论文预印本，论文发表于：2021 年欧洲移动机器人会议（ECMR）

用于提升操作员态势感知的虚拟全向视觉灵活框架

Martin Oehler 和 Oskar von Stryk

arXiv:2302.00362v1 [cs.RO] 1 Feb 2023

摘要—在移动机器人的远程操作过程中，确保操作员拥有良好的态势感知是一项重要任务，因为一个小小的失误都可能导致任务失败。远程操作广泛使用相机视频流，但其视野有限。本文提出一种灵活的虚拟投影框架，通过一种新颖的方法融合安装在机器人任意位置的多个摄像头，以提升态势感知。此外，我们还提出一种互补方法，通过融合相机图像与几何三维激光雷达数据，生成着色点云，从而改善场景理解。在紧凑型全向相机上的实现大幅降低了系统复杂度；与电动云台等传统方法相比，它能以小得多的占用空间满足多种使用需求。最后，我们将该方法应用于波士顿动力 Spot 的多摄像头系统，展示了其通用性。该软件实现以开源 ROS 软件包的形式发布于项目页面：https://tu-darmstadt-ros-pkg.github.io/omnidirectional_vision。

I. 引言

遥操作机器人常被用于搜救等危险环境中的任务。它们使操作员能够通过远程操作安全地探索环境并与之交互。然而，操作员失误是机器人任务失败的常见原因。导致这些失误的一个常见原因是态势感知不足 [1]。

通常，远程操作机器人配备有相机，用于将视频数据流传输给操作员。与人类的视场相比，标准透视相机的视野非常有限，因而会造成“隧道视野”。周边视觉的缺失会妨碍操作员建立关于机器人周围环境的心理模型。

在许多情况下，这一问题可通过以下一种或多种方法来解决：1）安装广角鱼眼相机，而非标准相机 [2]，可在一个方向上增大视场角（FOV）。这类图像存在严重畸变，操作员很难据此判断空间关系。此外，车辆两侧和后方仍然存在盲区。2）可以加装摄像头来覆盖这些盲区，但代价是

所有作者均来自德国达姆施塔特工业大学仿真、系统优化与机器人研究组{oehler, stryk}@sim.tu-darmstadt.de本文所述研究部分由德国联邦教育与研究部（BMBF）通过合作 A-DRZ 项目下的“地面机器人自主辅助功能”子项目（资助编号 13N14861）、德国联邦经济与技术部的“EXIST 研究转化计划”（资助编号 03EFLHE061）以及 AICO 协作实验室的 Nexplore 资助。本研究还获得了黑森州（德国）LOEWE 计划在 emergenCITY 中心框架内的共同资助。978-1-6654-1213-1/21/$31.00 ©2021 IEEE

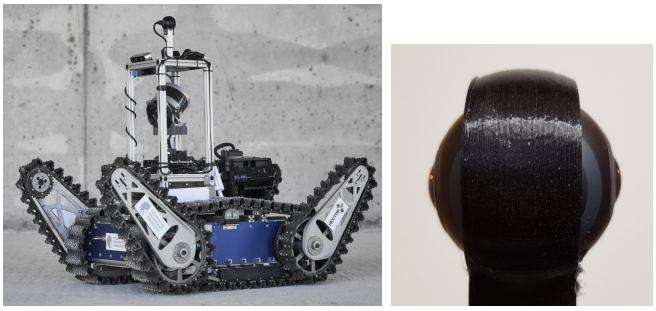


图 1：从全向相机（左上）传输两幅鱼眼图像，并将其转换为透视投影（左下）和墨卡托投影（右）。

系统复杂度增加。这会加重操作员的认知负担，因为操作员必须在不同视图之间切换注意力 [2]。实际操作中，操作员主要关注主摄像头，往往会错过辅助视图中的变化[3]。3）一种广泛采用的替代方案是使用机械式云台传感器头，将安装的摄像头转向所需的观察方向。与单独使用摄像头相比，这类系统结构复杂、空间需求更大，因此难以应用于小型机器人平台。驱动还会带来其他缺点。机械限制可能造成盲区，而执行器运动缓慢以及网络延迟则会降低控制响应速度。

