УДК 519.633.6

А.В. Лаптев

АВТОМАТИЗАЦИЯ РЕШЕНИЯ CAPTCHA В ТЕКСТОВОМ ФОРМАТЕ

Численное моделирование процессов в таких сложных образованиях, как многолетнемерзлые породы невозможно без привлечения современных многомасштабных методов. Например, для моделирования упругой деформации твердых тел может применяться гетерогенный многомасштабный метод конечных элементов (FE-HMM). Однако при решении практических задач необходимо иметь инструмент контроля вычислительной погрешности, а также априорные данные об технических ограничениях метода. В рамках данной работы проведены исследования о влиянии сеточных разбиений на разных уровнях иерархии на точность и скорость решения задачи упругой деформации для FE-HMM на полиэдральных носителях на макроуровне и тетраэдральных конечных элементах на микроуровне. Полученные оценки позволяют скорректировать стратегии планирования вычислительных экспериментов, которые во многом связаны с построением набора сеток различной точности. Например, получено, что измельчение микроуровневой сетки на ребрах макро-полиэдров позволяет увеличить точность решения до двух порядков при сохранении общего размера дискретизаций.

***Ключевые слова:****CAPTCHA, автоматизированное распознавание, нейронная сеть, TensorFlow, OCR, Tesseract, CRNN, Seq-to-Seq, Python.*

Введение

Структурная сложность геологических пород накладывает определенные ограничения на используемые вычислительные схемы. Например, прямые методы, такие как метод Галёркина [1], приводят к дискретизациям с большим числом параметром, что является критичным даже для современных вычислительных машин. Поэтому для моделирования процессов в гетерогенных разномасштабных средах применяются методы, основанные на декомпозиции исходной области на некоторые подобласти. Однако в таком случае необходимо отдельно формулировать специальные граничные операторы для обеспечения непрерывности решения на межфрагментарных границах. Наиболее общий и удобный в использовании математический аппарат для построения таких операторов предлагают многомасштабные конечноэлементые методы. Например, гетерогенный многомасштабный метод конечных элементов (FE-HMM) [2–4], в котором пространство решения является иерархическим и содержит два или более уровней, каждое из которых соответствует одному физическому или геометрическому масштабу рассматриваемой задачи. На каждом из уровней иерархии формулируются специальные подзадачи для построения локальных подпространств. Процедуры согласования уровней иерархии выбираются в соответствии со спецификой моделируемых процессов. Это обеспечивает достаточную гибкость и адаптивность метода.

Таким образом, сформулируем основные этапы решения задачи (1) – (7) в контексте применения двухуровневого FE-HMM (под верхним уровнем здесь и далее будем понимать макроуровень, отвечающий за эффективные свойства среды, нижний уровень – макроуровень, позволяющий детально учесть все неоднородности среды):

1. Построение сеточной иерархии, а именно разбиение всей области моделирования на подобласти (макроэлементы) без налегания и без строгого учета внутренней геометрической структуры, далее каждый из макроэлементов разбивается независимо на микроэлементы с учетом внутренней структуры.

2. Построение функциональной иерархии, то есть формирования макромасштабных функций формы, обеспечивающих гладкость решения, через серию подзадач на каждом из макроэлементов, используя микроэлементные разбиения.

Не смотря на достаточную представленность FE-HMM в литературе, в том числе для моделирования упругой деформации [5], некоторые математические и соответствующие им технологические аспекты данного метода должны разрабатываться под каждую конкретную физическую задачу. Кроме того, традиционно на верхнем уровне иерархии (будем далее называть его макроуровнем) используются конечноэлементные носители простой формы с равным числом граней (тетраэдры или параллелепипеды) для упрощения построения функциональных подпространств решения. Однако для задач упругой деформации гетерогенных объектов более эффективно использовать полиэдральные макроэлементы с нефиксированным числом граней. Нами были предложены и реализованы технологии автоматизирования построения сеточных дискретизаций на макроуровне и подходы к построению макромасштабных функций форм, обеспечивающих требуемую точность решения.

Постановка задачи и метод решения

Рассмотрим однородную область моделирования  (песчаник с модулем Юнга 5 Гпа и коэффициентом Пуассона 0.27).

