

Automated skeletal maturity assessment based on X-ray images: progress and challenges

CHEN Yi-lin, JIANG Tian-zi, LI Shu-yu, SHAO Wei-dong, ZHANG Shao-yan, LIU Li-juan

(National Laboratory of Pattern Recognition, Institute of Automation Chinese Academy of Sciences, Beijing 100080, China)

[Abstract] We summarize the development of the automated skeletal maturity assessment over the past decade and challenges. Then we analyze the main difficulties in implementing an automated skeletal maturity assessment system, and put forward some possible solutions to them.

[Key words] Bone age; Automated assessment system; Epiphysis; Metaphysis

基于 X 线影像的骨发育成熟度自动评价系统:研究进展和挑战

陈亿霖¹, 蒋田仔¹, 李淑宇¹, 邵伟东², 张绍岩², 刘丽娟²

(1. 中国科学院自动化研究所模式识别国家重点实验室, 北京 100080; 2. 河北省体育科学研究所)

[摘要] 本文综述近十年以来, 骨发育成熟度(骨龄)自动评价系统的国内外发展进程和现状, 分析了该系统在实现过程中所面临的挑战, 并提出一些可能的对策。

[关键词] 骨龄; 自动评价系统; 骨骺; 干骨后端

[中图分类号] R68; R814.41 **[文献标识码]** A **[文章编号]** 1003-3289(2004)01-0143-04

1 前言

近年来, 骨发育成熟度(骨龄)指标在预防医学、临床医学和体育科学等领域得到广泛的应用。国内外普遍认为, 按骨龄判定发育个体的成熟程度比较有效, 可以较全面地反映机体生理状态, 在很大程度上代表着生物学发育年龄。现今人们多采用拍摄手腕骨 X 线片的方法进行骨骼成熟测量和评价, 概括起来, 国内外判定骨龄的方法主要有 2 种: ①图谱法: 1937 年由 Todd 首创, 后经 Greulich 和 Pyle 等人的修改, 研制出手腕部骨骼成熟系列性 X 线图谱, 在国际上具有较高的权威性。图谱法判定骨龄简便、直观, 易于被人掌握, 因此目前各国应用较广泛。②计分法: Tanner 等人以英国普通家庭经济条件儿童少年为对象, 研

制出判定骨龄的计分方法(TW1, 1962), 后于 1972 年修改成 TW2 标准, 又于 2000 年进一步修改成为 TW3 标准。为了预测成年身高、使用的简便和简化评价标准, Tanner 等人还建立了以桡尺骨及掌指骨为基础的 RUS 骨龄计分法。计分法具有明确的量化概念, 判定结果较精确。但此方法比较繁琐, 工作人员要经过专门的训练方可进行有效判定。

目前在国际上沿用的 TW2 记分法和 GP 图谱法这两种骨龄评价方法, 存在着相同的缺点: 一是重复性差, 即同一评价人对同一受试者的多次评价以及不同评价人对同一受试者的评价结果存在较大差别; 二是难以普及应用, 评价人需经过长期的培训与经验积累, 才能掌握与使用。

由于儿童的生长发育受多种因素的影响, 上述两种方法制定的标准与我国儿童的骨龄发育速度不一致。1979 年, 我国的李果珍等人发表了对评定中国人骨龄更为合理与适用的《百分计数法》。为了实施中国汉族城市儿童运动员的科学选材, 国家体委于 1988 年修改了 TW2 法, 并于 1992 年制定了适合国情的中国

人骨发育标准 CHN 计分法。但此方法也同样存在问题, 因 CHN 法需评定 14 块骨, 相对照的标准片多达 81 张, 查得分表非常繁琐, 且 X 片上的有些标志点不易辨认, 所以骨龄的定量评价在临床上不易普及, 需要专家才能评定。

