

多目超高清机场视频拼接系统设计与软件实现

方诤言,陈贤富

(中国科学技术大学 微电子学院,安徽 合肥 230026)

摘要:图像拼接技术应用范围十分广泛,如全景图片、城市建模以及视频监控等工程应用领域。针对机场中多台超高清摄像机监控的场景,设计了一种全景多路超高清监控视频拼接系统。该系统利用具有重叠区域的多路视频进行实时拼接,从而形成具有更大视角的全景视频,给决策人员提供全面而有效的支持。系统将传统的图像拼接技术运用在视频拼接上,采用分布式的系统架构并设计了一种视频文件传输算法,保证了系统流畅、稳定地运行。首先介绍了该系统软硬件环境,给出了系统的总体设计;其次介绍了系统的关键功能设计以及主要界面实现;最后进行了实验测试,得出相关实验数据。经过测试,该系统实现了8路4K视频的拼接,显示流畅,无明显拼接痕迹,且对一帧的处理时间小于40 ms。

关键词:视频拼接;超高清视频;分布式系统;机场监控

中图分类号: TP391.41

文献标识码: A

DOI: 10.19358/j.issn.2096-5133.2020.12.011

引用格式: 方诤言,陈贤富. 多目超高清机场视频拼接系统设计与软件实现[J]. 信息技术与网络安全, 2020, 39(12): 61-66.

Design and implementation of multi-eye ultra-high-definition airport video stitching system

Fang Zhengyan, Chen Xianfu

(School of Microelectronics, University of Science and Technology of China, Hefei 230026, China)

Abstract: Image stitching technology is widely used in engineering fields, including panoramic pictures, city modeling and video surveillance. A panoramic multi-channel UHD(Ultra High Definition) surveillance video stitching system was designed for scenes monitored by multiple UHD cameras in the airport. Through real-time stitching multi-channel videos with overlapping areas, the system forms a panoramic video with a larger angle of view and provides comprehensive and effective support for decision makers. In addition, the system applies the traditional image stitching technology to video stitching, adopts a distributed system architecture and designs a transmission algorithm of video files, which ensures the smooth and stable operation of the stitching system. This paper firstly introduces the software and hardware environment of the system, and presents the overall design of the system; secondly, it introduces the design of key function of the system and the interface implementation of the main system; finally, it presents relevant experimental data obtained from an experimental test. After testing, this system has realized the stitching of 8-channel 4K video, with its features of smooth display, no-obvious-stitching traces and less than 40 ms of processing time per frame.

Key words: image stitching; ultra high definition; distributed systems; airport surveillance

0 引言

图像拼接技术是将数张有重叠部分的图像(可能是不同时间、不同视角或者不同传感器获得的)拼成一幅大型无缝高分辨率图像^[1]。日本的内藤诚^[2]采用多个PC平台的分布式架构来实现4K超高清实时视频的编码,同时基于MPI(信息传递接口)通

信协议实现多机通信。布朗^[3]采用不变的局部特征作为全局特征进行匹配,该方法适用于摄像机偏移忽略不计的理想状态。民航二所科研中心自主研发的全景增强监视系统,采用8台激光投影机与幕布,通过投影融合和视频墙控制显示全景视频。其优点是可以实现相对的无缝拼接,缺点是采用硬件的手

段构建全景视频,硬件成本高。国防科技大学的李静^[4]等人运用了一种贝叶斯特征细化模型来自适应地去除不正确的局部匹配点,提高匹配点的正确率,减少拼接后的图像出现重影。但对于实时性要求高的场景,该方案无法实现实时显示。电子科技大学的张浩^[5]实现了基于异构多核的DSP处理器进行高清1920×1080分辨率图像拼接。这种基于DSP处理器系统架构对于4K以上分辨率的图像处理实时性较差。

经过文献调研后,本文设计了一种全景多路超高清监控视频拼接系统。该系统运用了分布式架构,使用了多线程处理技术,并提出了一种基于SMB(Server Message Block)协议的跨服务器传输视频流算法,采用了基于预计算拼接参数的拼接模块,大大提高了拼接的实时性。最终实现了机场监控8路4K视频实时拼接。

1 开发环境和总体架构设计

1.1 软件开发环境

OpenCV是一个开源的跨平台计算机视觉库。OpenCV计算机视觉库作为常用的图像处理库之一,功能强大,支持目前先进的图像处理技术且体系十分完善,几乎涵盖了近十年内的主流算法^[6],可用于开发实时的图像处理、计算机视觉以及模式识别程序。

