



(12)发明专利申请

(10)申请公布号 CN 110276286 A  
(43)申请公布日 2019.09.24

(21)申请号 201910511712.2

(22)申请日 2019.06.13

(71)申请人 中国电子科技集团公司第二十八研究所

地址 210007 江苏省南京市苜蓿园东街1号

(72)发明人 陈杰 罗子娟 缪伟鑫 李友江  
荀智德 尚小东

(74)专利代理机构 江苏圣典律师事务所 32237  
代理人 胡建华 于瀚文

(51) Int. Cl.

G06K 9/00(2006.01)

G06K 9/46(2006.01)

G06K 9/62(2006.01)

G06T 3/40(2006.01)

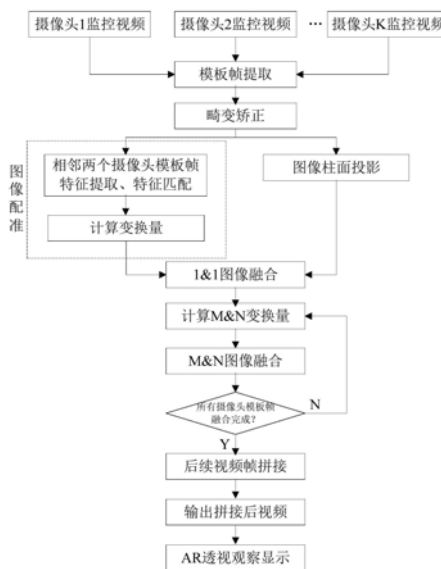
权利要求书2页 说明书8页 附图6页

(54)发明名称

一种基于TX2的嵌入式全景视频拼接系统

(57)摘要

本发明公开了一种基于TX2的嵌入式全景视频拼接系统,本发明能够用于对外部战场环境观察,拓展视野,并能够识别和判断周边环境威胁和异常,全方位提高作战人员感知战场环境的能力。在坦克等战车快速行驶过程中,车辆乘员在坦克内部能够利用本发明实时观察车周围360°信息,并且自动检测识别视频中的军事敏感目标,实时掌握车外战场环境军事目标信息,不需要将头或者身体探出车外,既扩大了观察视野,提高侦察效果,又能保证作战人员的人身安全。



1. 一种基于TX2的嵌入式全景视频拼接系统,其特征在于,包括视频采集分系统、视频拼接分系统、目标检测识别分系统和透视观察显示分系统;

所述视频采集分系统包括视频帧捕获模块和视频帧数据编码转换模块,所述视频帧捕获模块包括N路摄像头,视频帧捕获模块用于,实现对N路摄像头采集数据的捕获,并将捕获的视频数据传递给视频帧数据编码转换模块;

所述视频帧数据编码转换模块将视频转换为视频拼接分系统能够识别的RGB编码,并发送给视频拼接分系统;

所述视频拼接分系统利用OpenCV库进行实时的视频拼接融合,并将结果发送给目标检测识别分系统;

所述目标检测识别分系统用于,实现对视频中的军事敏感目标进行实时检测和识别,目标检测和识别结果将同步在透视观察显示分系统上进行显示。

2. 根据权利要求1中所述的系统,其特征在于,所述N路摄像头部署在战车周围或作为一个N路摄像头模组部署在车顶部,所述摄像头为非广角、非鱼眼的摄像头。

3. 根据权利要求2中所述的系统,其特征在于,所述视频拼接分系统部署在战车内部。

4. 根据权利要求3中所述的系统,其特征在于,所述视频拼接分系统利用OpenCV库进行实时的视频拼接融合,其中,对于第一帧视频帧图像的拼接融合处理包括如下步骤:

步骤1,畸变矫正:采用基于二维屏幕棋盘格模板的方法对图像进行矫正畸变;

步骤2,利用改进的SURF算法对视频帧图像提取特征点:在SURF特征矢量中增加了一个新的变量,即特征点的拉普拉斯属性,这两类特征点具有不同的属性,在匹配时只需对同类的特征点进行匹配,图像中两个特征点之间相似性的度量采用距离函数进行度量,通过相似性的度量就能够得到待配准图像间的潜在匹配点对,具体方法为:使用比值匹配法,即将一幅图像作为样本,在样本的特征点中寻找在另外一幅图像中与它距离最近的特征点和次近特征点,然后计算这两个特征点与样本点之间欧氏距离的比值,对于比值大于阈值的特征点,则判定是正确匹配的特征点;

步骤3,根据相邻两个摄像头采集画面的重叠区域的大小对匹配的特征点进行筛选,设定每个摄像头采集的图像的宽度为W,重叠区域比例为f,则重叠区域的宽度为 $f*W$ ,设定多个摄像头在水平方向呈圆形排列,则待融合的两幅源图像为左边图像、右边图像,左边图像的特征点为l-Points,右边图像的特征点为r-Points,特征点筛选方法为:如果,l-Points的宽度cols值大于 $f*W$ ,同时小于W,该匹配点保留,否则丢弃;如果,r-Points的宽度cols值小于 $(1-f)*W$ ,同时大于0,该匹配点保留,否则丢弃;

步骤4,采用柱面投影的方法对平面图像进行预处理,将在平面上提取的经过筛选的匹配特征点进行柱面投影变换;

