

3D 虚拟复原精度的差异对头骨测量数值的影响 ——以 Mimics 软件为例

张玄^{1, 2}; 张亚盟³; 吴秀杰^{1, 4}

1. 中国科学院脊椎动物演化与人类起源重点实验室, 中国科学院古脊椎动物与古人类研究所, 北京 100044; 2. 中国科学院大学, 北京 100049; 3. 山东大学青岛校区文化遗产研究院, 青岛 266237; 4. 中国科学院生物演化与环境卓越创新中心, 北京 100044

摘要: 近年来, 根据三维软件虚拟复原的头骨来获取测量数据的方法被越来越多地应用在古生物, 特别是古人类学的研究中。然而, 对于三维软件不同精度虚拟复原的头骨, 其测量数值是否有差异, 研究者并不是很清楚。本文以 Mimics 软件为例, 根据其复原模型简化规则, 选择未精简的最佳精度模型作为标准进行配对 t 检验或非参数检验, 通过对 43 例云南人头骨的顶骨矢状弦长、颅周长、头盖部面积、乳突小房表面积、颅容量、乳突小房体积等六个测量项目的对比和分析, 对 Mimics 软件低、中、高、最佳四种精度 3D 虚拟复原头骨间的测量差异进行了研究。结果显示: 颅周长、头盖部面积、颅容量、乳突小房体积四项的所有简化精度模型的测量数据均与最佳精度模型测量数据的差异具有显著性; 而除高精度组外, 顶骨矢状弦长及乳突小房表面积的其余精度组测量数据均与最佳精度组差异具有显著性; 此外, 顶骨矢状弦长、颅周长、头盖部面积、颅容量的简化精度与最佳精度的测量差异占比均小于 3%。而乳突小房表面积的低精度与最佳精度测量差异占比可超过 50%, 乳突小房体积的低精度与最佳精度测量差异占比可超过 120%。这一结果提示我们, 在测量 Mimics 复原的三维模型时, 体量大差异小的测量项可以在较低精度的复原模型上进行测量; 而对头骨内部腔室这样体积小表面粗糙的结构, 复原模型的精度选择及测量数据比较需要格外谨慎。

关键词: Mimics; 电子计算机断层扫描; 三维重建; 头骨测量; 差异检验; 精度

The influence of quality parameter selection on 3D virtual reconstruction model precision based on dry skull

ZHANG Xuan^{1, 2}; ZHANG Yameng³; WU Xiujie^{1, 4}

1. Key Laboratory of Vertebrate Evolution and Human Origin of Chinese Academy of Sciences, Institute of Vertebrate Paleontology and Paleoanthropology, Chinese Academy of Sciences, Beijing 100044; 2. University of Chinese Academy of Sciences, Beijing 100049; 3. Institute of Cultural Heritage, Shandong University, Qingdao 266237; 4. CAS Center for Excellence in Life and Paleoenvironment, Beijing 100044

收稿日期: 2019-01-28; 定稿日期: 2019-06-13

基金项目: 中国科学院战略性先导科技专项 B 类(XDB26000000), 中国科学院创新交叉团队(2016)和国家自然科学基金(416300102、41672020)资助

作者简介: 张玄(1993-), 女, 汉族, 中国科学院古脊椎动物与古人类研究所博士研究生。Email: zhangxuan115@mails.ucas.ac.cn

Citation: Zhang X, Zhang YM, Wu XJ, et al. The influence of quality parameter selection on 3D virtual reconstruction model precision based on dry skull[J]. Acta Anthropologica Sinica, 2019, 38e: 875-887

Abstract: In anthropological studies, craniometry is an important means of obtaining human skull dimensional information and other characteristics. Recent technological advances have resulted in computed tomography (CT) and three-dimensional (3D) reconstruction technologies that offer great advantages to craniometry specialists by facilitating both the collection of repeated measurements and analysis of inner structures without destroying specimens. However, the influence of 3D reconstruction precision based on dry skull measurements is unclear. Mimics, one of several commonly used 3D reconstruction software packages, provides users with the choice to select from four quality settings during the 3D model reconstruction process: low, medium, high, and optimal quality. Ultimately, lower quality corresponds with a smaller file size and faster modeling computing speed. In this study, four models were generated from a single skull using each of the four quality settings. Measurements were made of the parietal sagittal chord, cranial horizontal circumference, cranial surface area, cranial capacity, mastoid cell system surface area, and mastoid cell system volume of 43 reconstructed Yunnan modern human cranial specimens modeled using Mimics based on various quality settings. According to matrix simplification rules of Mimics 3D reconstruction, unsimplified optimal quality models were chosen as the standard for paired t-tests or non-parametric tests followed by calculation of the proportion of measurement difference (expressed as a percentage). The results indicated that the high-quality modeling group, including sagittal chord of parietal bone and mastoid cell system surface area measurements, exhibited no difference in optimal quality. Conversely, measurement data of the other six typical characteristics used to generate simplified quality models significantly differed from optimal quality model data. Notably, measurement differences between simplified and optimal quality models of sagittal chord, cranial horizontal circumference, surface area of cranium, and cranial capacity were below 3%, while absolute values of measurement differences between low and optimal quality measurements of mastoid cell system surface area and volume exceeded 50% and 120%, respectively. These results suggest that low-quality 3D reconstruction models can be useful for measurements of large-scale morphological features with smooth surfaces. As for small-scale morphological features with rough surfaces, such as the internal cavity sinus of the skull, three-dimensional reconstruction quality parameters must be selected very carefully.