VLC多媒体播放器(VLC media player)是可移植的、开源的支持多声道视频编码及各种流媒体协议的多媒体播放器。

OpenCV的优点在于灵活性强,可对视频帧进行处理。VLC获取视频流速度快,支持多种格式。

1.2 硬件开发环境

该系统通过千兆交换机连接8台索尼(SNC-VB770)4K网络摄像机与5台服务器。交换机为思科catalyst 3560千兆交换机。网络摄像机每秒最高输出30帧4K(3840×2160)视频。

服务器硬件参数如表1所示。

表1 计算机详细参数

类别	数量	主要参数
CPU	2	Intel至强6244:8核心最大频率4.4GHz
GPU	1	NIVEA P4000:显存8GB,CUDA核2304个,张量核288个
内存	4	16GB,频率2400MHz

1.3 总体架构设计

传统的图像处理技术对于4K分辨率图片处理,时间过长且无法保证实时性。表2是使用传统图像处理技术对4K图片操作进行测试,得出的花费时间。

表2 对4K图片处理时间

算法	花费时间/ms
获取视频帧	15
色差处理	25
视频拼接	大于100
视频显示	6

表2给出了4K图像获取视频帧、色差处理、视频拼接、视频显示所花费的时间。视频拼接这一操作,仅仅针对两张图片进行处理,花费的时间大于100ms,也就是每秒处理图像小于10帧。如果对视频处理速度小于每秒10帧,视频显示的流畅度则无法保证。故本文不仅选用了性能更强的设备,而且使用并行处理技术,还设计了如图1所示的分布式系统架构来加速视频拼接算法的运行速度。

该系统有5台服务器,包括4台下位机、1台上位机。每台下位机负责对2路4K视频流进行处理:获取视频帧、缩放视频帧、色差调整、传输到上位机,上位机获取8路视频进行拼接、存储以及显示。利用该分布式系统能够提高系统的处理速度,输出20帧/s的画面。

2 系统设计与软件实现

2.1 系统总体设计

本系统依据视频拼接逐帧拼接的特点,将整个视频拼接流程分配到上下位机5台机器上,上位机、下位机逻辑流程图如图2所示。

下位机使用VLC拉取视频流,获得视频帧后,由OpenCV进行视频帧缩放。这里的缩放是为了能够显示在4K显示器上,并且可以减少拼接视频花费的时间。若不缩放直接使用4K视频拼接会花费大量时间,无法满足实时性的要求,且会对后续的显示造成困难。故采用先对视频帧进行缩放,后进行拼接的处理过程。