步骤5,确定平移量:利用步骤3筛选过的高度匹配的特征点计算出图像之间的平移量,然后将源图像平移到目标图像上;图像拼接平移量的计算方法是取所有匹配特征点平移量的平均值作为图像的平移量,计算公式为:

$$\begin{cases} diff_x = \frac{\sum_i^N (W_t - P_{i,x} + P'_{i,x})}{N} \\ diff_y = \frac{\sum_i^N (P_{i,y} - P'_{i,y})}{N} \end{cases}$$

其中,  $W_t$  是目标图像宽度,  $P_{i,x}$  是目标图像上的特征点横坐标,  $P'_{i,x}$  是源图像上的对应特征点的横坐标,  $P_{i,y}$  是目标图像上的特征点纵坐标,  $P'_{i,y}$  是源图像上的对应特征点的纵坐标,  $N$  是匹配的特征点数量,  $diff_x$  是  $x$  方向平移量,  $diff_y$  是  $y$  方向平移量。

步骤6, 图像融合: 图像融合之前先分别进行柱面投影, 将圆形分布的两个以上的视频帧投影到柱面上, 取左边的图像为目标图像, 右边的图像为源图像, 源图像在平移量的作用下, 与目标图像融合。

5. 根据权利要求4中所述的系统, 其特征在于, 步骤6包括: 设定图像拼接结果为  $I_{result}$ ,  $W$  和  $H$  分别表示拼接结果的宽和高,  $W'_t$  和  $H'_t$  分别表示目标图像柱面投影后的宽和高,  $W'_s$  是源图像投影后的宽,  $W$  和  $H$  的计算方法为:

$$\begin{cases} W = W'_t + W'_s - diff_x \\ H = H'_t \end{cases} \quad (9)$$

6. 根据权利要求5中所述的系统, 其特征在于, 将圆形分布的摄像头采集的视频帧图像用步骤1至步骤6的方法进行图像融合操作, 获得  $360^\circ$  的全景视频。

7. 根据权利要求6中所述的系统, 其特征在于, 所述目标检测识别分系统部署在战车内部, 目标检测识别分系统包括目标定位跟踪模块, 目标检测识别模块、网络传输模块和视频信息发送模块;

其中, 目标定位跟踪模块, 用于, 对视频中的运动目标进行定位跟踪;

目标检测识别模块, 用于, 基于深度卷积神经网络的人工智能技术实现对视频中军事敏感目标进行检测识别, 并且将检测结果叠加到全景视频画面上;

网络传输模块, 用于, 实现全景视频的压缩编码、发送和接收;

视频信息发送模块, 用于, 实现叠加目标检测识别信息的全景视频的发送, 将视频帧信息转换为格式化的二进制信息, 并且调用网络传输模块, 将信息发送到车载显示终端或者AR增强现实眼镜上。

8. 根据权利要求7中所述的系统, 其特征在于, 所述透视观察显示分系统部署在战车内部, 用于, 实现对全景视频的接收和显示; 显示载体为车内显示终端或AR增强现实眼镜;

透视观察显示分系统包括视频信息接收处理模块和界面绘制显示模块;

其中, 视频信息接收处理模块, 用于, 将叠加了目标检测识别信息的全景视频从二进制转化为能够绘制显示的格式;

界面绘制显示模块, 用于, 将经过视频信息接收处理模块处理过的全景视频信息进行实时刷新显示。

## 一种基于TX2的嵌入式全景视频拼接系统

### 技术领域

[0001] 本发明属于图像/视频处理和人工智能领域,尤其涉及一种基于TX2的嵌入式全景视频拼接系统。

### 背景技术

[0002] 图像拼接是视频拼接的核心和基础,图像拼接问题由来已久。在现实生活中,常常想要获取大范围场景,但是由于摄像头拍摄角度有限,很难抓取到所期望得到的全部信息。虽然一些专业摄像头,广角摄像头和鱼眼摄像头的视角相对宽阔,但是一方面价格昂贵,另一方面图像失真严重,依然无法满足需求。因此,想要将多张有重叠区域的小视野图像对齐、构成大视野高分辨率图像,这就是图像拼接技术。图像拼接有效解决了单张图像分辨率低、拍摄视角小的问题。随着计算机技术的发展和计算机视觉的进步,也为图像拼接提供了广大应用空间,目前图像拼接技术被广泛应用于工业生产、日常生活、文化娱乐等多种领域。

[0003] 视频拼接技术在学术界已经研究较多,但是利用NVIDIA Jetson TX2嵌入式平台的视频拼接技术,实现在复杂作战条件下对战车周围视频进行实时360度拼接,并且对视频中敏感军事目标进行检测识别是一个新的应用方式,国内目前在这方面的应用较少。在军事应用方面,未来战场形势瞬息万变,坦克/装甲等战车乘员如何利用自身装备实时快速准确地获取战场作战信息是取得战斗胜利的关键所在,目前的坦克/装甲车辆等装备通常通过潜望镜或周视仪观察外部情况,易造成视野狭窄,难以快速识别周边威胁、定位目标,无法满足准确快速的作战侦察需求。

[0004] 目标检测与识别是计算机视觉领域中一个基础性的研究课题,相关研究工作一直是计算机视觉的研究热点。特殊类别的目标检测,例如人脸、行人、民用车辆,检测技术已经较为成熟。但是,目前的检测识别所用的视频源基本都是摄像头静止可控情况下采集的,在复杂的作战环境中,针对车辆快速颠簸运动情况下采集的视频,如何准确快速地对视频中的重点目标进行检测识别尚没有相关的研究。

