arXiv:2603.00990v1 [cs.CV] 1 Mar 2026

MLRecon：通过由粗到细位姿估计实现稳健的无标记自由手持三维超声重建

Yi Zhang[1],Puxun Tu[1],Kun Wang[1],Yulin Yan[2],Tao Ying[2],Xiaojun Chen[1,3]

[1] 上海交通大学机械工程学院生物制造与生命质量工程研究所，中国上海[2] 上海交通大学医学院附属上海市第六人民医院医学超声科，中国上海[3] 上海交通大学医疗机器人研究所，中国上海

摘要。 自由手持三维超声重建兼具标准二维探头的灵活性，并可实现体积成像；然而，现有跟踪范式面临严苛的三难困境：基于标记的系统成本高昂，由内向外方法需要侵入式安装传感器，而无传感器方法则存在严重的累积漂移。为克服这些局限，我们提出了**MLRecon**，一种稳健的无标记三维超声重建框架，利用单个商用RGB-D相机实现抗漂移的六自由度探头位姿跟踪。借助视觉基础模型的泛化能力，我们的流程可对探头进行连续的无标记跟踪，并结合视觉引导的发散检测器，自动监测跟踪完整性并触发故障恢复，以确保扫描不中断。尤为重要的是，我们进一步提出了一种双阶段位姿精化网络，能够显式分离高频抖动与低频偏差，在有效去噪轨迹的同时保持操作者操作的运动学保真度。实验表明，MLRecon显著优于其他无传感器及传感器辅助方法，在复杂轨迹上的平均位置误差低至0.88毫米，并生成平均表面精度达到亚毫米级的高质量三维重建。这为资源有限的临床环境中的低成本、易获取体积超声成像树立了新标杆。

关键词： 自由手持三维超声 · 无标记跟踪 · 位姿估计。

1 引言

超声（US）成像凭借实时能力、便携性和低成本，在图像引导介入中不可或缺。然而，传统二维采集对操作者的依赖性较高，且缺乏空间感知。因此，三维超声日益受到关注，因为它能提供更丰富的解剖背景，有助于体积评估、操作规划和诊断 [7, 14]。现有临床解决方案通常依赖专用三维探头或

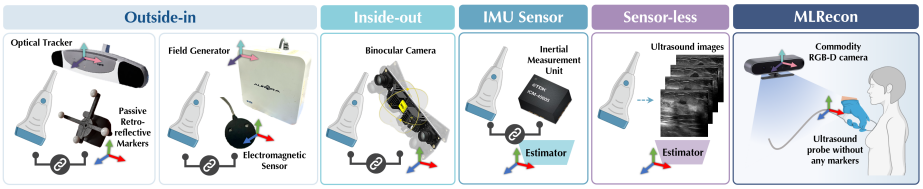


图 1. 自由手持三维超声中跟踪范式的示意对比。

机械扫描换能器，这些设备需要专用硬件，并限制了扫描范围。相比之下，自由手三维超声重建提供了一种颇具吸引力的替代方案：通过持续跟踪标准二维探头的六维位姿，并将 B 模式帧合成为连贯的体积表示，同时保留手持式超声检查的灵活性和易用性。

准确且稳健的探头位姿跟踪是实现高保真自由手持超声重建的基础，已有各种各样的解决方案（图 1）。基于标记的由外向内系统，包括光学 [18] 和电磁（EM） [6] 跟踪器，能够提供高精度，但成本高昂，且依赖专用基础设施。为消除外部跟踪器，由内向外范式将轻量级传感器直接安装在探头上。以惯性为核心的研究涵盖了从独立的惯性测量单元（IMU）方法 [15] 到用于位姿估计的深度 IMU+超声融合网络 [12,13,22]。结合视觉外感知，Buchanan等人 [1] 证明，来自探头安装式立体相机的视觉惯性里程计，其重建误差可与动作捕捉相媲美；He等人 [8]则通过李群上的无迹卡尔曼滤波器，将双目相机与 IMU 进行融合。Zhang 等人 [23]进一步提出了一种仿真在环视觉伺服方案，利用安装在探头上的单目相机观测带纹理平面工作空间。尽管这些由内向外方法无需外部跟踪设备，但需要在探头上实体安装传感器，并且在长距离轨迹中仍易受到增量漂移的影响。