使用视频流进行远程操作的另一个问题是缺乏深度信息。因此，估算与障碍物之间的距离颇具挑战性，使得在狭窄空间中导航困难重重且容易出错 [2]。目前有多种传感器模态可提供深度信息，其中激光雷达和深度相机是最常见的选择。激光雷达传感可提供高精度的距离信息和较大的视场，但其根本缺陷是缺乏颜色信息。仅凭几何距离数据，人们很难理解环境。例如，门很难与墙壁区分开来。大多数深度相机都能提供颜色信息，但与激光雷达相比，其测距范围、精度和视场角都有限。

本文提出两种互补的方法，利用全向视觉提升操作员态势感知。我们提出一个灵活的框架，用于计算安装在机器人任意位置的多个摄像头所生成的虚拟投影。由此可以生成新视图，为操作员提供不受限制的环境 360°视图（见图 1）。通过在线重新计算投影，可以追踪任意



(a) *Asterix* (b) Insta360 Air

图 2：*Insta360 Air* 安装在高机动性履带式平台*Asterix*的中心框架顶部。

机器人上或环境中的坐标系。这样便可实现虚拟云台传感器头。作为补充，第二种方法通过同时融合激光雷达数据与来自多个摄像头的颜色信息，获得着色点云，从而提升场景理解。

将两种方法整合到一个全面的用户界面中，可以实现更完善的情境感知遥操作。通过生成虚拟视图，操作员可以利用校正后的针孔投影进行导航，这比失真的鱼眼图像更容易理解。此外，在三维场景中显示着色点云，还能增强空间感知能力。两种方法的实现均以开源 ROS 软件包的形式提供[1]。

我们在两种不同的多视觉系统上展示了相较于传统方法的优势。首先，我们在安装于履带式机器人*Asterix* [4]上的*Insta360 Air* （图 2b）上实现了我们的方法（图 2a）。过去我们使用的是复杂的多摄像头系统，而改用单个紧凑型全向相机后，显著降低了系统复杂度，并以小得多的占用空间满足了不同用例的需求。我们通过在*Boston Dynamics Spot*的多摄像头系统上实现该方法，展示了投影框架的通用性。机器人机身上安装了五个标准摄像头，其摄像头配置与*Insta360 Air*显著不同。*Spot* 的标准配置不包含直接朝前的摄像头。因此，我们融合两个倾斜前置相机，生成适用于远程操作的视图。作为四足系统，Spot 可以向任意方向移动。在这种情况下，注意侧面的障碍物尤其重要。为提高对周围障碍物的感知能力，我们生成了一幅鸟瞰图，从上方展示环境。

II. 相关工作

人们提出了多种利用全向视觉提升操作员态势感知的方法。

[3]的作者比较了不同摄像头布置位置对遥操作性能的影响。他们发现

[1]https://tu-darmstadt-ros-pkg.github.io/ 全向视觉_

将机器人及其周围环境呈现在同一幅图像中，可以显著减少碰撞。俯视摄像头能够实现这种视角，但安装用于固定摄像头的支架并不实用。在 [5], 中，作者通过单应性变换将多台鱼眼摄像机拍摄的图像映射到地平面，从而生成虚拟鸟瞰图。这会导致障碍物发生失真，因为它们的高度与地平面不符。在 [6]中，作者通过融入深度传感器采集的数据解决了这一问题。障碍物会叠加在鸟瞰图上，以便准确呈现其位置。

文献 [2]概述了与用户界面设计相关的人类操作表现问题。研究指出，视场角较小会导致速度和距离判断错误，并造成周边视野损失。在 [7], 中，研究人员使用鱼眼相机来扩大操作员的视场角。作者提出了一种生成虚拟焦点与上下文视图的方法，该视图中心呈现透视视图，并逐渐过渡到边缘区域的畸变视图。这样既能确保中心区域的视图不失真，又能提供周边信息。