Нижнее основание области жестко зафиксировано, верхнее основание считается смещенным вверх по оси Z на 0.1 м. Боковая поверхность считается свободной. Сила тяжести в данном случае не учитывается. Таким образом, краевая задача упругой деформации имеет вид:

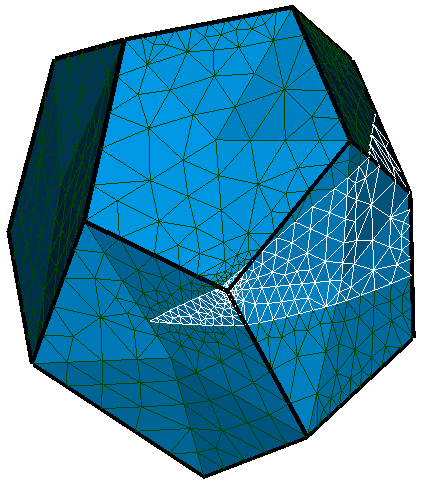
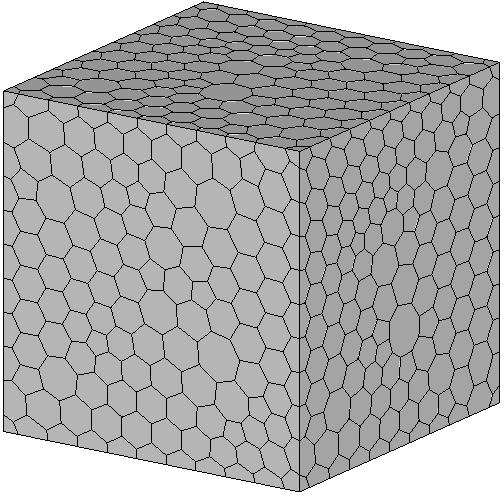








Рассмотрим полиэдральное согласованное нерегулярное разбиение всей области моделирования  на непересекающиеся полиэдры  (рис. 1.а), где  – число вершин полиэдра.



X

Y

Z

а) б)

**Рис. 1. Структура расчётной области и примеры сеточных разбиений: а – макроэлементное разбиение**  **на 1290 полиэдров; б – микроразбиение**  **одного из полиэдров.**

Введём на множестве  конечномерное подпространство:

,

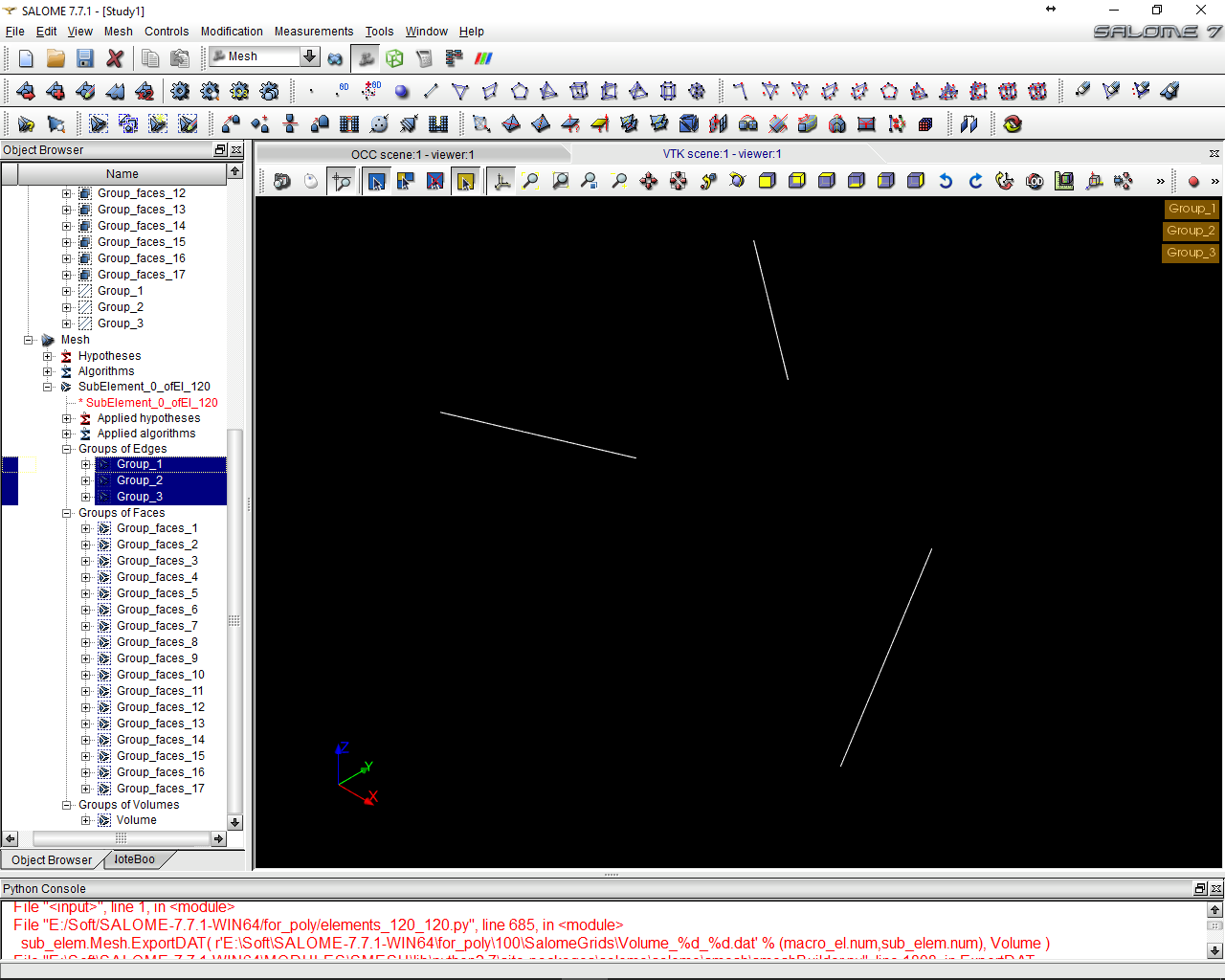
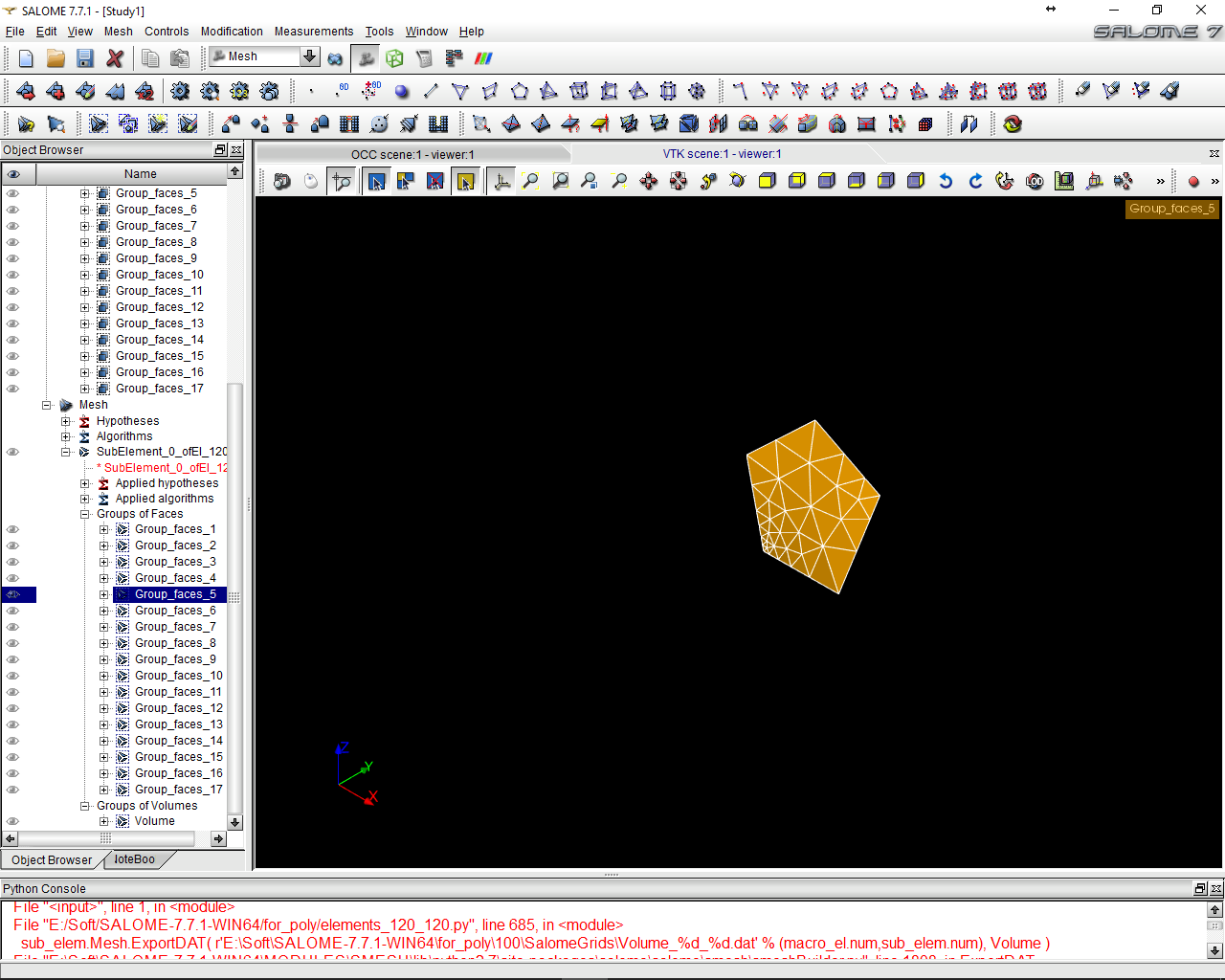
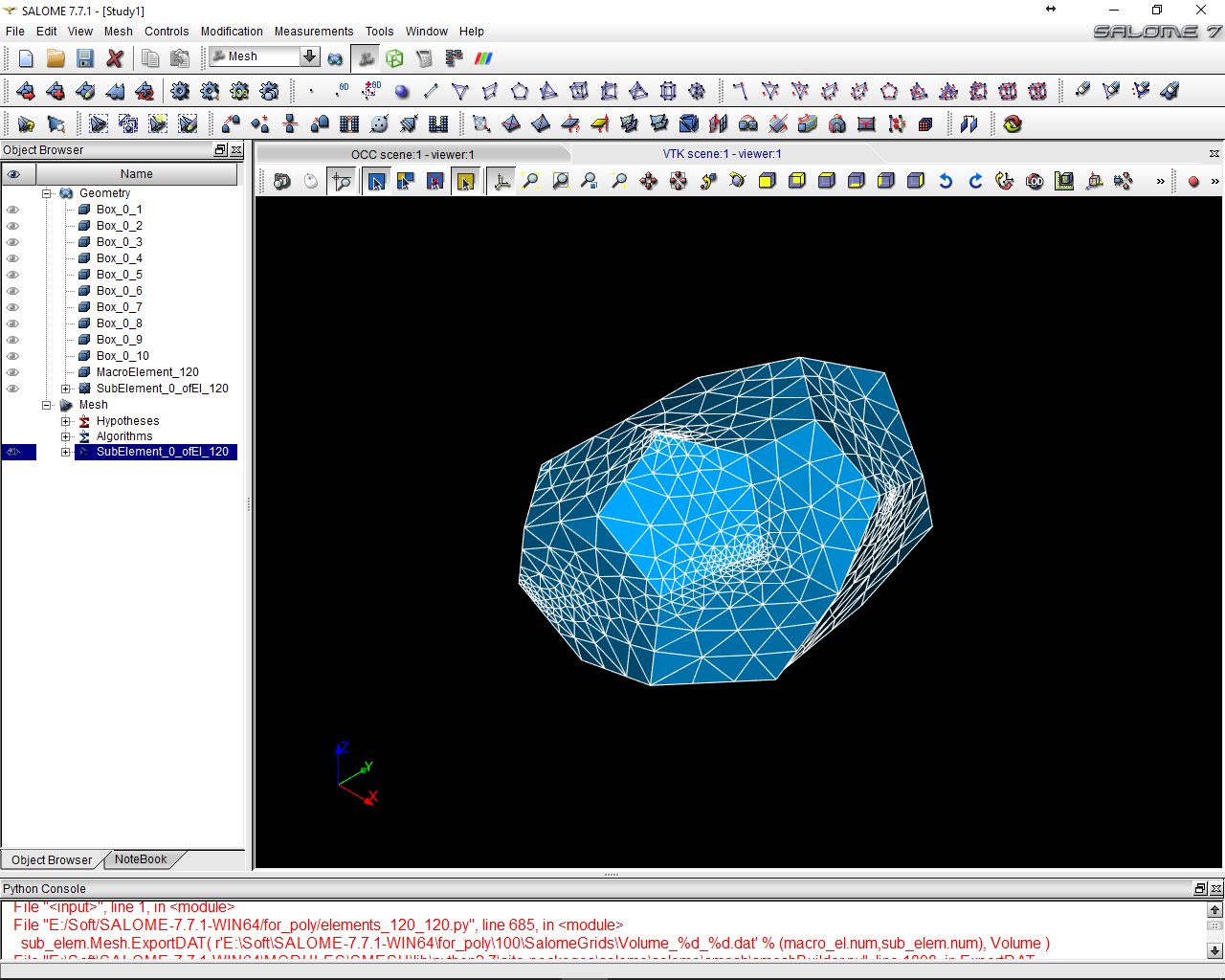
где  – пространство неполиномиальных финитных функций формы, ассоциированных с узлами  сетки . Поскольку функции  являются финитными и удовлетворяют условию «разбиения единицы» по требованию FE-HMM [6], то для их построения на каждом из макроэлементов-полиэдров  независимо построим тетраэдральные согласованные нерегулярные тетраэдральные разбиения ,  (рис. 1.б), и определим на них соответствующие конечноэлементные подпространства:

