

第2章 基于点跟踪的半自动连续勾画软件

2.1 引言

本章介绍如何使用 MATLAB 2019b 构建一款用户友好的 CEUS 视频连续勾勒的半自动跟踪软件,以实现在快速精准地在病变区域绘制边界框的目标。基于深度学习在处理 CEUS 视频数据方面广阔的前景,为了加快使用 DL 对 CEUS 数据端到端地自动检测、跟踪和分类任务,设计编写了该应用程序。尽管运动的矫正在 DCE-US 中应用历史已久,但都基于自主编写的 Matlab 脚本程序,灵活性很低,不懂 Matlab 编程的医生无法使用,且当运动的矫正效果不好时,也无法及时更改。大多数情况下,CEUS 数据预处理是为每个包含的帧手动绘制草图,或以短序列手动绘制一个帧,默认情况下,病变的位置在序列中不会更改。这些研究似乎没有意识到运动补偿算法的存在。这种现象反映出,当没有实用的工具,只有理论实验时,即使是聪明的研究者也不能利用这些算法来简化或进一步改进他们的工作。

目前,除了本研究公开的软件外,还没有其他的脚本程序允许交互式的勾画操作。该软件的使用非常简单,用户只需在视频中需要跟踪的位置画一个 **Bounding box**,软件就可以自动跟踪,软件允许用户拖动缩放每一帧的跟踪结果,也可以批量删除一些不好跟踪结果,跟踪长度可以设置。软件除了集成数据导入,工作空间保存,勾画结果导出,视频播放等基本的数据操作接口外,还针对 CEUS 的图像特性,如双幅显示,加入了镜像功能等自动分析方法。本章分三节分别介绍运动校正和不同跟踪算法的性能测试,软件搭建,软件使用,在我看来,项目的难度在与设计和编写交互式软件,由于没有参考,需要从零开始自行构想一个好用的软件界面和功能,并让软件能够响应多种用户操作,保证数据传递无误和时序不发生混乱。

图 2.1 显示了两种 CEUS 录制界面,展示了原始的未经处理的 CEUS 图像,可以看到,CEUS 图像的录制区域具有特定的形态,通过边缘提取和形态分析可以自动识别出图像的录制模式。与此同时,本研究提出了镜像概念,即将一个录制图窗的勾画结构自动映射到另一个图窗的对应区域,帮助医生在双幅模式时利用超声清楚的边界信息来勾勒 CEUS 中未完全增强的病灶和全增强时边界无法

确定的病灶。

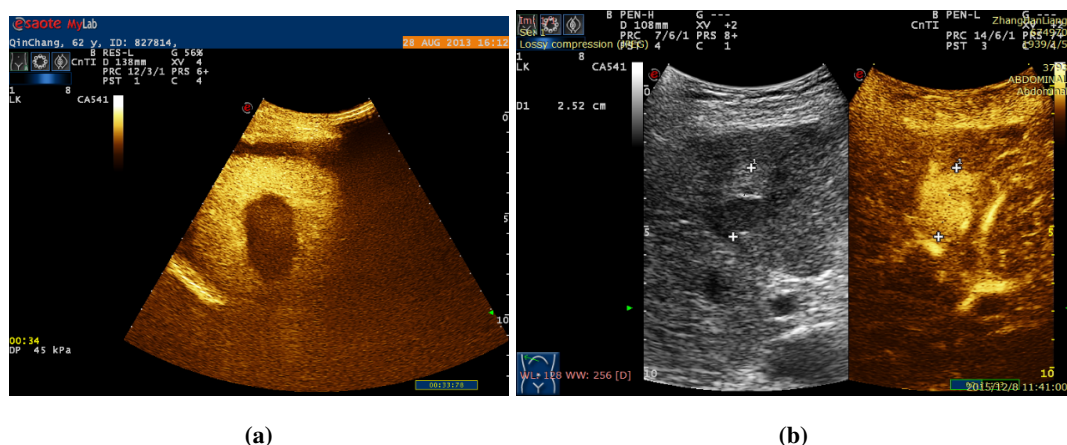


图 2.1 两种 CEUS 界面展示。(a) 单幅显示，图像中只有 CEUS 一个录制窗口，(b) 双幅显示，图像的一侧为 US 视频窗口，另一侧为 CEUS 视频窗口，两个窗口显示同一成像平面的不同信号分布。

Figure 2.1 Two CEUS interface displays. (a) single display with only one CEUS recording window in the image, (b) dual display with US video on one side of the image window, with a CEUS video window on the other side, and two windows showing different signal distributions from the same imaging plane.