另一方面，无传感器方法仅利用超声图像序列，通过深度学习预测帧间相对位姿 [9–11,21], ，从而无需任何辅助硬件。然而，这些方法存在显著的累积漂移，尤其是在临床实际应用中常见的非线性扫描路径上；而且，它们依赖学习到的散斑或纹理模式，天然限制了其在不同解剖部位和超声设备间的泛化能力。近期，Dai 等人 [5]通过将 N形尼龙标志物嵌入水凝胶耦合垫，使其在超声图像中留下空间标记，从而减轻了漂移；但该垫会对扫描表面造成物理限制，且仅适用于线阵探头。

理想的自由手三维超声跟踪方案应满足以下条件，以用于临床介入引导：(i)无标记，无需改装探头或在患者身上安装附件；(ii) 抗漂移，即便面对任意长且复杂的轨迹也能稳定工作；(iii)轻量且低成本，便于无缝集成到现有工作流程中。上述范式均无法同时满足这三项标准。

在本研究中，我们提出了**MLRecon**，这是一种完全无标记的自由手三维超声重建框架，旨在弥合这一差距。我们的主要贡献有两点：(i) 基于基础模型的探头位姿估计与跟踪流程，配备视觉引导的发散检测器，可实现自动故障恢复，从而在自由手持扫描过程中持续跟踪。(ii) 一种双阶段时序位姿优化网络，分别解耦并消除高频抖动与低频偏差，在忠实保留真实探头动态特性的同时，大幅降低位姿最大偏差。

2 方法

如图2所示，我们使用一台普通的 RGB-D相机（Astra 2，奥比中光，30 Hz）作为外部跟踪设备，在扫描过程中观测临床超声探头（迈瑞7L4P，5.0–10.0MHz）。所提出的流程包含一系列基于基础模型的模块，以实现鲁棒的探头位姿估计与跟踪。我们分别将 RGB-D相机、超声探头和超声图像的坐标系记为 $\{C\}$、$\{P\}$ 和 $\{I\}$。

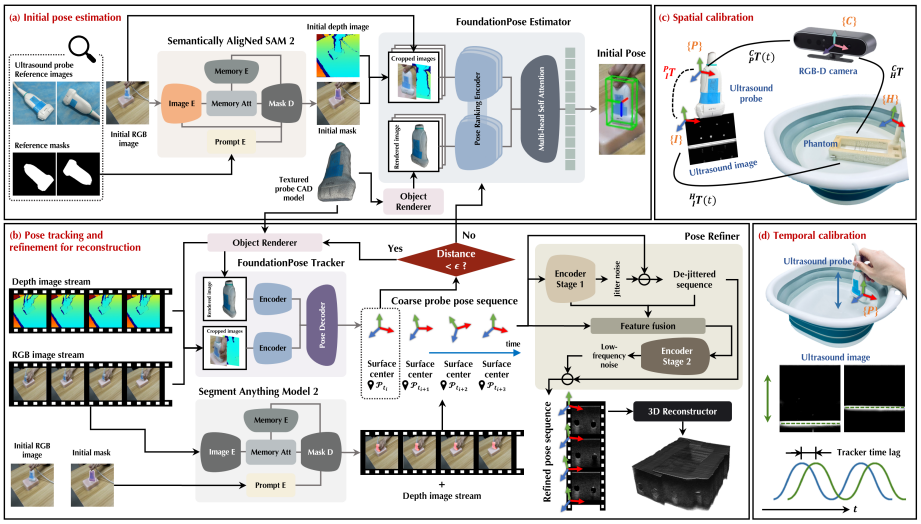


图2。MLRecon 概览。该框架集成了带故障恢复的鲁棒跟踪、双阶段频率感知位姿优化，以及经校准的三维复合。

2.1 基于基础模型的位姿估计与跟踪

为实现自由手持重建的无标记跟踪，我们的方法采用基于渲染与比较范式的方式，直接从 RGB-D视频流估计六维探头位姿 ${}^{C}_{P}\hat{\mathbf{T}} \in \mathrm{SE}(3)$。

