

# КОЛИЧЕСТВЕННАЯ ОЦЕНКА ОБЪЁМА ПОВРЕЖДЁННЫХ ОБЛАСТЕЙ ЛЁГКИХ НА ОСНОВЕ ЛИНЕЙНЫХ МЕТРИК

Н.А.Смелкина, А.Г. Храмов

Самарский государственный аэрокосмический университет имени академика С.П. Королёва (национальный исследовательский университет) (СГАУ), Самара, Россия

В данной работе предлагается метод измерения повреждённой области лёгких человека по сечениям компьютерной томографии, основанный на вычитании области здоровых тканей лёгких из общего объёма при отсутствии патологии. Общий объём лёгких вычисляется на основе линейной регрессионной модели с использованием комбинации характерных размеров, получаемых «вручную». Объём области здоровых тканей лёгких определяется с использованием пороговой сегментации функции яркости томографических сечений. Экспериментальные исследования показали, что погрешность рассмотренного метода не превышает 10%.

**Ключевые слова:** диагностические изображения, измерения на изображении, обработка изображений, интеллектуальный анализ данных, линейная регрессионная модель.

## Введение

В настоящее время основные методы сегментации лёгких [1] основаны на знании функции яркости, текстурных признаков и расположения на томографических сечениях. В любом из указанных методов является обязательным наличие контура лёгких. В случае же интерстициального поражения лёгких (пневмония, альвеолярная инфильтрация) или сильного механического повреждения контур лёгких часто является невыраженным или даже отсутствует. Существуют методы, позволяющие сегментировать лёгкие с данной патологией и вычислять их объём [1, 2, 3], но все они являются вычислительно трудоёмкими, при этом достигая погрешности в 10%, и требуют большого объёма исходной информации, такой, как: расположение повреждения, основные точки, характеризующие лёгкие и др.

Так или иначе, все указанные методы сегментации лёгких направлены на диагностирование заболевания и определение его количественных характеристик. Поскольку методы нахождения объёма повреждений являются очень сложными и медленными, достигая недостаточно высокой точности вычислений, был найден метод вычисления общей величины патологии лёгких, избегая сегментации органов. Так, в работе [4] был предложен метод определения величины плеврального повреждения, имеющего характерную форму, при помощи простейших метрик на плоскости и в пространстве, не сегментируя патологию. Точность данного метода варьируется в пределах 4,5- 5,1%, в зависимости от выбранного измерения или комбинации измерений.

В данной работе предлагается оценить общий объём лёгких с помощью простейших измерений, без использования инструментов сегментации. Так, измерение  $a$  (рис.1) характеризует пространственную координату, т.н. длину лёгких. Метрики  $b$  и  $c$  (рис.2, 3 соответственно) представляют собой «глубину» и «толщину» лёгких. Расстояние  $d$  (рис. 4) измеряется между крайними, наиболее выделяющимися точками лёгкого. Измерение  $e$  (рис. 5) содержит в себе диаметрную характеристику плевральной полости.



Рис.1. Пространственное измерение a

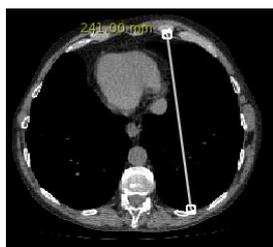


Рис.2. Измерение b

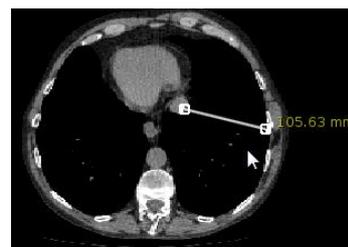


Рис.3. Метрика c

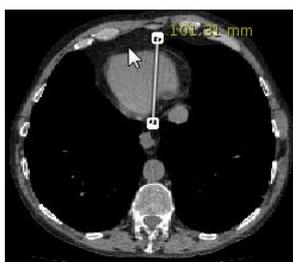


Рис.4. Расстояние d



Рис.5. Метрика e

Таким образом, оценивая объём здоровых лёгких, можно сделать вывод о величине патологических областей, исключая объём здоровой ткани лёгких, выделенный с помощью пороговой сегментации функции яркости и связности, из объёма здоровых лёгких. Необходимо также отметить, что анатомическая структура лёгких (положение, форма, текстура) уникальна, поэтому в данном случае невозможно воспользоваться знанием о форме лёгких для построения модели зависимости. Была построена линейная регрессионная модель на основе статистики, состоящей из 47 пациентов со здоровыми лёгкими, у каждого из которых были найдены соответствующие размеры и объём лёгких. Тестовая выборка состояла из 5 пациентов, двое из которых имели сильные механические повреждения лёгких в результате аварии, и один пациент с альвеолярной инфильтрацией в лёгких.

Были построены линейные регрессионные модели, зависящие как от одиночных измерений, так и от комбинации нескольких измерений. В качестве наилучшей была выбрана модель с минимальной остаточной суммой квадратов невязок.