Keyword: Mimics; CT; 3D virtual reconstruction; Measurement of dry skull; Difference test; Model quality

1 引 言

在人类学的研究中,研究者常常通过测量骨骼,尤其是头骨标志性解剖位置的角度、长度、面积、体积等,来获得标本的特征信息^[1-3]。传统的测量是在标本实体上进行手工测量,

如果标本的保存状况不佳^[4], 表面覆盖有胶结物不易清理, 或是标本数量过于庞大, 都会显著增加观测者的工作量。

目前, 随着技术的发展, 研究者可以通过计算机断层扫描 (Computed Tomography, CT) 或是三维激光扫描等方法获得标本信息, 在二维 (2D) 断层扫描影像上测量, 或是在计算机中重建为三维 (3D) 模型后再测量。这种方法不仅减少了对标本可能受到的损伤, 而且可以将大批量的标本实体压缩为数据资料, 使得处理过程可逆可重复, 为研究交流及数据积累创造了极大的便利^[5]。

在新方法与传统方法的测量兼容性问题, 有研究者利用带有金属标点的标本, 比较得到二维螺旋 CT 影像上线性测量与实际测量的差异范围在 0.1-3.0%^[6]; 与其他传统方法相比, 螺旋 CT (包括 2D 与 3D) 标志点定点的变异度略高^[7-9], 但同时变异度随熟练程度的提升而减小^[9]; 而对于针对口腔医学所使用的锥形束 CT (CBCT), 有比较研究显示 2D-CT 图测得的结果往往小于标点实际测量值, 但这种差异只在测量颅底内部时显著^[10]; 另一项探究 CBCT 的 3D 测量精确度与可靠性的研究对 29 项已确定标志点的线性测量项进行了分时分人分序测量, 发现虽有 5 项测量的单因素方差分析显示结果有差异, 但所有测量项的测量误差均小于扫描像素大小, 且没有测量者内部误差及测量者间误差^[11]; 此外, 也有研究者使用考古出土的 29 例人类头骨的 26 个测量点与 27 项测量项, 进行了传统手工测量法与三维激光扫描建立模型再测量法的比较, 结果发现这两种测量方法分别标定的两套测点的吻合度较高, 两种测量手段所获数据的绝对差值和相对差异程度均较小, 因此认为三维激光扫描的模型复原法与传统手工测量法之间的差异是可以接受的^[12]。

通过反复比对, 这种将标本数字化而后测量的方法的稳定性与精确性得到了研究者的认可, 尤其是 CT 技术带来的巨大变革更极度扩展了人类学及相关研究的思路。1984 年 Conroy 等人尝试使用 CT 扫描了一个 30 万年前的古哺乳动物头骨, 发现在 CT 影像下研究者更容易将密度不同的胶结、化石与空腔分辨开^[13], 以此为基础, 他们将 CT 技术应用到了人类化石上^[14,15]。在化石研究中, CT 可以帮助研究者排除胶结物的干扰, “穿透”和“放大”所研究的标本, 而高精度 CT 的诞生结合重建软件, 研究者得以在不破坏标本的情况下直接对内部结构与细微结构进行研究^[16,17]。基于以上优点, 国内外学者在人类学领域广泛使用 CT 扫描的方法进行研究, 在脑演化^[18]、创伤^[19]、头骨的内部结构^[20,21]、牙齿的内部结构^[22,23]等方面都做出了诸多成果。

在 CT 扫描获取了足够的信息后, 就需要三维模型复原技术来将数字信息可视化而在计算机中重建标本模型, 常用的软件有 Mimics, Avizo, VGStudio 等, 在此基础上进行进一步测量。然而在由标本实物到获得标本模型这个过程中, 模型重建方式的差异也可能导致最终模型与原标本的差异。以 Mimics 软件为例, 软件提供了多种计算模型的预设以满足不同的精度或文件体积需

求。根据矩阵简化 (matrix reduction) 程度不同, Mimics 的重建模型有低精度、中精度、高精度、最佳精度四种可选, 计算出的三维模型细节程度逐个提升。矩阵简化处理相当于将原本扫描分辨率下的单层合并为多层, 相当于定向降低了扫描精度。通过简化降低分辨率, 缩小模型的数字体积。对于不同精度的模型来说, 更高的精度模型细节更丰富, 模型文件的大小也会成倍增长。例如, 对同一原始现代人头骨的高精度 CT 数据, 最佳精度模型的可达 2GB 甚至以上, 而低精度模型大小往往不足 100MB; 计算机在计算最佳精度模型时耗费的时间也长于低精度模型。这给大体积标本的测量以及大样本量测量带来了麻烦。

在日常数据采集, 考虑到较大的样本量及高精度带来的巨大文件体积, 如果较低精度的模型就可以满足测量要求, 那么选择低精度计算重建人骨三维模型, 并在此基础上进行测量将适用于大批量数据采集。拓展开来, 扫描、复原精度的差异造成的测量差异是否显著, 也决定了不同精度的标本数据是否可以互相比较。

为了探究复原精度不同是否可能造成测量结果的显著差异, 以及这种差异在哪些测量项目中有所体现, 我们需要一系列有代表性的测量对比。此前有运动学研究者比较了医用 CT 获取的跗骨及周围软组织的运动学数据, 指出除了会改变 3D 模型体积和表面积外, 模拟精度、阈值选取与复原方式的变化并不会影响跗骨运动学指标 (以线性长度及角度为主) 的测量^[24], 为我们选择研究材料、测量指标提供了思路。考虑到 CT 分辨率、测量项目类型等的差异, 我们选择比较同一头骨标本进行不同精度的重建后测量得到的周弧线长、面积、体积等, 探究多种类别的测量项在 CT 重建中低精度复原模型测量的可行性。