因此, 如何找出新的评价方法, 尤其是实现评价的自动化, 使得临床骨龄评定达到简单、快速、准确、可重复性强的要求, 也就成为了国内外众多研究机构的研究重点。

2 国内外的发展和现状

2.1 国内近年来的发展与现状 CHN 法是现在我国统一的骨龄评定标准。参与评定的骨块一共有 14 块, 分别是桡骨、掌骨 I、掌骨 III、掌骨 V, 近节指骨 I、近节指骨 III、近节指骨 V、中节指骨 III、中节指骨 V, 远节指骨 I、远节指骨 III、远节指骨 V, 头状骨和钩骨。

由于 CHN 法选用 14 块手骨参与骨龄评定, 对于自动评价系统来说, 准确的分割出这些骨块, 并给出精确的等级描述, 难度很大。1994 年, 湖北省体育科学研究所能否减少判断骨块方面做出了一些有益的尝试^[1]。他们人工的对样本

[基金项目] 本文受国家自然科学基金(60172056 和 697908001), 国家体育总局备战 2008 年奥运会科技攻关项目(02046), 及中国科学院“引进海外杰出人才”基金资助。

[作者简介] 陈亿霖(1978—), 男, 陕西西安人, 在读硕士。E-mail: ylchen@nlpr.ia.ac.cn

[收稿日期] 2003-09-24

骨块判定出等级,并对这些骨块的等级和实际骨龄进行线性回归,根据得出来的曲线方程,得出以下结论:桡骨、掌骨Ⅰ、掌骨Ⅲ、近节指骨Ⅲ、中节指骨Ⅲ、远节指骨Ⅲ、头状骨、钩骨这8块骨块可以替代14块骨块进行骨龄评价,且误差很小,在0.37岁以内。

1995年,武汉测绘科技大学在此基础上进一步尝试,希望完全突破CHN法的框架,仅根据判断骨块的周长、面积和形状等信息对骨龄做出判断^[2]。借助大量的专家经验并进行了多次的试验后,最终选择了如下几种几何特征量:①桡骨:骨骺的周长、面积;②掌骨Ⅰ,掌骨Ⅲ,近节指骨Ⅲ,中节指骨Ⅲ,远节指骨Ⅲ:骨骺的宽度与长度比;③头状骨,钩骨:骨块的周长、面积、形状系数。通过判定,最终结果与使用传统的CHN法所做出的判断相差较小,从理论上和方法上均对传统的骨龄评定框架做出了突破。这项研究在建立骨龄评定的统计模型,为实现骨龄评定的计算机化,和为专家知识和经验的量化铺平了道路。

1998年第四军医大学^[3]及1999年上海第二医科大学^[4]分别独立研制出自己的计算机辅助评价系统。通过高清晰度扫描仪将X线图片扫描到计算机里,然后将标准样片自动显示扫描X线片旁边,医生就能很方便将其与标准的X线样片进行比较,判断发育等级,最后计算机根据医生评价的等级,自动查表给出总分。但是这些工作并没有完成自动评价,仅仅是提供了自动查表和计算的功能,避免了繁琐的翻阅手册的过程、因为疲劳而引起的计算错误以及计算查表的漫长时间,提高了骨龄计算的速度和准确性。

2000年,第四军医大学在1998年开发的辅助系统的基础上,进一步增强功能,但是本质依然是辅助系统^[5]。该系统的处理原理是,先对某些不清晰的图像进行伪彩色增强;然后对已分割好的骨块统计样本骨块的周长、面积、熵值、骨骺与骨干的横径之比、骨骺的长宽比、中心矩等信息,储存进知识库中。再对图像进行二值化,分割出各个骨块,对分割出来的骨块分别计算以上的参数,通过与知识库中的信息比较相似度得出骨块的发育等级。最后,根据CHN方法计

算得分,最终得到骨龄值。但是由于该系统是为儿童的牙齿进行诊断,所以更多考虑小年龄的儿童特点,该系统的应用范围主要为7、8岁以下年龄的儿童。

2002年上海大学也开发了类似于上海第二医科大学的辅助系统,并且增强了对图像进行快速旋转和通过骨块之间的位置关系来辅助定位等功能^[6]。第一点,对位图的快速旋转。大多数的采样图像不是十分规范,片中的手骨与片子的底边并不垂直。为克服这一问题,他们提供了一种快速的旋转算法,能使程序自动修正手骨的角度,使其处于图片的正中。第二点,骨块的定位。每个人的手掌骨都是不一样,因此不可能通过绝对位置来确定每块手骨的位置。他们采用动态的方法定位,因为每个人的手骨的比例是基本相似的,那么通过对大样片的统计,就可以获得了每一块手骨相对于其他骨块的相对位置。

