возможности восстановить сцену с известными среднестатистическими характеристиками и условиями съемки, компоненты критерия могут быть вычислены по формулам (10), (11). При отсутствии какой бы то ни было априорной информации работу предлагается строить следующим образом. Система восстановления трехмерного образа должна быть калибрована (тестирована) с целью определения фактической точности детектора. После восстановления сцены следует убедиться, что σ_3 больше, а σ_4 не превосходит по порядку величины шумовое значения σ_n , полученное на стадии калибровки.

Принцип построения критерия разрешимости лег в основу построения оценок погрешностей восстановления формы объекта (13) и ориентаций камер (23). На основе численного моделирования было показано, что 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. Оценки погрешностей являются завышенными примерно в два раза по сравнению со среднеквадратичным отклонением восстановленного образа от исходной модели.

Применение оценок (13), (23) на стадии априорного анализа требует использования также и оценок (10), (11), поэтому по результатам настоящей работы может быть произведено только для сцен малой глубины. 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.

Использование (13), (23) в условиях апостериорного анализа не требует никакой дополнительной информации о сцене кроме обычных исходных данных для алгоритмов факторизации и не накладывает на нее дополнительных ограничений.

Таким образом, в результате проведенного анализа алгоритмов, основанных на факторизации матриц, получены аналитические выражения, позволяющие как априорно, так и апостериорно с высокой точностью оценить достоверность и погрешность результатов восстановления.

7. 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/~ynk/>,
<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-navar-tele-iuwl-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. Деммель. *Вычислительная линейная алгебра. – М.: Мир, 2001.*
- [8] 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

Об авторах

Свешникова Наталья Владимировна – студентка 5-го курса ФФКЭ Московского Физико - Технического Института.
E-mail: oduduka@cos.ru

Юрин Дмитрий Владимирович, к.ф.-м.н., ФГУП НПП «ОИП-ЭКС», нач. отдела.
E-mail: yurin@cos.ru