文献 [1]的作者通过DARPA机器人挑战赛决赛的案例研究，比较了不同的人机交互方法。他们得出结论：在虚拟环境中可视化数据，并允许操作员控制视点，能够提升态势感知。在 [8],中，研究人员将四台鱼眼摄像机拍摄的图像投影到半球网格上。由于网格与环境不匹配，这种方法会造成畸变。此外，比例取决于网格尺寸，可能令人不适。文献 [9]通过将环境纳入考量改进了这一方法。研究人员利用激光测距仪近似构建环境网格，并以多台鱼眼摄像机拍摄的图像为其贴图。尽管这种方法在某些环境中能提供良好的态势感知，但其假设环境中只有平坦墙面，因此会造成障碍物畸变。文献 [10]的作者采用了类似的方法，将两台鱼眼相机拍摄的图像投影到球体上。他们将虚拟相机置于球体内部，以渲染透视视图，并将图像显示在配有姿态跟踪功能的头戴式显示器上，从而实现直观的可视化。在我们的评估中，我们将这种方法与自己的方法进行比较。此外，他们还提出了第二种方法，将激光雷达点投影到球体上，以融合颜色信息与深度。文献 [11]进一步研究了点云着色。作者提出了一种利用外接相机为激光雷达点数据着色的方法。他们融合多次观测获得的颜色信息，并将可见性和时间偏移纳入考量。

III. 方法

在下一节中，我们将介绍如何通过多个源相机生成投影。随后，我们将说明如何利用同一组相机为点云数据着色。

A. 相机与激光雷达标定

所述方法要求对所有摄像头的内参和外参进行精确标定，并对

激光雷达进行外参标定。

单个相机的通用投影模型 C_i 如下:

$$\pi_{c_i}\left({}^{C_i}T_R\mathbf{p}\right)=\mathbf{u} \tag{1}$$

相机外参标定描述了每个相机相对于公共参考坐标系R的六维位姿，例如第一个相机或公共相机头。该位姿由 SE(3)变换 ${}^{R}T_{C_i}$给出。其逆变换 ${}^{C_i}T_R$将点 $\mathbf{p}=(x,y,z)$ 从参考坐标系转换到相机坐标系。相机内参标定得到投影函数 $\pi_{c_i}:\ \mathbb{R}^3\rightarrow\mathbb{R}^2$，该函数将一个三维点投影到二维图像坐标 $\mathbf{u}=(u,v)$ 上。请注意，我们使用通用投影模型，不对特定镜头类型作任何假设。因此，我们的方法可推广到不同的投影模型，例如标准针孔相机模型和广角镜头模型。所有相机的外参标定和内参标定均通过标定框架*kalibr*[2] [12]联合完成。

第二步是估计相机与激光雷达之间的外参标定结果。它描述了激光雷达坐标系 L 相对于相机参考坐标系 R 的六维位姿，由变换 ${}^{R}T_L$给出。我们采用 [13]中所述的方法进行无靶标标定。

B. 多相机图像投影

投影由函数 $P:\ \mathbb{R}^2\rightarrow\mathbb{R}^2$ 定义，该函数将坐标从输出投影图像映射到输入相机图像。对于多个摄像头，每个摄像头都定义一个投影函数:

$$P_{c_i}(\mathbf{u}_p)=\mathbf{u}_{c_i} \tag{2}$$

其中，$\mathbf{u}_P=(u_p,v_p)$ 为投影图像坐标，$\mathbf{u}_{c_i}=(u_{c_i},v_{c_i})$ 则分别为相机 C_i 中的图像坐标。投影函数用于查找输出图像中每个像素对应的相机图像坐标。

从几何角度看，投影可以用三维空间曲面表示。曲面的形状取决于投影类型和参数，参数应根据预期应用进行选择。用户通过变换 ${}^{R}T_P$ 指定投影曲面坐标系 P 相对于相机参考坐标系 R 的六维位姿。