该范式基于 FoundationPose [19]。需要解决两个关键挑战：在无需人工标注的情况下获取可靠的初始位姿，以及检测并从不可避免的跟踪失败中恢复。

初始位姿估计。给定从不同视点拍摄的 K 探头 {带标注参考图像 $\{(\mathbf{I}_r^k, \mathbf{a}_r^k)\}_{k=1}^K$，我们采用 SANSA（ [4], 一种经过语义对齐的 SAM 适配方法 2 [16], ），将目标掩码传播至实时输入的第一帧 RGB 图像 $\mathbf{I}_{r,\ t_0}$ 无需人工标注。随后，将预测的掩码 $\mathbf{S}^{t_0}$、初始 RGB-D 观测以及探头的预扫描 CAD 模型 $\mathcal{M}$ 输入 FoundationPose。该模型先进行全局位姿采样，再通过迭代式渲染与比较优化，得到初始位姿 ${}^{C}_{P}\hat{\mathbf{T}}_{t_0}$（图2a）。

结合发散检测与恢复的鲁棒跟踪。初始化后，FoundationPose 以 30 Hz 的频率在输入的 RGB-D视频流中逐帧跟踪探头位姿。然而，刚性物体的视觉跟踪容易因大范围遮挡、深度传感器噪声或探头突然运动而暂时失效。因此，我们引入一种视觉引导的跟踪发散检测器，用于监测跟踪位姿与独立视觉线索之间的一致性，并在两者不一致时触发自动重新初始化。

在运行 FoundationPose 跟踪的同时，我们以较低的频率（$\sim$3Hz）运行 SAM2，以分割当前 RGB 帧中的探头，生成一个二值掩码 $\mathbf{S}^{t_i}$。根据此掩码和对齐深度图 $\mathbf{D}^{t_i}$，我们将掩码内的深度像素反投影到三维空间，计算视觉质心 $\mathbf{c}_{t_i}^{\mathrm{vis}}$:

$$\mathbf{c}_{t_i}^{\mathrm{vis}} = \frac{1}{|\mathcal{P}|} \sum_{(u,v)\in\mathcal{P}} \mathbf{D}_{t_i}(u,v)\, \mathbf{K}^{-1} \begin{pmatrix} u \\ v \\ 1 \end{pmatrix}, \tag{1}$$

其中，$\mathcal{P} = \{(u,v) \mid \hat{\mathbf{S}}^{t_i}(u,v) > 0, \mathbf{D}^{t_i}(u,v) \in [D_{\min}, D_{\max}]\}$ 表示有效掩码像素的集合，$\mathbf{K}$ 为深度相机内参矩阵。同时，根据对应时间戳处的 ${}^{C}_{P}\hat{\mathbf{T}}_{t_i}$ 得到跟踪探头质心$\mathbf{c}_{t_i}^{\mathrm{trk}}$。

当欧氏差异超过自适应的、考虑物体特性的阈值时，即判定跟踪发生偏离:

$$\|\mathbf{c}_{t_i}^{\mathrm{vis}} - \mathbf{c}_{t_i}^{\mathrm{trk}}\|_2 \ > \ \epsilon \ = \ \frac{\|\mathbf{d}_{\mathcal{M}}\|_2}{2}\,\eta \ + \ \delta_0, \tag{2}$$

其中，$\mathbf{d}_{\mathcal{M}} \in \mathbb{R}^3$ 表示 $\mathcal{M}$ 的包围盒尺寸，η 为几何缩放因子，δ_0 为固定容差。在 t_k 检测到偏离后，跟踪器将暂停；利用当前 SAM 2 掩码和 RGB-D 帧重新初始化位姿，随后跟踪便会无缝恢复。t_k 与前一次检查之间的位姿将标记为缺失。这一闭环机制对超声检查医师透明，扫描期间无需人工干预。

2.2 双阶段位姿精化

即使经过鲁棒跟踪，粗略位姿序列 ${}^{C}_{P}\hat{\boldsymbol{\mathcal{T}}} = \{{}^{C}_{P}\hat{\mathbf{T}}_{t_1}, \ldots, {}^{C}_{P}\hat{\mathbf{T}}_{t_L}\}$仍不可避免地呈现出两类会降低重建质量的不同误差：(i)高频抖动 源于逐帧深度噪声和独立的网络预测；(ii) 低频偏差 主要来自

