

基于二维医学影像推算三维人体姿态

胡海¹, 周平¹, 王恒¹, 徐圆圆^{2*}

¹贵阳朗玛信息技术股份有限公司, 贵州 贵阳

²贵州大学大数据与信息工程学院, 贵州 贵阳

收稿日期: 2022年7月25日; 录用日期: 2022年8月15日; 发布日期: 2022年8月24日

摘要

针对三维人体姿态估计模型复杂, 计算量偏大的问题, 提出一种基于二维医学影像推算三维人体姿态的方法。在特定姿态下采集人体医学影像并进行二维影像姿态估计, 得到人体各个骨骼的医学影像的长度信息。再在人体位置不发生改变的情况下, 对任意姿态的人体采集图像并进行二维医学影像的人体姿态估计, 得到骨骼在与相机透镜主光轴垂直平面上的投影的影像的近似长度, 反推骨骼与该平面的夹角, 最后从投影的影像位置反推骨骼的三维空间朝向。最后将得到的各个骨骼的空间朝向组合在一起, 便得到了三维人体医学影像。使用此方法只需要对医学图像进行二维人体姿态估计, 再加上一些三角函数、反三角函数的计算就可以从二维医学影像去推算三维人体姿态。与ICCV 2019中代表当今最高水平(State of the Art, SOTA)的三维人体姿态模型的方法相比, 大大减少了计算量。

关键词

计算机视觉, 人体姿态估计, 医学影像, 理想透镜, 高斯成像公式, 弥散圆, 合焦

Estimation of Three-Dimensional Human Posture Based on Two-Dimensional Medical Images

Hai Hu¹, Ping Zhou¹, Heng Wang¹, Yuanyuan Xu^{2*}

¹Longmaster Information & Technology Co., Ltd, Guiyang Guizhou

²College of Big Data and Information Engineering, Guizhou University, Guiyang Guizhou

Received: Jul. 25th, 2022; accepted: Aug. 15th, 2022; published: Aug. 24th, 2022

*通讯作者。

Abstract

Considering the complexity of 3-D human pose estimation model and the computation of its training and very high inference, the estimation of three-dimensional human posture based on two-dimensional medical images is proposed. By capturing a human's medical image with a certain pose and doing 2-D human pose estimation on it, the result can be used to get the bones' medical image length. And then acquiring an image of a human body of an arbitrary posture and estimating the human body posture of a two-dimensional medical image without changing the human body position. As to every bone, an approximate length of the image of the bone's projection on a plane that is perpendicular to the main axis of the camera lens can be got. Furthermore, with this approximate length of the medical image, the angle between the bone and the plane can be inferred. Finally, the spatial orientation of the bone can be inferred with this angle and that image position. With the orientations of bones of human skeleton, a 3-D human pose can be composed. Using our method, through a 2-D human medical image pose estimation, some trigonometric calculation and some inverse trigonometric calculation, a 3-d human pose estimation can be figured out. Compared with the three-dimensional human posture model in ICCV 2019, which represents the highest state of the art (SOTA), the calculation amount is greatly reduced.

Keywords

Computer Vision, Human Pose Estimation, Medical Imaging, Ideal Lens, Gaussian Thin Lens Formula, Circle of Confusion, In Focus

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

近年来,随着人工智能的迅速发展,医学影像识别领域诞生了很多优秀的算法,不断提高着影像识别的水平。而在影像识别领域,人体姿态估计是一个很热门的子领域,各种算法不断提高着人体姿态估计的精度。随着人体姿态估计的发展,姿态估计的精细化程度也从二维提升到了三维,而要进行精确的三维人体姿态估计,一般需要使用人体医学影像的三维信息进行训练,这样的信息包括体素(三维空间中的点,对应于二维空间的像素)、点云或多重二维影像等,这些数据类型的的数据量比单张二维影像高出了一个数量级,因此训练和估计的计算量也都比二维影像的姿态估计多出许多[1]-[14]。为了降低三维人体姿态估计的计算量,使得在一些计算能力有限的设备上也能较快地进行三维人体姿态的估计,甚至实时地对人体姿态视频进行估计,本文提出了一种使用二维人体姿态估计结果推算三维人体姿态的方法。

2. 利用二维人医学影像识别得到人体骨骼像的长度

首先,我们利用特定姿态下人体姿态的估计得到人体在距离相机一定距离时人体各个骨骼的像的长度。当人体正对相机镜头、抬头挺胸、双腿直立、双手自然垂于人体两侧时,可近似看作人体处于一个与相机镜头的主光轴垂直的一个平面内。这时,经相机镜头所成的像也位于一个平面内,也即在相机的成像平面内,见图1。