2.2 跟踪算法

2.2.1 运动校正

在 CEUS 中，有许多研究集中在通过整个序列的自动运动估计方法来补偿观察到的肝脏实质性病变（focal liver lesions, FLLs）运动，这实质上是一个连续帧上特定区域的跟踪任务。运动估计方法分为直接方法和间接方法，直接方法测量图像强度变化来估计运动矢量，对噪声和光照变化敏感 (Makela 等, 2002)，间接方法通过一组稀疏的图像特征进行连续帧的匹配。具体来说，基于点的跟踪技术（point-based registration techniques, PBRTs）非常适用于在 CEUS (Bakas 等, 2012) 或 US 序列 (Luca 等, 2015) 中跟踪 FLLs。

PBRTs 主要由两部分组成：(1) 图像特征提取，即提取突出区域的特征 (2) 匹配和利用结果估计几何变换。经典的点描述器有 Harris 角检测器 (Harris 和 Stephens, 1988)、利用特征值检测角点的 Kanade-Lucas-Tomasi (Lucas 和 Kanade, 1981; Jianbo Shi 和 Tomasi, 1994)、HOG (Dalal 和 Triggs, 2005)、SIFT (Rosten 和 Drummond, 2005)、FAST (Rosten 和 Drummond, 2006)、SURF (Bay 等, 2008)、BRISK (Leutenegger

等, 2011)、ORB(Rublee 等, 2011)、CARD(Ambai 和 Yoshida, 2011)。基于Mikolajczyk 和 Schmid (2003), 统计描述符的匹配策略有三种选择: 基于阈值的方法、基于最近邻的方法和最近邻距离比。一般的距离指标有欧氏距离、点积和 Hamming 距离。有时, 匹配过程中可以通过估计几何变换来修正, 如 M-estimator SAmple Consensus(Kalal 等, 2010), 它利用前向和后向误差反复寻找最佳匹配和几何变换矩阵。

软件的跟踪算法的流程为: 手动绘制一个边界框进行初始化; 提取当前帧和下一帧的点特征; 将点匹配后计算变换矩阵; 通过 M-estimator SAmple Consensus 检验跟踪结果, 输出更稳健的跟踪点, 再重新估计变换矩阵; 应用变换矩阵更新边界框的位置。该算法将进入下一帧, 并进行相同的程序, 直到匹配点数目少于三或达到跟踪长度。在跟踪过程中, 由于不断删未匹配的点, 可用的点会不断减少, 因而每隔 10 张图像通过刷新点集来重新初始化需要跟踪的点, 重新初始化会降低运行速度。由于 CARD 逐帧提取图像各个子区域的特征, 特征向量个数由图像自身决定, 不需要重新初始化。

MATLAB 2019b 图像处理工具箱提供了多种关键点检测算法, 如特征值、HOG、FAST、BRISK、MSER、ORB。根据Bakas 等 (2019) 的研究, 基于 HOG 的跟踪算法比较耗时, 所以我们在软件中排除了这个功能。目前, 在 MATLAB 中导入 SIFT 特征检测器仍然是一个需要其他特定库的重任, 所以没有使用 SIFT 特征。(Bakas 等, 2019) 也报道, 虽然 SIFT 有比较高的运行速度和跟踪精度, 但没有 CARD 好。由于 CARD 有[官方开源](#), 且库的安装和调用都很方便, 所以我们将描述长度为 128 维和 256 维的 CARD 纳入跟踪方法中。

2.2.2 跟踪性能测试

由于 PBRTs 是一个比较成熟的系统, 我们的测试主要针对该软件中不同跟踪算法的量化跟踪精度和速度, 算法的超参数都使用函数提供的默认值 (通过预实验发现默认值的效果最好)。测试数据为专业医生选取的两例典型的 HCC 动脉期 CEUS 片断。其中, 病例一的病灶边界不清晰, 体积相对较小, 呼吸运动明显; 例二的病变边界相对清晰, 体积相对较大, 扫描过程相对稳定。帧率均为 25 帧/秒。医师使用跟踪软件生成病灶区域, 再逐帧精细调整, 作为本次实验的 Ground-truth。