根据投影面计算投影函数。首先，使用基于网格的采样将曲面离散化，采样分辨率为所需的目标分辨率（$W_p\times H_p$）。每个三维采样点 $\mathbf{x}_P=(x_p,y_p,z_p)$ 都对应输出图像（u_p,v_p）中的一个像素。二维投影图像与三维投影面之间的映射由表面投影函数 $S:\mathbb{R}^2\rightarrow\mathbb{R}^3$ 给出:

$$S(u_p,v_p)=\mathbf{x}_p=\begin{pmatrix}x_p\\y_p\\z_p\end{pmatrix} \tag{3}$$

以下详细介绍三种不同投影的投影面函数。

[2]https://github.com/ethz-asl/kalibr

透视投影：透视投影面函数描述一个平面。针孔模型的焦距 f和垂直视场角 ϕ_h 由用户指定。基于这些参数，计算像素尺寸 m_p：

$$m_p=\frac{2f\tan\left(\phi_h/2\right)}{W_p} \tag{4}$$

由此可得:

$$S\left(u_p,v_p\right)=\begin{pmatrix}(u_p-W_p/2)\cdot m_p\\(v_p-H_p/2)\cdot m_p\\f\end{pmatrix} \tag{5}$$

墨卡托投影：墨卡托投影以圆柱体表示，可生成环境的全景视图。用户指定所需的垂直视场角 ϕ_v 和圆柱半径 c_r。据此计算角像素尺寸 α_p 和圆柱高度 c_h：

$$\alpha_p=\frac{2\pi}{W_p} \tag{6}$$

$$c_h=2c_r\tan\left(\phi_v\right) \tag{7}$$

表面投影函数如下:

$$S\left(u_p,v_p\right)=\begin{pmatrix}c_r\cos(-u_p\alpha_p)\\c_r\sin(-v_p\alpha_p)\\c_h(0.5-\frac{v_p}{H_p})\end{pmatrix} \tag{8}$$

球面投影：球面投影在几何上可用球体表示。视场角 ϕ 和球体半径 s_r 由用户指定。首先，将投影图像坐标（u_p，v_p）转换为归一化极坐标 $r\in[0;\ 1]$ 和 $\gamma\in(-\pi,\ \pi]$。极坐标系的原点位于图像中心:

$$p_x=\frac{u_p}{W_p}-0.5 \tag{9}$$

$$p_y=0.5-\frac{v_p}{H_p} \tag{10}$$

$$r=\sqrt{p_x^2+p_y^2} \tag{11}$$

$$\gamma=\arctan 2(p_y,p_x) \tag{12}$$

为了将一条射线投射到三维空间中，我们还计算第二个角度θ，它对应于纬度，而 γ 对应于球体的经度:

$$\theta=r\frac{\phi}{2} \tag{13}$$

根据纬度 θ 和经度 γ，计算出球体:

$$S\left(\theta,\gamma\right)=\begin{pmatrix}s_r\sin\theta\cos\gamma\\-s_r\sin\theta\sin\gamma\\s_r\cos\theta\end{pmatrix} \tag{14}$$

鱼眼镜头会生成球面图像。理想鱼眼镜头的图像中心 r 与纬度 θ 之间呈线性关系，如式（13）所示。真实的鱼眼

镜头通常并非线性，因此可以通过投影对图像进行“校正”。

使用表面投影函数 S，为输出图像中的每个像素计算一个表面采样点 $\mathbf{x}_P$。

接下来，使用用户指定的表面位姿 ${}^{R}T_P$ 和相机外参标定 ${}^{C_i}T_R$，将每个点 $\mathbf{x}_P$ 变换到每个摄像头的相机坐标系中：

$$\mathbf{x}_{c_i} = {}^{C_i}T_R \, {}^{R}T_P \mathbf{x}_p \tag{15}$$

最后，使用每个摄像头的内参标定函数 $\pi(\cdot)$，将该点投影到所有相机图像中：

$$\pi_i(\mathbf{x}_{c_i}) = \mathbf{u}_{c_i} \tag{16}$$

通过这一过程，我们为输出图像 $\mathbf{u}_P$ 中的每个像素以及每个摄像头 C_i 计算映射 P_{c_i}。如果某个像素存在多个有效映射，也就是说，投影面被多个摄像头观测到，我们会优先选择 $\mathbf{u}_{c_i}$ 最接近相应图像中心的映射。另一种方法是融合颜色，以综合利用所有观测结果。