这时成像平面内的人体骨骼的像的长度与人体骨骼的真实长度的比值应等于像距与物距的比值。

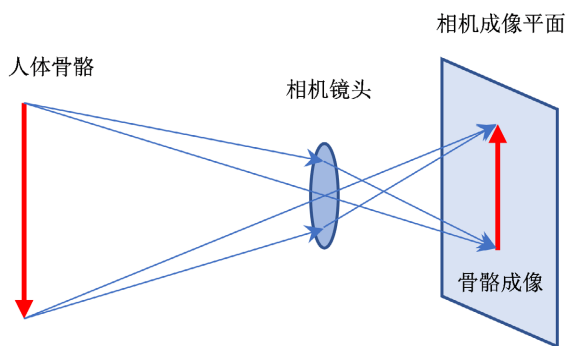


Figure 1. Imaging of human body in specific posture
图 1. 人体特定姿态成像

3. 通过人体骨骼的像在与主光轴垂直平面上的投影推算人体骨骼的朝向

当同一个人在不改变位置的情况下改变人体姿态。在改变姿态后，我们再通过相机进行拍摄，考虑拍摄是合焦的，由于人体处于相机的景深范围内，我们可以知道各个骨骼经相机的镜头所成的像的准确位置就在相机成像平面的附近[15] (焦深范围内)，一个点光源在相机成像平面上的像可能不是一个点，而是一个“圆” (弥散圆)，只是因为这个圆的直径很小，人眼无法分辨其与点的区别，所以看到的像依然是“清晰”的。考虑到相机的入射光线是近轴光线(与相机镜头的主光轴夹角很小)，所以经相机镜头折射后的光线也应该与主光轴的夹角也很小(先考虑过光心的光线，该光线与主光轴的夹角很小，而由于此时相机是“合焦”的，所以其他光线在像平面上的入射点应该与过光心的光线的入射点很接近，间隔很小，可以近似看作：过光心的光线的入射点就是光源在像平面上所成的图像)。因而可以近似地认为光线与成像平面垂直，可以把成像平面内所形成的图像近似看作该“空间的像”在成像平面内的投影，见图 2。

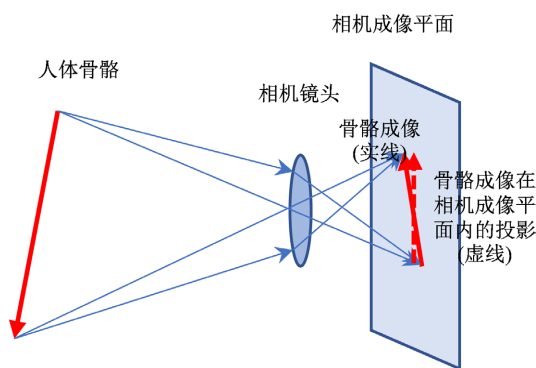


Figure 2. Imaging of human body in any posture
图 2. 人体任意姿态成像

可以证明，直线段经理想薄凸透镜所成的实像还是直线段，并且直线段的组合的像等于直线段的像的组合。现在我们把骨骼简化为一个线段，我们来看一下，一个直线段经相机(可以近似看作一种理想薄凸透镜)成像的解析，见图 3。

设有一个线段 AB ， A 是离透镜光心 O 较远的一点， AB 可以看作其在过 A 点与主光轴垂直的平面上一个投影 AC 与其在过 A 点与主光轴平行的直线上的投影 AD 的向量和。由于“向量和的像等于像的向量和”，所以 AB 的像 $A'B'$ 应该是 AC 的像 $A'C'$ 和 AD 的像 $A'D'$ 的向量和。其中 $A'C'$ 在过点 A' 与主光轴垂直的平面内， $A'D'$ 与 AD 、主光轴都在一个平面内， $A'D'$ 可分解为其在过点 A' 与主光轴垂直的平

面内的投影 $A'E$ 和其在过点 A' 与主光轴平行的直线上的投影 $A'F$ 。现在来看一下 $A'E$ 的长度。

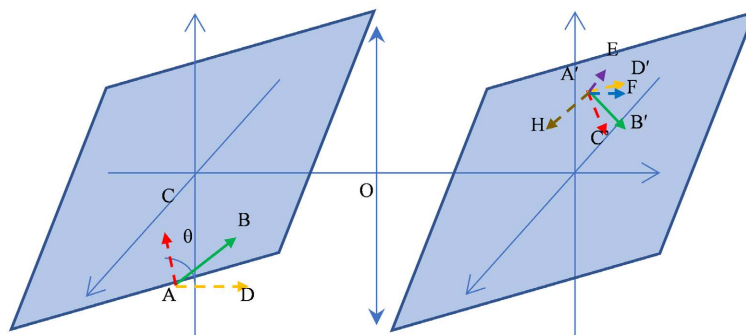


Figure 3. Image of bone projection
图3. 骨骼投影的像

$A'E$ 其实就是 A' 与主光轴的距离与 D' 与主光轴的距离之差，也就是点 A 的“像高”与点 D 的“像高”之差。设过点 A 的物距为 u ，像距为 v ，透镜的焦距为 f ，向量 AB 与 AC 的夹角为 θ ，向量 AB 的长为 l ，则 AC 的长为 $l \cos \theta$ ， AD 的长为 $l \sin \theta$ 。由于 AD 平行于主光轴，所以 A 和 D 的“物高”应该是一样的，记为 h ，则点 A 的“像高” h_A 应满足：

