

黄氏春秋

本纪 其壹 三维重建

凡历肆载有余

- SFM 者，稀疏重建也，己亥年始，余为 VPS 之际亦始工于 SFM 与 Reloc。当是时，余乃深研 Colmap 之 triangulator、batch_mapper、mapping 模块，此后乃定余等之法。及辛丑年，又始再深工 BA 之迭代次数而成速度之精进。及于此间，余又以 BA 为机，将 ceres 之 Autodiff 换为解析求导而成速度之升。辛丑年末，余再深研 PixSFM 并精其 desc 与 reference 之模块，又以 Featuremetric BA 入于框架。至壬寅年初，复将 Camera-Rig 用于 Geometric 与 Featuremetric BA 之中。其后数月，乃分而精进 track 组织、outer iteration、pose graph 诸多模块，乃至此，余已甚谙 Colmap。壬寅年末，余又精进 BA，以 imu 之约束入于 Batch BA 之中。癸卯年，余亦深研 DROID-Slam，乃悟此法之精度甚高，乃以其与后作 DPVO 相合，共为吾等短序列建图之器。
- MVS 者，稠密重建也，始于己亥年。余初鉴 Kinect-Fusion 等乃为较好之重建。及辛丑年，余于 Colmap 之 Patch_Match_Stereo 模块又加 PMVS 之 3D Patch Model 乃得更佳之果。阳月，余乃比深度方法 neural recon、mobile recon 与现有传统方法异处，后乃以 depth 恢复协同为重建之功。壬寅年二月，又精进 sgmm 算法。四月，成多帧深度图之融合，此亦借传统与深度方法之相融。
- Relocalization 者，实乃高精度重定位也。己亥年，余乃于为 VPS 时便同开 SFM 与 Reloc 之研学。彼时，以 Hloc 为基，余乃构 Reloc 之框架并精进其重复匹配、一对多、pose cluster 部分而令其可行动态调整。及至辛丑年，为消抖动之患，余乃设计一精巧之 Filter，以此方除此患。十一月，余又深研 PixSFM，除以 fba 与 ka 入于建图模块，亦于 PnP 之际更添 qka 与 qfba 并精之。壬寅年，余所为者共二，一者于深研现有之 PnP 及 RanSaC 法后乃定余之新法。余所为者乃多目 CPnP+VSaC_MagSaC+NeFSaC 之法，其非线性优化乃以 MagSaC 得 inlier 之概率乃为 BA 之权重。二者乃于 postprocess 谋精进之法，余以多帧信息而使 pose 平滑、准确更甚，复于定位时加 Camera-Rig。癸卯年，余又重设定位之框架，此法乃充分利用所有性能，故提速数倍。
- NerF 者，实为余所触之最添趣味者之一也。壬寅年十月，余试以 BarF 为 Photogrammetry App。方是时，余忽悟可以 NerF 优化 SFM pose。癸卯年，及深研 DROID-Slam 之际乃习 DIM-Slam、Nicer-Slam 并初做复现。
- 太史公曰：凡三维重建者，无论何域，余皆有所涉猎，最深谙者自为 Reloc 与 SFM。余既可为各模块协同之协调，又可于其一者深研而获可观之收益。

本纪 其贰 Retrieval & Matching(特征匹配)