### 发明内容

[0005] 本发明为战术战斗级作战车辆作战侦察装备的构建问题,提出一种基于TX2的嵌入式全景视频拼接系统。

[0006] 本发明系统利用NVIDIA Jetson TX2核心板作为小型轻量化的嵌入式平台,该平台是基于NVIDIA Pascal™架构的AI单模块超级计算机,配置256CUDA cores,具有超强的GPU计算能力,可进行基于深度学习的人工智能技术开发。本发明主要基于该平台提供的NVIDIA Jetson SDK接口和高性能运算能力进行图像视频拼接和视频目标检测识别。

[0007] 本发明系统包括视频采集分系统、视频拼接分系统、目标检测识别分系统和透视观察显示分系统;

[0008] 所述视频采集分系统包括视频帧捕获模块和视频帧数据编码转换模块,所述视频

帧捕获模块包括N路摄像头,视频帧捕获模块用于,实现对N路摄像头采集数据的捕获,捕获帧率设置为30fps,捕获的数据为UYVY 4:2:2编码的视频,并将捕获的视频数据传递给视频帧数据编码转换模块;

[0009] 所述视频帧数据编码转换模块将视频转换为视频拼接分系统能够识别的RGB编码,并发送给视频拼接分系统;

[0010] 所述视频拼接分系统利用OpenCV库进行实时的视频拼接融合,并将结果发送给目标检测识别分系统;

[0011] 所述目标检测识别分系统用于,实现对视频中的军事敏感目标进行实时检测和识别,目标检测和识别结果将同步在透视观察显示分系统上进行显示。

[0012] 本发明采用基于柱面投影的方法实现全景视频拼接。

[0013] 所述N路摄像头部署在战车周围或作为一个N路摄像头模组部署在车顶部,所述摄像头为非广角、非鱼眼的普通摄像头。

[0014] 所述视频拼接分系统部署在战车内部。

[0015] 所述视频拼接分系统利用OpenCV库进行实时的视频拼接融合,其中,对于第一帧视频帧图像的拼接融合处理包括如下步骤:

[0016] 步骤1,畸变矫正:摄像头提取的原始画面存在一定的畸变,采用基于二维屏幕棋盘格模板的方法对图像进行矫正畸变,可最大限度地还原真实画面(参考文献:Zhang Z Y.A Flexible new technique for camera calibration[J].Transactions on Pattern Analysis and Machine Intelligence,2000,22(11):1330-1334.)。本步骤中获取的摄像头参数和畸变系数可保留用于后续视频帧。

[0017] 步骤2,利用改进的SURF(加速稳健特征,Speeded Up Robust Features,SURF)

[0018] 算法对视频帧图像提取特征点:在SURF特征矢量中增加了一个新的变量,即特征点的拉普拉斯属性,这两类特征点具有不同的属性,在匹配时只需对同类的特征点进行匹配,图像中两个特征点之间相似性的度量采用距离函数进行度量,通过相似性的度量就能够得到待配准图像间的潜在匹配点对,具体方法为:使用比值匹配法,即将一幅图像作为样本,在样本的特征点中寻找在另外一幅图像中与它距离最近的特征点和次近特征点,然后计算这两个特征点与样本点之间欧氏距离的比值,对于比值大于阈值(一般取值0.8)的特征点,则判定是正确匹配的特征点;

[0019] 本发明在SURF特征矢量中增加了一个新的变量,加快匹配过程,即特征点的拉普拉斯相应正负号,这两类特征点具有不同的属性,在匹配时只需对同类的特征点进行匹配,这样会大大地提高匹配速度和精度。图像中两个特征点之间相似性的度量一般采用欧氏距离等距离函数进行度量,通过相似性的度量就可以得到待配准图像间的潜在匹配点对。具体方法为:使用比值匹配法,即将其中的一幅图像作为样本,在样本的特征点中寻找在另外一幅图像中与它距离最近的特征点和次近特征点,然后计算这两个特征点与样本点之间欧氏距离的比值。对于比值小于阈值范围内的特征点,则判定是正确匹配的特征点。

[0020] 步骤3,根据相邻两个摄像头采集画面的重叠区域的大小对匹配的特征点进行筛选,设定每个摄像头采集的图像的宽度为W,重叠区域比例为f,则重叠区域的宽度为 $f*W$ ,设定多个摄像头在水平方向呈圆形排列,则待融合的两幅源图像为左边图像、右边图像,左边图像的特征点为l-Points,右边图像的特征点为r-Points,特征点筛选方法为:如果,l-

Points的宽度cols值大于 $f*W$ ,同时小于 $W$ ,该匹配点保留,否则丢弃;如果,r-Points的宽度cols值小于 $(1-f)*W$ ,同时大于0,该匹配点保留,否则丢弃;