在4K图片拼接测试中,针对大场景的图像进行拼接,得到的全景图会出现色差,人肉眼就能感受到拼接的痕迹,影响拼接效果。图3所示为截取的拼接

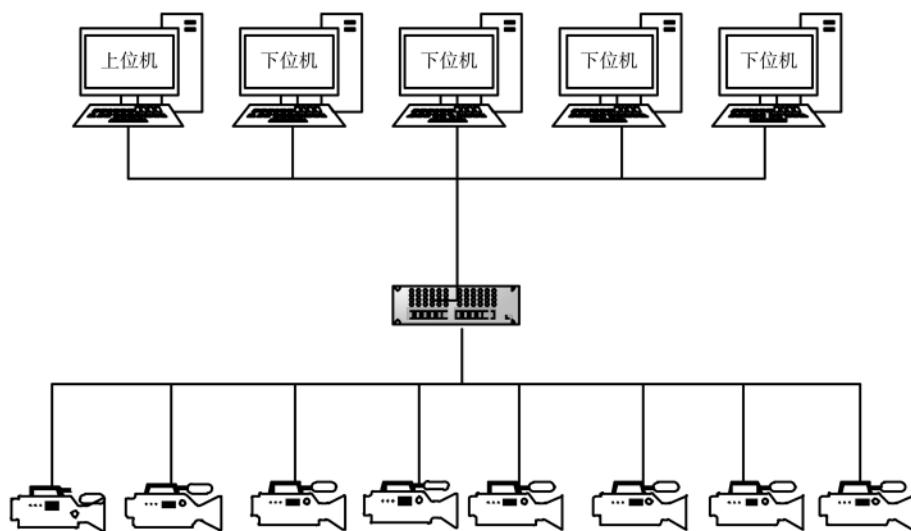


图 1 系统架构图

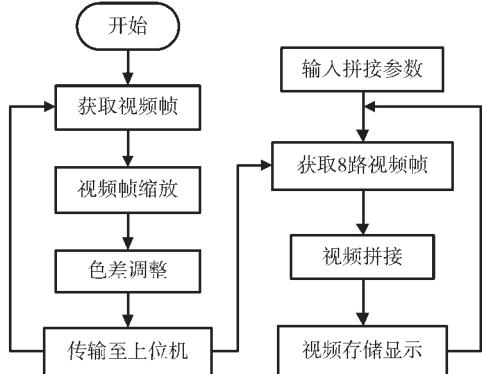


图 2 逻辑流程图



图 3 全景色差图

全景图的部分图像。

故在传输之前,下位机对视频帧进行色差校正后再对视频帧进行缩放、色差校正等一系列处理,之后传输到上位机。上位机进行拼接,最后存储并显示。

2.2 系统关键功能设计

2.2.1 视频传输模块

在传输中,全景视频的超大分辨率带来诸多困难,包括带宽不足和实时性不佳等问题。这些问题为算法设计与软件实现提出了挑战^[7]。

SMB(Server Message Block)协议是一种基于TCP/IP协议实现的网络文件共享传输协议。利用该协议可将视频文件写入上位机中。本文提出了一种交替读取写入策略。以一路视频流为例:

(1) 初始化:上位机开辟缓冲区建立文件夹V1,下位机向上位机V1文件夹写入第一个文件1.avi与第二个文件2.avi。同时置标志文件1.txt,2.txt为0。视频帧数量控制为50帧左右。

(2) 上位机循环读取:上位机读取标志文件1.txt。此时1.txt标志位为0,读取1.avi。读取完毕之后释放1.avi,置标志文件1.txt为1。上位机读取标志文件2.txt。如此反复。

(3) 下位机循环写入:下位机读取标志文件,查看标志位。若标志位为1,则写入视频帧;反之等待。如此反复。

视频传输流程如图4所示。

通过这种传输机制,实现了将下位机初步处理后的8路视频写入上位机进行视频拼接。该机制初始化时,设置初始文件大小为50帧左右,是考虑到初始化时间过长会导致拼接图像与实际图像存在时延。

2.2.2 视频拼接模块

视频拼接模块是该系统的核心,也是该系统的

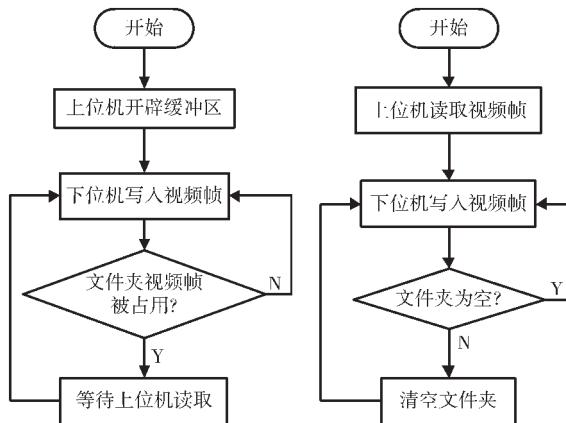


图 4 视频传输流程图

设计难点。考虑一种特殊情况：视频拼接的速度小于视频写入的速度，该情况下会导致拼接图像与实际图像出现越来越大的时延。此种情况说明该系统的瓶颈是视频拼接的时间。

如何有效降低视频拼接的耗时是该系统的一个重点。本文结合系统的使用场景和硬件环境，提出了一个基于预算算拼接参数的拼接模块。

传统图像拼接根据相邻图片之间存在的重叠区域，找出图片的特征点进行匹配，之后利用匹配好的特征点计算出单应矩阵即像素点坐标的变换矩阵。而上述流程在拼接过程中花费大量时间。针对监控摄像头有位置相对固定的特点，以机场这一场景为例：摄像机是布置在位于跑道一侧的中间位置，在该位置修建了监控高塔用以布置摄像机。摄像机在塔顶呈扇形布置，并在外侧设置保护罩防止外界因素对摄像机造成干扰。摄像机位置是固定保持不变的，利用图像配准算法一次计算出单应矩阵作为固定参数，可直接用于后续的图像拼接中，将大大减少耗时。设原图中某一像素点的坐标为 (x, y) ，变换后的坐标为 (x', y') 。单应矩阵为：

$$\mathbf{H} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \quad (1)$$

则 (x, y) 与 (x', y') 的变换关系为：

$$\begin{bmatrix} u \\ v \\ z \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (2)$$