$$\frac{h_A}{h} = \frac{v}{u}$$

而 $\frac{v}{u} = \frac{f}{u-f}$ ，所以：

$$h_A = \frac{f}{u-f} h$$

同理，点 D 的“像高”为： $h_D = \frac{f}{u-l \sin \theta - f} h$ ，所以：

$$\begin{aligned} h_D - h_A &= \frac{f}{u-l \sin \theta - f} h - \frac{f}{u-f} h \\ &= l \sin \theta \frac{f}{u-f} \cdot \frac{h}{u-l \sin \theta - f} \end{aligned}$$

而 $A'C'$ 的长度与 AC 的长度之比应为 $\frac{v}{u}$ ，即 $\frac{f}{u-f}$ ，而 AC 的长为 $l \cos \theta$ ，所以 $A'C'$ 的长度为 $l \cos \theta \frac{f}{u-f}$ 。所以当 θ 小于 $\frac{\pi}{2}$ 时， $A'E$ 的长度与 $A'C'$ 的长度之比为

$$\frac{l \sin \theta \frac{f}{u-f} \cdot \frac{h}{u-l \sin \theta - f}}{l \cos \theta \frac{f}{u-f}}$$

即：

$$\tan \theta \frac{h}{u-l \sin \theta - f}$$

我们考虑先让 AB 在 AB 、 AC 所在的平面绕端点 A 转至过 A 点与主光轴垂直的平面内，则此时 AB 的像 $A'B'$ 的长度为 $l' = l \frac{f}{u-f}$ ，如前所述，当人体正对相机镜头、抬头挺胸、双腿直立、双手自然垂于人体两侧时，可近似看作人体处于一个与相机镜头的主光轴垂直的一个平面内，此时我们可以通过相机所拍的照中对应的骨骼长度即是 $A'B'$ 的长度 l' 。再考虑 B 沿反方向重新转回原来的位置，此时的 $A'C'$ 的长度与 l' 的比就是 $\cos \theta$ 。所以只要我们能得到 $A'C'$ 的长度信息，就能得到 $\cos \theta$ ，进而通过反余弦函数得到 θ 的值。当得到 θ 以后，先将 $A'C'$ 在过点 A' 与主光轴垂直的平面内，沿点 A' 旋转 180° 得到 $A'C''$ ，过点 A' 可作一个向量 $A'H$ ，其在过点 A' 与主光轴垂直的平面内的投影为 $A'C''$ ，在与主光轴平行的方向上的投影与 AD 同向，长度为 $l' \sin \theta$ ，则该向量 $A'H$ 相对于 A' 的朝向与 AB 相对于 A 的朝向一致，长度与 AB 的长度之比为 $\frac{f}{u-f}$ 。所以问题的关键在于求出 θ 。我们知道 $A'C'$ 的长度与 l' 的比是 $\cos \theta$ ，而 l' 可通过特定姿态的骨骼成像获得，所以只要知道 $A'C'$ 的长就可以得到 $\cos \theta$ 。

下面我们讨论把 $A'C'$ 和 $A'E$ 的向量和当做 $A'C'$ 来求 θ 以及 $A'H$ 的方向，误差在可接受的范围之内（最大相对误差不超过 10%）。我们知道在过点 A' 与主光轴垂直的平面内， $A'C'$ 与 $A'E$ 的向量和可看作是 AB 的像 $A'B'$ 在该平面内的投影，而由于近轴光线可近似看作与主光轴平行，所以 $A'B'$ 在成像平面上的投影可近似看作 AB 在成像平面内所拍摄成的图像。因而 AB 在成像平面内所拍摄成的图像的长可以当作 $A'B'$ 在过点 A' 与主光轴垂直的平面内投影的长（成像平面与过点 A' 与主光轴垂直的平面平行，所以 $A'B'$ 在这两个平面内的投影平行且长度相等）。先考虑 $\theta = \frac{\pi}{2}$ 时。当 $\theta = \frac{\pi}{2}$ 时， AB 就是 AD ，此时因为 $A'C'$ 长度为 0，所以 $A'C'$ 与 $A'E$ 的向量和就是 $A'E$ ，其长度为：