投影映射函数只需计算一次。每当摄像头收到新图像时，就使用为每个摄像头计算出的映射填充输出图像。颜色通过对相机图像进行双线性插值得到。

C. 点云着色

激光雷达点云通过在收到新的扫描点云后立即为每个点分配颜色，实现在线着色。为此，将扫描中的每个点投影到最新的相机图像中。这与前文所述的投影方法遵循相同的思路，只是将三维投影表面替换为从激光雷达扫描中获取的三维点。

激光雷达扫描点 $\mathbf{x}_L$ 以激光雷达传感器坐标系 L 为参照给出。使用激光雷达—相机外参标定 ${}^{R}T_L$ 以及每个摄像头的相机外参标定 ${}^{C_i}T_R$，将这些点变换到各个相机坐标系中：

$$\mathbf{x}_{c_i} = {}^{C_i}T_R \, {}^{R}T_L \mathbf{x}_L \tag{17}$$

利用内参标定，根据公式（16）将这些点投影到所有相机图像中。如果某个点在多个摄像头中均可见，则优先采用距离相应图像中心最近的观测值。颜色采用双线性插值。对扫描中的每个点以及每个摄像头重复这一过程。

IV. 结果

本节将通过多个应用场景，重点介绍与传统方法相比，我们所提出方法的优势。我们使用两种不同的多摄像头系统进行演示：安装在我们的履带式平台 *Asterix* 上的紧凑型全向相机 *Insta360 Air*，以及四足机器人 波士顿动力 *Spot*。

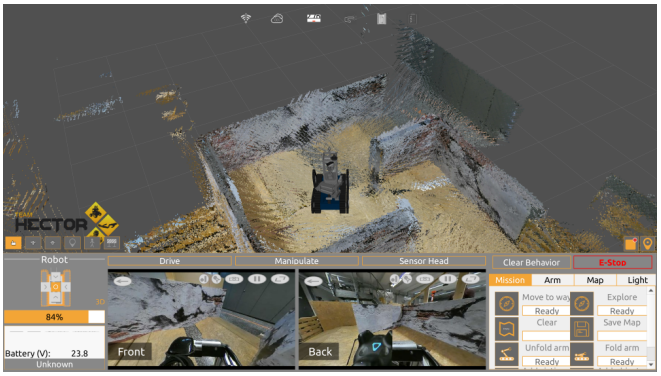


图3：操作员用户界面。主视图以虚拟表示的形式显示环境的着色点云。主视图下方显示投影后的前视和后视摄像头画面。可通过操纵杆在线调整其观察方向。

A. Insta360 Air

Insta360 Air（图 2b）是一款紧凑型全向相机，由两个 210° 鱼眼镜头组成，安装在直径为 4 cm 的球形机身上。该相机安装在我们的高机动性履带式机器人 *Asterix* [4]（图 2a）上，并置于中央框架顶部，以获得最大的视野。用全向相机替换此前的多相机系统后，我们在实现相同应用场景的同时，大幅降低了系统复杂度和空间需求。此外，标定工作也减少为只需标定一台带有两个镜头的相机。