自回归初始化会传播校正不足的残差。经典滤波器仅能处理其中一种情形，因而可能过度平滑真实运动，或使相关偏差得不到解决。

为解决这一问题，我们提出了位姿精化器，这是一种卷积时序网络，通过两阶段残差架构（图2b）显式解耦并消除这两种误差分量。每个位姿 $\overset{C}{P}\hat{\mathbf{T}}_{t_i}$ 都参数化为 $\mathbf{x}_{t_i} = [\mathbf{r}_{t_i}; \mathbf{p}_{t_i}] \in \mathbb{R}^9$，其中 $\mathbf{r} \in \mathbb{R}^6$ 是连续旋转表示 [25]，$\mathbf{p} \in \mathbb{R}^3$ 是平移。原始序列通过仿射映射 $\tilde{\mathbf{X}} = \boldsymbol{\Lambda}^{-1}(\mathbf{X} - \bar{\mathbf{x}}_{\mathrm{ref}})$进行归一化，其中 $\boldsymbol{\Lambda} = \mathrm{diag}(\mathbf{I}_6, s_p\mathbf{I}_3)$，偏移量为 $\bar{\mathbf{x}}_{\mathrm{ref}} = [\mathbf{0}_6; \mathbf{p}_{\mathrm{ref}}]$，其中 $\mathbf{p}_{\mathrm{ref}}$是首帧位置，$s_p$ 为空间缩放因子，从而得到 $\tilde{\mathbf{X}} \in \mathbb{R}^{9\times L}$。

阶段 1 采用膨胀时序卷积编码器 $\mathcal{E}_1$，并限制膨胀率 $\{1, 2, 4, 8, 16\}$ 以分离局部抖动模式：

$$\hat{\mathbf{n}}^{\mathrm{hf}} = \mathcal{E}_1(\tilde{\mathbf{X}}), \quad \tilde{\mathbf{X}}^{(1)} = \tilde{\mathbf{X}} - \hat{\mathbf{n}}^{\mathrm{hf}}. \tag{3}$$

随后，轻量级融合层将 $\tilde{\mathbf{X}}$、去抖后的信号 $\tilde{\mathbf{X}}^{(1)}$ 和 Stage1 语义特征拼接起来，形成 $\mathbf{F}_{\mathrm{fuse}}$，为后续阶段提供互补的多尺度线索。

Stage2 采用另一个编码器 $\mathcal{E}_2$，其膨胀率大幅提高$\{1, 2, \ldots, 128\}$，感受野覆盖整个序列，从而能够捕捉残余的低频偏差。通过对称的反归一化恢复精炼后的输出：

$$\hat{\mathbf{n}}^{\mathrm{lf}} = \mathcal{E}_2(\mathbf{F}_{\mathrm{fuse}}), \quad \mathbf{X}^{\star} = \boldsymbol{\Lambda}\big(\tilde{\mathbf{X}}^{(1)} - \hat{\mathbf{n}}^{\mathrm{lf}}\big) + \bar{\mathbf{x}}_{\mathrm{ref}}. \tag{4}$$

训练目标。该网络使用来自上海第六人民医院医学超声科的 243 次扫描（∼360k 帧）进行训练，将模拟的含噪 RGB-D 姿态与干净的光学跟踪序列配对。组合损失对$\mathbf{X}^{\star}$ 和$\mathbf{X}^{(1)}$进行联合监督：

$$\mathcal{L}_{\mathrm{comp}} = \lambda_g\mathcal{L}_{\mathrm{geo}} + \lambda_l\mathcal{L}_{\ell_1} + \lambda_v\mathcal{L}_{\mathrm{vel}} + \lambda_f\mathcal{L}_{\mathrm{freq}}, \tag{5}$$

其中，$\mathcal{L}_{\mathrm{geo}}$ 是三维特殊正交群 SO(3) [17]上的测地距离，$\mathcal{L}_{\ell_1}$ 是旋转和平移表示的 L1误差，$\mathcal{L}_{\mathrm{vel}}$ 惩罚时间导数差异，以保持运动动力学；$\mathcal{L}_{\mathrm{freq}}$ 则通过实数快速傅里叶变换确保频谱幅值一致性，防止真实运动频率过度平滑。

2.3 标定与三维超声复合

为将每个B型超声像素映射到三维空间，我们通过改进的N线体模标定方案确定从超声图像坐标系 $\{I\}$ 到探头坐标系 $\{P\}$ 的空间变换$\overset{P}{I}\mathbf{T}$ （图2c）。时间标定通过最大化准周期往复运动的互相关来对齐超声和相机数据流 [18] （图2d）。在三维复合中，每个B型超声像素$p=(u, v)$都会映射到物理坐标$\mathbf{x}_{t_i} = \overset{C}{P}\hat{\mathbf{T}}^{\star}_{t_i}\,\overset{P}{I}\mathbf{T}\,(u, v, 0, 1)^T$——其中$\overset{C}{P}\hat{\mathbf{T}}^{\star}_{t_i}$ 表示第2.2节中的精炼姿态——并通过分箱填充分配至最近的体素，空体素则通过梯度感知孔洞填充进行图像修复 [20]。

