****

硕士学位论文开题报告及论文工作计划书

**课题名称 基于事件相机进行人体运动模型的预测**

**学 号**

**姓 名 \_\_**

**专 业 应用统计学\_\_\_**

**学 院**

**导 师**

**副导师**

**选题时间 2023 年 9 月 20 日**

**东北大学研究生院**

**2023年 12 月 6 日**

填表说明

1、本表一、二、三、四、五项在导师指导下如实填写。

2、学生在通过开题后一周内将该材料交到所在学院、研究所。

3、学生入学后第三学期应完成论文开题报告，按有关规定，没有完成开题报告的学生不能申请论文答辩。

一、立论依据

|  |
| --- |
| 课题来源、选题依据和背景情况、课题研究目的、理论意义和实际应用价值   1. **课题来源**   以导师确定的3D人体姿态估计为大方向，在此基础上大致了解了基于RGB传统相机进行3D人体姿态估计和基于事件相机进行3D人体姿态估计的研究现状，最终经过与导师讨论决定，将基于事件相机进行人体运动模型的预测作为课题方向。   1. **课题背景**   近年来，随着计算机视觉领域的发展，有越来越多的深度学习算法与人们的实际生活联系到一起。人体运动合成是计算机图形学中的一个重要问题，它是一系列重要任务的根本任务，解决了上述问题，就可以把它应用在诸多计算机视觉任务中，为人机动作交互，人体动作分析，人体运动规划，运动辅助系统，动作识别等领域提供数据支撑。因此，人体运动模型的建立任务是一项十分有意义的基础研究任务，事实上，建立人体运动模型，关键是通过深度学习算法，准确的推测出人体运动的各个关节点的2D和3D位置信息，并使用SMPL位姿估计算法，还原人体的真实框架参数，最终合成真实的人体运动框架。  最初，学者们尝试使用传统的数学建模的方法解决人体姿态估计的问题。然而，这些方法不仅需要较大的计算资源，而且所得预测结果的准确性也十分有限。随着深度学习技术的进步与发展，诸如卷积神经网络（CNN）和transformer等模型结构能够隐式地学习人体关节点的特征及其关系，从而得到更出色的结果。过去，人们主要研究了基于传统摄像机获取的RGB图像进行人体姿态估计，并取得了显著的成就。然而，传统摄像机由于成像原理和硬件技术上的缺陷，其感光元件可能无法同时捕捉到亮度差异较大的区域细节，导致图像中的过曝或低曝情况。其次，当拍摄运动物体时，传统相机可能会出现运动模糊问题，因为它们在获取图像时需要保持相对较长的快门速度，而快门速度较低会导致物体在图像中产生模糊效果。而事件相机是一种更为先进的技术，它使用异步传感器来记录每个像素的光照强度变化，并将其存储为事件数据，它可以更好地解决传统相机对于光照强度的需要和对于快速运动的捕捉的需要，因此利用事件相机去进行人体运动模型的建模是很有价值的方向。   1. **课题研究目的**   虽然基于RGB相机进行人体姿态估计的方法已经应用广泛而且出现较大突破，但是由于其获取数据原理的缺陷性，在对于快速运动的人体进行建模方面仍然没有较为满意的结果，事件相机的出现和使用为这一问题的解决提供了新的思路，其独有的特性可以使得我们在光照条件较差、目标人物快速移动的场景中仍然能够获取有价值的信息，用于人体姿态估计，从而让我们更好地建立3D人体运动模型。我们希望使用事件相机为主，灰度图像相辅助的多模态数据进行人体姿态估计，恢复其在2D和3D中的位置和姿势信息，为人体运动分析等后续工作做好准备。   1. **理论意义和实际应用价值**   （1）事件相机进行人体姿态估计的理论意义：  由于事件相机是利用光子的强度变化异步地记录光线进入相机的时间，并将这些信息存储为事件数据的，他在每一个像素位置达到一定的光照强度变化阈值时会触发事件，形成一个包括像素位置信息、极性信息和时间信息的四元组数据格式，这种获取图像信息的方式能够在遇到光照强度变化剧烈的情况下快速地产生事件，并且具有非常高的时间分辨率和帧率。相比传统相机，事件相机可以实现几百万帧每秒的帧率，这意味着事件相机可以捕获快速运动的物体，从而获取更高质量的图像。  其次，事件相机产生的事件图像的灰度值代表的并不是光照信息，他是二值的极性信息，记录的是事件发生的时间戳，因此，不同的像素点可以具有不同的光照感受区域，这使得事件相机能够在高动态范围的场景下，更好地捕捉物体的边缘和细节信息，为快速运动的人体进行建模。  最后，事件相机只在光照强度发生变化时记录事件，因此他不会考虑颜色和纹理的信息，对于估计运动人体的关节点和骨架而言，其具有更少的噪声信息。  （2）事件相机进行人体姿态估计的实际应用价值：  由于事件相机具有更高的动态范围和快门速度，这就使得事件相机在各种光照条件下捕捉快速运动物体中能够提供高精度的图像信息，能够更好地学习光照条件欠缺或者快速运动的人体的人体姿态信息，有助于在更广泛的环境中分析运动中的人体运动姿态，以便于更好的分析和指导人体的运动。这对于更广泛场景下的人机交互是十分有意义的。 |