接下来，我们介绍如何将该方法集成到一个全面的用户界面中（见图 3）。该界面由主窗口中的虚拟三维场景以及下方的虚拟视图组成。

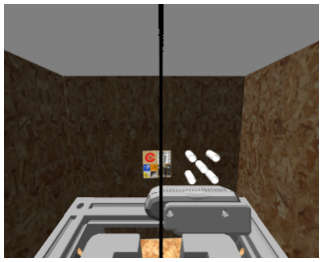
虚拟视图：我们使用两个针孔投影来呈现虚拟的前向和后向驾驶摄像机，使操作员能够快速切换行驶方向，同时还能增强其对机器人后方情况的态势感知，正如 [3] 中所建议的那样。投影的视场角高达 130°，以提高侧方的可见性。通过投影，我们可以在实现大视场角的同时避免类似鱼眼镜头所产生的失真。为确保投影准确并减少拼接缝，用户可以根据标准操作观看距离调整投影面。

此外，操作员可以自由旋转和缩放投影，以检查机器人的周围环境。此功能通过根据操纵杆输入重新计算投影来实现。这实际上实现了一个虚拟传感器头，无需使用受驱动云台，也避免了其带来的机械复杂性和空间需求。此外，我们还可以跟踪机器人或环境中的任意坐标系，例如机械臂末端执行器。

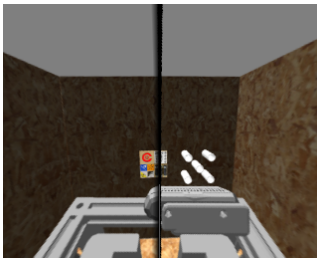
投影直接在机器人上生成，以降低带宽需求。传输的仅是感兴趣区域的图像，而非高分辨率球面源图像。



(a) Vunder 等人



(b) 我们的方法



(c) 经我们的方法校正后的 Vunder 等人方法

图 4：在具有两个镜头的模拟全向相机上，对比 Vunder 等人 [10] 的方法与我们的方法，其中两个镜头绕光轴偏移了 3° 。在 (a) 中，墙面出现了错位，而我们的方法没有这一问题 (b)。通过在预处理步骤中使用我们的方法校正鱼眼图像，墙面得以对齐 (c)。

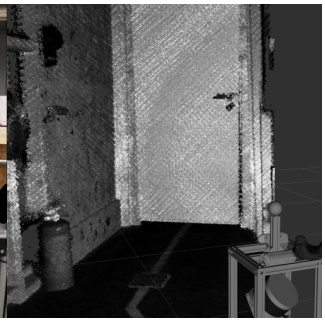
创建虚拟投影使我们能够同时将相机用于多个目的。除了用作操作员视图外，我们还使用相机进行 360°物体检测。系统会基于墨卡托投影生成环境的全景视图（见图 1），用于检测环境中的危险品标志。

接下来，我们将我们的方法中的针孔投影与 Vunder 等人 [10] 描述的方法进行比较（图 4）。两种方法存在根本差异：我们的方法直接生成针孔投影，而不是先映射到球体。这个中间步骤依赖于这样一个假设：使用两幅理想的球面图像，且相机中心重合。现实中的相机很少满足这一假设，因此会产生伪影，如图 4a 中墙面的错位所示。我们的方法会考虑每个相机的外参和内参标定，因此不会出现这些伪影（图 4b）。我们可以使用我们的方法校正投影：首先根据原始鱼眼图像计算理想的球面投影（图 4c）。

三维视图： 机器人配备了一台旋转倾斜的*Velodyne VLP-16* 激光雷达，可生成环境的点云。点云通过安装在上方的全向相机进行在线着色。机器人部件遮挡的点云点会被过滤，以防机器人自身部件被投影到环境中。操作员可在虚拟相机视图上方查看三维场景（图 3）。根据 [1], 空间感知