,

где  – пространство полиномов первой степени (в общем случае полиномов степени ). Тогда сформулируем дискретные вариационные постановки на микроуровне для , , 



где  является значением функции  на внешней границе полиэдра, и может быть найдено из решения аналогичной задачи, но меньшей мерности, то есть из решения квазидвумерной задачи по грани (рис. 2.б). Для этого так же потребуется значения искомой функции по ребрам (рис. 2.в), которые могут быть получены простой интерполяцией от  до , если ребро не пересекается включением, в обратном случае строятся одномерные подзадачи по ребрам с краевыми условиями  или  в зависимости от индекса функции, то есть в результирующей функции  в весе для степени свободы, соответствующей вершине с индексом  должно быть значение , а во всех остальных весах, ассоциированных с вершинами макроэлемента значения равны .



а) б) в)

**Рис. 2. Иерархическая система сеточных разбиений одного макроэлемента для серии вложенных подзадач построения неполиномиальных функций формы: а – тетраэдральная сетка во всем макроэлементе; б – треугольные сетки по граням (выделяются из тетраэдральной сетки); в – одномерные сетки по ребрам (выделяются из треугольных сеток)**

Также по определению FE-HMM должно выполняться свойство



Тогда дискретная вариационная формулировка на макроуровне имеет вид



где  – гомогенизированный коэффициент среды.

Таким образом, решение задачи – может быть представлено в каждой точки области  в виде разложения по соответствующим функциям формы:

,

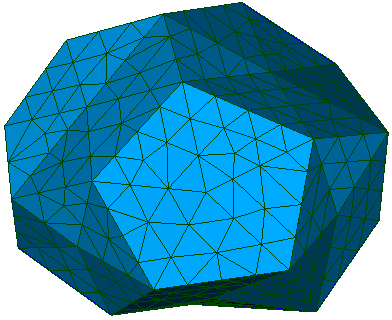
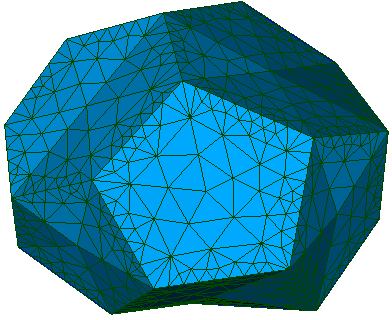
где  и  – значения весов степеней свободы, получаемые из решения и , соответственно,  – базисные функции из пространства, определенного в .

Численная оценка точности FE-HMM

В качестве точного решения  для всех вариантов использовалось решение задачи – , полученное классическим методом конечных элементов на подробной сетке (более 10 млн. тетраэдров). Тогда относительная погрешность z-компоненты решения гетерогенным многомасштабным методом конечных элементов в норме  имеет вид:



В таблице приведено сравнение двух подходов к построению разбиений на микроуровне: без адаптации и с адаптацией. В варианте с адаптацией выполнялось дополнительное подразбиение тетаэдральной сетки на тех ребрах полиэдра, где при стандартном шаге построения оказывалось менее 5 узлов тетраэральной сетки (рис. 3).