### Измерение общего объёма легких

Статистическая линейная регрессионная модель зависимости объёма от линейных измерений.

Классическая модель линейной относительно параметров  $\theta$  регрессии имеет вид [5]:

$$Y = F\theta + \xi, \quad (1)$$

где  $F$  – матрица эксперимента,  $Y$  – вектор наблюдения,  $\xi$  – вектор шума наблюдения.

Всего было построено более 40 моделей линейной регрессии, кубических и квадратичных относительно как одиночных измерений, так и их произведений. Оптимальность модели выбиралась исходя из критерия минимизации суммы квадратов остатков (метод наименьших квадратов):

$$R_0 = \sum_{i=1}^N \left[ y^i - (f_1^i \theta_1 + f_2^i \theta_2 + f_3^i \theta_3 + \dots + f_m^i \theta_m) \right]^2 \rightarrow \theta_1, \theta_2, \theta_3, \dots, \theta_m.$$

Наилучшей признана модель, зависящая от произведения трёх наиболее информативных измерений и имеющая следующий вид:

$$\begin{aligned} Y &= 0,2788 + 0,2013 \times (a \times b \times e) + 0,0117 \times (a \times b \times e)^2, \\ R_0 &= 0,4934. \end{aligned} \quad (2)$$

График данной модели представлен на рис. 6.

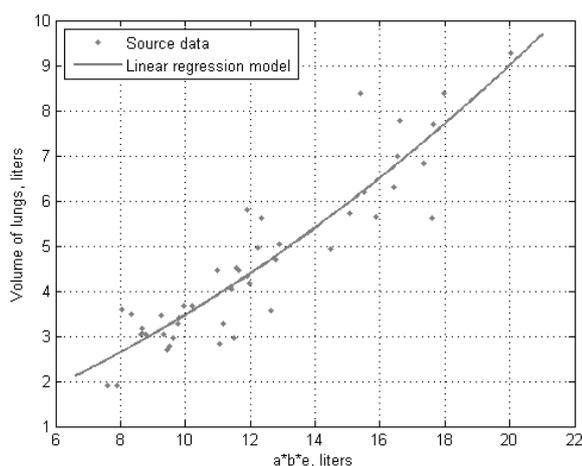


Рис.6. График линейной регрессии модели (2)

Критерием информативности различных метрик является также метод наименьших квадратов. Таким образом, наиболее информативными оказались измерения  $a, b, e$ , наиболее полно характеризующие анатомическую форму лёгких. Модели, содержащие информацию о «толщине» лёгких (метрика  $c$ ) и расстояние между крайними точками лёгкого (измерение  $d$ ), показали неточный результат.

Наилучшая модель, зависящая от одного измерения, имеет существенно большую сумму квадратов невязок и выглядит следующим образом:

$$\begin{aligned} Y &= 2,0360 - 0,3829 \times (b \times b) + 0,2771 \times (b \times b)^2, \\ R_0 &= 1,3410. \end{aligned} \quad (3)$$

## Измерение объёма области здоровых тканей

### Метод Оцу

Метод Оцу (*Otsu's method*), благодаря своей универсальности, широко применяется в обработке изображений и распознавании образов. Алгоритм основан на разделении пикселей полутонового изображения на «полезные» и «фоновые» путём нахождения порога, уменьшающего внутриклассовую дисперсию, которая определяется как взвешенная сумма дисперсий двух классов [6]:

$$\sigma_w^2(t) = w_1(t)\sigma_1^2(t) + w_2(t)\sigma_2^2(t),$$

где веса  $w_i$  - это вероятности двух классов, разделенные порогом  $t$ ,  $\sigma_i^2$  - дисперсия этих классов.

Полученное бинарное изображение имеет следующий вид:

$$y(i, j, k) = \begin{cases} 1, & x(i, j, k) \in D \\ 0, & x(i, j, k) \notin D \end{cases},$$

где  $D$  – область лёгких и воздуха,  $x(i, j, k)$  - пиксель исходного полутонового изображения,  $y(i, j, k)$  – пиксель полученного бинарного изображения.

### Определение связной области

Далее ставится точка в области лёгких, тем самым определяя область интереса. После этого обрабатывается область, связанная с этой точкой. Таким образом, полученное изображение имеет вид:

$$y(i, j, k) = \begin{cases} 1, & x(i, j, k) \in D \cup x(i, j, k) \in O \\ 0, & x(i, j, k) \notin D \cap x(i, j, k) \notin O \end{cases},$$

где  $O$  – связная область с заданной начальной точкой.

Далее вычисляется объём полученного бинарного изображения легких по формуле:

$$V = \frac{1}{10^6} \sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^l x_1(i, j, k) \times size(i) \times size(j) \times size(k),$$

[V] = liters,

где  $x_1(i, j, k)$  – пиксели со значением 1,  $size(i, j, k)$  – размеры соответствующих вокселей в мм.