[0021] 本发明提出一种改进的特征点筛选方法,减少特征点误匹配。本发明根据相邻两个摄像头之间重叠区域的大小对匹配的特征点进行高效筛选。假设摄像头采集的源图像的宽度为 $W$ ,重叠区域比例为 $f$ ,那么重叠区域的宽度为 $f*W$ ,本发明中多个摄像头在水平方向呈圆形排列,因此,假设两幅待拼接融合的源图像分别为左边图像和右边图像。左边图像的特征点为l-Points,右边图像的特征点为r-Points,特征点筛选方法为:如果,l-Points的cols值大于 $f*W$ ,小于 $W$ ,该匹配点保留,否则丢弃;如果,r-Points的cols值小于 $(1-f)*W$ ,大于0,该匹配点保留,否则丢弃。该方法对重叠区域之外的匹配特征点对剔除,提高了特征匹配的精准度。避免了重叠区域之外的误匹配特征点对,同时减少了程序的耗时。

[0022] 步骤4,为了在全景视频拼接时保持图像的空间约束与视觉一致性,采用柱面投影的方法对平面图像进行预处理,将在平面上提取的经过筛选的高度匹配特征点进行柱面投影变换,摄像头采集的画面都是平面的(相对于柱面而言);

[0023] 本发明采用柱面投影的方法,将在平面上提取的经过筛选的高度匹配特征点进行柱面投影变换,以便准确计算两幅图像在后期融合时需要的变换量。

[0024] 步骤5,确定平移量:利用步骤3筛选过的高度匹配的特征点计算出图像之间的平移量,然后将源图像平移到目标图像上;图像拼接平移量的计算方法是取所有匹配特征点平移量的平均值作为图像的平移量,计算公式为:

$$[0025] \quad \begin{cases} diff_x = \frac{\sum_i^N (W_t - P_{i,x} + P'_{i,x})}{N} \\ diff_y = \frac{\sum_i^N (P_{i,y} - P'_{i,y})}{N} \end{cases}$$

[0026] 其中, $W_t$ 是目标图像宽度, $P_{i,x}$ 是目标图像上的特征点横坐标, $P'_{i,x}$ 是源图像上的对应特征点的横坐标, $P_{i,y}$ 是目标图像上的特征点纵坐标, $P'_{i,y}$ 是源图像上的对应特征点的纵坐标,, $N$ 是匹配的特征点数量, $diff_x$ 是x方向平移量, $diff_y$ 是y方向平移量。本步骤中获取的平移量可保留用于后续视频帧。

[0027] 步骤6,图像融合:图像融合之前先分别进行柱面投影,将圆形分布的两个以上的视频帧投影到柱面上,取左边的图像为目标图像,右边的图像为源图像,源图像在平移量的作用下,与目标图像融合。

[0028] 步骤6包括:设定图像拼接结果为 $I_{result}$ , $W$ 和 $H$ 分别表示拼接结果的宽和高, $W'_t$ 和 $H'_t$ 分别表示目标图像柱面投影后的宽和高, $W'_s$ 是源图像投影后的宽, $W$ 和 $H$ 的计算方法为:

$$[0029] \quad \begin{cases} W = W'_t + W'_s - diff_x \\ H = H'_t \end{cases} \quad (9)$$

[0030] 本发明中多个摄像头在水平方向圆形分布,将圆形分布的摄像头采集的视频帧图像用步骤1至步骤6的方法进行图像融合操作,可获得 $360^\circ$ 的全景视频。

[0031] 本发明解决了不同朝向摄像头拍摄画面不满足视觉一致性的问题。经过柱面投影后的图像通过简单的平移就可以实现图像对齐融合。图像融合的具体方法为:取左边的图像为目标图像,右边的图像为源图像,源图像在平移量的作用下,可以与目标图像完美融

合。

[0032] 后续视频帧图像融合过程中所需的摄像头参数和畸变系数、图像平移量在模板帧即第一帧视频帧图像处理中就已得到,所以在后续视频帧的融合过程中免于进行摄像头标定、特征检测和匹配以及平移量计算,这对算法速度的提升起到很大作用。

[0033] 另外,柱面投影变换和图像融合过程都使用了NVIDIA GPU的CUDA平台并行加速所以后续视频帧的处理速度快,可以做到实时拼接。实验结果表明,在NVIDIA 1080Ti GPU上运行该算法,可以实时地生成无缝清晰的全景视频。

[0034] 所述目标检测识别分系统部署在战车内部,目标检测识别分系统包括目标定位跟踪模块,目标检测识别模块、网络传输模块和视频信息发送模块;

[0035] 其中,目标定位跟踪模块,用于,利用麻省理工学院人工智能实验室提出的双目视觉技术实现对视频中的运动目标(士兵、战车、无人机等)进行定位跟踪;

[0036] 目标检测识别模块,用于,基于深度卷积神经网络的人工智能技术实现对视频中军事敏感目标进行检测识别,并且将检测结果叠加到全景视频画面上;目标检测识别主要的难点是对视频图像检测的实时性问题。本发明针对基于深度学习的回归方法(YOLO/SSD/RNN等算法),调整优化参数,并且适当的减少RPN网络输出的候选框数量(减少30%),提升检测识别速度。

[0037] 网络传输模块,用于,实现全景视频的压缩编码、发送和接收;

[0038] 视频信息发送模块,用于,实现叠加目标检测识别信息的全景视频的发送,将视频帧信息转换为格式化的二进制信息并且调用网络传输模块,将信息发送到车载显示终端或者AR增强现实眼镜上,以便于后期观察显示。视频帧绘制显示支持的格式为:bmp,png,jpg等。