二、文献综述

|  |
| --- |
| 国内外研究现状、发展动态描述（不少于1000字）；所阅文献的查阅范围及手段，附参考文献（不少于10篇，其中近3年文献不少于5篇，英文文献不少于3篇，全部按照标准格式列出，并在文中顺序标注）   1. **国内外研究现状**    1. **3D人体形状建模**   在早期的最初工作中，研究团队研究了通过使用各种几何基元去逼近估计身体形状的方法。受脸部建模的成效性进展的启示，研究团队使用3D 扫描的数据集中得到人体的形状参数，然后使用统计学的角度去创建人体的3D模型。因而根据建模的方法，人体建模大致可以分为以下两个类别:一是用几何基元法表示人体，二是使用特定主干的身体扫描数据去构建人体网络。  在2005年，SCAPE【1】(Shape Completion and Animation of People)模型提出，他是最早的3D人体建模框架，这个模型提出的目的是为了创建具有各种身体形状和姿势的3D人体模型。他是通过三角形变换表现人体姿态和形状的，是一种可变形的3D人体模型。后续有许多模型都是在SCAPE模型的基础上构建的。Dyna【2】是SCAPE模型的扩展，它是将软组织变形与运动和身体形状相关联，并能够产生广泛的逼真软组织运动。SPM【3】（Stitched puppet Model） 是一种结合了运动学和几何表示优点的人体模型。它旨在为人体动作捕捉和动画提供高水平的真实感和准确性。该模型将身体表示为缝合在一起的部件的集合，每个部件都有自己的变形属性。这允许更逼真和更自然的运动，以及更好地控制身体的形状和位置。  2015年，SMPL【4】(Skinned Multi-Person Linear)模型提出，它是一种广泛使用的3D人体姿态模型，它的目的是使用紧凑的方式去获得人体的形状和姿势变化的参数。它使用线性混合蒙皮技术来表示与姿势相关的变形，并使用主成分分析(PCA) 来模拟身体形状。SMPL 广泛用于计算机图形学、计算机视觉和机器学习应用程序。SMPL 是一种基于顶点的线性模型，它是目前研究界最广泛使用的人体模型之一，并与现有的渲染引擎兼容。与 SCAPE 类似SMPL将变形分解为形状和姿势变形。同时，SMPL也可以经过扩展来捕捉软组织动力学，得到DMPL模型。  2019年，SMPL-X【5】模型问世，它的提出解决了原始SMPL模型的诸如非局部变形、忽略身体和姿态形变相关关系以及难以再现高度非线性变形等局限性。SMPL-X模型采用的是同时对人体、面部和手进行建模的办法，他们将FLAME【6】头部模型和MANO【7】手模型扩展到SMPL上，分别用几组参数来表示身体、手和面部。SMPL-X 使用更灵活的混合形状基础来更好地捕捉身体形状变化，并引入额外的形状参数来捕捉身体脂肪和肌肉等特征。   * 1. **基于传统相机的2D和3D关节点预测**   目前，研究人员对于人体姿态的估计大多数还是以使用RGB相机获得的数据进行的，而且几乎都是采用有监督的深度学习方法，去学习二维平面的人体姿态，对于真值的表述，主流方法主要有两个，其一是关键点的数值坐标，其二是概率热图的形式，因此，根据真值的不同表示形式，对人体姿态估计的研究方法大致也可以分为两类，即基于坐标回归算法和基于热图检验算法。  2014年，研究者A.Toshev及其同事开创性地提出了首个针对个体姿态估计而设计的深度学习模型——Deep Pose【8】。此模型选用了AlexNet【9】作为其核心网络架构，旨在对关节点的位置坐标进行预测。然而，由于该模型采用了递归策略，这增加了计算的复杂性，并且限制了它在精确预测复杂人体动作方面的能力。后续的Carreira等人【10】通过分阶段直接回归的技术来提取输入空间的详尽表征，尽管这样，他们在回归坐标的过程中并未充分利用人体各部分间的关联性，忽视了姿态的内在结构，这导致了这类方法在稳健性方面的不足。  在2016年，CPM【11】序列化卷积姿态机模型提出，该模型通过采用序列化的卷积结构来同时表达图像和空间信息。与传统的需要显式构建空间模型或递归神经网络的方法不同，CPM通过扩大网络的感受野来隐式地学习人体各部分间的空间关系。同时，Rafi等人【12】设计了一个多分辨率网络，能够捕捉并有效融合不同尺度的特征，从而提高了姿态定位的精确度。基于此，Newell等人【13】受启发提出了一种沙漏网络结构，该结构能够自下而上和自上而下地进行重复推理，由多个残差模块【14】组成，用于融合不同分辨率下学习到的特征，并通过中继监督来限制沙漏模块的学习内容，从而隐式地学习到人体姿态的空间结构关系。与CPM相比，沙漏模型在算法复杂度方面具有优势。此后，许多研究者也针对这一经典模型进行了改进。  2017年，随着Transformer【15】算法的提出，凭借它在长距离依赖上的天然优势，其在NLP（Natural Language Processing）领域大展身手，因此研究人员不断尝试将其从NLP领域迁移学习到CV（Computer Vision）领域， Zheng【16】等研究者率先将Transformer架构应用于三维人体姿态的预测，并推出了PoseFomer模型。该模型通过描绘二维人体关键点间的连续空间联系，有效提升了姿态估计器的性能。随后，Zhang【17】等人在PoseFomer模型的基础上，通过在空间Transformer的框架中穿插时序Transformer，进一步刻画了各个关节点之间的动态联系。  2021年，Yang等学者【18】采用HRNet【19】作为主网络对图像进行初步的特征抽取，随后应用Transformer架构在较高层次上提取全局依赖关系，取得了速度与性能的平衡。而Li等研究者【20】提出的TokenPose方法则摒弃了传统视觉任务中常见的二维空间结构，直接将特征图信息融入一维向量中。