则有：

$$x' = \frac{u}{z} = \frac{h_{11}x + h_{12}y + h_{13}}{h_{31}x + h_{32}y + h_{33}} \quad (3)$$

$$y' = \frac{v}{z} = \frac{h_{21}x + h_{22}y + h_{23}}{h_{31}x + h_{32}y + h_{33}} \quad (4)$$

若直接利用单应矩阵进行图像映射和拼接，会花费大量时间。因此本文采用通过单应矩阵得出像素点的映射关系，并存储在矩阵中，具体步骤如下：

- (1) 构造一个指针矩阵，矩阵大小为合成图后的像素大小；
- (2) 将对应像素点的映射关系写入指针矩阵中；
- (3) 合成图像时遍历指针矩阵，即可得到对应关系。

变换前的原坐标能够用单应矩阵与变换后的坐标表示，推导如下：

由式(3)、(4)可得：

$$\begin{cases} x(h_{11}-x'h_{31})+y(h_{12}-x'h_{32})=x'h_{33}-h_{13} \\ x(h_{21}-y'h_{31})+y(h_{22}-y'h_{32})=y'h_{33}-h_{23} \end{cases} \quad (5)$$

最后可得 x, y ：

$$x = \frac{x'h_{33}-h_{13}}{\begin{bmatrix} h_{11}-x'h_{31} & h_{12}-x'h_{32} \\ h_{21}-y'h_{31} & h_{22}-y'h_{32} \end{bmatrix}} \quad (6)$$

$$y = \frac{y'h_{33}-h_{23}}{\begin{bmatrix} h_{11}-x'h_{31} & x'h_{33}-h_{13} \\ h_{21}-y'h_{31} & y'h_{33}-h_{23} \end{bmatrix}} \quad (7)$$

采用该方法进行拼接，对下位机传输的 8 路视频，拼接一帧全景图只需要 14~15 ms。

2.3 系统软件界面

根据系统功能实际需要，使用 C# 语言构建系统软件。该软件客户端由四个部分组成，分别是视频拼接、视频回放、日志管理、系统设置。

图 5 展示了视频拼接界面，该界面可以完成拼接后显示全景视频，而且保留传统监控平台对多路视频进行播放显示的功能。

图 6 展示了视频回放界面，监控视频通常需要保存一个月以上，该模块能够通过时间检索视频，并且实现了倍速播放，方便快速查找具体某个时间点的视频片段。

日志管理界面用于记录所有操作记录及错误信息，配合系统设置中的用户管理，对客户端进行系统权限控制。如普通用户无法查看后台视频回放和系统设置中的权限管理功能；超级管理员用户拥有最高管理权限，能够使用系统所有功能，并能分配系统账号及权限。



图 5 视频监控界面



图 6 视频回放界面

3 系统测试

3.1 系统实时性与稳定性测试

由于机场场景特殊,无法进入机场内搭建测试环境进行测试,故自行搭建了测试环境。在楼顶架设摄像机平台,安装摄像机并保证视频间具有一定重叠区域。摄像机平台如图 7 所示。



图 7 摄像机平台

该系统要实现对 8 路 4K 视频进行拼接并输出 4K 拼接后全景视频。下位机完成了旋转图像、色差调整、图像缩放、图像写入。上位机完成了 8 路视频图像拼接、视频显示、视频存储。表 3 是系统实时性测试结果。测试了上下位机处理 1 000 帧图像的平

表 3 实时性测试结果
(ms)

测试机器	平均处理时间	
	运行 5 min 后	运行 1 h 后
上位机	35.82	36.39
下位机	38.35	37.04

均时间。测试结果表明该系统实时性较高,处理速度达到预期,速度达到 25 帧/s。

调试系统后连续不间断运行 24 h,未发现视频有明显延迟,运行速度及拼接效果无明显下降,说明系统稳定性良好。

3.2 系统拼接效果

图 8 是未拼接前的部分视频图像,由于后期调整了摄像机位置,因此视角有所缩小;图 9 展示了系统实时运行拼接效果图,该图无明显拼接痕迹与畸变,图像质量良好,达到了设计的预期效果。