2001年台湾国立清华大学^[7]开发出了自动评价系统,相对于大量的辅助判读系统来说,已经有了不小的突破。他们采用的基本原理如下:第一步,对原图像中值滤波,然后使用Sobel算子,再在得到的边缘图像中去除孤立点。第二步,寻找cROI(感兴趣的腕骨区域)。先基本定位出所有骨块的粗略位置。再根据腕骨各骨块之间的灰度直方图信息,得出腕骨各骨块区域之间粗略分割线。第三步,骨块提取。在原图像上进行阈值化,通过图像的灰度信息得到实际的骨块的边缘。第四步,特征提取。对分割出来的每个骨块计算面积、周长和长宽比,跟知识库中已有的统计信息进行比较判别,就可以剔除不需要的骨块。最后,骨龄计算。分别通过最小距离、贝叶斯分类器和神经网络三种方法进行分类,均取得了不错的效果。

该系统的缺点是由于在分割骨块的时候使用种子点向外扩张的方法,需要各腕骨骨块之间不能存在融合,否则,算法将失败,所以该系统只能用于7、8岁以下的小年龄段儿童。

2.2 国外近年来的发展与现状 1992年,TW记分法的作者,英国的Tanner报告了一种依据手腕骨X线片来评价TW-RUS骨成熟度得分的计算机辅助方法。但该方法需操作者将视频摄像机依

次移动到每块骨骺片的大概位置上方,然后在计算机屏幕上观察骨骺片图像,给出得分。虽然采用数学模型得到了连续的骨发育等级,但是仍需拍摄X线片和人工操作而达不到自动化的要求。

由于面临对腕骨骨块和指骨骨骺片的分割、提取并选取利于快速简便判读的骨块特征,以及由于受到法律限制而造成样本的不足等等的困难,国外现在真正对系统进行完整的设计和实现的,只有加州大学圣迭戈分校和加州儿童医院的由Pietka和Huang领导下的一个小组,以下,我们将主要介绍这个小组的工作^[8-10]。他们一共取得560张样片,分为黑人男女孩和白人男女孩四组,对0~9岁的儿童,每年龄段每组采集5例,对10~18岁的儿童每年龄段每组采10例,图片大小为 $2000 \times 2000 \times 10$ bytes。他们的实现过程如下:

第一步,预处理:①图像的统一定位。②去除背景。③对图像进行各项异性滤波。

第二步,骨块的提取。

(1)对指骨的分割。此处主要指的是中指指骨的分割,主要采用了两种方法来进行:

方法一:①根据确定出来的指尖,以及图像中手腕最下部分的中点,可以得到整个图像中手腕骨的垂直方向的中轴线。②在原图像上使用Sobel算子,得到原图像的梯度图像,并根据该图像的直方图进行阈值化。③沿中指的中轴线计算每一个垂直于中轴线的行的灰度均值,然后在此基础上计算灰度均值的一阶导数,其中导数变化最大的几处的极小值,分别对应着近节指骨,中节指骨和远节指骨的骨缝位置,如果在奇异点处存在有几个极小值,那么最靠近指尖部分的那个极小值就被认为是骨缝所在的位置。

方法二:在图像分割中,由于使用Active Contour Model(Snake)^[11]可以十分准确分割图像,故此得到大量应用,但由于它受初始值的影响很大,对于分割精度要求很高的骨骺片来说,初始值稍微有偏差,就会引起分割的不准确。而另一种方法Active Shape Model由于使用了统计信息,并根据统计信息来设定初始值,所以对于初始值不是很敏感,可