通过使用Transformer模型来建立关节语义向量和图像像素特征之间的联系，最终生成热图并达到了高精度的成果，为后续研究提供了新的思路。   * 1. **使用事件相机进行三维人体姿态估计**   事件相机【21】作为一种新兴的仿生成像传感器，它在很多方面是与传统相机不同的，其最重要的概念就是事件，与传统的帧相机相比，事件流在空间上也要稀疏得多，由于其高时间分辨率、低延迟、高动态范围和低功耗的独特优势，事件相机已在越来越多的计算机视觉任务中得到应用，包括相机姿态估计、特征跟踪、光流、多视图立体、手势识别、运动去模糊等  与此同时，很少有研究人员研究基于事件相机的人体姿态和形状的3D关节估计，Calabrese 等人【22】制作了首个基于事件数据的大型 3D 人体姿态估计数据集，并且提出了一种处理事件数据的方法，即：将固定数目的事件数汇集在一张图片上，形成数据帧的形式，然后再使用CV领域的相关方法进行人体姿态估计，Enrico Calabrese等人在DHP19【23】和的研究中首次提出了用于人体姿态估计的公开数据集，并且使用CNN去进行人体姿态估计，但是它不能输出3D姿势和人体形状，Lan Xu 等人提出的EventCap【24】模型，在对于三维手信息的重建任务中，实现了实时性。   1. **所阅文献的查阅范围及手段** 2. 通过CVPR，ICCV，ECCV以及Nips等顶级会议的官网查询每年会议收录的论文。 3. 使用paperwithcode网站查看各个数据集上评价指标结果位于前列的论文。 4. 使用知网搜索相应类别的文献。 5. 在优秀论文参考文献中阅读相关研究领域的论文。 6. **参考文献** 7. Anguelov D，Srinivasan P， Koller D, et al. Scape: shape completion and animation of people[M]//ACM SIGGRAPH 2005 Papers.2005:408-416. 8. Pons-MollG，Romero J, Mahmood N, et al. Dyna: A model of dynamic human shape in motion[J] ACM Transactions on Graphics (TOG), 2015, 34(4):1-14. 9. Zuffi S, Black M J. The stitched puppet: A graphical model of 3d human shape and pose[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015: 3537-3546. 10. Loper M ,Mahmood N , Romero J ,et al. SMPL: A Skinned Multi-Person Linear Model[J] Acm Transactions on Graphics,2015,34(6cd):248.DO1:10.1145/2816795.2818013. 11. Pavlakos G， ChoutasV，Ghorbani N，et al. Expressive body capture: 3d hands， face， and body from a single image[C]//Proceedings of the IEEE/CVF Conference on computer Vision and PatternRecognition.2019:10975-10985. 12. LiT, Bolkart T, Black MJ, et al. Learning a model of facial shape and expression from 4D scans[J].ACM Trans.Graph,2017, 36(6):194:1-194:17. 13. Romero J, Tzionas D, Black M J. Embodied hands: Modeling and capturing hands and bodies together[J]. arXiv preprint arXiv:2201.02610,2022. 14. Toshev A, Szegedy C. Deep pose: Human pose estimation via deep neural networks[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,2014: 1653-1660. 15. Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[C]//International Conference on Neural Information Processing Systems, 2012:1097-1105. 16. Carreira J, Agrawal P, Fragkiadaki K, et al. Human pose estimation with iterative error feedback[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,2016:4733-4742. 17. Wei S E, Ramakrishna V, Kanade T, et al. Convolutional pose machines[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,2016:4724-4732. 18. U. Rafi, B. Leibe, J. Gall, et al. An efficient convolutional network for human poseestimation [C].