3 实验与结果

在我们的实现中，MLRecon已集成到现有的超声数据采集与重建软件中 [24], , 其中FoundationPose和SAM 2通过C++ TensorRT加速管线部署在配备NVIDIA RTX 3080 Ti的笔记本电脑GPU上，支持最高140 Hz的跟踪。用于三维复合的超声图像-位姿对以20 Hz采集。SANSA初始化所用的参考图像采集自 K=2 个视点。对于发散检测，我们设置 η=0.8和 δ_0=30毫米。该检测器在涉及探头短暂完全不可见或快速运动的10次专门试验中均成功触发重新初始化，证明了恢复机制的可靠性。所提出的位姿精化器使用AdamW训练1000个周期（学习率 $= 10^{-3}$，余弦退火），批量大小为32（s_p=100, λ_g=1.0, λ_l=5.0, λ_v=3.0, λ_f=0.1）。位姿估计与重建的实验设置如图3a所示。

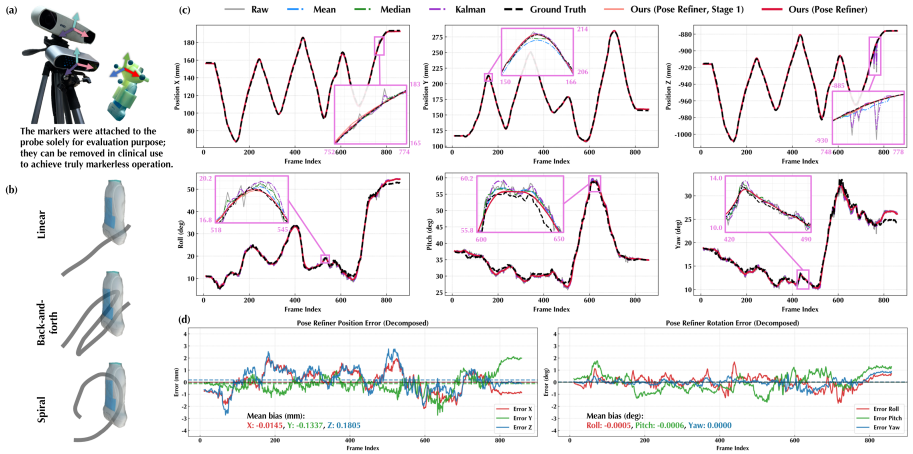


图3. (a) 实验设置。(b) 三种扫描轨迹模式。(c) 消融研究中的代表性平移（X/Y/Z）和旋转（滚转/俯仰/偏航）时间序列。(d) 双阶段位姿优化后各自由度（DOF）的位姿误差。

位姿精度。 我们在三种具有代表性的徒手轨迹（图3b）上评估了 MLRecon, 并与现有方法进行了比较。评估指标包括最终漂移率（FDR）、平均漂移率（ADR）、最大漂移（MD）、最大旋转误差（MRE）、平均位置误差（APE）、平均旋转误差（ARE）和总距离（TD）。在同等范围内，针对线性扫查比较了无传感器方法和 IMU+US 融合方法 [5, 11, 12]，而针对往返轨迹和螺旋轨迹比较了其他传感器辅助方法 [1, 2, 8, 23]。所有 MLRecon 结果均为 5 次试验的平均值。

如表1所示，在 7.6× 更长的轨迹上，与表现最佳的其他无传感器方法相比, MLRecon在线性扫查中实现了 7.6× 更低的FDR和 12.4× 更低的ADR,

表1. 三种扫描轨迹模式下的位姿精度均值（标准差）。

Mode	方法	FDR(%)	ADR(%)	均值距离(mm)	均值旋转误差(deg)	平移距离(mm)
线性	RecON [11]	25.90(18.33)	30.36(13.95)	–	–	∼80
	MoNet [12]	12.75（9.05）	19.05（11.46）	–	–	∼80
	Dai 等人 [5]	2.74(2.98)	3.35(3.24)	2.52(2.25)	–	∼80
	MLRecon	**0.36(0.15)**	**0.27(0.05)**	**1.85(0.33)**	**1.12(0.14)**	∼250