а) б)

**Рис. 3. Микроразбиение макроэлемента без использования адаптации (а) и с использованием (б) дополнительного подразбиения при равных исходных критериях мелкости разбиений**

Таблица

Численная характеристика вычислительных сеток ( – суммарное число тетраэдров на микроуровне) для однородного образца и сравнение двух подходов к построению сеток на микроуровне

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Число полиэдров | Критерий мелкости микроразбиения | Без адаптации | | С адаптацией | |
|  | Относительная погрешность для |  | Относительная погрешность для |
| 1 | 2 | 3 | 4 | 5 | 6 |
| 26 | 0.03 | 136 235 | 1.44E-02 | 135 650 | 1.48E-02 |
| 0.02 | 490 329 | 1.42E-02 | 534 517 | 1.40E-02 |
| 0.015 | 957 375 | 1.40E-02 | 1 131 719 | 1.40E-02 |
| 0.01 | 4 106 514 | 1.40E-02 | 4 517 391 | 1.41E-02 |
| 76 | 0.03 | 142 146 | 5.74E-02 | 242 252 | 8.70E-03 |
| 0.02 | 454 084 | 7.65E-03 | 549 335 | 7.88E-03 |
| 0.015 | 1 194 948 | 7.38E-03 | 1 300 279 | 7.54E-03 |
| 0.01 | 4 017 742 | 7.35E-03 | 4 237 923 | 7.63E-03 |
| 1 | 2 | 3 | 4 | 5 | 6 |
| 135 | 0.03 | 143 046 | 3.45E-01 | 431 854 | 6.82E-03 |
| 0.02 | 522 917 | 4.37E-02 | 628 431 | 6.68E-03 |
| 0.015 | 1 225 841 | 5.69E-03 | 1 275 817 | 5.94E-03 |
| 0.01 | 4 551 570 | 5.46E-03 | 4 527 933 | 5.75E-03 |
| 217 | 0.03 | 155 259 | 4.81E-01 | 686 404 | 6.73E-03 |
| 0.02 | 493 548 | 1.78E-01 | 746 568 | 6.27E-03 |
| 0.015 | 1 271 845 | 5.61E-03 | 1 398 700 | 5.65E-03 |
| 0.01 | 4 279 962 | 5.35E-03 | 4 336 714 | 5.11E-03 |
| 575 | 0.03 | 375 650 | 8.00E-01 | 2 282 913 | 5.76E-03 |
| 0.02 | 534 636 | 4.93E-01 | 1 831 935 | 4.16E-03 |
| 0.015 | 1 122 026 | 3.13E-01 | 2 517 785 | 3.98E-03 |
| 0.01 | 4 290 063 | 3.68E-03 | 4 290 063 | 3.68E-03 |
| 1290 | 0.03 | 756 602 | 8.44E-01 | 5 156 084 | 1.33E-02 |
| 0.02 | 769 786 | 8.33E-01 | 3 618 022 | 3.79E-03 |
| 0.015 | 1 159 803 | 5.53E-01 | 4 396 185 | 4.00E-03 |
| 0.01 | 3 884 906 | 3.16E-01 | 5 883 967 | 3.69E-03 |

Как видно из анализа результатов таблицы сходимость FE-HMM является иерархической, то есть относительная погрешность решения уменьшается, как при увеличении числа полиэдров на макроуровне, так и при дроблении сетки на микроуровне. Однако в варианте микросетки без адаптации есть предел мелкости на макроуровне. А именно, на сетке с 1290 полиэдрами относительная погрешность выше, чем на сетках с меньшим числом полиэдров при сопоставимой мелкости микроразбиений. Это связано с тем, что при заданном критерии микроразбиений не достигается необходимая точность при построении макромасштабных функций формы даже в однородной среде. Введение дополнительного критерия адаптации сетки по ребрам позволяет это контролировать, однако может существенно повысить суммарное число степеней свободы на микроуровне.

Масштабируемость вычислительных схем

С точки зрения реализации алгоритм гетерогенного многомасштабного метода конечных элементов состоит из следующих этапов (в данном случае будем опускать процедуры построения иерархических конечноэлементых сеток, поскольку оценить их с точки зрения требуемых ресурсов затруднительно, поскольку привлекается сторонний программный комплекс SALOME, однако для каждого макроэлемента микросетка сроится также независимо):

1. Считывается и обрабатывается (например, формируется отношение соседства) полиэдральная макросетка.

2. Для каждого из макроэлементов независимо строятся независимо неполиномиальные функции формы, для этого

2.1) считывается и обрабатывается тетраэдральная сетка макроэлемента (в том числе выделяются подсетки по каждой из граней и по каждому ребру);

2.2) выполняется анализ ребер и формирование следов функции формы по одномерной границе: если на ребре нет однородностей, то подставляется аналитическое соотношение, если есть, то решаются подзадачи;

2.3) формируются следы функций формы на квазидвумерных гранях (в общем случае грани могут быть не плоскими) через решение серии подзадач;

2.4) формируются и решаются системы линейных алгебраических уравнений (СЛАУ) для каждой из функций формы с учетом соответствующих следов на границах.