2 研究材料与方法

2.1 研究材料

头骨测量中主要包括了长度测量、角度测量、面积测量、体积测量等^[2,24]。考虑到模型的差异主要体现在模型的表面变化, 我们选择了顶骨矢状弦长、颅周长、头盖部面积、乳突小房表面积、颅容量、乳突小房体积 6 项测量指标作为检测性状, 分别代表直线测量, 弧线测量, 较大单位的面积测量, 较小单位的面积测量, 较大单位的体积测量, 和较小单位的体积测量。

本文的研究材料为在云南墓葬考古遗址出土的距今约 300 年左右的现代人头骨 43 例。由于头骨保存状况限制, 部分标本不能进行全部六个测量项目。由于此前研究发现乳突小房的体积、表面积的侧别差异不显著^[26], 因此进行乳突小房双侧混合统计。具体各项使用标本数量见表 1。

标本保存在中国科学院古脊椎动物与古人类研究所 (以下简称古脊椎所)。

2.2 头骨 CT 扫描及模型重建

表 1 测量项与对应标本数

Tab.1 Measurement item and the number of the specimens

测量项	对应标本数量n
顶骨矢状弦长	30
颅周长	30
头盖部面积	30
乳突小房表面积	60
颅容量	30
乳突小房体积	60

头骨 CT 数据的采集使用的是古脊椎所的高分辨率工业 CT 机 (450kV- 工业 CT), 扫描电压为 450kV, 空间分辨率为 $160\mu\text{m}$, 所得 CT 扫描的原始数据首先使用高能所研发中心开发的重建软件转换为二维影像后, 在 Dell 图形工作站上使用可视化三维图像处理软件 Mimics17.0 对二维断层扫描影像进行重建。

Mimics 通过使用矩阵简化算法, 为建模提供四种精度: 低精度、中精度、高精度、最佳精度。精度越高, 模型越精细, 但相应的文件体积和处理时间也会大幅增加。根据软件本身提供的参数, 我们知道低精度的简化为 xy 方向上 $6\times$ 分辨率, z 方向上 $2\times$ 分辨率, 具体到本研究就是在 xy 方向上分辨率简化为 $960\mu\text{m}$, 在 z 方向上简化为 $320\mu\text{m}$; 中精度为 xy 方向 $3\times$ ($480\mu\text{m}$), z 方向同样 $2\times$ ($320\mu\text{m}$); 高精度为 xy 方向 $2\times$ ($320\mu\text{m}$), z 方向 $1\times$, 即 z 方向不简化; 而最佳精度在 xyz 方向均不简化, 因此我们一般认为最佳精度复原模型为同一 CT 扫描处理后能得到的最接近真实的模型。为了对简化后的模型进行比较, 我们在同一阈值选取结果的基础上选择不同的计算精度导出不同精度的模型, 进行后续测量。

2.3 头骨模型测量

为了减少测量者在测量不同精度模型时造成的误差, 测量中将带四个不同精度模型一同原位导入逆向工程软件 Rapidform XOR 中, 进行测量

2.3.1 顶骨矢状弦长的测量

顶骨矢状弦长为前凶点至人字点的直线距离^[2]。通过在最佳精度模型上点出前凶点及人字点后, 过标志点作模型表面垂线, 按照投影的方式得到其他三个精度上对应的点位, 然后测量两点间的距离即可得到不同精度模型的顶骨矢状弦长。

2.3.2 颅周长的测量

颅周长为通过眉间点和颅后点的颅骨水平周长^[2]。测量时, 先在最佳精度模型上确定鼻根点、上齿槽前点及枕外隆凸点而做出头骨的正中矢状面后确定眉间点及颅后点, 通过这两点作与正中矢状面垂直的平面, 这个平面与不同精度头骨模型相交形成的外周长即为不同精度模型的颅周长 (图 1A 中黑色线条所示)。

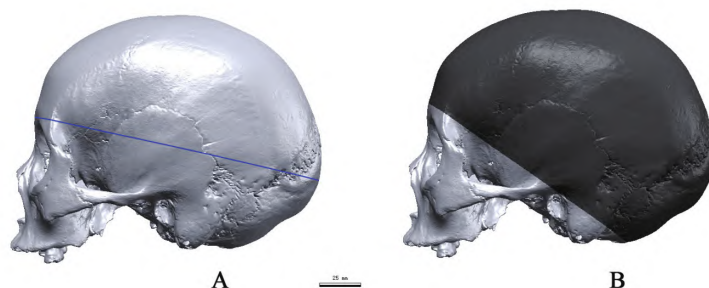


图 1 颅骨三维重建模型在逆向工程软件 Rapidform 中的颅周长 (A) 及头盖部面积 (B) 测量方法
Fig.1 The measuring method of cranial horizontal circumference (A) and surface area of cranium (B) of 3D reconstruction model of skull in software

2.3.2 头盖部面积的测量

根据参考文献^[25], 本研究将过眉间上点及左右耳门上点的平面所切割的头骨上部外表面定义为头盖部面积。同样在最佳精度下作出统一切割平面后由软件计算不同精度头骨模型上部面积, 在图 1B 中示为黑色。

2.3.3 乳突小房表面积、体积及颅容量的测量

通过阈值选取的方法选择出乳突小房、颅内模区域后重建获得三维模型。乳突小房的表面积、体积及颅容量由Rapidform软件直接计算复原乳突小房模型及颅内模的体积等得到。