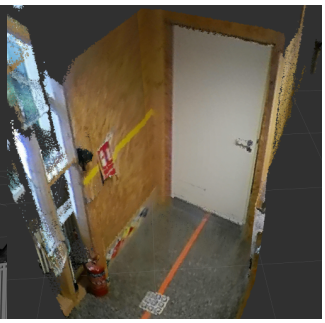
(a) 参考照片



(b) 按强度着色



(c) 提出的方法



(d) 不同透视

图 5：着色后的点云通过使物体能够按颜色识别，提升了语义理解能力。

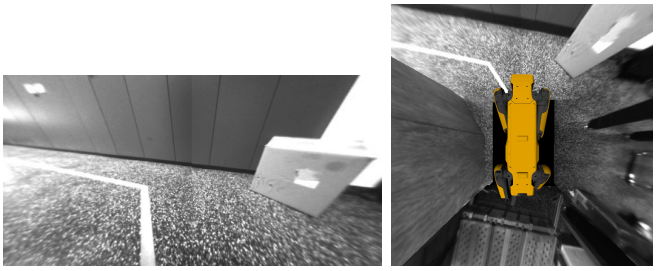
通过在环境的虚拟表示中显示着色后的点云，空间感知能力得以增强，操作员可以自由更改视点。仅凭激光雷达的几何数据，人很难形成对机器人周围环境的语义理解。请看图 5 中的对比。图 5a 给出了场景的参考照片。图 5b 展示了从相同角度观察并按回波强度着色的激光雷达数据。虽然该视图已能帮助我们识别障碍物和一些物体，但在着色后的点云中（图 5c），白色的门、灭火器以及左侧的标志更容易辨认。因此，该方法将精确的激光雷达点云所提供的空间感知与彩色相机数据所提供的语义信息结合起来。除了有利于远程操作外，颜色信息还可用于改进自动三维物体检测。

B. 波士顿动力 *Spot*

波士顿动力 Spot 是一款四足机器人，配备由五个灰度针孔相机构成的相机系统，可实现全向视觉。与前一个全向相机应用示例相比，这种摄像头配置有显著差异，也凸显了所提方法的通用性。图 6a 展示了五幅源相机图像。该机器人没有朝前的相机。因此，我们采用透视投影融合两台倾斜前置相机，生成适用于远程操作的视图（图 6b）。在水平视场角为 130° 的情况下，生成的投影比单台相机提供了更好的周边视觉。作为一款



（a）Spot 相机图像



（b）虚拟前视图　　（c）虚拟鸟瞰图

图 6：Spot 的五幅源相机图像如（a）所示。利用这些图像可计算出前视图（b）和叠加了机器人模型的鸟瞰图（c），从而增强态势感知。

足式平台，Spot 可以向任意方向移动。若要同时掌握各个方向的运动情况，操作员就需要在不同相机之间切换，这既费力又容易出错 [2]。为了解决这一问题，我们生成了虚拟鸟瞰图（图 6c），并叠加渲染的机器人模型；按照 [3] 的建议，该视图在单幅图像中展现了机器人及其环境。此视图通过透视投影实现：将虚拟相机置于机器人上方，并将投影面设在地面高度。与 [5]，相比，该投影无需任何中间步骤即可直接生成。虽然这种可视化方式能提升近距离范围内的态势感知能力，但操作员需要注意，由于物体高于地面，它们会产生变形。

C. 性能评估

我们在英特尔 *i7- 4710HQ@2.50GHz* 移动处理器和*Insta360 Air* 摄像头上评估了性能。我们测试了不同的投影和目标分辨率。两个处理步骤所需的计算时间如表 I 所示。仅当投影发生变化时才需要生成投影映射，例如在启动时或旋转视图时。计算时间与目标分辨率呈线性关系。与将映射应用于相机图像以生成新投影帧的映射操作相比，这一步的开销相当大。即使采用较高的目标分辨率，也可以实现实时应用。我们使用*VLP-16* 激光雷达测试了点云着色性能。在我们的设置中，它以 10 Hz 的频率运行，每次扫描约生成 18000 个点。处理单次扫描平均耗时 57.72 ms，因此可以满足实时使用需求。

表 I：图像投影性能

投影	分辨率	投影映射（毫秒）	Map 运算（毫秒）
透视	512x256	23.78	12.74
透视	1024x512	86.27	12.5
透视	2048x1024	334.62	23.36
墨卡托	1024x512	141.67	12.33
球面	512x512	92.26	10.9