共历伍载有余

- Image Matching 者，亦称特征匹配，乃三维重建最要之处，于此学之研乃始于己亥年。六月，余于 TCL 实习，乃于 feature 实时性为深度改进，终成彼时业内最快之 feature。七月，与恩师章公国峰共探 feature，章公虽为前辈，亦存未知之惑，故而于吾深为赞叹。庚子年，余于 Jmgo 坚果深研 Superpoint 与 SIFT，乃成鱼眼灰度-原图之匹配，期间亦于线特征初有所学。辛丑年，余于 XREAL（彼时 Nreal）亲躬乃成两套 Image Matching 之框架。基传统之法，乃将 NN 匹配之 Ratio Test 改为 Blob 匹配，更加 FGINN 修正，而 SIF 亦为 DOG+SOSNet+AffNet+OriNet，此法既实时又准确。而于深度方法所研，余又加 homo_adapt 及 crop 匹配而使 SuperGlue 之匹配能力亦更强更准。此外，余亦得 Matching 之跨域匹配，可用鱼眼灰度低清图匹高清图 RGB 图。壬寅年，四月，余又评测而今数十 feature，乃明各法之优劣所在，故此方提自适应 feature。六月，余幸得特征匹配赛全球亚军。十月，余又引混合特征、homo_aug、keypoint adjustment 而提其精度。十二月，余又完 SuperPoint 与 SuperGlue 之重训，引 hard-mining、linear-attention、flash attention 而用之。癸卯年，余于 E2E 之法中又引定位模块与建图模块而成 Matching 大一统。癸卯年，于 GlueStick 与 limap 未现前，余亦提出思路，用线特征引入方法。太史公曰：忆往昔五载，余常于 Image Matching 而奋加研学，故得持续精进之效。有赖余之精学，于前沿方法之深谙，于 trick 之运用，故深度学习方法抑或传统方法，华夏出于吾右者不足十人也。
- Image Retrieval 者，VPS 之基础也。庚子年始余乃深研此学，彼时余尝调研现有之方法。及至辛丑年，余始研 TopK 之取值方式，更言动态 TopK 之法，此法既可提速，其 Recall 亦甚高。壬寅年，余始于 Retrieval 加入 trick，余用 Query Expansion、Retrieval 合并之法，desc concatenate，TTA 等诸多之法。此外，余亦成量化、部署等加速之操作。癸卯年，余始接触实时之 Retrieval，方于 dbow 作深悉之解，更重训 freak、orb 等 desc 之码本，亦成降维之功。

- 丙申年：如月，结 orbslam 之习，及上庠第二载粗通 BA、三角化、EPnP，彼时余即象牙塔最通 slam 之人，故蒙副院之青睐得实习契机。凉月，于荷福以 gmapping 促 lidar 建立地图之成，故机器人队以其算法于省闹夺魁。
戊戌年：凉月，以本科生入独角兽之司驭势实习，为此司无二，久成 slam 图之续更，可汇逾十场景之图，且每图之成率可堪 70%。于其间深研 orbslam 地图点之存取及 essential graph 等，由解地图点自融之难。此法后为此司所采之框架而沿用数年。
- 己亥年：焦月于 TCL 实习，深晓特征点及其所匹，成远逾 fast+brisk 之 feature。仲秋，习 vins，习得预积分、边缘化及初始化，因得优秀学员之誉。后于仲秋两参 slam 盛会，习 ice-ba。
- 庚子年：皋月习 lidar 并初晓 ndt 与 icp。及焦月，始任于坚果 jingo，手书 slam 测距，其差数不足毫厘，于其间精晓线特征并三位坐标之变换。至仲秋，受任筹光机-相机双目系统，竣超短焦投影仪自动对屏，深谙单应矩阵，遂解众人莫展之难。此法如切如磋，如琢如磨，时至今日，仍远胜别家。
- 辛丑年：杏月，余统率诸君设计 XREAL 之 VPS 建图及定位框架。荔月乃成 7-DoF 对齐，此法甚胜 Orb-Slam、Colmap、EVO 诸法。菊月，余群览多家之法乃深悉 Co-Slam 框架，后新更地图并之 Slam。此法比于 Orbslam3、MapLab，则更添高效准确之意。
- 壬寅年，余更深研 Benchmark。及悟数十精度 1cm 内之数据集并与 Benchmark 系统相连，余乃言多种获取误差于 1cm 内之自采数据 GT pose 法并以实试验之。其后，乃得 VPS Reloc 与 Image Processing 模块之 Benchmark。

世家 其肆 深度学习

历伍载

- 深度学习之学始自戊戌年，幸而习得吴恩达 cs229 及周志华《机器学习》。庚子结业之际，成 DenseNet 对抗攻击之作，粗通 DenseNet 及 FGSM 对抗攻击算法。后习三维点云处理，统涉 PointNet++，聚类（dbscan、kmeans）并 PCA。庚子仲秋月，深攻自动对屏算法之深度学习特征点及其部署。既辛丑年，余于深度研学更有全新之大得。于 feature，余以 superpoint 的重练而告功成，复以 R2D2、SuperPoint 及 DISK 为基成全新 feature 之框架。后余又深精 Image Retrieval 之 Deep Image Retrieval，成远逾 NetVLAD 之成效。更有 PV-Net 等 work 之复现。
- 壬寅年，余以微薄之力研训 SuperPoint 及 SuperGlue，此二者比于 pretrained，快、准、稳更胜之，亦可于任一 feature 研训 SG。再者，余更尝以深研之法试于 linear attention、Quadtree attention，以图 attention 层之精进。
- 吴公克艰曾赞，“工 slam 之人最善深研治学者”