实验希望量化不同算法在不同跟踪长度设置下的精度和速度。跟踪长度范

围为 30~270 帧，步长为 30。测试方法为，在特定的跟踪算法和跟踪长度下，用视频中的每一帧和相应人工标定的边界框来初始化多次跟踪过程。每次跟踪后，计算该跟踪序列中与 Ground-truth 的重叠比（交集面积除以并集面积），并通过跟踪时间除以跟踪长度（frame per second, FPS）计算运行速度，再对所有跟踪结果计算平均值。我们对 CEUS 和 US 跟踪模式下的性能进行测试，分析同一算法在两种成像模式是否存在性能差异，以及跟踪精度随时间的衰减情况。测试结果见表 2.1、2.2、2.3、2.4，其中蓝色字体表示本列性能最好，红色字体表示最差。

综合表 2.1、2.2、2.3、2.4，可以发现，随着跟踪长度的增加，所有基于关键点特征的算法都呈现出精度急剧下降的现象，而 CARD 的算法性能稳定且表现较好，但速度相对较低。在研究案例和跟踪帧长度固定的情况下，在 CEUS 视频和 US 视频中，US 的跟踪精度略微好于 CEUS。对比案例 1 和案例 2，可以发现案例 2 的跟踪效果更好，但时间也更长，从一定程度上表明边界清晰的大病灶因为能检测到更多特征点，所以表现更好。综合来看，我们建议使用 CARD 128 执行跟踪任务。

表 2.1 案例 1 的 CEUS 视频跟踪速度和精度

Table 2.1 Case 1's CEUS video tracking speed and accuracy

跟踪方法		FPS	跟踪长度								
			30	60	90	120	150	180	210	240	270
Key Points	KTL	13.63	0.93	0.803	0.77	0.695	0.622	0.523	0.562	0.473	0.356
	FAST	14.15	0.894	0.804	0.771	0.695	0.625	0.524	0.562	0.475	0.364
	BRISK	11.24	0.893	0.803	0.77	0.693	0.622	0.523	0.564	0.511	0.359
	SURF	15.28	0.892	0.802	0.767	0.693	0.621	0.521	0.563	0.471	0.356
	MSER	12.11	0.892	0.801	0.768	0.692	0.621	0.522	0.559	0.471	0.356
	ORB	10.55	0.927	0.799	0.764	0.69	0.618	0.518	0.558	0.465	0.347
	CARD 256	4.35	0.852	0.906	0.895	0.889	0.894	0.877	0.903	0.843	0.865
CARD 128		5.25	0.852	0.906	0.895	0.889	0.894	0.877	0.903	0.843	0.865

表 2.2 案例 1 的 US 视频跟踪速度和精度

Table 2.2 Case 1's US video tracking speed and accuracy

跟踪方法		FPS	跟踪长度								
			30	60	90	120	150	180	210	240	270
Key Points	KTL	10.51	0.929	0.801	0.769	0.693	0.623	0.523	0.569	0.473	0.36
	FAST	11.28	0.895	0.807	0.773	0.701	0.628	0.527	0.575	0.519	0.371
	BRISK	9.06	0.894	0.805	0.772	0.701	0.627	0.525	0.571	0.516	0.367
	SURF	12.31	0.891	0.8	0.766	0.692	0.623	0.522	0.566	0.47	0.354
	MSER	10.88	0.894	0.805	0.772	0.7	0.627	0.525	0.57	0.517	0.36
	ORB	9.09	0.888	0.797	0.764	0.69	0.618	0.519	0.557	0.465	0.347
	CARD 256	3.81	0.919	0.923	0.911	0.892	0.892	0.894	0.873	0.901	0.851
CARD 128	4.58	0.919	0.923	0.911	0.892	0.892	0.894	0.873	0.901	0.851	