[0039] 所述透视观察显示分系统部署在战车内部(包含显示载体硬件和显示软件),用于,实现对全景视频的接收和显示;显示载体为车内显示终端或AR增强现实眼镜(观察显示需要显示载体,例如头显或者平面显示器等,是透视观察显示分系统的硬件部分);

[0040] 透视观察显示分系统包括视频信息接收处理模块和界面绘制显示模块;

[0041] 其中,视频信息接收处理模块,用于,将叠加了目标检测识别信息的全景视频从二进制转化为能够绘制显示的MJPG4等格式;

[0042] 界面绘制显示模块,用于,将视频信息接收处理模块处理过的全景视频信息进行实时刷新显示。

[0043] 本发明首次将TX2平台引入到军用作战车辆作战侦察装备的构建中,与现有技术相比,其显著优点为:

[0044] (1)体积小。军事作战车辆内部空间及其有限,本发明使用中的TX2核心板面积大小只相当于一张银行卡,极大限度地降低了处理平台在车辆内占用的物理空间。

[0045] (2)低功耗。军事作战车辆在实际运动作战中,无法携带大功率电源,所以装备的供电是一个难题。本平台功耗仅为9W,对电源的依赖性低,可以长时间在野外工作。

[0046] (3)效率高。利用该平台的CUDA并行计算能力,在图像进行柱面投影时采用CUDA加速,可使视频拼接帧率达到近实时,同时可以快速地进行目标定位、跟踪和识别。检测识别所花时间是毫秒级。

[0047] (4)智能化:能够根据用户的第一人称视角,智能地显示用户关注的军事目标等信

息。

[0048] (5) 便携化:解放战车乘员双手,可在不影响本身工作的基础上对战车周围环境进行360度透视观察,并且自动检测识别视频中的军事敏感目标。这一优点在需要用双手操控坦克/装甲装备的战术用户身上体现尤为明显。

[0049] (6) 安全性。车辆乘员不需要将头或者身体探出车外,既扩大了观察视野,提高侦察效果,又能保证作战人员的人身安全。

[0050] 本发明借助布置于战车周围的多个摄像头进行周边环境的拍摄,进而对这些拍摄的视频帧进行特征点提取匹配融合,实现360度无缝的全景视频拼接处理。本技术综合了利用图像融合、虚拟与增强现实、通信导航等技术,将战场环境的360度视频投影在车载显示终端或装甲/坦克车辆乘员佩戴的增强现实头盔显示器上,使战车乘员能够在封闭甚至无窗的乘员舱内“透过”车辆装甲,对车外的环境进行实时观察,并且基于人工智能方法自动、高效检测识别车外士兵、坦克、战车和飞机等敏感目标,使乘员获得对车外环境的实时情况的感知能力,大幅提高索敌速度和准确度,增加歼敌几率。用于支撑战术级坦克/装甲车辆以及其他战车的作战侦察水平。

## 附图说明

[0051] 下面结合附图和具体实施方式对本发明做更进一步的具体说明,本发明的上述或其他方面的优点将会变得更加清楚。

[0052] 图1是一种基于TX2的嵌入式全景视频拼接系统柱面全景视频拼接及AR透视观察流程图。

[0053] 图2是一种基于TX2的嵌入式全景视频拼接系统视频采集分系统摄像头模组图。

[0054] 图3a是图像畸变矫正前效果图。

[0055] 图3b是图像畸变矫正后效果图。

[0056] 图4a是特征点筛选前示意图图。

[0057] 图4b是特征点筛选后示意图。

[0058] 图5a是一种基于TX2的嵌入式全景视频拼接系统摄像头排序方式柱面投影图。

[0059] 图5b是一种基于TX2的嵌入式全景视频拼接系统柱面投影图。

[0060] 图6a是图像投影前效果图。

[0061] 图6b是图像投影后效果图。

[0062] 图6c是图像融合后效果图。

[0063] 图7是一种基于TX2的嵌入式全景视频拼接系统8路摄像头视频全景拼接过程图。

[0064] 图8是一种基于TX2的嵌入式全景视频拼接系统多路摄像头全景拼接效果图。

[0065] 图9是一种基于TX2的嵌入式全景视频拼接系统全景视频目标检测识别效果图。

## 具体实施方式

[0066] 下面结合附图及实施例对本发明做进一步说明。

[0067] 基于TX2的嵌入式全景视频拼接系统,利用NVIDIA Jetson TX2核心板作为小型轻量化的嵌入式平台,该平台是基于NVIDIA Pascal™架构的AI单模块超级计算机,配置256CUDA cores,具有超强的GPU计算能力,可进行基于深度学习的人工智能技术开发。本发



明主要基于该平台提供的NVIDIA Jetson SDK接口和高性能运算能力进行图像视频拼接和视频目标检测识别。

[0068] 基于TX2的嵌入式全景视频拼接系统由视频采集分系统、视频拼接分系统、目标检测识别分系统和透视观察显示分系统组成。本发明采用基于柱面投影的方法实现全景视频拼接,图1是柱面全景视频拼接及AR透视观察流程图。

[0069] 所述视频采集分系统可部署在战车周围或作为一个N路摄像头模组部署在车顶部,采用的摄像头为非广角/鱼眼的普通摄像头。图2是视频采集摄像头模组。视频采集分系统包括:视频帧捕获模块,视频帧数据编码转换模块等2个模块。每个模块的具体功能如下:

[0070] 视频帧捕获模块,实现对多路摄像头采集数据的捕获,捕获帧率设置为30fps,捕获后传递给视频帧数据编码转换模块。

[0071] 视频帧数据编码转换模块,本发明采用的摄像头捕获的视频格式为UYVY 4:2:2编码,本模块将其转换为视频拼接输入能够识别的RGB编码。

[0072] 所述视频拼接分系统可部署在战车内部。视频拼接分系统利用OpenCV库进行实时的视频拼接融合。视频拼接分系统中第一个视频帧图像的处理主要分为以下几个步骤:

[0073] (1) 畸变矫正。摄像头提取的原始画面存在一定的畸变,本发明采用基于二维屏幕棋盘格模板的方法矫正畸变,可最大限度地还原真实画面。本步骤中获取的摄像头参数和畸变系数可保留用于后续视频帧。图3a是图像畸变矫正前效果图。图3b是图像畸变矫正后效果图。

[0074] (2) 提取特征点。利用改进的SURF算法对视频帧图像提取特征点。本发明在SURF特征矢量中增加了一个新的变量,加快匹配过程,即特征点的拉普拉斯相应正负号,这两类特征点具有不同的属性,在匹配时只需对同类的特征点进行匹配,这样会大大地提高匹配速度和精度。图像中两个特征点之间相似性的度量一般采用欧氏距离等距离函数进行度量,通过相似性的度量就可以得到待配准图像间的潜在匹配点对。具体方法为:使用比值匹配法,即将其中的一幅图像作为样本,在样本的特征点中寻找在另外一幅图像中与它距离最近的特征点和次近特征点,然后计算这两个特征点与样本点之间欧氏距离的比值。对于比值小于某一阈值范围内的特征点,则认为正确匹配的特征点。

[0075] (3) 特征点筛选。本分系统提出一种改进的特征点筛选方法,减少特征点误匹配。本发明根据相邻两个摄像头之间重叠区域的大小对匹配的特征点进行高效筛选。假设图像的宽度为W,重叠区域比例为f,那么重叠区域的宽度为 $f*W$ ,左边图像的特征点为l-Points,右边图像的特征点为r-Points,。特征点筛选方法为:如果,l-Points的cols值大于 $f*W$ ,小于W,该匹配点保留,否则丢弃;如果,r-Points的cols值小于 $(1-f)*W$ ,大于0,该匹配点保留,否则丢弃。该方法对重叠区域之外的匹配特征点对剔除,提高了特征匹配的精准度。避免了重叠区域之外的误匹配特征点对,同时减少了程序的耗时。图4a是特征点筛选前示意图。图4b是特征点筛选后示意图。

[0076] (4) 柱面投影。本发明采用柱面投影的方法,将在平面上提取的经过筛选的高度匹配特征点进行柱面投影变换,以便准确计算两幅图像在后期融合时需要的变换量。图5a是一种基于TX2的嵌入式全景视频拼接系统摄像头排序方式柱面投影图。图5b是一种基于TX2的嵌入式全景视频拼接系统柱面投影图。图6a是图像投影前效果图。

[0077] (5) 确定平移量。经利用步骤3筛选过的高度匹配的特征点集计算出图像之间的平

移量,然后将源图像平移到目标图像上即可。本步骤中获取的平移量可保留用于后续视频帧。

[0078] (6) 图像融合。图像平移之后就可以进行融合。图像融合之前先分别进行柱面投影,将圆形分布的多个视频帧投影到柱面上,解决了不同朝向摄像头拍摄画面不满足视觉一致性的问题。图6b是图像投影后效果图。经过柱面投影后的图像通过简单的平移就可以实现图像对齐融合。图像融合的具体方法为:取左边的图像为目标图像,右边的图像为源图像,源图像在平移量的作用下,可以与目标图像完美融合。图6c是图像融合后效果图。

[0079] (7) 后续视频帧图像融合过程中所需的摄像头参数和畸变系数、图像平移量在模板帧处理中就已得到,所以在后续视频帧的融合过程中免于进行摄像头标定、特征检测和匹配以及平移量计算,这对算法速度的提升起到很大作用。

[0080] 另外,柱面投影变换和图像融合过程都使用了NVIDIA GPU的CUDA平台并行加速所以后续视频帧的处理速度快,可以做到实时拼接。实验结果表明,在NVIDIA 1080Ti GPU上运行该算法,可以实时地生成无缝清晰的全景视频。图7是8路摄像头视频全景拼接过程图。图8是全景视频效果图。

[0081] 所述目标检测识别分系统部署在战车内部,实现对视频中的军事敏感目标进行实时检测和识别,目标检测识别结果将同步在车内显示终端和增强现实眼镜上叠加显示。目标检测识别分系统包括:目标定位跟踪模块,目标检测识别模块、网络传输模块,视频信息发送模块等4个模块。4个模块的功能描述如下:

[0082] 目标定位跟踪模块,利用麻省理工学院人工智能实验室提出的双目视觉技术实现对视频中的运动目标(士兵、战车、无人机等)进行定位跟踪。

[0083] 目标检测识别模块,基于深度卷积神经网络的人工智能技术实现对视频中军事敏感目标进行检测识别。目标检测识别主要的难点是对视频图像检测的实时性问题。本发明针对基于深度学习的回归方法(YOLO/SSD/RNN等算法),调整优化参数,并且适当的减少RPN网络输出的候选框数量,提升检测识别速度。图9是全景视频目标检测识别效果图。