列传 其伍 开发经验

捌年有余

- 余尤擅 C++ 之法，既往之六载，余尝历十余千行项目，亦于各 oj 网站恬留微名。于 python 万余行代码量，深度学习、数据处理与爬虫，余更深有所得，因尤擅之也。至 ide，余于 vim 得 emacs，余尤嗜纯粹之代码指令，更视 emacs 为吾之上帝编译器。
- 余亦尝粗研安卓及 unity，因余所涉之 vps，故余常深研 opengl 坐标与 slam 坐标系对齐之问题，亦时以 unity 导新建之 map 及 mesh 入余之手机。余深谙其机制及各大框架组件，亦曾以一己之力钻研 app。
- 既入余所供职之所，余时于闲暇之时以爬虫与网络协议为乐，于吾电脑之上自制按键精灵与抓包工具，遂可抢购华为手机与贵州茅台，更曾借抖音评论察女友出轨与否，并以网络图片为己之壁纸。

表 其陆 表述之能及所得之奖

- 2022 **CVPR** Image Matching Challenge(IMC) top 2(Public)/total 643
- 2023 **CVPR** Image Matching Challenge(IMC) top 11/total 495(余兴参赛为之)
- **Kaggle** top 0.5%*1, top 2.5%*1
- Visual localization benchmark - Robocar v2 dataset top 1(Private)
- 英吉利语: Duolingo 120 分(合雅思 7 分, 托福 100 分)。GRE 330。

黄凯

教育经历

2015. 09 - 2020. 06	电子科技大学 格拉斯哥大学	电子信息工程专业 电子与电气工程	成都 格拉斯哥
---------------------	------------------	---------------------	------------

奖项

2022 CVPR Image Matching Challenge(IMC)	top 2(Public)/total 643
2023 CVPR Image Matching Challenge(IMC)	top 11/total 495 (兼职参赛)
Kaggle	top 0.5%*1, top 2.5%*1
Visual Localization Benchmark - Robocar v2 dataset	top 1(Private)

工作经历

2023. 10 - 至今	清华大学	科研助理	北京
---------------	------	------	----

用 nerf 进行环境重建, 并以此进行交互式仿真用于训练。现已实现将 nerf 按物体解耦、进行实时交互和物体自动放置, 以及对应的碰撞、遮挡等规则。

- **Paper1:** gaussian splatting 的 slam 及对应的 editor 仿真交互系统 (一作, 投稿预期: RSS)
- **Paper2:** 更准的 hloc: gaussian splatting 的 depth、featuremetric 渲染及重定位算法 (一作, 投稿预期: RSS)
- **Paper3:** 更快更准的 lightglue (一作, 投稿预期: ECCV)
- 二作三作若干, 包括: gaussian splatting active mapping、gauss 的机械臂抓取、diffusion 等
- **Idea 若干:** 包括: 基于 omnimotion 的 sfm、nerf 融合、nerf-slam 全局 ba 的稀疏化等

2020. 12 - 2023. 09	XREAL (Nreal)	算法工程师	北京
---------------------	---------------	-------	----

带领三人小组, 开发大规模建图的重定位系统 (visual positioning system-VPS), 使用 480p 建图, 同期效果好于河图在内的所有友商, 单帧精度仅次于河图, 时间单卡可达 5fps。项目中, 本人代码贡献量超 50%

算法改进

- **Retrieval:** 1. 速度上, 使用了量化剪枝等操作, 同时探索 TopK 的边界并提出自适应框架, 较好的提速。2. Recall 上, 评估现有 Desc 并混合, 同时增加额外 trick 提高召回率。
- **Matching:** 1. 重训了 SuperPoint 及 SuperGlue, 既比 pretrained 更快更准更稳, 也能在任意 feature 上训 SG。2. 将 End2End 方案加入框架中并混合使用 feature, 提升了性能。3. 调研线特征方案, 充分实验并在加入现有框架。4. 在端侧 NN 及云侧 SG 及 feature 均有小改进, 优化了效果。5. 对 Matching 的过程加入多种 trick, 优化了复杂环境下的表现及本身的精度。
- **Localizer:** 1. 在 HLoc 框架基础上进行了多次改进, 明确各问题边界, 获得比原生 HLoc 更好的基础框架。2. 在 pose 求解上, 探索所有 PnP (L) 及 RanSaC 方法, 得到目前 SOTA 方案, 同时也改进后续非线性 pose refinement 模块。3. 在 query 部分引入 PixSFM, 并作出此前未有的改进, 提升精度。
- **Post process:** 1. 引入多帧与多目信息, 并以此优化 2d-2d、2d-3d 匹配对, 提升鲁棒性和精度。2. 加入 filter, 效果显著, 能完全消除每一帧 query 的抖动且跟随延迟不会超过 1 秒。
- **建图:** 1. 引入了 FBA, 并分别对 Geometric-BA 和 Featuremetric-BA 加入了 Camera-Rig, 并合并使用达到 SOTA 效果。2. 对 Colmap 本身的几个模块进行改进, 使精度和实时性均有保证。3. 设计了全新的建图流程, 在短序列的建图上较大的提高了精度和时间。