表 2.3 案例 2 的 CEUS 视频跟踪速度和精度

Table 2.3 Case 1's CEUS video tracking speed and accuracy

跟踪方法		FPS	跟踪长度								
			30	60	90	120	150	180	210	240	270
Key	KTL	4.44	0.936	0.831	0.753	0.718	0.64	0.518	0.621	0.559	0.498
Points	FAST	7.87	0.881	0.829	0.752	0.718	0.641	0.519	0.624	0.558	0.5
	BRISK	6.9	0.883	0.831	0.754	0.718	0.641	0.518	0.623	0.559	0.499
	SURF	7.55	0.885	0.829	0.752	0.717	0.639	0.517	0.62	0.559	0.498
	MSER	7.75	0.883	0.828	0.75	0.714	0.637	0.516	0.619	0.556	0.497
	ORB	6.36	0.884	0.829	0.753	0.716	0.638	0.516	0.618	0.558	0.497
CARD 256		2.72	0.962	0.958	0.95	0.936	0.96	0.907	0.955	0.915	0.967
CARD 128		3.13	0.962	0.958	0.95	0.936	0.96	0.906	0.955	0.915	0.968

表 2.4 案例 2 的 US 视频跟踪速度和精度

Table 2.4 Case 1's US video tracking speed and accuracy

跟踪方法		FPS	跟踪长度								
			30	60	90	120	150	180	210	240	270
Points	FAST	8	0.878	0.825	0.749	0.716	0.64	0.517	0.623	0.556	0.5
	BRISK	6.88	0.88	0.828	0.752	0.718	0.641	0.519	0.624	0.558	0.501
	SURF	7.93	0.886	0.831	0.754	0.719	0.642	0.519	0.62	0.561	0.499
	MSER	8.01	0.884	0.83	0.753	0.718	0.64	0.519	0.622	0.561	0.499
	ORB	6.68	0.887	0.831	0.755	0.718	0.641	0.518	0.618	0.56	0.497
CARD 256		2.99	0.962	0.958	0.95	0.936	0.96	0.906	0.955	0.915	0.967
CARD 128		3.49	0.958	0.963	0.942	0.938	0.931	0.868	0.946	0.913	0.968

2.3 软件构建

由于软件的构建是一个细节比较丰富的环节，涉及很多 Matlab 使用的专业知识，而功能的介绍在下一节展开，所以本节只说明了一些重要的问题，更多细节请阅读开源代码。

- 软件的算法部分集中在图像处理和计算机视觉两个领域，主要使用 Matlab 的 Image Processing Toolbox 和 Computer Vision Toolbox，视频文件的读取使用 VideoReader 函数，只需传入视频名称即可加载，可以更改 curtime 成员变量再调用成员函数 readFrame 访问视频的任意帧。由于 1÷ 帧频往往除不尽，而随意设置的时间会使算法自动插值生成对应的图像，该图像不是真实存在的，所以在加载视频后，遍历视频的 frame 并保存这些时间位置，建立字典，实际应用访问图像在视频中的次序，然后通过字典转化为时间。

- 软件搭建基于 MATLAB 图形用户界面创建应用图形用户界面（Graphical User Interface, GUIs），GUI 提供了软件应用程序的点击式控制，无需其他人学习 Matlab 语言或命令即可运行应用程序。为了对软件设计和开发执行灵活的控制，程序的主体基于编程方式创建，使用 MATLAB 函数来定义应用的布局 and 响应，通过函数确定每个组件在图窗中的位置。

- 作界面信息保存，交互界面的所有变量统一存放在一个字典变量中，但由于这些变量部分是系统的默认设置，部分可以通过其他变量经过标准流程计算得到，因而，为了节省存储空间，只导出字典的部分 Key 到“.mat”格式的文件中，包括视频文件路径，成像图窗在图像中的位置，图窗模式，造影图窗位置，当前帧，和所有 Bounding-Box 信息（跟踪序列的编号，所在帧在视频中的位置，跟踪框在图像中的位置），Mat 文件一般只有几 KB 大小。

- 响应事件的能力基于调用交互控件的回调函数（Callback Function），软件中交互式菜单基于 uimenu，文件载入使用 uigetfile，文件保存过使用 uisave，上下文菜单基于 uicontextmenu，使用交互控件有数字滑块、复选框、编辑字段和按钮，Bounding-Box 基于 imRect 实现，整个界面是一个 uifigure，视频播放图窗是一个 uiaxes。