图 8 未拼接视频图像



图 9 全景视频图像展示

4 结论

视频监控是安全防范的重要组成部分,监控的第一要务是用最短时间从被监控的地方获取尽可能多的信息反馈^[8]。传统监控摄像头的布置较为分散,监控人员需要目不转睛地专注于多个摄像头的内容来保证对某个区域的监控。超出一定数量的摄像头,监控人员难以保持长时间的注意力。本设计针对这一问题,设计并实现了一个基于分布式系统的全景视频拼接软件,在一定程度上减少了监控人员的工作压力,也给决策人员提供了更加全面、有

效的支持。

经多次调试验证后,该全景拼接软件系统稳定性和实时性满足设计要求,并顺利通过验收。但是对于视频拼接的质量,如对图像色差进一步调整以实现更大对比度下的流畅过渡、整个系统的可靠性、处理速度的提升以实现30帧以上的视频处理等问题,还有待进一步研究。

参考文献

- [1] 王娟,师军,吴宪祥.图像拼接技术综述[J].计算机应用研究,2008(7):1940–1943,1947.
- [2] PERAZZI F,SORKINE-HOMUNG A,ZIMMER H,et al.Panoramic video from unstructured camera arrays[J].Computer Graphics Forum,2015,34(2):57–68.
- [3] BROWN M,LOWE D G.Automatic panoramic image stitching using invariant features[J].International Journal of Computer Vision,2007(74):59–73.
- [4] LI J,WANG Z,LAI S,et al.Parallax-tolerant image

(上接第60页)

- [5] SUN X D,WU P C,HOI S C H,et al.Face detection using deep learning:an improved Faster RCNN approach[J].Neurocomputing,2018,299:42–50.
- [6] REDMON J,DIVVALA S K,GIRSHICK R,et al.You only look once: unified, real-time object detection[C].Computer Vision and Pattern Recognition,2016:779–788.
- [7] LIU W,ANGUELOV D,ERHAN D,et al.SSD:single shot multibox detector[C].European Conference on Computer Vision,2016:21–37.
- [8] 朱敏.基于SSD的行人检测方法[J].科学技术创新,2017(36):75–76.

stitching based on robust elastic warping[J].IEEE Transactions on Multimedia,2018,20(7):1672–1687.

- [5] 张浩.基于DSP的高清全景图像拼接的研究与实现[D].成都:电子科技大学,2018.
- [6] BRADSKI G,KAEHLER A.Learning OpenCV[M].[s.l.]:O'Reilly Media, Inc., 2008.
- [7] 罗莹,宋利,解蓉,等.全景媒体的系统架构研究综述[J].电信科学,2018,34(2):88–98.
- [8] 黄凯奇,陈晓棠,康运锋,等.智能视频监控技术综述[J].计算机学报,2015,38(6):1093–1118.

(收稿日期:2020-07-14)

作者简介:

方诤言(1997-),男,硕士,主要研究方向:智能信息处理。

陈贤富(1963-),男,副教授,主要研究方向:复杂系统及认知计算。

2017(36):75–76.

- [9] 吕俊奇,邱卫根,张立臣,等.多层卷积特征融合的行人检测[J].计算机工程与设计,2018,39(11):3481–3485.

(收稿日期:2020-07-09)

作者简介:

冯婷婷(1993-),女,硕士研究生,主要研究方向:深度学习、目标检测等。

葛华勇(1976-),女,硕士生导师,副教授,主要研究方向:数字图像处理、信号与信息处理、信息安全等。

孙家慧(1997-),女,硕士研究生,主要研究方向:深度学习、目标检测等。

版权声明

经作者授权，本论文版权和信息网络传播权归属于《信息技术与网络安全》杂志，凡未经本刊书面同意任何机构、组织和个人不得擅自复印、汇编、翻译和进行信息网络传播。未经本刊书面同意，禁止一切互联网论文资源平台非法上传、收录本论文。

截至目前，本论文已经授权被中国期刊全文数据库（CNKI）、万方数据知识服务平台、中文科技期刊数据库（维普网）、JST 日本科技技术振兴机构数据库等数据库全文收录。

对于违反上述禁止行为并违法使用本论文的机构、组织和个人，本刊将采取一切必要法律行动来维护正当权益。

特此声明！

《信息技术与网络安全》编辑部
中国电子信息产业集团有限公司第六研究所