$$l \frac{f}{u-f} \cdot \frac{h}{u-l-f}$$

用此长度除以 l' （即 $l \frac{f}{u-f}$ ），得到的值为 $\frac{h}{u-l-f}$ ，由于 h 、 l 、 f 都远小于 u ，假设 h 、 l 、 f 都小于 $\frac{u}{30}$ ，那么这个值小于 $\frac{1}{28}$ ，用反余弦函数求得的 θ 的计算值 $\hat{\theta}$ 大于 $\arccos\left(\frac{1}{28}\right)$ ，即 87.953° ，此时求得的 θ 值与真实值的误差小于 3° 。用前述方法作图得到的向量 $A'H$ 与理论上的 $A'H$ 的偏差（在与主光轴垂直的平面内的向量误差 $A'E$ 加上在与主光轴平行方向上的向量误差即是最终的偏差）应小于：

$$\sqrt{\left(l' \sin \theta \frac{h}{u-l \sin \theta - f}\right)^2 + \left(l' \sin \theta - l' \sin \hat{\theta}\right)^2}$$

因为 $\theta = \frac{\pi}{2}$ ，所以也即：

$$\sqrt{\left(l' \cos \hat{\theta}\right)^2 + \left(l' - l' \sin \hat{\theta}\right)^2}$$

也即：

$$\sqrt{2} \cdot l' \sqrt{1 - \sin \hat{\theta}}$$

小于：

$$\sqrt{2} \cdot l' \sqrt{1 - \sin(87.953^\circ)}$$

约等于 $\sqrt{2} \cdot \sqrt{0.0006} \cdot l'$ ，其与 l' 的比值为： $\sqrt{2} \cdot \sqrt{0.0006} \approx 0.035$ ，可以看到这个相对误差是很小的。注意到 θ 为 0 的情况其实已经讨论过了（即是 AB 在与主光轴垂直的平面内，这种情况 $A'C'$ 的长度等于 l' ，用投影长度除以 l' 等于 1，求反余弦等于 $\frac{\pi}{2}$ ，结果没有误差）。下面只考虑 $\theta \in \left(0, \frac{\pi}{2}\right)$ 的情况。

当 $\theta \in \left(0, \frac{\pi}{2}\right)$ 时，设 $A'C'$ 与 $A'E$ 的向量和的长度为 p' ，设 $A'C'$ 的长度为 p ，那么 p' 应在区间 $\left[\left|p - p \frac{h \tan \theta}{u - l \sin \theta - f}\right|, p + p \frac{h \tan \theta}{u - l \sin \theta - f}\right]$ 之内。

现在我们把 $\frac{h}{u - l \sin \theta - f}$ 记为 $k(\theta)$ 。并记 $k_1 = \frac{h}{u - l - f}$ ， $k_2 = \frac{h}{u - f}$ 。因此有 $k_2 \leq k(\theta) \leq k_1$ 。我们注意到 p' 的区间为 $\left[\left|p - p \frac{h \tan \theta}{u - l \sin \theta - f}\right|, p + p \frac{h \tan \theta}{u - l \sin \theta - f}\right]$ ，如果 h 为 0，那么该区间退化为一个值 p ，以此值求 θ 不存在误差。下面只考虑 h 不为 0 的情况。对应的求得的 θ 的余弦值的区间为

$\left[\left|\frac{p}{l'} - \frac{p}{l'} \cdot \frac{h \tan \theta}{u - l \sin \theta - f}\right|, \frac{p}{l'} + \frac{p}{l'} \cdot \frac{h \tan \theta}{u - l \sin \theta - f}\right]$ ，把 $\frac{h}{u - l \sin \theta - f} = k(\theta)$ 代入，得到：

$$\left[\left|\frac{p}{l'} - \frac{p}{l'} \tan \theta k(\theta)\right|, \frac{p}{l'} + \frac{p}{l'} \tan \theta k(\theta)\right]$$

注意到 $\frac{p}{l'} = \cos \theta$ ，所以也就是：

$$\left[\left|\cos \theta - \sin \theta k(\theta)\right|, \cos \theta + \sin \theta k(\theta)\right]$$

现在我们对 θ 分段来分析，先来看区间上确界 $\cos \theta + \sin \theta k(\theta)$ ：

1. 当 $0 < \theta \leq \arccot\left(\frac{1 - k_2^2}{2k_2}\right)$ 时： $\cot \theta \geq \frac{1 - k_2^2}{2k_2}$ ，而 $\frac{1 - k_2^2}{2k_2} \geq \frac{1 - k^2(\theta)}{2k(\theta)}$ ，所以 $\cot \theta \geq \frac{1 - k^2(\theta)}{2k(\theta)}$ ，两边同时乘以 $2k(\theta) \sin^2(\theta)$ ，得到：

$$2k(\theta) \sin \theta \cos \theta \geq (1 - k^2(\theta)) \sin^2 \theta$$

即：

$$2k(\theta) \sin \theta \cos \theta + k^2(\theta) \sin^2 \theta + \cos^2 \theta \geq 1$$