//British Machine Vision Conference(BMVC),2016,2. 19. Newell A, Yang K, Deng J. Stacked hourglass networks for human pose estimation[C]// Proceedings of European Conference on Computer Vision,2016:483-499. 20. He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,2016:770-778.. 21. A. Vaswani, N. Shazeer, N. Parmar, et al. Attention is all you need[C].//Conference on Neural Information Processing Systems(NeurIPS),2017:5998-6008. 22. ZHENG C, ZHU S, MENDIETA M, et al. 3d human pose estimation with spatial and temporal transformer[C//] Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, Oct 10-17,2021. Piscataway: IEEE, 2021:11656-11665.. 23. ZHANG J, TU ZYANG J, et al. Mixste: Seo2seg mixed spatio-temporal encoder for 3d human pose estimation in video[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, New Orleans, Jun 19-24, 2022. Piscateway: IEEE, 2022:13232-13242. 24. S. Yang, Z. Quan, M.Nie, et al. Transpose: Towards explainable human pose estimation by transformer[C].//IEEE International Conference on Computer Vision(ICCV),2021:11802-11812. 25. Sun K, Xiao B ,Liu D, et al. Deep high-resolution representation learning for human pose estimation[C]//Proceedings of the 2019 IEEE Conference on Computer Vision and Pattern Recognition, 2019:5693-5703. 26. Y. Li, S. Zhang, Z. Wang, et al. Token pose: Learning key point tokens for human pose estimation[C].//IEEE International Conference on Computer Vision(ICCV),2021:11313-11322. 27. Guillermo Gallego, Tobi Delbruck, Garrick Orchard, Chiara Bartolozzi, Brian Taba, Andrea Censi, Stefan Leutenegger, Andrew Davison, Jorg Conradt, Kostas Daniilidis, and Da-vide Scaramuzza. Event-based vision: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence, pages 1–30, 2020. 28. Calabrese E, Taverni G, Awai Easthope C, et al. Dhp19: Dynamic vision sensor 3d human pose dataset[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops. 2019: 0-0.   [23] Enrico Calabrese, Gemma Taverni, Christopher Awai Easthope, Sophie Skriabine, Federico Corradi, Luca Longinotti, Kynan Eng, and Tobi Delbruck. Dhp19: Dynamic vision sensor 3d human pose dataset. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2019.  [24] Lan Xu, Weipeng Xu, Vladislav Golyanik, Marc Habermann, Lu Fang, and Christian Theobalt. Eventcap: Monocular 3d capture of high-speed human motions using an event camera. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 4968–4978,2020. |