分割精度又不够。于是有人做了大胆尝试,将 Snake 和 Active Shape Model 结合起来使用,开始时使用 Active Shape Model 进行变换分割,然后将某个中间结果作为 Snake 的初始值,最后通过 Snake 的迭代变换,分割出期望得到的骨块,这取得了不错的效果,可以应用于设计实际的系统中。

(2)对腕骨的分割。首先根据生理信息分割出腕骨部分的粗略区域,然后使用一个滑动窗口(一般为 7×7 或是 33×33),从图像的上沿开始逐行进行局部阈值化,得到一个二值图像。再在这个图像上划分出来一个较精确的腕骨区域,从刚才分割出的区域边界开始,沿着骨块向内部进行腐蚀,最后对剩余的骨块分别计算参数,并将参数与已经统计得出来的参数进行比较,如果相差太多,则为虚假骨块,将其剔除,最后得到真实腕部的七块骨块。

第三步,特征提取。由于现在国外受到法律限制,所以不能得到大量的样本进行试验和验证。所以 Pietka 和其他很多小组,分别根据已有的样本提出了可能成为自动评价系统使用的几个特征:①对中指部分的近节、中节和远节指骨的长度进行统计;②对近节指骨/中节指骨和远节指骨/中节指骨的比例进行统计;③对(骨骺宽度/干骨宽度)的比值进行统计;④对(骨骺宽度/骨骺与干骨之间的距离)的比值进行统计。

此外,他们小组还做了一些有益的尝试。比如,对桡骨骨骺进行小波变化,看其 VHFC 分量占总能量的百分比,VHFC 指的是垂直方向上的高频分量。通过上述的百分比可以大致看出骨骺的等级是处于完全没有与干骨融合(等级 1~3)、开始融合(等级 4~6)、基本融合(等级 7~8)、完全融合(等级 9~10)四个阶段,且精度比较高,误差不会超过一岁。

最后,他们的小组还做出一个完善的网上交互系统,这样全美国的医院都可以通过这个网上系统跟他们进行交互,一方面可以不断的扩充他们的图片库,另一方面也可以对他们的系统进行测试和做出修正,具有很大的实际意义。

除了 Pietka 的小组以外,还有很多的小组也积极的探寻可能的特征和适用

于自动评价的技术。其中,利兹大学的 Efford 教授在 1994 年发表了一个技术报告^[12],并在以后的时间对这个报告进行不断的补充修改。他提出了以下技术路线:

(1)提取手的轮廓。对于在 X 片存在标签的情况下,通过求取 X 片中每个完整图像的面积,取面积最大的那个作为手骨图像。

(2)根据图像确定手的中轴线。将图像往水平和垂直两个方向上分别投影,并且画出直方图,取直方图上最高的那个点所在的行(或列)作为水平和垂直方向上的中轴线。

(3)提出运用曲率的概念来分析手的轮廓曲线,以及描述指骨部分的等级。根据不同年龄阶段儿童骨块的标准样片画出它的曲率变化,可以发现在曲率变化很大的地方,也即奇异点,均对应着不同阶段的形态的不同,也即不同的等级。并且对于手形区域来说,曲率变化很大的地方就是手指尖或两个相邻手指相连的部分,可以作为 Landmark 进行使用,为手指的中轴线以及其他骨块的定位提供信息。

(4)使用 Landmark 对其他一些生理特性对骨块的位置做出限制。从手骨的解剖结构和生理意义可以得到,从小指的中轴线的延长线基本上可以确定出来头骨的大概位置,而通过中指、无名指和小指三者的中轴线延长线就可以确定头骨和钩骨的大概位置。

(5)提出了一些等级的计算机化的描述方法。例如对于桡骨骺的第三等级来说,它的语言描述为骨化中心为楔形,桡侧端厚而圆,尺侧端薄而尖,而用数学描述则可以定义为骨块的重心位置在几何中心的右下方,也即如果将几何中心看成是原点,桡侧端为 x 轴的正方向,那么重心就位于第四象限。