3. Ассемблирование СЛАУ по макроэлементам с учетом глобальных краевых условий и решение.

4. Анализ и вывод результатов.

Как не трудно видеть, на каждом из этапов возможна, как алгоритмическая, так и программная оптимизация. Однако в любом случае на втором этапе порождается большое число дополнительных подзадач (на полиэдральном носителе в нашем случае количество функций формы совпадает с количеством вершин, которых может быть более 30). Тем не менее, обработка каждого из макроэлементов в FE-HMM может выполняться полностью независимо от соседних элементов. В рамках данной работы была реализована параллельная версия алгоритма на базе технологии OpenMP. На рисунке 4 приведены замеры времени решения задачи – для однородного образца для разных вариантов сеток с 26, 76 и 135 полиэдрами и сохранением мелкости микроразбиения (около 1 млн. тетраэдров). Эксперименты выполнялись на компьютере с процессором AMD Ryzen Threadripper PRO 3975WX 32-Cores (3.50 GHz).

**Рис. 4. Зависимость времени решения задачи от числа используемых процессоров**

Заключение

Для решения задачи об упругой деформации гетерогенного твердого тела предложена параллельная модификация гетерогенного многомасшатбного метода конечных элементов на полиэдральных носителях.

Разработанные алгоритмы и вычислительные схемы проанализированы с точки зрения численной сходимости и требований к вычислительным ресурсам. Экспериментально показана специфичная для иерархических конечноэлементных методов «двухуровневая» сходимость, то есть точность решения можно контролировать, как варьированием дискретизации на макроуровне, так и на микроуровне. При этом для достижения наилучшей точности решения необходимо соблюдать баланс при измельчении сеток как при построении неполиномиальных функций формы на уровне микроразбиения, так и на уровне макроразбиения. Предложенная в рамках работы автоматическая адаптация сетки на микроуровне вблизи ребер полиэдра, позволяет получить более точное решение (до двух порядков) при сопоставимых размерах дискретизаций. Использованный в рамках данного этапа подход к параллельной реализации вычислительных схем позволил сократить время решения до 10 раз на 32х потоках, однако в дальнейшем планируется модифицировать технологию распределения нагрузки по процессорным потокам для повышения эффективности.

*Работа выполнена при финансовой поддержке Гранта Президента РФ МК-3230.2022.1.5 (задача упругой деформации).*

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Zienkiewicz O.C., Bahrani A.K., Arlett P.L. Numerical solution of 3-dimensional field problems // Proc. IEE (London). 1968. Vol. 16. P. 367-369.
2. E W., Ming P., Zhang P. Analysis of the heterogeneous multiscale method for elliptic homogenization problems // J. Am. Math. Soc. 2003. Vol. 8. P. 121-156.
3. Abdulle A., E W., Engquist B., Vanden-Eijnden E. The heterogeneous multiscale methods // Acta Numerica. 2012. Vol. 21. P. 1-87.
4. Epov M.I., Shurina E.P., Itkina N.B., Kutishcheva A.Y., Markov S.I. Finite element modeling of a multi-physics poro-elastic problem in multiscale media // Journal of Computational and Applied Mathematics. 2019. Vol. 352. P. 1-22.
5. Eidel B., Fischer A. The Heterogeneous Multiscale Finite Element Method for the Homogenization of Linear Elastic Solids and a Comparison with the FE2 Method // Computer Methods in Applied Mechanics and Engineering. 2017. Vol. 329(1). P. 332-368.
6. Shurina E.P., Epov M.I., Kutishcheva A.Y. Numerical simulation of the percolation threshold of theelectric resistivity // Computational technologies. 2017. Vol. 22(3). P. 3-15.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