2.4 数据统计

完成测量后, 将数据记录并导入到 SPSS 20.0 软件中。对于每个测量指标, 本研究以最佳精度复原模型的测量结果作为标准, 首先对四个精度组整体进行正态性检验及方差同质性检验, 根据实际数据选择参数(两两配对的 t 检验)或非参数检验(相关样本的 Wilcoxon 符号秩检验), 同时比较不同精度模型的测量差异占比。在这里测量差异占比 = (目标精度模型测量值 - 最佳精度模型测量值) / 最佳精度模型测量值。

3 结果

3.1 不同精度下测量值

根据确定的测量方法, 我们对同一标本采用不同计算精度重建模型的顶骨矢状弦长、颅周长、头盖部面积、颅容量、乳突小房表面积以及体积进行了测量(表 2)。

对于顶骨矢状弦长、颅周长、头盖部面积、颅容量, 其各精度的测量结果分布如图 2 所示, 整体差异不大。顶骨矢状弦长及头盖部面积测量值随精度变化不明显, 颅周长测量值随精度提升有一定提升, 颅容量测量值则随精度提升有所下降。

对于乳突小房的表面积和体积, 其各精度的测量结果分布如图 3 所示, 整体变化较大: 乳突小房表面积测量平均值从低精度到高精度有明显升高, 数据分布范围更分散, 但高精度与最佳精度的数据整体分散情况相近; 乳突小房体积测量平均值从低精度到最佳精度显著降低, 测量结果分布范围也缩小。

表 2 不同精度下测量结果 (均值 ± 标准差)
Tab.2 Measuring data in different qualities(average±SD)

	低精度	中精度	高精度	最佳精度
顶骨矢状弦长(mm)	110.52±6.25	110.25±6.32	110.19±6.34	110.09±6.33
颅周长(mm)	504.98±15.23	505.64±15.34	507.19±15.20	509.63±15.20
头盖部面积(cm ²)	564.30±32.44	564.73±32.08	566.73±31.54	567.77±31.92
乳突小房表面积(cm ²)	37.94±15.43	51.10±23.76	58.83±29.09	60.03±31.15
颅容量(cm ³)	1319.92±110.65	1303.99±109.80	1298.72±109.34	1293.11±109.22
乳突小房体积(cm ³)	5.68±3.13	4.60±2.72	4.11±2.49	3.61±2.25

3.2 不同精度间测量值的差异比较

顶骨矢状弦长、颅周长、头盖部面积、颅容量这四项测量值, 方差同质性检验 p 值分别为 0.999、0.997、1.000 和 1.000, 均大于 0.05, 即对于这四项测量值来说, 不同精度组之间具有方差同质性; 但乳突小房表面积及乳突小房体积的方差同质性检验 p 值分别为 0.000 和 0.056, 说明乳突小房表面积四个精度组之间方差不具有同质性, 但体积的四个精度组之间具有一定的方差同质性。因此我们考虑将没有进行简化的最佳精度复原测量结果作为基准值, 进行配对 t 检验 (满足方差同质性) 或相关样本的 Wilcoxon 符号秩检验 (不满足方差同质性), 来比较针对同一测量项时低精度、中精度、高精度与最佳精度模型之间的测量差异, 并据此分析对不同测量指标精度造成差异的影响。

以最佳精度复原测量结果作为基准值, 将其他精度复原测量结果与基准值进行配对 t 检验, 计算出低精度、中精度、高精度与最佳精度的检验 p 值, 结果见表 3。可知颅周长、头盖部面积、颅容量、乳突小房体积四项的所有简化精度模型的测量数据均与最佳精度模型测量数据的差异具有显著性; 而除高精度组外, 顶骨矢状弦长及乳突小房表面积的其余精度组测量数据均与最佳精度组差异具有显著性。这一结果在箱图 (图 2、图 3) 分布中也有所体现。

计算各测量项的测量差异占比, 统计得到表 4, 我们发现, 顶骨矢状弦长、颅周长、头盖骨、颅容量这四项测量差异占比小, 多在 2% 以下; 但乳突小房表面积的低精度测量差异占比绝对值最高可达 53.571%, 而体积的低精度测量差异占比绝对值可达 124.359%。