框架设计

- 设计并开发了 VPS Reloc 与建图模块的全新框架, 模块解耦, 并能完全利用所有性能从而提速数倍。
- 设计并开发了独立的 Image Processing 模块, 统一了 E2E 和 2-Stage 的特征, 并融合 Retrieval 和预处理, 建立大一统框架。

Benchmark

- 设计并开发了 VPS Reloc 与 Image Processing 模块的 Benchmark，并对现有所有方法评测。
- 提出多种获取误差在 1cm 以内的自采数据 GT pose 方法，并部分验证。
- 调研公开数据集，找到精度 1cm 以内的数据集十余个并接入 Benchmark 系统。

兼容性

- 支持 VPS 的跨相机迁移，如手机相机建图-鱼眼 cv 定位，广角 cv-手机-专业微单混合建图等。
- 支持的相机列表：鱼眼 cv 相机、广角 cv 相机、手机相机、专业微单相机、全景相机。

2020.06 - 2020.12 Jmgo 坚果投影仪 算法工程师 深圳

无感自动对屏 (2020.8 - 2020.11)

- 完成安卓平台的超短焦投影仪的自动对屏及梯形校正算法并写入 framework 层，在单目且无需预设值的情况下能够使对屏达到 5mm 的误差，梯形校正 2mm 误差。
- 主要难点：1. 自底向上拍照形状无法确定 2. 拍照越往上越不精确 3. 特征点不好找 4. 相机和光机位置固定，无法生成 mappoint 5. 由于幕布与墙的深度不一致，打在幕布的光无法与打在墙面的光三维坐标完全不一致 6. 幕布边缘过细 7. 鱼眼镜头与广角镜头间的取舍
- 主要方案：梯形校正：利用光机-相机的伪双目系统，通过深度学习特征点找出相机-光机的单应矩阵，并迭代优化，分解得到墙面的法向量，并由此推出相机-用户的单应矩阵。自动对屏：用直线聚类找出幕布边框，将相机-用户的单应矩阵加入幕布的约束条件。

自动梯形校正结果检测 (2020.6 - 2020.8)

- 手写简易视觉 slam 松耦合系统完成 windows 平台测距软件，误差达到 0.1 毫米级，且在奔腾 4400 上能在 2s 内完成。
- 主要难点：1. 投影像素（棋盘格，圆点）的不精确 2. 画面虚焦及色差 3. 相机画面会实时变化 4. 投影画面边缘会随焦距改变而产生严重误差
- 利用 aruco 棋盘格+圆点的方式检测特征点，并用整体单应矩阵的方式增加棋盘格边缘的精度从而可以检测特征线进行匹配，中间用李代数单应矩阵最小化对称转移误差，并用 BA 优化空间坐标。

实习经历

2018.07 - 2018.09 驭势科技 视觉 slam 工程师 北京

- 负责 vslam 的地图更新，能增量式的存储并更新地图从而适应季节天气和光照变化，并保持地图容量在一定大小内。
- 根据初始化数十帧的数据在服务器端的总地图中选择适当的地图点和关键帧组合出最适合当前场景的地图，用组合出的 mappoint 和 keyframe 形成新地图以定位。并且用凸优化将总地图的容量限定在范围内。

2019.06 - 2019.08 tcl 工业研究院港研所 算法工程师 深圳

- 负责传统视觉算法的代码实现，在手机上实现实时视频降噪以及实时视频抠图。
- 在降噪任务中实现了用特征点固定视频的任务。利用 agast 配合哈希描述子进行快速特征匹配，从而整个降噪过程能实时运行高于 bm3d 的效果。其中所使用的特征点为行业内最快 feature，速度远高于 fast。