

РАЗРЕШИМОСТЬ МАЛОЙ ГЛУБИНЫ СЦЕНЫ В ПРИБЛИЖЕНИИ МАСШТАБИРУЕМОЙ ОРТОГРАФИЧЕСКОЙ ПРОЕКЦИИ.

Н.В. Свешникова, Д.В. Юрин.

Московский физико-технический институт

ЦОС и ВТ МФТИ, ФГУП НПП ОПТЭКС

Одной из наиболее востребованных областей среди задач восстановления трехмерных сцен по цифровым изображениям является задача восстановления рельефа Земной поверхности по кадрам, полученным со спутника. Специфика этой задачи состоит в малой глубине восстанавливаемой сцены по сравнению с ее поперечными размерами. В настоящей работе представлены результаты исследования алгоритма, основанного на факторизации матриц [1] – приближения масштабируемой ортографической проекции (МОП).

Рассмотрим математическое представление моделей и основные принципы метода факторизации матриц. 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 (объекте) выделено $p = 1..P$ характеристических точек с координатами \mathbf{s}_p в МСК, u_{fp} , v_{fp} – их координаты на изображениях в пикселях. Тогда, из подобия треугольников можно записать [1]:

$$u_{fp} = g_f \frac{\mathbf{i}_f(\mathbf{s}_p - \mathbf{t}_f)}{\mathbf{k}_f(\mathbf{s}_p - \mathbf{t}_f)}, \quad v_{fp} = g_f \frac{\mathbf{j}_f(\mathbf{s}_p - \mathbf{t}_f)}{\mathbf{k}_f(\mathbf{s}_p - \mathbf{t}_f)}, \quad g_f = \frac{l_f}{\tilde{u}_{fp}^e N} = \frac{1}{2 \operatorname{tg}(\beta_{\max f}/2)}, \quad (1)$$

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 – фокусное расстояние камеры, $u_{fp}, v_{fp} \in [-1/2, 1/2]$; $\tilde{u}_{fp}^e, \tilde{v}_{fp}^e$ – размеры одного пикселя на фоточувствительной матрице; N – ширина изображения в пикселях, $\beta_{\max f}$ – максимальный угол зрения камеры. Вводя обозначение $z'_f = -(\mathbf{k}_f, \mathbf{t}_f)$, $z''_f = z'_f/g$ и полагая $(\mathbf{k}_f, \mathbf{s}_p)/z'_f \ll 1$, получим формулы для приближения МОП:

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

Уравнения (2) можно представить в матричной форме: $\mathbf{W} = \mathbf{MS} + \mathbf{T}$, где \mathbf{W} – матрица, размера $2F \times P$, пары строк которой соответствуют координатам точек на изображении, а

$$\mathbf{T} = (x_1 \quad y_1 \quad \dots \quad x_F \quad y_F)^T, \quad \mathbf{S} = (\mathbf{s}_1 \quad \dots \quad \mathbf{s}_P), \quad \mathbf{M} = (\mathbf{m}_1 \quad \mathbf{n}_1 \quad \dots \quad \mathbf{m}_F \quad \mathbf{n}_F)^T,$$

Метод факторизации [1], основан на том, что ранг матрицы $\mathbf{W}^* = \mathbf{W} - [1..1]\mathbf{T}$ не может быть больше 3, так как она представляется в виде произведения двух матриц ранга 3 (если предмет объемный и 2, если он плоский; в этом случае ранг \mathbf{W}^* также не превышает двух). На реальных данных матрица \mathbf{W}^* всегда содержит шум (по крайней

мере шум оцифровки $\pm\delta \sim 1/N$), поэтому в общем случае ее ранг равен $\min(2F, P)$, где F – количество кадров, P – количество точек. Такую матрицу можно представить в виде: $W^{*real} = W^* + \omega$, где ω – матрица шума. Поэтому при факторизации матрицы W^* выбираются 3 (или 2 в плоском случае) наибольших сингулярных числа, а остальные принимаются за шум и отбрасываются. Этот принцип лег в основу построения критерия разрешимости глубины почти плоского объекта – двух параллельных плоскостей, $z = h/2$, $z = -h/2$, расстояние h между которыми мало.