三、研究内容

|  |
| --- |
| 1．研究构想与思路、主要研究内容及拟解决的关键技术（不少于1000字）  目前使用事件相机进行人体姿态估计是一个很有潜力的方向，截止目前为止，能公开使用的数据集只有DHP19和MMHPSH数据集，DHP19 数据集是第一个公开用于研究人体姿态估计的事件相机数据集，它是使用四台同步的 DAVIS 相机去进行数据收集，每个像素平面的大小为346\*260，一共邀请了17 个模特，每名模特进行 33 组动作，每组动作重复十次，存储的原始的事件流约有145GB。MMHPSD数据集是一个多模态数据集，该数据集的采集包含了四种不同的成像模式，他使用了一个事件相机、一个偏振相机和五个RGB-D相机去进行数据收集，事件相机使用的是CeleX-V型号，具有1280×800的分辨率。采集过程中，帧相机捕获的图像与事件相机同步，事件相机捕获的事件与灰度图像同步，进而同步收集事件。数据集中包含了15名参与者，其中11名为男性，4名为女性，每个参与者进行了3组动作，共有21个不同的动作。此外，数据集中还包含了SMPL形状和姿态的注释，用于对齐和评估人体姿态估计的结果。针对MMHPSD数据集的中的事件流数据，研究员通常是这么处理的，他们是将固定数量的事件流转换成数据帧，并记录下事件发生的时间窗口信息，如图（1）所示，然后文章使用2层的CNN网络，提取事件帧的光流特征，再使用一个shapenet（这也是一个CNN网络）去提取事件帧的图像流特征，设计使用一个光流图形流相干损失函数进行模型训练，最终使用GRU门控网络得到人体的姿态估计结果。    图（1）将事件流数据通过固定数量的操作转换成事件帧的过程  上述办法已经取得不错的实验结果，但是第一个阶段提取光流的网络结构过于简单，提取到的特征过于粗糙，因而会对模型最终的预测结果产生一定的影响，其次就是并没有充分考虑别的相机的特征，单使用事件相机进行人体姿态估计也存在一定的问题，事件相机只针对运动的部分产生事件，如果某一身体部分处于静止，那对这一部分的预测效果并不会很理想，再者，没有有效的考虑帧之间的信息，因此，在上述提到的部分，我们可以提出某些改进，以取得更好地预测结果。  可以针对模型中的特征提取模块，通过设计卷积注意力网络，提高光流特征的表示能力，然后在实验过程中，利用多模块的数据进行辅助工作，并合理融合事件数据特征和RGB灰度图像特征，通过适当位置融合图像帧，更好地解决事件相机对于人体静止部位获得较少的问题，以更准确的预测3维人体关节点的坐标，最后通过使用SMPL模型，实现三维人体运行建模，构建人体的形状和三维姿态。  此外，还可以考虑设计专门训练事件流数据的网络，例如脉冲神经网络，他的运行模式更加贴合事件相机获取数据的格式，用这种类型的网络去处理事件流数据是很有潜力的，这样就可以不进行事件流到事件帧的转化，直接将事件流输入到模型中进行训练，以达到仅使用事件相机获得的数据进行人体姿态估计的效果，这也是一个研究方向，只是对于SNN的使用需要进行一定的设计才有可能实现。 |
| **2．拟采取的研究方法、技术路线、实施方案及可行性分析**  首先下载好公开的事件流 3D 人体姿态估计数据集MMHPSD，然后分析数据集的构成，并用python对原始的数据流进行预处理，将事件流转换成RGB格式的事件帧，然后再使用事件帧的方式（或者自己设计更加恰当的事件流处理方式）生成神经网络需要的数据集与标签，接着对训练、验证以及测试需要使用的数据进行清晰地划分。由于MMHPSD数据集有30G的大小，必须进行足够的分析，选定合适的样本进行训练。  准备好数据集后，我选择通过使用 Pytorch 深度学习框架搭建骨干网络，并且根据数据的特点尝试设计合适的深度学习网络结构，并在合适的地方添加空间注意力机制。此外，我们还计划使用SMPL中的f和m参数模型直接重构出人体的三维形状和状态。在搭建好模型训练与测试的整个框架后，我将代码和使用的数据集上传到实验服务器上，然后在服务器中训练模型。在模型训练完毕之后，需要和改进之前的方法进行性能的对比，同时也需要在别的数据集上进行测试，验证一下我们的方法是否适用于别的事件相机的数据集，包括预测值与真值的平均关节接近程度误差（MPJPE）、预测的关键点与真实关键点之间的比例一致性（PCK）的检验、模型预测的顶点与真实顶点之间的可见性差异（PVE）的计算、模型的训练速度以及模型推理速度，这些代表着模型的准确性指标和实时性指标。为了验证我们模型设计的有效性，和保证结果的可信度，我们需要在相同的实验环境下进行对比实验，同时对自身设计的网络中的一部分环节做消融实验，以便于验证我们提出的方法的有效性，展示网络设计的性能，如果性能有明显的提升，再从理论上分析内在原因。  实验过程中比较难以进行的部分就是网络结构的设计和代码的整体实现，我计划使用开源的深度学习框架进行创新，学习和改进已有的方法，明白其运行机制，然后在此基础上进行修改和创作，这样既可以高效完成网络框架，又可以大幅度的避免逻辑错误，能让实验进行的更加顺利，在对比实验中，直接选择开源的代码，这样在调试程序阶段是可行的，再者就是进行消融实验，这个环节中只要清晰的知道自己改动和添加的模块，如何来控制变量，就可以顺利进行实验，还有就是设置合适的事件数据与RGB数据结合的形式，是选择首帧插入还是中间某些位置进行插入，是隔几段进行插入，还是没有RGB的图像对效果影响不大，这些都需要实验来验证。 |