- 视频播放。视频播放通过计时器 timer 实现，当计时器每隔一个间隔，就执行一次界面刷新（显示下一帧图像、Bounding-Box 等），通过调节 time 的刷新周期可以轻易调节播放速度，比如 ×2 或者 ×0.5，timer 可以轻松开启和暂停。

视频播放的需要注意在于刷新功能在响应时，计时器会继续工作，如果刷新操作不能再下一次 timer 启用时结束，就会导致大量任务累积未处理，容易发生系统错误，因此 timer 的最小刷新时间需要比一次刷新时间长。

- 处理关联的事件。当用户对对一个控件做出更改时，一些更改的变量会在其他控件中复用，这是这些控件也需要做出响应。

- 该软件可以自动定位录音和判断显示模式（双或单）。大多数 CEUS 机器都支持单屏成像模式和双屏成像模型，同时具有 B 模式成像功能。基于单屏 CEUS 一般为扇形的观察，基于传统的数字图像开发了自动分析算法。我们汇总了连续帧的差异图像，然后使用阈值、Sobel 边缘检测、霍夫线检测来获取记录屏幕的轮廓和形状。当轮廓为单个扇形时，成像模式将被视为单幅，否则屏幕将在水平中均匀分割。当 cine-loop 处于双显示模式时，软件将显示双边界框，用户需要通过手动选择软件哪一侧是 CEUS，软件默认为左侧。

- 默认打开文件位置自动更改。由于数据集往往在同一个大的文件夹下，为了减少用户每次加载保存文件时，文件夹路径的更改操作，软件自动保存最后一次用户加载数据和保存过数据操作的文件夹的上一层位置并且选中该文件夹，保证下次操作可以直接进入数据库目录，并且可以轻易知道上一次处理的数据是哪一个，如果操作者按顺序处理数据，那么他可以很快知道选定文件夹的下一个文件夹就是他需要处理的文件。

- 限制跟踪区域。由于超声视频的录制窗口并非铺满整个图像，跟踪结果需要限定在录制区域内。

- 灰度化和图像精度改变。由于 CEUS 和 US 都是伪彩色图像，为了降低颜色带来的不确定性和高内存占用，读入后将彩色图像转为灰度图像，并将 0-255 的像素值转为 0-1 之间的双精度数据。

- 多线程加速。由于跟踪过程以人工勾画所在位置为中心向前后两个方向跟踪，互相独立，所以使用 parfeval 执行多线程操作，再将每个线程的结果使用 fetchNext 读出，整合。需要注意，当勾画图像为第一张和最后一张时，要关闭多线程。如果电脑有显卡，也可以使用 GPU 加速。

- 处理中断。由于用户很可能密集操作，一个响应在没有结束的时候发起另一个指令，需要设置优先级，决定是按时间顺序响应，还是停止目前的操作执行新的操作，或者直接恢复到本操作之前的状态再执行下一个操作。本软件除了视

频播放允许新的中断插入，其他都按时间顺序响应。

- 处理错误。用户操作难免会发生不符合软件使用规范的操作，比如加载文件格式不支持，此时会发生错误，为了保证用户使用体验，对于错误操作自动恢复成执行前的状态，同时告知错误原因。

- 处理不稳定的造影区段。由于造影数据的录制是人为控制的，在实际操作中，难免会比较灵活或者不规范，比如扫描肝脏查找病灶，或者暂停录制但同时记录屏幕内容。由于一个造影视频的时间可能较长，而且跟踪方法也不能再剧烈运动时工作良好，所以加入光流分析接口 `opticalFlowLKDoG`，分析运动是否发生且是否剧烈，设置固定的高低阈值，通过对图像光流幅值求平均给出判断。由于光流的计算基于局部一小段视频得到，所以可以使用多线程将视频分为多段处理，并使用 **GPU** 加快计算。但光流计算对电脑配置要求较高，建议根据实际情况决定是否使用。