Очевидно, что при $h = 0$ задача становится плоской и $\text{rank} W^* = 2$, поэтому, разложив матрицу W^* на две составляющие: не содержащую h и содержащую, можно разделить задачу на плоскую и добавку: $W^* = MS_1 + h^*MS_2$, где S_1 состоит из координат точек, принадлежащих плоскости $z = 0$, а S_2 – из столбцов вида $(0, 0, \pm 1/2)^T$. Учитывая малость добавки и то, что ее ранг равен 1, можно оценить единственное сингулярное число матрицы h^*MS_2 , принять его за оценку σ_3 исходной матрицы W^{*real} , сравнить его с оценкой шумового сингулярного числа σ_n и проверить, выполняется ли условие $\sigma_3 > \sigma_n$. В противном случае при восстановлении поверхности будет выбрано не то сингулярное число и восстановленная форма будет определяться шумовыми собственными векторами. Для нахождения оценки было использовано понятие Эрмитовой нормы прямоугольной матрицы A , выражаемой следующим образом: $\|A\| = \sqrt{\lambda}$, где λ – наибольшее сингулярное число матрицы AA^T . Используя свойство сингулярных чисел (собственные значения λ_i матрицы AA^T есть квадраты сингулярных чисел матрицы A , $\lambda_i = \sigma_i^2$) и свойство квадратных матриц B ($\text{trace}(B) = \sum_{i=1}^n b_{ii} = \sum_{i=1}^n \lambda_i$), получаем следующий вид критерия разрешимости малой глубины сцены:

$$\sigma_3 = \frac{h}{2} \sqrt{P} \sqrt{\sum_{f=1}^F \left(\left(\frac{g_f}{z_{cf}} \right)^2 (i_{zf}^2 + j_{zf}^2) \right)} > \frac{\delta}{2} \sqrt{2FP} = \sigma_n \quad (3)$$

где i_{zf}, j_{zf} – z -компоненты векторов $\mathbf{i}_f, \mathbf{j}_f$, задающих правую ортонормированную тройку координат, δ – определяет верхний порог шума. Критерий проверялся на модельных данных. Из рис. 1 видно, что оценка σ_3 совпала с опытным значением.

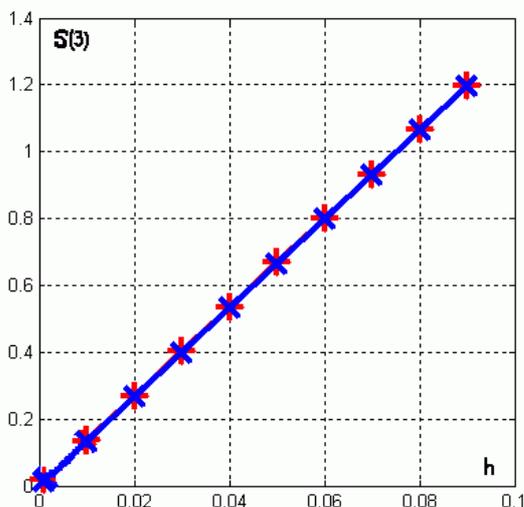


Рис. 1. Зависимость σ_3 от расстояния между плоскостями h .

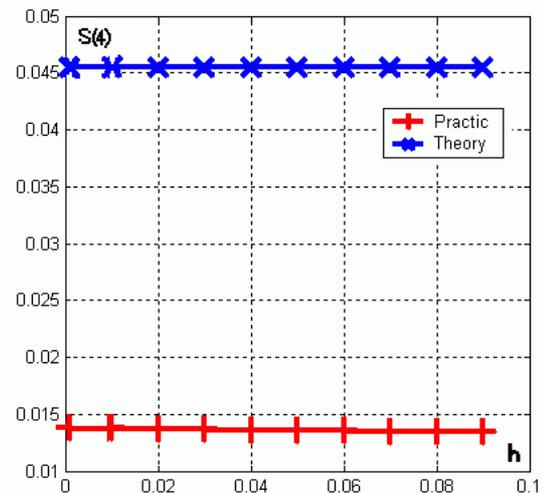


Рис. 2. Зависимость σ_n от расстояния между плоскостями h .