Mode	方法	绝对位置误差(mm)	绝对旋转误差(deg)	均值距离(mm)	均值旋转误差(deg)	平移距离(mm)
Back and 向前	Buchanan等人 [1]	∼4.00	–	–	–	∼360
	He等人 [8]	3.78	**0.36**	–	–	∼800
	ORB-SLAM [2]	2.65(0.74)	1.99(1.99)	–	–	–
	Zhang 等人 [23]	1.46(0.70)	1.37(0.57)	2.99	2.28	∼520
	MLRecon	**0.88(0.45)**	0.48(0.23)	**2.37(0.25)**	**1.33(0.22)**	∼560
螺旋	Zhang 等人 [23]	1.98(1.43)	1.57(0.82)	5.72	3.25	∼400
	MLRecon	**1.44（0.69）**	**0.92（0.34）**	**3.29（0.52）**	**1.75（0.15）**	∼480

此外，它在往返轨迹和螺旋轨迹上的位置误差最低，优于所有参与比较的由内向外方法，同时旋转精度也具有竞争力。最大偏差大幅降低，进一步证明了我们方法的稳定性。

位姿精化消融研究。表 2 对比了位姿精化器与经典信号处理基线方法（卡尔曼滤波器、均值滤波器和中值滤波器），以及仅第一阶段变体；所有方法均应用于相同的原始 FoundationPose 序列，并进行了 5 次自由扫描试验（三种模式的随机组合）。

原始序列已经取得了可接受的平均误差，但偶尔会出现较大的尖峰。均值滤波器和卡尔曼滤波器均减小了均值距离（MD），但由于离群值扩散和刚性统计先验，平均位置误差或旋转误差有所升高。仅使用阶段 1 就带来了显著改善，加入阶段 2 后进一步将 MD 降至 3.73 mm，并实现了最佳的旋转稳定性（均值旋转误差（MRE）为 1.88°），表明两阶段分解至关重要。图 3c–d 展示了轨迹保真度，证实在全部六个自由度（DOFs）上均接近零均值偏差。

表 2. 位姿精化消融研究的均值（标准差）结果。

	绝对位置误差（APE）（mm）	绝对旋转误差（ARE）（deg）	均值距离（MD）（mm）	均值旋转误差（MRE）（deg）
Raw	1.46(0.10)	0.75(0.09)	25.92(9.34)	2.63(0.40)
卡尔曼	1.46(0.10)	0.78(0.09)	18.82(6.08)	2.42(0.33)
Mean	1.79(0.16)	**0.70(0.09)**	6.05(1.05)	2.21(0.42)
中位数	1.45(0.06)	0.71(0.09)	4.91(0.69)	2.23(0.31)
位姿精化器 (S1)	1.41(0.07)	0.71(0.08)	3.90(0.24)	2.04(0.14)
位姿精化器 (S1+S2)	**1.39(0.06)**	0.71(0.07)	**3.73(0.37)**	**1.88(0.19)**