四、研究条件

|  |
| --- |
| 1．所需实验手段、研究条件和实验条件  （如果没有，可以空白）  硬件环境： RTX 3090  操作系统：Windows 10 64bit  开发语言：Python  Python版本: 3.8  深度学习框架：Pytorch  Torch版本：torch 1.16  2．所需经费，包含经费来源、开支预算（工程设备、材料须填写名称、规格、数量）  （如果没有，可以空白） |

五、工作计划

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 序号 | 阶段及内容 | 工作量估计  （时数） | | 起止日期 | 阶段研究成果 |
| 1 | 查找文献，阅读与事件相机和人体姿态估计有关的论文，学习并实操开源代码。 | 500 | | 2023年10月～2024年1月 | 相关论文学习和理解代码结构。 |
| 2 | 复现经典算法，并分析经典模型的优缺点。 | 300 | | 2024年1月～2024年4月 | 记录并总结各个方向，寻找创新点。 |
| 3 | 根据最新研究，设计实验，提出创新方案，完成自己的网络设计，继续阅读相关论文。 | 600 | | 2024年4月～2024年6月 | 完成网络流程设计，并制定实验计划。 |
| 4 | 在公开数据集上进行实验测试，调试自己的网络。 | 400 | | 2024年6月～2024年9月 | 获得实验结果，并证明自己方法发的有效性。 |
| 5 | 撰写毕业论文初稿。 | 800 | | 2024年9月～2025年4月 | 完成毕业论文初稿。 |
| 6 | 修改定格毕业论文，并准备毕业答辩。 | 400 | | 2025年4月～2025年6月 | 完成毕业论文终稿，并完成毕业答辩。 |
|  |  |  | |  |  |
|  |  |  | |  |  |
|  |  |  | |  |  |
| 合计 | 3000 |