Оценка σ_n получилась завышенной по сравнению с опытом, но описываемая зависимость верна (см. рис. 2). Сильно завышенная оценка может привести к тому, что оценка четвертого сингулярного числа всегда будет примерно равной или больше оценки третьего сингулярного числа, т.е. критерий будет утверждать неразрешимость задачи, в то время как в реальности она разрешима (см. рис 3.). Тем не менее, некоторое завышение оценки σ_n должно сохраниться для большей строгости неравенства (3). Поправка к σ_n^2 в оценке шумового сингулярного числа была найдена методами численного моделирования в виде билинейной функции по \sqrt{F} и \sqrt{P} , коэффициенты оценивались методом наименьших квадратов. С учетом этой поправки, правая часть (3) принимает вид:

$$\sigma_n = \frac{\delta}{2} (0.57\sqrt{2FP} - 0.27\sqrt{F} + 0.18\sqrt{P} + 35)$$

На рис. 4 представлено отношение σ_4/σ_3 после внесения добавки, которое можно сравнить с изображенным на рис. 3. Фрагменты в углах рисунков – увеличенное изображение одинаковых по диапазону параметров фрагментов больших графиков.

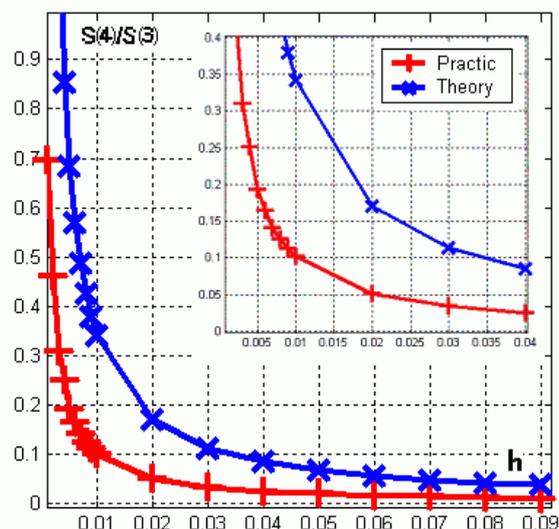


Рис. 3. Зависимость отношения σ_4/σ_3 от расстояния между плоскостями h без внесения добавки.

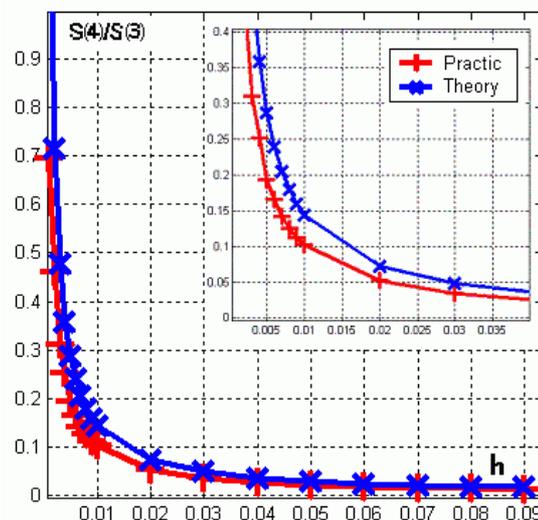


Рис. 4. Зависимость σ_4/σ_3 от расстояния между плоскостями h после внесения добавки.

Подстановка параметров в построенный критерий показала, что для типичных параметров спутниковой съемки земной поверхности (высота орбиты 1000 км, диапазон углов к надиру в диапазоне $\pm 60^\circ$, разрешение $N=2048$ пикселей в строке) минимальная разрешимая высота h соизмерима с наземным разрешением изображения при съемке в надиру. К примеру при разрешении 1 м. в 1 пикселе минимальная разрешимая высота составляет 2 м (см. рис. 4, h выражается в км.). Важно отметить, что рассматриваемый метод восстановления [1] вообще не требует знания положения и пространственной ориентации камер, требования к точности которых часто бывают неоправданно завышены, что позволяет использовать его для легких спутников.

Список литературы.

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*