2.4 软件使用

本节通过图 2.2和表 2.5的方式展示软件的界面和介绍各个组件的功能。

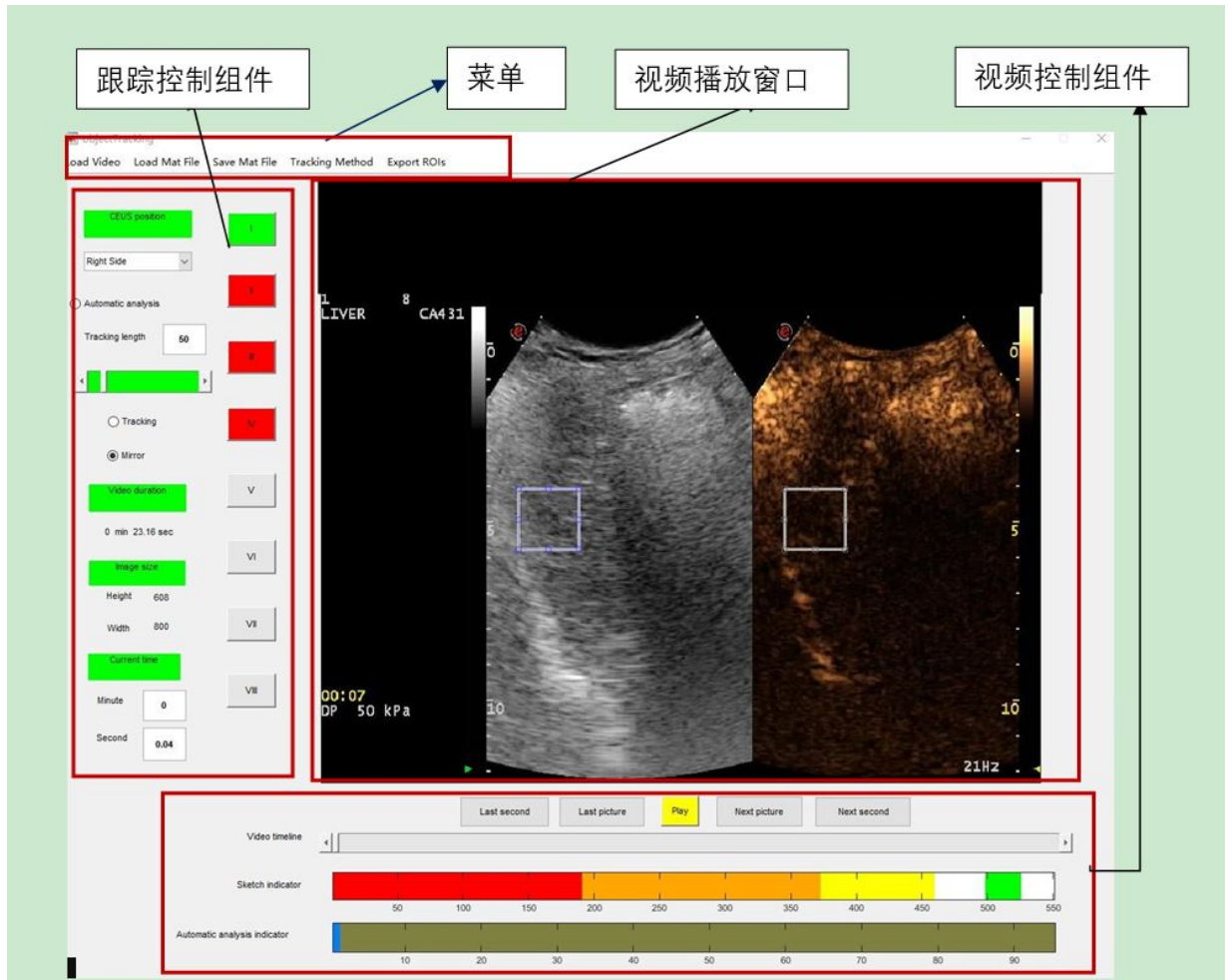


图 2.2 软件运行效果图。软件总体分为四个区域，菜单，跟踪控制组件，视频播放窗口和视频控制组件，图像显示了一例勾画完的视频在软件中的显示效果，图下方区域的彩色标记代表着不同跟踪任务在视频中的位置。软件界面简洁紧凑

Figure 2.2 A rendering of how the software works. The software can be divided into four areas: the menu, the tracking control component, the video playback window and the video control component. The image shows an example of how the finished video is displayed in soft armor, and the colored markers in the lower area of the figure represent the different tracking tasks in the Location in video. The interface is simple and compact.