三维重建精度。为评估我们方法的实用性，我们对三个仿体进行了三维超声重建：两个矩形组织仿体（仿体 A 和 B，分别包含 8 个病灶和 4 条血管）

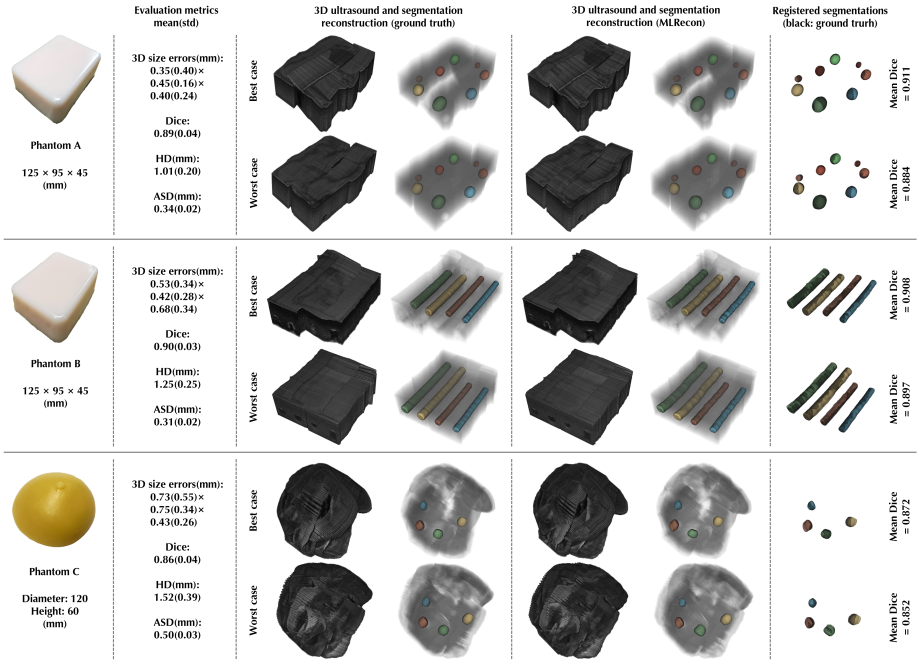


图4. 三种仿体的重建定量与定性比较。

以及一个乳房形体模（体模 C，包含 4 个病灶），每个仿体均进行了 3 次重复试验。所有重建结果均与通过 NDI 光学跟踪获取的真实体积进行比较，采用 [24] 中详述的基于配准的方案。评估指标包括沿主轴方向的三维尺寸误差、配准后的病灶或血管之间的 Dice系数、豪斯多夫距离（HD）和平均表面距离（ASD）。

如图4所示，在最佳和最差情况下，我们的方法重建的三维体积和分割结果均与真实形态高度吻合，所有仿体的 Dice 系数均在 0.85 至 0.91 之间。即使对于表面不平整的体模C，三维尺寸误差和平均表面距离（ASD）仍保持在亚毫米级精度，这表明扫描不同身体几何形状对重建质量的影响很小。

4 结论

在本研究中，我们提出了MLRecon，一种融合视觉基础模型与单台商用 RGB-D相机的无标记徒手三维超声重建框架。通过结合闭环发散检测与频率感知位姿精化，该系统在复杂轨迹上达到了最先进的精度，并在多个仿体上提供高保真体积表示。至关重要的是，其无标记、轻量校准工作流程有助于

无缝临床集成，且无需改变常规扫描实践，从而将其定位为光学或电磁跟踪的稳健且经济高效的替代方案，在实现可及的即时诊断方面具有重要前景。

参考文献

1. Buchanan, R., Tu, S.J., Camurri, M., Mellon, S.J., Fallon, M.：利用视觉惯性和深度惯性里程计进行自由手持三维超声成像，以测量髌骨跟踪。在：2024年IEEE医疗测量与应用国际研讨会（MeMeA）。第1–6页。IEEE（2024） 2. Busam, B., Ruhkamp, P., Virga, S., Lentes, B., Rackerseder, J., Navab, N., Hennersperger, C.：用于三维超声复合的无标记内向外跟踪。在：即时超声国际研讨会。第56–64页。施普林格（2018） 3. Carbajal, G., Lasso, A., Gómez, Á., Fichtinger, G.：改进基于n线模体的自由手持超声标定。《国际计算机辅助放射学与外科学杂志》**8**，1063–1072（2013） 4. Cuttano, C., Trivigno, G., Averta, G., Masone, C.：SANSA：释放SAM2中隐藏的语义信息，实现小样本分割。arXiv预印本 arXiv:2505.21795（2025） 5. Dai, L., Zhao, K., Li, Z., Zhu, J., Liang, L.：利用新型耦合垫推进无传感器徒手三维超声重建。在：医学图像计算与计算机辅助干预国际会议。第559–569页。施普林格（2024） 6. Daoud, M.I., Alshalalfah, A.L., Awwad, F., Al-Najar, M.：采用电磁跟踪的自由手持三维超声成像系统。在：2015年开源软件计算国际会议（OSSCOM）。第1–5页。IEEE（2015） 7. Fenster, A., Downey, D.B., Cardinal, H.N.：三维超声成像。《医学与生物学物理》**46**(5)，R67–R99（2001） 8. He, W., Zhao, B., Zhou, Y., Wu, R., Wu, G., Li, Y., Lu, M., Zhu, L., Gao, Y.：基于探头搭载视觉与惯性测量单元系统的自由手持三维超声成像。《医学与生物学超声》**50**(8)，1143–1154（2024） 9. Lee, S., Kim, S., Seo, M., Park, S., Imrus, S., Ashok, K., Lee, D., Park, C., Lee, S., Kim, J., 等：利用深度学习增强自由手持三维光声与超声重建。《IEEE医学影像汇刊》（2025） 10. Li, Q., Shen, Z., Li, Q., Barratt, D.C., Dowrick, T., Clarkson, M.J., Vercauteren, T., Hu, Y.：在无外部跟踪器的情况下实现自由手持超声三维重建的长期依赖关系。《IEEE生物医学工程汇刊》**71**(3)，1033–1042（2023） 11. Luo, M., Yang, X., Wang, H., Dou, H., Hu, X., Huang, Y., Ravikumar, N., Xu, S., Zhang, Y., Xiong, Y., 等：RecON：用于无传感器徒手三维超声重建的在线学习。《医学图像分析》**87**，102810（2023） 12. Luo, M., Yang, X., Wang, H., Du, L., Ni, D.：用于自由手持三维超声重建的深度运动网络。在：医学图像计算与计算机辅助干预国际会议。第290–299页。施普林格（2022） 13. Luo, M., Yang, X., Yan, Z., Cao, Y., Zhang, Y., Hu, X., Wang, J., Ding, H., Han, W., Sun, L., 等：Monetv2：增强型自由手持三维超声重建运动网络。《IEEE神经网络与学习系统汇刊》（2025）