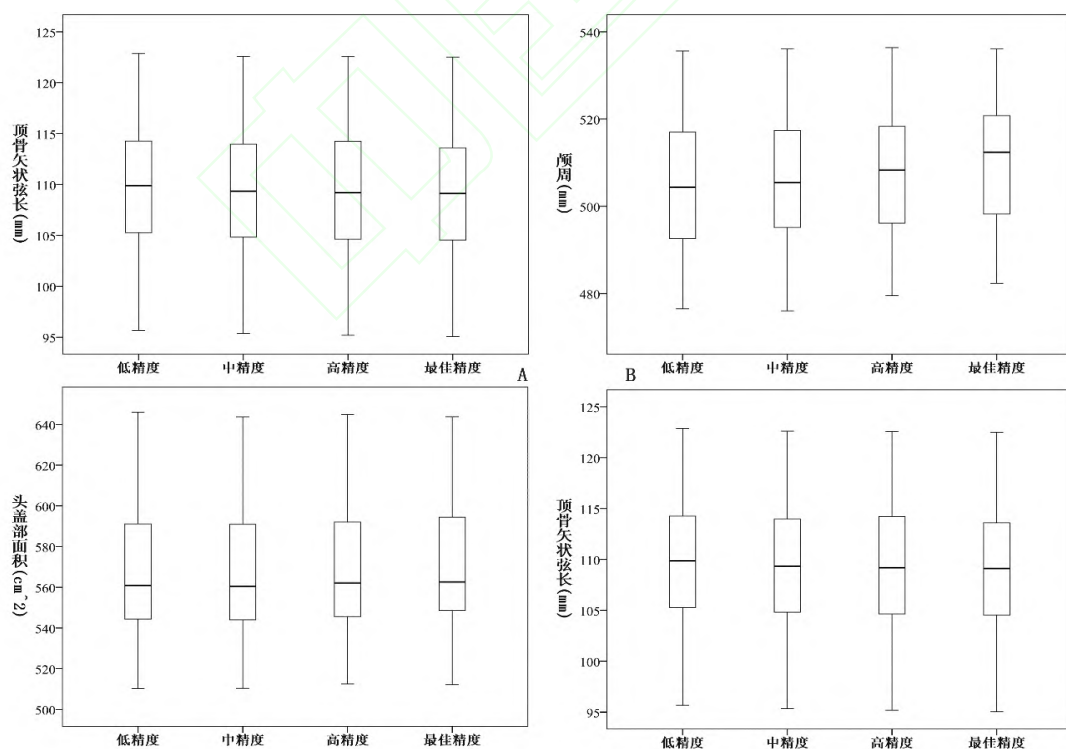


图 2 顶骨矢状弦长 (A)、颅周长 (B)、头盖部面积 (C)、颅容量 (D) 不同复原精度测量结果分布箱图
Fig. 2 Box-plot of cranial sagittal chord (A), cranial horizontal circumference (B), cranium area (C), cranial capacity (D) in different reconstructing qualities

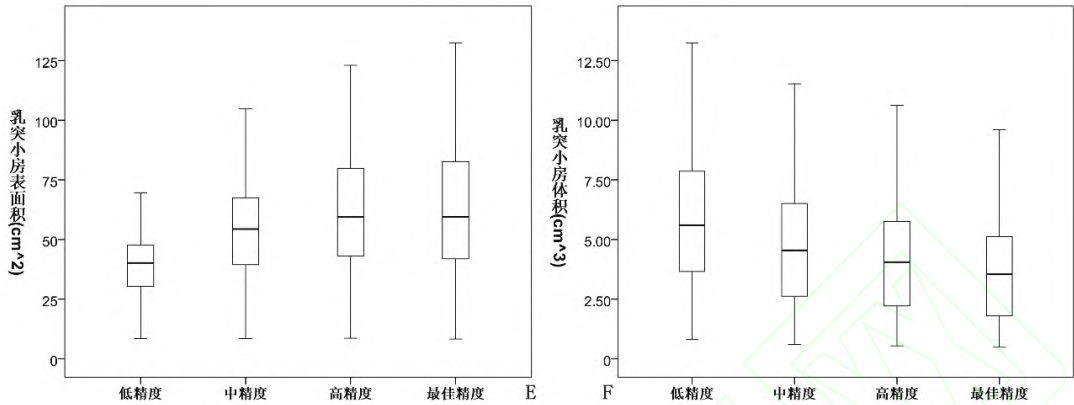


图 3 乳突小房表面积 (E) 及乳突小房体积 (F) 不同复原精度测量结果分布箱图

Fig. 3 Box-plot of surface area (E) and volume (F) of mastoid cells in different reconstructing qualities

表 3 低、中、高精度与最佳精度测量结果配对 t 检验或相关样本 Wilcoxon 符号秩检验 p 值

Tab.3 P-value of paring t-test or related samples Wilcoxon signed rank test between simplified and optimal reconstructing qualities

	低精度	中精度	高精度
顶骨矢状弦长	0.000*	0.000*	0.024
颅周长	0.000*	0.000*	0.000*
头盖部面积	0.000*	0.000*	0.002*
颅容量	0.000*	0.000*	0.000*
乳突小房表面积 [#]	0.000*	0.000*	0.324
乳突小房体积	0.000*	0.000*	0.000*

因不满足方差同质性，乳突小房表面积项列出结果为相关样本 Wilcoxon 符号秩检验 p 值

* 由于共有三组数据进行平行比较，进行 Bonferroni 校正后，p 值小于 0.0167 为显著

表 4 低、中、高精度与最佳精度测量结果的测量差异占比

Tab.4 Measurement difference proportion of different reconstructing qualities

	低精度		中精度		高精度	
	均值 标准差SD	最小值 最大值	均值 标准差SD	最小值 最大值	均值 标准差SD	最小值 最大值
顶骨矢状弦长	0.404% 0.400%	-0.285% 1.973%	0.151% 0.135%	-0.110% 0.535%	0.092% 0.192%	-0.042% 1.057%
颅周长	-0.911% 0.843%	-2.486% 0.268%	-0.783% 0.726%	-2.209% 0.336%	-0.479% 0.536%	-1.655% 0.187%
头盖部面积	-0.616% 0.820%	-2.671% 0.438%	-0.534% 0.463%	-1.781% 0.096%	-0.181% 0.289%	-1.133% 0.349%
颅容量	2.079% 0.085%	1.942% 2.317%	0.844% 0.039%	0.776% 0.953%	0.437% 0.055%	0.394% 0.706%
乳突小房表面积	-28.551% 19.880%	-53.571% 24.703%	-10.032% 12.934%	31.805% 17.313%	0.105% 6.365%	-11.577% 11.963%
乳突小房体积	65.815% 23.365%	31.450% 124.359%	29.694% 8.572%	15.669% 51.025%	14.872% 3.910%	8.163% 24.343%

4 讨论及结论

4.1 复原精度的差异对各测量项造成的影响

本研究中主要选择了最佳精度模型作为预设的基准值, 即可获取的相对接近真实值的测量值, 在此基础上, 我们比较得到的差异也可以认为是由软件输出不同精度所模拟的近似差异。我们的实验数据表明, 对于我们选择的顶骨矢状弦长、颅周长、头盖部面积、颅容量、乳突小房的表面积和体积的测量, Mimics 提供的简化复原精度与最佳复原测量精度结果的差异大都具有显著性。