(6)对 Snake 方法进行改进,使用 Rigidly-adjusted Template 方法。也即在 Snake 的变形过程中,每一次迭代计算后,每一个 Landmark 的点都向其法线方向移动,而不是原来的进行重新量化后,根据所受到的力的方向进行变形。

还有很多小组都在致力于研究开发自动评价系统,但基本上都是按照计算机辅助设计系统在设计,人工标记出骨

块的区域,再进行训练和分类。例如,德国 Heidelberg 大学提出骨块的发展等级可以按照骨骺和干骨之间的距离和各自大小来区分,但需要大样本训练和先验知识。现阶段他们根据已有的数据进行训练,可以达到允许误差为 1 岁时,准确率为 93%。

3 面临的挑战

国外近年来对研发骨龄自动评价系统的热情日益升温,但是由于所面临的困难很大,所以发展的速度并不是很快,主要的挑战有以下几点:

一是图像定位。因为影响骨骼发育的主要是指骨、指骨之间及掌骨间的骨骺,因此判断被测人的骨龄主要是基于对手掌上的特定的骨骼的分析和判断,所以在图像预处理后的首要任务就是确定要分析的手骨的准确位置。如不能将图像统一定位于同一坐标系下,那么对同年龄段的儿童的不同部分骨骺等级和几何信息将无法准确统计,从而导致今后的工作困难。

二是骨块的分割。由于儿童在 7、8 岁左右会出现骨骺与干骨融合的现象,对于低年龄段的儿童(指 7、8 岁以下年龄段的儿童)相对来说,腕骨的各个骨块比较清晰,利于分割,可以使用 Snake、Active Shape Model 以及 Level Set 等方法。在此可以考虑沿用国外将 Deformable Model 和 Active Shape Model 结合起来使用进行分割的思路,预计可以取得不错的效果。同时也可以对指骨进行分割处理,作为腕骨骨龄得有益补充,将无名指和食指的骨骺也加入进来进行考虑,精度也可以得到进一步的提高。

对于大年龄段的儿童(指 7、8 岁以上年龄段的儿童)而言,腕骨和指骨已经出现关节面,即骨骺与干骨后端已经开始融合,相邻骨块凹凸不齐,在图像上显示为钙化线相互交叉,不能反映出骨块边缘的真实情况。这也是现今国内外普遍在做低年龄段儿童的系统,而高年龄段儿童的自动评价系统迟迟没有进展的重要原因。在此可以考虑将全局的灰度信息综合考虑,国外已有类似的结果,可精度有待提高,但是这也提供了一个很新颖的思路。

三是手腕骨生物特征的提取和选择,包括手腕部各骨块的形态特征以及

钙化特征等。这些特征必须是以专家知识为指导的、能够描述青少年骨骼生长发育过程各阶段本质的特征。而且由于以往的方法均为通过语言描述来表达一个等级,那么如何将这些语言描述进行计算机化的表达也就成为了一个难点,通过计算机化的表达不但要保证等级的连续性,和描述的准确性,还不能出现等级描述的交叉、跳跃和丢失。那么如何寻找更为有效的,便于计算机评价的特征,即用有限的特征尽可能准确的描述少儿骨骼生长发育的连续过程,同样成为了很大的难点,并且也是现在人们普遍关注和重点研究一个方向。

4 总结与展望

由于骨龄评价在预防和临床医学、体育科学和司法等领域有着广泛的意义,但由于以往的方法是根据图谱判读和根据标准样片等级进行打分,受到的主观影响很大,如何避免这种主观性,尽量减少这种误差,使得整个评价的过程和结果更加客观、真实,也就成为了国内外同行的焦点,近年,已经有大量的成果技术出现,本文仅就这些成果做一介绍和总结。

以往,在做骨龄评价的时候,由于专家已经具备很多的先验知识,那么在进行等级评价的时候并不需要对骨块做出准确的分割,但是如果要想实现自动评价的时候,不同的等级描述有着准确的定义,在将以往等级描述翻译成计算机化的等级表达以后,骨块的准确提取也就成为了最大的难点,如果没有准确的骨块提取,那么在下面的工作中,很容易造成等级评价的不准确,或是等级交叉等,从而影响整个系统的精度。