也即：

$$(\cos \theta + \sin \theta k(\theta))^2 \geq 1$$

而 $\cos \theta + \sin \theta k(\theta) > 0$ ，所以 $\cos \theta + \sin \theta k(\theta) \geq 1$ ，这时可能超出了余弦值的范围，所以只能取 1，通过反余弦函数得到的 θ 的计算值 $\hat{\theta}$ 为 0，所以 θ 与 $\hat{\theta}$ 的差值为 θ ，其大小属于 $\left(0, \arccot\left(\frac{1 - k_2^2}{2k_2}\right)\right]$ 区间，小于等于 $\arccot\left(\frac{1 - k_2^2}{2k_2}\right)$ ，假设 h, l, f 都小于 $\frac{u}{30}$ ，则 $k_2 < \frac{1}{29}$ ， $\frac{1 - k_2^2}{2k_2} > \frac{420}{29}$ ，所以 $\arccot\left(\frac{1 - k_2^2}{2k_2}\right) \leq \arccot\left(\frac{420}{29}\right) \approx 3.9499^\circ$ ，计算得到的 $A'H$ 与真实的 $A'H$ 之间的差异（向量差）的长度不超过：

$$\sqrt{\left(l' \sin \theta \frac{h}{u-l \sin \theta-f}\right)^2 + \left(l' \sin \theta - l' \sin \hat{\theta}\right)^2}$$

小于:

$$l' \sqrt{k_1^2 + \sin^2 \theta}$$

小于:

$$l' \sqrt{\left(\frac{1}{28}\right)^2 + \sin^2(3.9499^\circ)} \approx 0.0776l'$$

2. 当 $\operatorname{arccot}\left(\frac{1-k_2^2}{2k_2}\right) < \theta < \operatorname{arccot}\left(\frac{1-145k_1^2}{2k_1}\right)$ 时: 如果 $\cos \theta + \sin \theta k(\theta) \geq 1$, 那么通过反余弦函数得到的 θ 的计算值 $\hat{\theta}$ 为 0, $\hat{\theta}$ 与 θ 的差值小于 $\operatorname{arccot}\left(\frac{1-145k_1^2}{2k_1}\right)$, 假设 h, l, f 都小于 $\frac{u}{30}$, 则 $k_1 < \frac{1}{28}$, $\frac{1-145k_1^2}{2k_1} \geq \frac{639}{56}$, 所以

$$\operatorname{arccot}\left(\frac{1-145k_1^2}{2k_1}\right) \leq \operatorname{arccot}\left(\frac{639}{56}\right) \approx 5.0084^\circ$$

计算得到的 $A'H$ 与真实的 $A'H$ 之间的差异(向量差)的长度不超过:

$$\sqrt{\left(l' \sin \theta \frac{h}{u-l \sin \theta-f}\right)^2 + \left(l' \sin \theta - l' \sin \hat{\theta}\right)^2}$$

小于:

$$l' \sqrt{k_1^2 + \sin^2 \theta}$$

小于:

$$l' \sqrt{\left(\frac{1}{28}\right)^2 + \sin^2(5.0084^\circ)} \approx 0.0943l'$$

如果 $\cos \theta + \sin \theta k(\theta) < 1$, 因为 $\hat{\theta}$ 小于 θ , 所以 $\hat{\theta}$ 与 θ 的差也不会超过 θ , 所以计算得到的 $A'H$ 与真实的 $A'H$ 之间的差异(向量差)的长度也不会超过 $0.0943l'$ 。

3. 当 $\theta \geq \operatorname{arccot}\left(\frac{1-145k_1^2}{2k_1}\right)$ 时: $\cot \theta \leq \frac{1-145k_1^2}{2k_1}$, 因此 $\sqrt{1-k_1^2} - 2k_1 \cot \theta > 12k_1$ 。反余弦函数 $\arccos(x)$ 的导数是 $\frac{1}{-\sin(\arccos(x))}$, 所以求得的 $\hat{\theta}$ 与 θ 的差值应为 $\frac{\sin \theta k(\theta)}{-\sin \theta'}$, 其中 $\hat{\theta} \leq \theta' \leq \theta$, 所以 $\frac{\sin \theta k(\theta)}{\sin \theta'} \leq \frac{\sin \theta k(\theta)}{\sin \hat{\theta}}$, 而 $\cos \hat{\theta} = \cos \theta + \sin \theta k(\theta)$, 所以

$$\frac{\sin \theta k(\theta)}{\sin \hat{\theta}} = \frac{\sin \theta k(\theta)}{\sqrt{1-(\cos \theta + \sin \theta k(\theta))^2}}$$