14. Mozaffari, M.H., Lee, W.S.：自由手持三维超声成像：系统综述。《医学与生物学超声》**43**(10)，2099–2124（2017） 15. Rahni, A.A.A., Yahya, I., Mustaza, S.：基于六自由度微机电惯性测量单元的姿态进行自由手持三维超声扫描的二维转换。在：第四届吉隆坡国际生物医学工程会议2008：BIOMED 2008，2008年6月25–28日，马来西亚吉隆坡。第699–702页。施普林格（2008） 16. Ravi, N., Gabeur, V., Hu, Y.T., Hu, R., Ryali, C., Ma, T., Khedr, H., Rädle, R., Rolland, C., Gustafson, L., Mintun, E., Pan, J., Alwala, K.V., Carion, N., Wu, C.Y., Girshick, R., Dollár, P., Feichtenhofer, C.：SAM 2：分割图像和视频中的任意对象。arXiv预印本 arXiv:2408.00714（2024） 17. Salehi, S.S.M., Khan, S., Erdogmus, D., Gholipour, A.：采用测地线损失进行实时深度姿态估计，以实现图像到模板的刚性配准。《IEEE医学影像汇刊》**38**(2)，470–481（2018） 18. Treece, G.M., Gee, A.H., Prager, R.W., Cash, C.J., Berman, L.H.：高清自由手持三维超声。《医学与生物学超声》**29**(4)，529–546（2003） 19. Wen, B., Yang, W., Kautz, J., Birchfield, S.：FoundationPose：新颖物体的统一六维位姿估计与跟踪。在：IEEE/CVF计算机视觉与模式识别会议论文集。第17868–17879页（2024） 20. Wen, T., Zhu, Q., Qin, W., Li, L., Yang, F., Xie, Y., Gu, J.：一种用于自由手持三维超声重建的准确有效的基于快速行进法的方法。《生物医学信号处理与控制》**8**(6)，645–656（2013） 21. Wilson, P.F., Ronchetti, M., Göbl, R., Markova, V., Rosenzweig, S., Prevost, R., Mousavi, P., Zettinig, O.：DualTrack：无传感器三维超声需要局部与全局上下文。在：医学超声简化技术进展国际研讨会。第3–12页。施普林格（2025） 22. Yan, Z., Yang, X., Luo, M., Chen, J., Chen, R., Liu, L., Ni, D.：用于自由手持三维超声重建的细粒度上下文与多模态对齐。在：医学图像计算与计算机辅助干预国际会议。第340–349页。施普林格（2024） 23. Zhang, Y., Huang, D., Meng, M.Q.H., Navab, N., Jiang, Z.：自由手持三维超声成像：通过视觉伺服实现仿真在环探头姿态优化。《IEEE/ASME机电一体化汇刊》（2025） 24. Zhang, Y., Yan, Y., Wang, K., Cai, M., Xiang, Y., Guo, Y., Tu, P., Ying, T., Chen, X.：一种导航引导的三维乳腺超声扫描与重建系统，用于多病灶自动空间定位与诊断。《医学图像分析》**110**，103965（2026） 25. Zhou, Y., Barnes, C., Lu, J., Yang, J., Li, H.：论神经网络中旋转表示的连续性。在：IEEE/CVF计算机视觉与模式识别会议论文集。第5745–5753页（2019）