五、结论

我们提出了一种新颖且灵活的框架，可利用机器人上任意位置的多个摄像头为远程机器人操作员提供虚拟投影；同时还提出了一种互补方法，可根据图像数据为激光雷达扫描着色。这两种方法的软件实现均以开源 ROS 软件包的形式提供。将这两种方法结合到一个综合用户界面中，可以显著提升远程操作期间的态势感知能力。

作为未来工作，我们计划在分辨率更高的全向摄像机上评估我们的方法。此外，我们预计 GPU 加速将显著缩短计算时间，因为数据处理的结构非常适合并行化。

参考文献

[1] A. Norton、W. Ober、L. Baraniecki、E. McCann、J. Scholtz、D. Shane、A. Skinner、R. Watson 和 H. Yanco，“DARPA机器人挑战赛决赛中的人机交互分析”，*IJRR*，第 36 卷，第 5–7 期，第 483–513 页，2017 年。[2] J. Y. Chen、E. C. Haas 和 M. J. Barnes，“遥操作机器人的人类操作表现问题与用户界面设计”，*IEEE Trans Syst ManCybernCApplRev*，第 37 卷，第 6 期，第 1231–1245 页，2007 年。[3] B. Keyes、R. Casey、H. A. Yanco、B. A. Maxwell 和 Y. Georgiev，“遥控机器人操作的摄像机布置与多摄像机融合”，载于 *IEEESSRR*，2006 年，第 22–24 页。[4] M. Schnaubelt、T. Ullrich、M. Torchalla、J. Diegelmann、M. Hoffmann 和 O. von Stryk，“面向自主性的高机动性地面机器人系统研发，用于灾害应对”，载于 *VDI-MECHATRONIK*，2021 年，第 20–25 页。[5] T. Sato、A. Moro、A. Sugahara、T. Tasaki、A. Yamashita 和 H. Asama，“使用多台鱼眼摄像机生成时空鸟瞰图像”，载于 *IEEE/SICE SII*，2013 年，第 753–758 页。[6] Y. Awashima、R. Komatsu、H. Fujii、Y. Tamura、A. Yamashita 和 H. Asama，“使用深度传感器在遥控机器人鸟瞰图中可视化障碍物”，载于 *IWAIT*，2017 年。[7] J. Vaughan、S. Kratz 和 D. Kimber，“看清前进方向：机器人遥操作的视觉界面”，载于 *IEEE RO-MAN*，2016 年，第 273–280 页。[8] S. Iwataki、H. Fujii、A. Moro、A. Yamashita、H. Asama 和 H. Yoshinada，“用于工程机械遥操作的三维空间 3DCG模型中周围环境与操作部件的可视化”，载于 *IEEE/SICE SII*，2015 年，第 81–87 页。[9] R. Komatsu、H. Fujii、Y. Tamura、A. Yamashita 和 H. Asama，“用于室内机器人遥操作的鱼眼摄像机与激光测距仪自由视点图像生成系统”，*ROBOMECH*，第 7 卷，第 1 期，第 1–10 页，2020 年。[10] V. Vunder、R. Valner、C. McMahon、K. Kruusamae 和 M. Pryor，“在 ROS 中使用全景视觉与虚拟现实提升态势感知”，载于*IEEE/IES HSI)*，2018 年，第 471–477 页。[11] P. Vechersky、M. Cox、P. Borges 和 T. Lowe，“使用独立摄像机为点云着色”，*IEEE RA-L*，第 3 卷，第 4 期，第 3575–3582 页，2018 年。[12] P. Furgale、J. Rehder 和 R. Siegwart，“多传感器系统的统一时间与空间标定”，载于 *IEEE/RSJ IROS*，2013 年，第 1280–1286 页。[13] G. Pandey、J. McBride、S. Savarese 和 R. Eustice，“通过最大化互信息对三维空间激光雷达与相机进行自动无目标外参标定”，载于 *AAAI*，第 26 卷，第 1 期，2012 年。