差异不显著的仅有顶骨矢状弦长及乳突小房表面积的高精度组与最佳精度组, 表明高精度的简化在一些测量项上造成的组间差异小于了组内差异。不过这可能并不意味着差异显著的项目在未来测量中就可采用简化精度, 也并不意味着一些项目在差异不显著的精度上就可以随意选择, 仍然需要根据具体数值要求来判断。

为了尽可能将测量者造成的精度间差异降到最低, 除了软件自动计算的颅容量、乳突小房表面积和体积外, 其他各项在测量时均将四个精度模型做同位导入, 颅周长和头盖部面积的都由统一平面切割后由软件自动测量, 唯一可能产生测量者内部误差的测量项为顶骨矢状弦长。

为了排除测量者内部误差造成不同精度模型顶骨矢状弦长测量差异的可能性, 首次测量一段时间后测量者进行了再次测量, 得到表 5。

无论是第一次测量还是第二次测量, 得到的简化精度与最佳精度的平均测量差异占比都小于 1%, 低精度及中精度与最佳精度测量结果的配对 t 检验差异显著。两次测量在高精度与最佳精度的配对 t 检验结果上有差异: 第一次测量中得到顶骨矢状弦长高精度与最佳精度差异不显著, 第二次测量中发现高精度与最佳精度差异显著。

为了寻找造成统计结果不同的原因, 我们将第二次测量得到的结果与第一次测量结果进行配对 t 检验, 得到 p 值为 0.003, 说明两次测量差异显著。为了衡量测量时的变异性, 我们针对每个标本的顶骨矢状弦长测量结果计算标准差, 得到首次测量时同一标本不同精度测量结果的标准差平均为 0.226 ± 0.195 (0.024~1.018), 而再次测量时标准差平均为 0.273 ± 0.206 (0.060~1.978), 比较同精度的两批测量结果, 得到下表:

表 5 顶骨矢状弦长第二次测量结果及不同精度比较
Tab.5 The second measurement result of cranial sagittal chord and comparison among different reconstructing qualities

	测量结果(mm)	与最佳精度的测量差异占比		与最佳精度配对t检验p值
		均值±标准差	最小值-最大值	
低精度	110.88±6.10	0.55%±0.40%	0.12%~1.79%	0.000*
中精度	110.55±6.19	0.25%±0.29%	0.03%~1.46%	0.000*
高精度	110.38±6.26	0.09%±0.12%	-0.03%~0.46%	0.000*
最佳精度	110.28±6.27			

我们发现，对于顶骨矢状弦长的测量，同一标本同批测量不同精度的标准差是小于同精度两批次测量结果的差异的，也就是说不同精度造成的离散是小于标志点选取差异造成的离散的。同时，这种差异具有统一倾向性：随复原精度提升，顶骨矢状弦长、颅容量及乳突小房的体积逐渐减小，而颅周长、头盖部面积、乳突小房表面积逐渐增加，整体上差异绝对值越来越小。

这种差异的统一倾向性来自于精度不足造成的模型细节损失。对于顶骨矢状弦长，模型表面的标志点位置因低精度模型在骨缝处膨胀而产生垂直于表面的位移，导致最终测量值的增大；对于颅周长测量，颅骨表面精细变化的丢失造成了测量过程中的以直代曲，低精度测量值相对较高精度的测量值是缩小的；对于头盖部面积与乳突小房面积的测量，低精度情况下模型表面精细变化的丢失同样造成了测量值的缩小；而对颅容量、乳突小房体积这样的体积类测量项，低精度表面信息的丢失反而增加了模型的体积，使低精度下测量值增大。

对于顶骨矢状弦长、颅周长、头盖部面积和颅容量，虽然不同复原精度测量结果差异存在统计学上的显著性，但本批样本的测量差异占比的绝对值均在 3% 以下，即使用简化复原精度的模型测量以上四项性状得到的与最佳精度测量值差异小于 3%，对于一般的测量要求，这一差异是可以接受的。但对于乳突小房的表面积，其低精度与最佳精度间的差异可能达到 50% 以上，中精度造成的差异可达 30% 以上，而数据分布与最佳精度类似

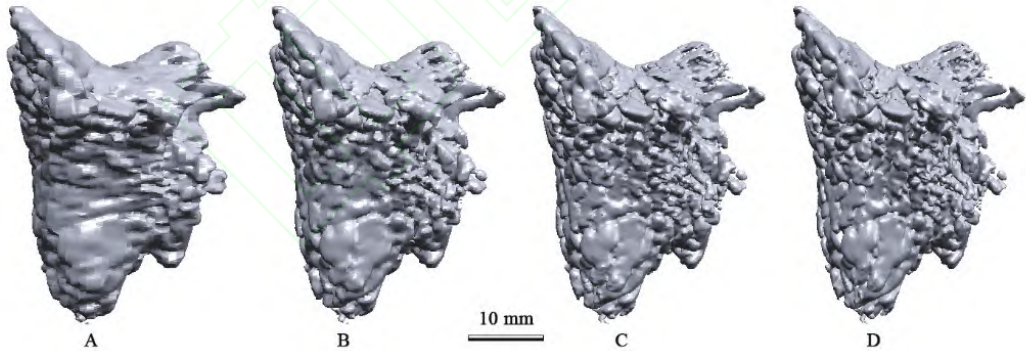


图 4 不同精度计算 3D 虚拟重建的同一乳突小房对比。

Fig.4 3D models in different qualities of the same mastoid air cell system

A: 低精度 B: 中精度 C: 高精度 D: 最佳精度 A: low quality; B: medium quality; C: high quality; D: optimal quality