|  |  |
| --- | --- |
| Кутищева Анастасия Юрьевна  c.н.с лаб. Математического моделирования многофизичных процессов в нативных и искусственных многомасштабных гетерогенных средах, ФГБУН Институт нефтегазовой геологии и геофизики им. А.А. Трофимука СО РАН, 630090, Россия, г. Новосибирск, пр-т Академика Коптюга, 3,  доц. каф. Геофизических систем, ФГБОУ ВО Новосибирский государственный технический университет, 630073, Россия, г. Новосибирск, пр-т К.Маркса 20,  Эл. почта:KutischevaAY@ipgg.sbras.ru | **Марков Сергей Игоревич**  Ассистент каф. Вычислительных технологий, ФГБОУ ВО Новосибирский государственный технический университет, 630073, Россия, г. Новосибирск, пр-т К.Маркса, 20,  зав. лаб. Математического моделирования многофизичных процессов в нативных и искусственных многомасштабных гетерогенных средах, ФГБУН Институт нефтегазовой геологии и геофизики им. А.А. Трофимука СО РАН,  630090, Россия, г. Новосибирск, пр-т Академика Коптюга, 3,  Эл. почта: www.sim91@list.ru |

*A.Yu. KUTISHCHEVA, S.I. MARKOV*

**NUMERICAL ESTIMATION OF THE ACCURACY OF THE HETEROGENEOUS MULTISCALE FINITE ELEMENT METHOD ON POLYHEDRAL SUPPORTS**

**Numerical simulation of processes in such complex formations as permafrost is impossible without using modern multiscale methods. For example, the heterogeneous multiscale finite element method (FE-HMM) can be used to simulate the elastic deformation of solids. However, when solving practical problems, it is necessary to have a tool for controlling the computational error, as well as a priori data on the technical limitations of the method. As part of this work, studies were conducted on the effect of mesh decomposition at different hierarchy levels on the accuracy and speed of solving the elastic deformation problem for FE-HMM on polyhedral supports at the macrolevel and tetrahedral finite elements at the microlevel. The estimates obtained allow us to adjust the strategies for planning computational experiments, which are largely related to the construction of a set of meshes of varying detail. For example, it was obtained that refinement of the microlevel mesh on the edges of macro-polyhedrons allows us to increase the solution accuracy to two orders while preserving the total size of the discretizations.**

***Keywords:*** *heterogeneous multiscale finite element method, polyhedra, elastic deformation of solids, natural parallelism.*

REFERENCES

1. Zienkiewicz O.C., Bahrani A.K., Arlett P.L. Numerical solution of 3-dimensional field problems. *Proc. IEE (London)*. 1968. Vol. 16. P. 367-369.
2. E W., Ming P., Zhang P. Analysis of the heterogeneous multiscale method for elliptic homogenization problems. *J. Am. Math. Soc*. 2003. Vol. 8. P. 121-156.
3. Abdulle A., E W., Engquist B., Vanden-Eijnden E. The heterogeneous multiscale methods. *Acta Numerica*. 2012. Vol. 21. P. 1-87.
4. Epov M.I., Shurina E.P., Itkina N.B., Kutishcheva A.Y., Markov S.I. Finite element modeling of a multi-physics poro-elastic problem in multiscale media. *Journal of Computational and Applied Mathematics*. 2019. Vol. 352. P. 1-22.
5. Eidel B., Fischer A. The Heterogeneous Multiscale Finite Element Method for the Homogenization of Linear Elastic Solids and a Comparison with the FE2 Method. *Computer Methods in Applied Mechanics and Engineering*. 2017. Vol. 329(1). P. 332-368.
6. Shurina E.P., Epov M.I., Kutishcheva A.Y. Numerical simulation of the percolation threshold of theelectric resistivity. *Computational technologies*. 2017. Vol. 22(3). P. 3-15.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

|  |  |
| --- | --- |
| **Anastasiya Yu. Kutishcheva**  Senior researcher, Laboratory for Mathematical Modeling of Multi-Physical Processes in Native and Artificial Multiscale Heterogeneous Media, The Trofimuk Institute of Petroleum Ge-ology and Geophysics, SB RAS, 3, Koptug ave., Novosibirsk, Russia, 630090,  Associate Professor, Department of Geophysical Systems, Novosibirsk State Technical University, 20, Prospekt K. Marksa, Novosibirsk, Russia, 630073,  E-mail:KutischevaAY@ipgg.sbras.ru | **Sergey I. Markov**  Header of the Laboratory for Mathematical Modeling of Multi-Physical Processes in Native and Artificial Multiscale Heterogeneous Media, The Trofimuk Institute of Petroleum Ge-ology and Geophysics, SB RAS, 3, Koptug ave., Novosibirsk, Russia, 630090,  Assistant at the Department of Computational Technologies, Novosibirsk State Technical University, 20, Prospekt K. Marksa, Novosibirsk, Russia, 630073,  E-mail: www.sim91@list.ru |