表 2.5 软件操作界面模块组成和功能表

Table 2.5 Software operator interface module composition and function

所在地区	组件	功能
菜单	加载视频	通过调用 MATLAB 内置功能 VideoReader 将视频导入软件。支持格式包括 MPEG-4, H.264 编码视频（.mp4,.m4v,.mov）或其他微软媒体播放器支持的格式。DICOM 格式的文件需要使用其他软件（例如 RADIAUT DICOM 查看器）转化为可读取的视频格式。
	加载 Mat 文件	在软件保存的.mat 文件中导入数据。它可以在保存文件时恢复软件的状态。
	保存 Mat 文件	将数据保存在当前软件状态的.mat 文件中，其中包括所有变量和变量的值。在软件使用期间，不会更改原始视频。
	ROI 边界框导出	导出视频中所有勾画的边界框。
	跟踪方法	默认方法是 CARD 128。
轨道控制区域	造影所在位置选择栏	选择栏有两个值（左和右），当 CEUS 录制为双幅模式，该值有效。左值表示左侧的为 CEUS 窗口，而右侧值表示右侧为 CEUS 窗口。在处理新视频之前，应设置该值。
	自动分析	当您单击该选项时，软件将对导入视频进行自动分析。它将分析记录窗口位置和录制模式。
	跟踪长度	跟踪长度可以通过手动输入数字或拖动下面的进度条来设置一次跟踪的帧数。
	跟踪	单击该选项时，当前绘图边界框将以当前帧的 Bounding Box 为初始值，执行一次跟踪。
	镜像	该选项是自动分析后的默认设置。如果分析错误，则可以设置该选项，点击后，双幅模式下单侧的 Bounding Box 会复制到另一侧图像窗口中。
	视频基本信息文本栏	这些文本表示加载视频的时间以及帧宽度和高度，以使用户了解正在处理的视频的空间分辨率和总时长。
	当前时间	文本指示整个视频中当前帧的时间。文本框可以手动输入以快速准确地跳到特定时刻。
	跟踪控制按钮组	这 8 个按钮分别表示 8 个不同的跟踪序列。每个按钮对应的跟踪可以从录制屏幕中的任何帧和任何位置开始。它们使用灰色、绿色、红色表示没有跟踪序列，已有跟踪序列和正在处理。当按钮为灰色时，按下按钮可以绘制当前帧上的边界框。当按钮为绿色时，按下按钮可以跳转到相应的跟踪区域

表 2.5 续表。

Table 2.5 Continue Table.

所在地区	组件	功能
视频播放窗口	播放上下文菜单	单击当前帧中的图像时，将显示上下文菜单。选项包括“下一张图片”、“前一张图片”、“下一张图片”、“最后一张图片”。它们设置为检查或按帧微调边界框。
	删除上下文菜单	单击当前框架的边界框中的区域时，将显示上下文菜单。右键单击将显示“删除当前帧的 Bounding Box”、“删除当前帧之前的所有 Bounding Box”、“删除当前帧之后的所有 Bounding Box”和“删除整个序列 Bounding Box”
	镜像上下文菜单	单击当前框架的边界框中的区域时，将显示上下文菜单。右键单击将显示“镜像当前帧的 Bounding Box”、“镜像当前帧之前的所有 Bounding Box”、“镜像当前帧之后的所有 Bounding Box”和“镜像整个序列 Bounding Box”。这些操作仅影响当前跟踪序列。使用镜像意味着边界框的位置将切换到另一个图窗。
视频控制区域	视频播放控制按钮组	控制视频播放的按钮，从左到右分别为“上一秒”、“上一帧”、“播放/停止”、“下一帧”、“下一秒”，用于精细检查跟踪结果是否需要调整。
	视频播放进度条	允许手动拖动，进度条的当前值图像在整个视频中的时间。
	勾画指示进度条	用不同颜色表示视频进度条上每个位置对应的图像是否有 Bounding Box。默认的颜色为白色，其他不同的颜色，分别代表不同的跟踪进度。
	自动分析进度条	以不同颜色指示视频进度条相应位置的图像模式。绿色表示帧处于单模式，棕色表示双幅显示模式，灰色表示未使用自动分析功能或此时间段的录制图像没有改变。