表 6 两次顶骨矢状弦长测量中各精度测量结果标准差

Tab.6 Standard deviation of two measurement in different reconstructing qualities

低精度(mm)		中精度(mm)		高精度(mm)		最佳精度(mm)	
均值	最小值	均值	最小值	均值	最小值	均值	最小值
标准差SD	最大值	标准差SD	最大值	标准差SD	最大值	标准差SD	最大值
0.492	0.001	0.546	0.011	0.478	0.000	0.504	0.013
0.405	1.185	0.487	2.373	0.474	2.077	0.425	1.752

的高精度组其差异也可以达到 10% 以上; 对于乳突小房的体积, 其低精度与最佳精度间的差异可能达到 120% 以上, 中精度造成的差异可达 50% 以上, 高精度造成的差异可达 20% 以上, 这样的差异在实际过程中很难忽视。

4.2 不同复原精度对乳突小房表面积、体积测量的影响

乳突小房的表面积及体积测量不同于头盖部面积或是颅容量的测量。因为相对于头骨, 乳突小房体积小, 总面积小。对于 Mimics 的矩阵简化算法, 低精度的简化为在 x, y 方向上 $6\times$ 分辨率, z 方向上 $2\times$ 分辨率, 本研究使用的 CT 分辨率为 $160\mu\text{m}$, 因此低精度简化造成的基础误差为 $960\mu\text{m}$, 接近 1mm , 对于总长度及高度分别在 3.5cm 及 2.0cm 左右, 单一气腔可能小于 1mm 的乳突小房系统, 虽然绝对误差小, 但相比于整体体积的计算得到的相对误差(也就是测量差异占比)会远高于头盖部面积、颅容量等基础数值大的测量性状, 甚至造成倍数于最佳复原精度模型的测量结果。

乳突小房是颞骨中由众多小的气腔组合形成的一个完整系统, 这一系统表面变化剧烈, 因此我们认为复原精度变化对乳突小房的两个测量项造成的巨大差异是可以理解的。如图 4 所示, 不同复原精度得到的乳突小房模型差别明显: 在计算重建乳突小房这样的多泡结构时, 不同精度会影响到每个气腔的具体形态, 甚至会影响到气腔与气腔之间的连接表现。在低精度下, 小而独立的气腔在重建的模型中互相连接, 会极大地影响到表面积以及体积的测量。

根据表 3, 我们发现, 乳突小房的表面积及体积的测量差异占比具有相当大的标准差, 这说明对于乳突小房标本来说, 复原精度变化的影响在个体间差异较大。乳突小房内部结构的复杂性就在于, 其表面变化不仅复杂于头骨脑区内外表面, 且个体与个体之间乳突小房也可能存在很大差异, 具有极大的随机性。因此复原精度的变化对每个乳突小房的影响可能都是不同的。

4.3 实际研究中的测量模型的选择

除了考虑不同精度模型本身的机械误差, 我们在实际研究中也要考虑模型低精度带来的标志点位不明晰问题。根据本研究所选取的代表性性状的测量结果, 我们认为, 使用复原的模型来进行测量时, 依据标志点来进行测量的弦长可以在保证标志点准确的情况下根据精度要求选择不同的复原精度; 对于表面光滑, 不同精度模型差异小的区域, 复原精度造成的差异影响较小, 选择时主要考虑的是数据的精度要求; 而对于体量较大的测量项, 如颅骨表面的弧长测量、面积测量或是体积测量, 不同的复原精度造成的绝对差异在较大体量基础上不会造成测量结果的巨大差异; 因此, 在一般研究中可以使用低、中、高等简化复原精度模型进行测量, 得到的结果与使用大体积、长处理时间的最佳精度模型的测量结果差异较小。

而对于以乳突小房为代表的小体量、表面粗糙、个体变异大的头骨内部腔窦, 不同复原精度都可能造成测量结果的质的改变。因此为了获得乳突小房等更准确的数据, 复原精度越高越好, 同时数据的横向比较只可在相同简化精度间进行。在对这些性状进行测量与比较时, 需要尽量保证测量使用的 CT 参数、复原参数、测量人员及方法的前后一致性。

4.4 不足与展望

本研究中只讨论了单一情况下不同复原精度影响,此外,在扫描及复原过程中还可能存在着其他的一些影响因素,例如:由于 Mimics 的简化模式是 z 轴与 x, y 轴不等价的,因此在重建过程中, z 轴方向的规定不同可能会对简化精度模型的测量产生不同影响。

对于像乳突小房的这样总量小、个体变异大、易受影响的性状,除了复原软件(如 Mimics)中复原精度的选择可能会影响到 CT 数据重建模型的测量,其他条件也可能影响测量结果^[27]:使用的 CT 的分辨率会对重建精度造成基础性的影响;复原过程中原则上边缘阈值的选择遵循最大半值法(half-maximum height protocol, HMH),但由于复原材料的状况不同,以半阈值复原出的模型往往并不理想,因此实际操作中微调十分常见,微调的大小也会影响到模型的重建结果^[28];复原软件对选取区域的计算方式的差异以及光滑预处理也可能造成测量结果偏差^[29]。可能的影响因素非常多,而现在我们缺少对这些因素造成偏差的综合控制分析,在未来的研究中如果可以得到系统的差异比较,那将对标本数据的共享与处理起到更多帮助。