小于:

$$\frac{\sin \theta k_1}{\sqrt{1 - (\cos \theta + \sin \theta k_1)^2}}$$

即:

$$\frac{k_1}{\sqrt{1 - k_1^2 - 2k_1 \cot \theta}}$$

因为 $\sqrt{1 - k_1^2 - 2k_1 \cot \theta} > 12k_1$, 所以:

$$\frac{k_1}{\sqrt{1 - k_1^2 - 2k_1 \cot \theta}} < \frac{1}{12} \approx 4.7746^\circ$$

计算得到的 $A'H$ 与真实的 $A'H$ 之间的差异(向量差)的长度不超过:

$$\sqrt{\left(l' \sin \theta \frac{h}{u - l \sin \theta - f}\right)^2 + (l' \sin \theta - l' \sin \hat{\theta})^2}$$

小于:

$$l' \sqrt{k_1^2 + (\theta - \hat{\theta})^2}$$

小于:

$$\sqrt{\left(\frac{1}{28}\right)^2 + \left(\frac{1}{12}\right)^2} l' \approx 0.0907l'$$

现在再来看区间下确界

$$|\cos \theta - \sin \theta k(\theta)|:$$

我们注意到 $\cos \theta - \sin \theta k(\theta)$ 是 θ 的单调递减函数, 设某个值 θ_1 使

$$\cos \theta_1 - \sin \theta_1 k(\theta_1) = 0$$

4. 当 $0 < \theta \leq \theta_1$ 时: 因为 $\cos \theta_1 - \sin \theta_1 k(\theta_1) = 0$, 所以 $\cot \theta_1 = k(\theta_1) \geq k_2$, 所以

$$\theta_1 \leq \operatorname{arccot}(k_2)。$$

另外, 此时 $\cos \theta - \sin \theta k(\theta) \geq 0$, $|\cos \theta - \sin \theta k(\theta)| = \cos \theta - \sin \theta k(\theta)$ 。

$$\hat{\theta} = \arccos(|\cos \theta - \sin \theta k(\theta)|)$$

设 $\arccos(|\cos \theta - \sin \theta k(\theta)|) = \hat{\theta}$, 则 $= \arccos(\cos \theta - \sin \theta k(\theta))$

$$= \arccos(\cos \theta) - \left(-\frac{\sin \theta k(\theta)}{\sin \theta'}\right)$$

其中, θ' 介于 $\hat{\theta}$ 与 θ 之间。

因而有:

$$\hat{\theta} - \theta = \frac{\sin \theta k(\theta)}{\sin \theta'}$$

而

$$\frac{\sin \theta k(\theta)}{\sin \theta'} \leq \frac{\sin \theta k(\theta)}{\sin \theta} = k(\theta) \leq k_1, \quad k_1 \leq \frac{1}{28} \approx 0.0357 \approx 2.0455^\circ$$

计算得到的 $A'H$ 与真实的 $A'H$ 之间的差异(向量差)的长度不超过:

$$l' \sqrt{k_1^2 + (\theta - \hat{\theta})^2} \leq l' \sqrt{\left(\frac{1}{28}\right)^2 + \left(\frac{1}{28}\right)^2} \approx 0.0505l'$$

5. 当 $\theta > \theta_1$ 时: $\cos \theta - \sin \theta k(\theta) < 0$, $|\cos \theta - \sin \theta k(\theta)| = -\cos \theta + \sin \theta k(\theta)$, 设 $\arccos(|\cos \theta - \sin \theta k(\theta)|) = \hat{\theta}$, 则:

$$\begin{aligned} \hat{\theta} &= \arccos(|\cos \theta - \sin \theta k(\theta)|) \\ &= \arccos(-\cos \theta + \sin \theta k(\theta)) \\ &= \arccos(-\cos \theta) - \frac{\sin \theta k(\theta)}{\sin \theta'} \\ &= (\pi - \theta) - \frac{\sin \theta k(\theta)}{\sin \theta'} \end{aligned}$$

其中 θ' 介于 $\hat{\theta}$ 与 $\pi - \theta$ 之间, 因此

$$\frac{1}{\sin \theta'} \leq \max \left\{ \frac{1}{\sin(\pi - \theta)}, \frac{1}{\sin(\arccos(-\cos \theta + \sin \theta k(\theta)))} \right\}$$

所以:

$$\frac{\sin \theta k(\theta)}{\sin \theta'} \leq \max \left\{ \frac{\sin \theta k(\theta)}{\sin(\pi - \theta)}, \frac{\sin \theta k(\theta)}{\sin(\arccos(-\cos \theta + \sin \theta k(\theta)))} \right\}$$