## Экспериментальные исследования

### Вычисление объёма повреждённой области

Тестовая выборка состоит из 5 пациентов. Объём патологии в лёгких вычисляется как разность общего объёма здоровых лёгких, определяемого по формуле линейной регрессии, и объёма здоровой ткани лёгких:

$$V_{nam} = V_{общ} - V_{зд}$$

У пациентов со здоровыми лёгкими разность должна быть близка к 0. В табл.1 приведены результаты исследования модели (2).

Табл.1. Результаты экспериментального исследования модели (2)

Номер пациента	1	2	3	4	5
<b>Наличие патологии</b>	нет	нет	есть	есть	есть
<i>a, dm</i>	2,319	2,412	2,613	2,302	2,722
<i>b, dm</i>	1,768	1,960	1,563	1,446	1,756
<i>e, dm</i>	2,387	3,100	2,623	2,392	2,607
$V_{зд}, l$	3,333	5,335	3,178	2,062	4,056
$V_{общ}, l$	3,370	5,742	3,778	2,623	4,605
$V_{nam}, l$	-	-	0,407	0,524	0,584
<b>Ошибка</b>	0,011	0,076	0,054	0,014	0,007

Как видно из табл. 1, погрешность рассмотренного метода не превышает 10%, а в некоторых случаях достигает 0,7%. Объём патологии рассчитывался вручную для сравнения с полученными результатами.

Модель, основанная на одном измерении, менее точна по сравнению с моделью, зависящей от произведения различных метрик. Однако результаты исследования модели (3), приведенные в табл.2, выявили также достаточно хорошую оценку объёма патологии.

Табл.2. Результаты экспериментального исследования модели (3)

Номер пациента	1	2	3	4	5
<b>Наличие патологии</b>	нет	нет	есть	есть	есть
<i>b, dm</i>	1,768	1,960	1,563	1,446	1,756
$V_{зд}, l$	3,333	5,335	3,178	2,062	4,056
$V_{общ}, l$	3,547	4,654	2,755	2,446	3,491
$V_{nam}, l$	-	-	0,407	0,524	0,584
<b>Ошибка</b>	0,064	0,128	0,231	0,054	0,248

Согласно табл. 2, погрешность модели (3) значительно превышает погрешность модели (2) и достигает почти 25% в отдельных случаях. Это объясняется недостаточностью знания о форме лёгких и отсутствии пространственных измерений.

## Заключение

Данная работа призвана определять объём патологии с целью диагностирования тяжести повреждений. Исследовался метод простейшего вычисления объема патологической области лёгких при помощи основных линейных метрик лёгких. Была достигнута погреш-

ность вычислений, не превышающая 10%, при этом метод является вычислительно простым и дающим хорошую производительность. Погрешность моделей в подавляющем случае обусловлена двумя факторами - человеческим вмешательством (расстановкой наиболее точно описывающих лёгкое измерений) и анатомической структурой лёгких. Также точность регрессионной модели в некоторой степени снижается из-за несоответствия качества изображений в обучающей выборке – количество срезов варьируется от 49 до 298. В настоящее время ведётся разработка автоматической расстановки измерений для снижения погрешности вычислений.

## Литература

1. Mansoor, A. Segmentation and Image Analysis of Abnormal Lungs at CT: Current Approaches, Challenges, and Future Trends / Awais Mansoor, Ulas Bagsi, Brent Foster, Ziyue Xu, Georgios Z. Papadakis, Les R. Folio, Jayaram K.Udupa, Daniel J. Mollura // *RadioGraphics*. – 2015. – Vol. 35(1). – P. 1056-1076.
2. Birkbeck, N. Robust Segmentation of Challenging Lungs in CT Using Multi-stage Learning and Level Set Optimization / Neil Birkbeck, Michal Sofka, Timo Kohlberger, Jingdan Zhang, Jens Wetzl, Jens Kaftan, S. Kevin Zhou // *Computational Intelligence in Biomedical Imaging*. – Springer New York, 2014. – P. 185-208.
3. Mansoor, A. A Generic Approach to Pathological Lungs Segmentation / Awais Mansoor, Ulas Bagsi, Brent Foster, Ziyue Xu, Georgios Z. Papadakis, Les R. Folio, Jayaram K.Udupa, Daniel J. Mollura // *IEEE Transactions on Medical Imaging*. – 2014. – Vol. 33(12). – P. 2293 – 2310. – ISSN 0278-0062.
4. Hazlinger, M. Quantification of pleural effusion on CT by simple measurement / Martin Hazlinger, Filip Ctvrtlik, Katerina Langova, Miroslav Herman // *Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub.* – 2014. – Vol. 158(1). – P. 107 – 111.
5. Дрейпер Н. Прикладной регрессионный анализ/ Н. Дрейпер, Г. Смит; пер. с англ. под ред //Адлера ЮП, Горского ВГМ: Статистика. – 1973.
6. Otsu N. A threshold selection method from gray-level histograms / Nobuyuki Otsu // *Transactions on systems, man, and cybernetics*. – 1979. – Vol. 9(1). – P. 62-66.