[0084] 军事目标数据包含以下字段信息:

[0085] 目标属性:序号、类别、方位、距离、敌我属性、行动意图、威胁程度;

[0086] 军队关系属性:序号、类别、隶属部队、编制级别;

[0087] 进行测试的目标信息字段如下表1所示:

[0088] 表1

[0089]

属性名称	长度(字节)	类型
序号	4	UINT
类别	20	char[ ]
方位	25	char[ ]
距离	4	UINT
敌我属性	8	char[ ]
行动意图	254	char[ ]
威胁程度	32	char[ ]

[0090] 网络传输模块,实现全息视频信息的压缩编码、发送和接收。

[0091] 视频信息发送模块,实现叠加目标检测识别信息的全景视频的发送,将视频帧信息转换为格式化的二进制信息并且调用网络传输模块,将信息发送到车载显示终端或者AR增强现实眼镜上,以便于后期观察显示。视频帧绘制显示支持的格式为: bmp, png, jpg等。

[0092] 所述透视观察显示分系统部署在战车内部,实现对全景视频流的接收和显示。显示载体可为车内显示终端或AR增强现实眼镜。透视观察显示分系统包括: 视频信息接收处理模块和界面绘制显示模块2个模块。2个模块的功能描述如下:

[0093] 视频信息接收处理模块,将叠加了目标检测识别信息的全景视频从二进制转化为可以绘制显示的MJPG4等格式。

[0094] 界面绘制显示模块,将经过编码的全景视频信息进行实时刷新显示。

[0095] 本发明提供了一种基于TX2的嵌入式全景视频拼接系统,具体实现该技术方案的方法和途径很多,以上所述仅是本发明的优选实施方式,应当指出,对于本技术领域的普通技术人员来说,在不脱离本发明原理的前提下,还可以做出若干改进和润饰,这些改进和润饰也应视为本发明的保护范围。本实施例中未明确的各组成部分均可用现有技术加以实现。