而 $\frac{\sin \theta k(\theta)}{\sin(\pi - \theta)} = k(\theta)$,

$$\begin{aligned} &\frac{\sin \theta k(\theta)}{\sin(\arccos(-\cos \theta + \sin \theta k(\theta)))} \\ &= \frac{\sin \theta k(\theta)}{\sqrt{1 - (-\cos \theta + \sin \theta k(\theta))^2}} \\ &= \frac{\sin \theta k(\theta)}{\sqrt{(1 - k^2(\theta)) \sin^2 \theta + 2 \sin \theta \cos \theta k(\theta)}} \\ &\leq \frac{\sin \theta k(\theta)}{\sqrt{(1 - k^2(\theta)) \sin^2 \theta}} \\ &= \frac{k(\theta)}{\sqrt{1 - k^2(\theta)}} \end{aligned}$$

而 $k(\theta) < \frac{k(\theta)}{\sqrt{1 - k^2(\theta)}}$

所以:

$$\frac{\sin \theta k(\theta)}{\sin \theta'} \leq \frac{k(\theta)}{\sqrt{1-k^2(\theta)}} \leq \frac{k_1}{\sqrt{1-k_1^2}}$$

而由于 $\cos \theta_1 - \sin \theta_1 k(\theta_1) = 0$, 可知 $\cot \theta_1 = k(\theta_1) \leq k_1$, 因而

$$\frac{\pi}{2} > \theta > \theta_1 \geq \operatorname{arccot}(k_1)。$$

因而

$$\begin{aligned} |\hat{\theta} - \theta| &= \left| \pi - 2\theta - \frac{\sin \theta k(\theta)}{\sin \theta'} \right| \\ &< \max \left(\pi - 2\theta, \frac{\sin \theta k(\theta)}{\sin \theta'} \right) \\ &\leq \max \left(\pi - 2 \operatorname{arccot}(k_1), \frac{k_1}{\sqrt{1-k_1^2}} \right) \end{aligned}$$

$$\text{而 } \frac{k_1}{\sqrt{1-k_1^2}} < \frac{\frac{1}{28}}{\sqrt{1-\left(\frac{1}{28}\right)^2}} \approx 0.0357 \approx 2.0455^\circ$$

$$\pi - 2 \operatorname{arccot}(k_1) < \pi - 2 \operatorname{arccot}\left(\frac{1}{28}\right) \approx 0.0714 \approx 4.0909^\circ$$

所以, $|\hat{\theta} - \theta| \leq 4.0909^\circ$ 。

计算得到的 $A'H$ 与真实的 $A'H$ 之间的差异(向量差)的长度不超过:

$$l' \sqrt{k_1^2 + (\theta - \hat{\theta})^2} \leq l' \sqrt{\left(\frac{1}{28}\right)^2 + (0.0714)^2} \approx 0.0798l'$$

综上所述, 用 $A'E$ 与 $A'C'$ 的向量和(即 $A'B'$ 的像在过点 A' 上的投影)代替 $A'C'$ 来求 θ , 并作 $A'H$ 最终得到的骨骼缩小后的位置与真实位置之间的差都在可接受的范围内(最大相对误差不超过 10%)。

对于人体所有的关键骨骼都采用此方法计算其朝向, 然后把所有骨骼的朝向组合起来便得到了三维人体姿态。

4. 与三维人体姿态模型的计算速度对比

使用谷歌的 PoseNet 二维人体姿态模型, 使用默认参数对 480P 分辨率的图像, 进行人体姿态识别, 并使用上述方法(以下简称 PoseNet-to-3D)推算人体的三维姿态, 在 MacBook Pro (13-inch, 2017, Four Thunderbolt 3 Ports)的平台上耗时平均大约为 0.083 秒。而使用《Learnable Triangulation of Human Pose》[1]中所述的方法(以下简称 LTH)进行姿态估计, 在相同平台下, 对 480P 分辨率的图像进行姿态估计, 耗时大约为 2.5 秒。而使用《Cross View Fusion for 3D Human Pose Estimation》[2]中所述的方法(以下简称 CVF)进行姿态估计, 在相同平台下, 对 480P 分辨率的图像进行姿态估计, 耗时大约为 2.7 秒, 见表 1。

Table 1. Comparison of three-dimensional human posture estimation time
表 1. 三维人体姿态估计耗时对比

模型	图像分辨率	耗时(秒)
LTH	480P	2.5
CVF	480P	2.7
PoseNet-to-3D	480P	0.083

5. 结语

通过对人体医学影像特定姿态的估计得到了骨骼的像的长度。在位置不变的前提下,任意改变影像姿态,通过骨骼的像在与主光轴垂直的平面上的投影以及像的长度推算骨骼的朝向。通过一些二维和三维空间的几何关系以及对高斯成像公式的使用,巧妙地把医学影像三维姿态估计转换为二维姿态估计加上一些代数运算,大大降低了三维人体影像姿态估计的计算复杂度。