致谢:感谢中国科学院古脊椎动物与古人类研究所的侯叶茂、潘雷和邢松在 CT 扫描、三维虚拟复原和性状测量等方面的指导和帮助。

参考文献

- [1] 刘武, 杨茂有. 现代中国人颅骨测量特征及其地区性差异的初步研究 [J]. 人类学学报, 1991, 10(2): 96-106.
- [2] 邵象清. 人体测量手册 [M]. 上海辞书出版社, 上海, 1985: 57-110
- [3] 席焕久, 陈昭. 人体测量方法 [M]. 科学出版社, 北京, 2010
- [4] Holloway R L. New Endocranial Values for the East African Early Hominids[J]. Nature, 1973, 243(5402): 97-99.
- [5] Tattersall I. Handbook of paleoanthropology[M]. Springer Berlin Heidelberg, 2007: 787-789.
- [6] Waitzman A A, Posnick J C, Armstrong D C, et al. Craniofacial skeletal measurements based on computed tomography: Part II. Normal values and growth trends[J]. The Cleft Palate-Craniofacial Journal, 1992, 29(2): 118-128.
- [7] Kragstov J, Bosch C, Gyldensted C, et al. Comparison of the reliability of craniofacial anatomic landmarks based on cephalometric radiographs and three-dimensional CT scans[J]. The Cleft palate-craniofacial journal, 1997, 34(2): 111-116.
- [8] Williams F L E, Richtsmeier J T. Comparison of mandibular landmarks from computed tomography and 3D digitizer data[J]. Clinical Anatomy: The Official Journal of the American Association of Clinical Anatomists and the British Association of Clinical Anatomists, 2003, 16(6): 494-500.
- [9] Valeri C J, Cole III T M, Lele S, et al. Capturing data from three - dimensional surfaces using fuzzy landmarks[J]. American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists, 1998, 107(1): 113-124.
- [10] Lascaia C A, Panella J, Marques M M. Analysis of the accuracy of linear measurements obtained by cone beam computed tomography (CBCT-NewTom) [J]. Dentomaxillofacial Radiology, 2004, 33(5): 291-294.
- [11] Berco M, Rigali Jr P H, Miner R M, et al. Accuracy and reliability of linear cephalometric measurements from cone-beam computed tomography scans of a dry human skull[J]. American Journal of Orthodontics and Dentofacial Orthopedics, 2009, 136(1): 17. e1-17. e9.
- [12] 惠家明, 贺乐天, 王明辉. 基于三维激光扫描的颅骨测量与手工测量的比较 [J]. 人类学学报, 2019, 38(2):254-264.
- [13] Hounsfield G N. Computerized transverse axial scanning (tomography): Part I. Description of system[J]. British Journal of Radiology, 1973, 46(552):1016.
- [14] Conroy G C, Vannier M W. Noninvasive three-dimensional computer imaging of matrix-filled fossil skulls by high-resolution computed tomography[J]. Science, 1984, 226(4673):456-458.
- [15] Conroy G C, Vannier M W, Tobias P V. Endocranial features of Australopithecus africanus revealed by 2- and 3-D computed tomography[J]. Science, 1990, 247(4944):838-841.

- [16] Conroy, G. C. Endocranial Capacity in an Early Hominid Cranium from Sterkfontein, South Africa[J]. *Science*, 1998, 280(5370):1730-1731.
- [17] Spoor F, Hublin J J, Braun M, et al. The bony labyrinth of Neanderthals[J]. *Journal of Human Evolution*, 2003, 44(2):141-165.
- [18] Gilmor R L, Yune H Y, Holden R W. Computed Tomography of the Temporal Bone[J]. *Acta Oto-Laryngologica*, 1984, 103(sup434):1-31.
- [19] Falk D, Clarke R. Brief communication: New reconstruction of the Taung endocast[J]. *American Journal of Physical Anthropology*, 2007, 134(4):529-34.
- [20] Wu X J, Schepartz L A, Liu W, et al. Antemortem trauma and survival in the late Middle Pleistocene human cranium from Maba, South China[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2011, 108(49):19558.
- [21] Balzeau A, Grimaudhervé D. Cranial base morphology and temporal bone pneumatization in Asian Homo erectus. [J]. *Journal of Human Evolution*, 2006, 51(4):350-359.
- [22] Wu X J, Crevecoeur I, Liu W, et al. Temporal labyrinths of eastern Eurasian Pleistocene humans[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2014, 111(29):10509-13.
- [23] Liu W, Jin C Z, Zhang Y Q, et al. Human remains from Zhirendong, South China, and modern human emergence in East Asia[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2010, 107(45):19201.
- [24] Plotino G, Grande N M, Pecci R, et al. Three-dimensional imaging using microcomputed tomography for studying tooth macromorphology[J]. *Journal of the American Dental Association*, 2006, 137(11):1555-1561.
- [25] 潘雷, 魏东, 吴秀杰. 现代人颅骨头面部表面积的纬度分布特点及其与温度的关系 [J]. *中国科学: 地球科学*, 2014, 44(8):1844-1853.
- [26] 张玄, 吴秀杰. 颞骨乳突小房的 3D 虚拟复原及形态变异——以现代云南人为例 [J]. *第四纪研究*, 2017, 37(4): 747-753
- [27] Byun S W, Lee S S, Jin Y P, et al. Normal Mastoid Air Cell System Geometry: Has Surface Area Been Overestimated? [J]. *Clinical & Experimental Otorhinolaryngology*, 2016, 9(1): 27-32.
- [28] Coleman M N, Colbert M W. CT thresholding protocols for taking measurements on three - dimensional models[J]. *American Journal of Physical Anthropology*, 2007, 133(1): 723-725.
- [29] Spradley J P, Pampush J D, Morse P E, et al. Smooth operator: The effects of different 3D mesh retriangulation protocols on the computation of Dirichlet normal energy[J]. *American Journal of Physical Anthropology*, 2017, 163(1): 94-109.