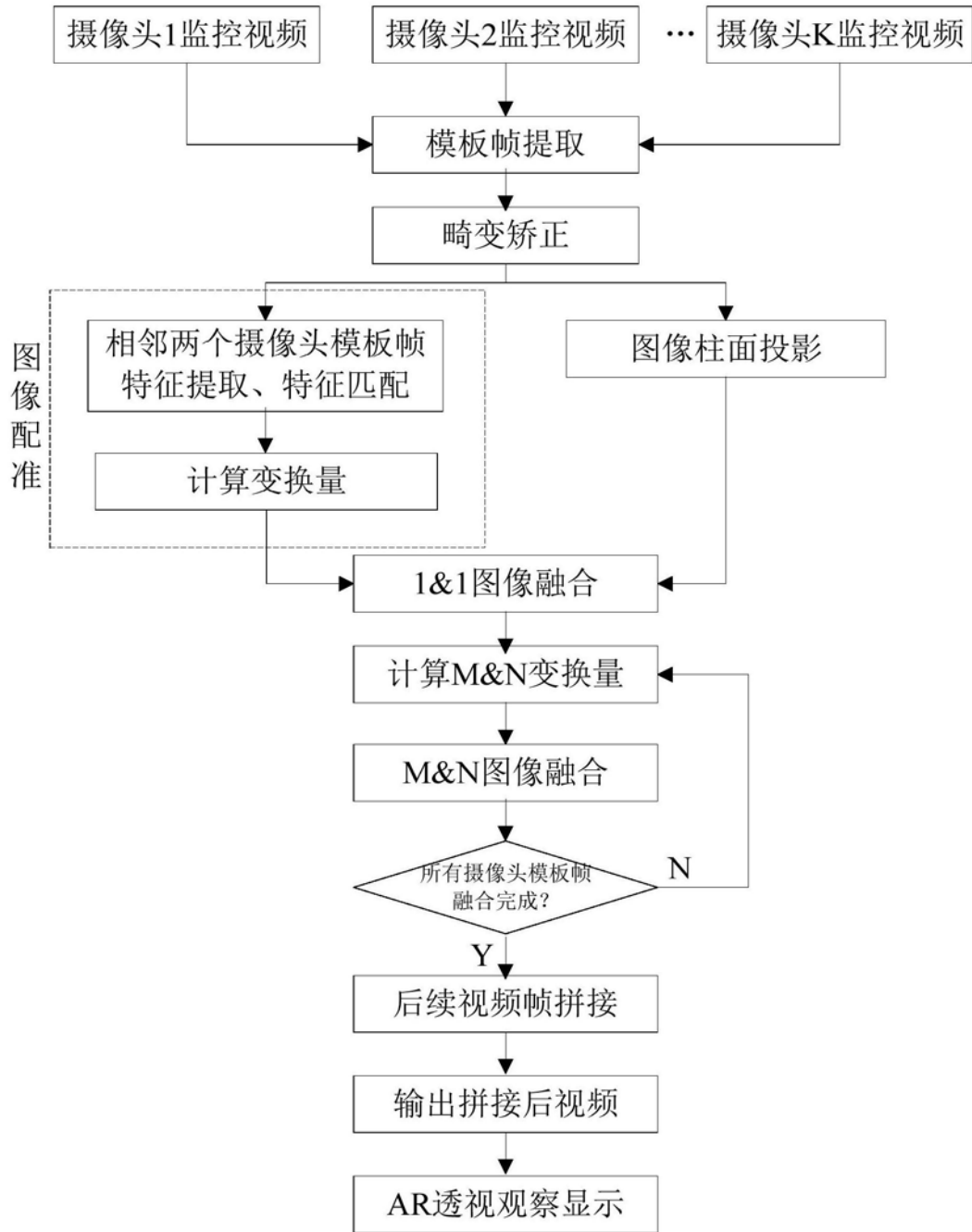


图1



图2



图3a



图3b



图4a

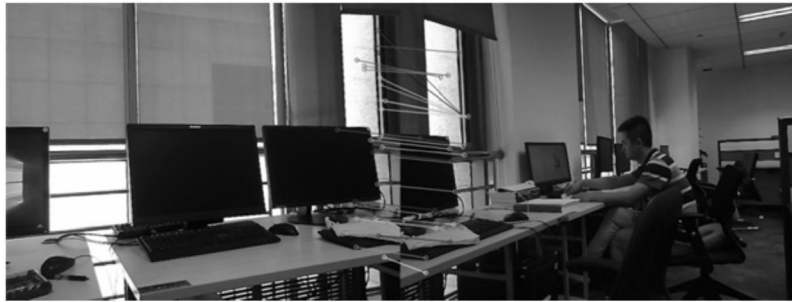


图4b

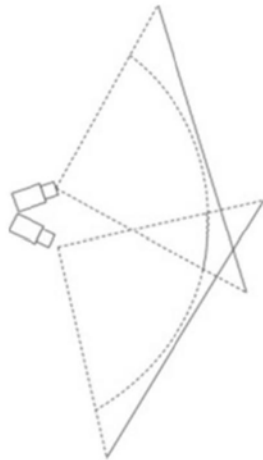


图5a

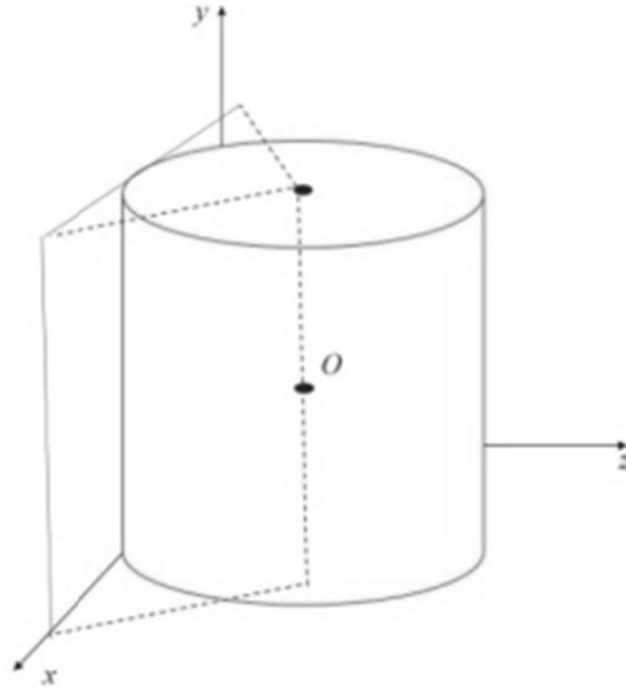


图5b



图6a



图6b



图6c

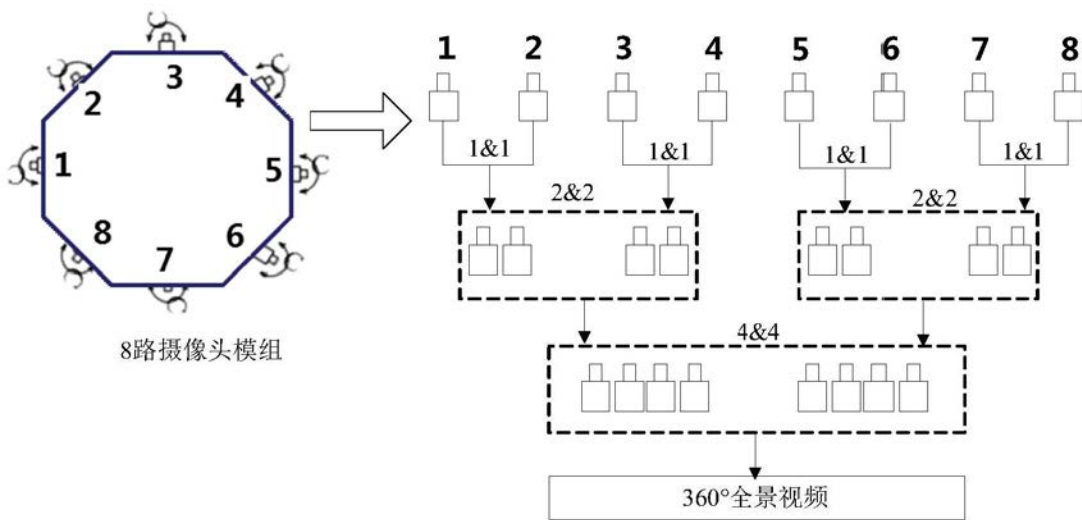


图7





图8



图9