由于以往的方法均为直观的观测,很多描述并不能准确的用计算机来表达,而且由于使用的特征过多,也造成了自动评价时的困难。所以现在人们普遍希望自动评价突破了以往方法的框架,

试图寻找其他有用的特征,例如,骨块众多的几何特性,那么如何寻找简单、有效而又准确的特征也就成为了人们的另一个热点。

以上两个困难也是现今自动评价所面临最大的两个挑战,通过这些年来研究,人们已经找到了很多比较可靠的规律和方法,将这些规律和方法进一步准确化和细化,也就将成为人们下一步的工作,突破了这两个难点,自动评价系统的开发成功也就指日可待了。

[参考文献]

[1] Cai R, Shen BH. Research on simplified CHN bone age assessment methods[J]. Journal of Hubei Sports Science, 1994, 2(1): 6-8.
蔡睿,沈邦华. 简化CHN骨龄评定方法的研究[J]. 湖北体育科技科学选材, 1994, 2(1): 6-8.

[2] Huang YC, Liu YG, Xu YM, et al. Establishment of bone age assessment statistic model[J]. Journal of Wuhan Mapping Science and Technology University, 1995, 20(3): 246-250.
黄幼才,刘友光,徐亚明,等. 骨龄评定统计模型的建立[J]. 武汉测绘科技大学学报, 1995, 20(3): 246-250.

[3] Luo JY, Lu RX, Lin Z, et al. Development of the computer-assisted bone age assessment system[J]. J Fourth Mil Med Univ, 1998, 19(1): 115-116.
罗家燕,卢瑞祥,林珠,等. 计算机化骨龄评价系统的研制[J]. 第四军医大学学报, 1998, 19(1): 115-116.

[4] Huang SX, Zhang YQ. Comparison research of several major bone age assessment methods[J]. Fujian Medical Journal, 2002, 24(1): 77-78.
黄松雄,张玉琴. 几种主要骨龄测定法的对照研究[J]. 福建医药杂志, 2002, 24

(1): 77-78.

[5] Luo JY, Lin Z, Lu RX, et al. A computer system for bone age measurement of hand radiographs[J]. Chinese Journal of Stomatology, 1998, 33(6): 378-380.
罗家燕,卢瑞祥,林珠,等. 手腕骨骨龄测量的计算机系统[J]. 中华口腔医学杂志, 1998, 33(6): 378-380.

[6] Song AP, Miao HK. Design and implementation of bone age assessment system[J]. Computer Engineering, 2002, 28(11): 84-87.
宋安平,缪淮扣. 骨龄评测系统的设计与实现[J]. 计算机工程, 2002, 28(11): 84-87.

[7] Bo-Chun Fan, Chi-Wen Hsieh, Tai-Lang Jong, et al. Automatic bone age estimation based on carpal-bone Image-a preliminary report[J]. Chinese Medical Journal (Taipei), 2001, 64(): 203-208.

[8] Pietka E, McNitt-Gray MF, Kuo ML, et al. Computer-assisted phalangeal analysis in skeletal age assessment[J]. IEEE Transactions on Medical Imaging, 1991, 10(4): 616-620.

[9] Pietka E, Kaabi L, Kuo ML, et al. Feature extraction in carpal-bone analysis[J]. IEEE Transactions on Medical Imaging, 1993, 12(1): 44-49.

[10] Pietka E, Gertych A, Pospiech S, et al. Computer-assisted bone age assessment: image preprocessing and epiphyseal/ metaphyseal ROI extraction[J]. IEEE Transactions on Medical Imaging, 2001, 20(8): 715-729.

[11] Nopola T, Jarvi A, Svedstrom E, et al. Segmenting bones from wrist-hand radiographs[R]. Turku Centre for Computer Science TUCS Technical Report, 2000.

[12] Efford ND. Knowledge-Based Segmentation and Feature Analysis of Hand-Wrist Radiographs[R]. University of Leeds School of Computer Studies Research Report Series, 1994.