参考文献

- [1] Iskakov, K., Burkov, E., Lempitsky, V. and Malkov, Y. (2019) Learnable Triangulation of Human Pose. 2019 *International Conference on Computer Vision (ICCV)*, Seoul, 27 October-2 November 2019, 7717-7726. <https://doi.org/10.1109/ICCV.2019.00781>
- [2] Qiu, H.B., Wang, C.Y., Wang, J.D., Wang, N.Y. and Zeng, W.J. (2019) Cross View Fusion for 3D Human Pose Estimation. 2019 *International Conference on Computer Vision (ICCV)*, Seoul, 27 October-2 November 2019, 4341-4350. <https://doi.org/10.1109/ICCV.2019.00444>
- [3] He, Y.H., Yan, R., Fragkiadaki, K. and Yu, S.-I. (2020) Epipolar Transformers. 2020 *Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, 13-19 June 2020, 7779-7788.
- [4] Huang, F.Y., Zeng, A.L., Liu, M.H., Lai, Q.X. and Xu, Q. (2021) DeepFuse: An IMU-Aware Network for Real-Time 3D Human Pose Estimation from Multi-View Image. 2020 *IEEE Winter Conference on Applications of Computer Vision (WACV)*, Snowmass, 1-5 March 2020, 418-427. <https://arxiv.org/pdf/1912.04071v1.pdf>
<https://doi.org/10.1109/WACV45572.2020.9093526>
- [5] Liang, J.B. and Lin, M.C. (2019) Shape-Aware Human Pose and Shape Reconstruction Using Multi-View Images. 2019 *International Conference on Computer Vision (ICCV)*, Seoul, 27 October-2 November 2019, 4351-4361. <https://doi.org/10.1109/ICCV.2019.00445>
- [6] Martinez, J., Hossain, R., Romero, J. and Little, J.J. (2017) A Simple Yet Effective Baseline for 3d Human Pose Estimation. 2017 *International Conference on Computer Vision (ICCV)*, Venice, 22-29 October 2017, 2659-2668. <https://doi.org/10.1109/ICCV.2017.288>
- [7] Tu, H.Y., Wang, C.Y. and Zeng, W.-J. (2020) VoxelPose: Towards Multi-Camera 3D Human Pose Estimation in Wild Environment. 2020 *European Conference on Computer Vision (ECCV)*, Glasgow, 23-28 August 2020, 1-17.
- [8] Dong, J.T., Jiang, W., Huang, Q.-X., Bao, H.J. and Zhou, X.W. (2019) Fast and Robust Multi-Person 3D Pose Estimation from Multiple Views. 2019 *Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, 15-20 June 2019, 7784-7793. <https://doi.org/10.1109/CVPR.2019.00798>
- [9] Belagiannis, V., Amin, S., Andriluka, M., Schiele, B., Navab, N. and Ilic, S. (2014) 3D Pictorial Structures for Multiple Human Pose Estimation. 2014 *Conference on Computer Vision and Pattern Recognition (CVPR)*, Columbus, 23-28 June 2014, 1669-1676. <https://doi.org/10.1109/CVPR.2014.216>
- [10] Ershadi-Nasab, S., Noury, E., Kasaei, S. and Sanaei, E. (2018) Multiple Human 3D Pose Estimation from Multiview Images. *Multimedia Tools and Applications*, **77**, 15573-15601. <https://link.springer.com/article/10.1007/s11042-017-5133-8>
- [11] Belagiannis, V., Amin, S., Andriluka, M., Schiele, B., Navab, N. and Ilic, S. (2021) 3D Pictorial Structures Revisited: Multiple Human Pose Estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **38**, 1929-1942. <http://campar.in.tum.de/pub/belagiannis2016pami/belagiannis2016pami.pdf>
- [12] Belagiannis, V., Wang, X.C., Schiele, B., Fua, P., Ilic, S. and Navab, N. (2021) Multiple Human Pose Estimation with Temporally Consistent 3D Pictorial Structures. *Computer Vision—ECCV 2014 Workshops*, Zurich, 6-7, 12 September 2014, 742-754. <http://campar.in.tum.de/pub/belagiannis2014eccvChalearn/belagiannis2014eccvChalearn.pdf>

-
- [13] Veges, M. and Lorincz, A. (2021) Temporal Smoothing for 3D Human Pose Estimation and Localization for Occluded People. *Neural Information Processing 27th International Conference, ICONIP 2020*, Bangkok, 23-27 November 2020, 557-568. <https://arxiv.org/pdf/2011.00250v1.pdf>
- [14] Mehta, D., Sotnychenko, O., Mueller, F., Xu, W.P., Elgharib, M., Fua, P., Seidel, H.-P., Rhodin, H., Pons-Moll, G. and Theobalt, C. (2021) XNect: Real-Time Multi-Person 3D Motion Capture with a Single RGB Camera. <https://arxiv.org/pdf/1907.00837v2.pdf>
- [15] 王清英. 景深公式的推导[J]. 南阳师范学院报, 2003, 2(3): 24-26.