

文章编号: 1002-0446(2004)03-0218-04

# JMF 视频传输技术在 Web 机器人中的应用\*

高玉华, 苏剑波

(上海交通大学自动化系, 上海 200030)

**摘要:** 本文提出了一种 JMF(Java Media Framework) 视频传输技术在基于网络遥操作机器人中的应用, 实现了图像的实时传输. 实验表明了这个解决方案的先进性和有效性, 它为基于 Internet 的机器人的远程控制提供了强有力的支撑技术.

**关键词:** 机器人; 遥操作; JMF; Java

**中图分类号:** TP24      **文献标识码:** B

## JMF based Video Transmission for Web Robot

GAO Yu-hua, SU Jian-bo

(Department of Automation, Shanghai Jiaotong University, Shanghai 200030, China)

**Abstract:** This paper presents an application of JMF-based video transmission for Web robot control that realizes the real time video transmission. Experiments show that this method is valid and effective, and it provides a strong technical support for remote control of Internet-based robots.

**Keywords:** robot; remote control; JMF; Java

### 1 引言 (Introduction)

目前, Internet 已经成为人们工作、生活和娱乐的一部分. 但由于目前网络的终端设备形态单一、功能简单, 所以提供的都是单一的文本、媒体和数据的获取和查询服务. 把机器人引入网络, 作为网络的智能终端, 实现对终端环境的操作, 可以为网络用户提供更多更新的服务. 它不仅可以提供传统的静态的信息服务, 还可以提供动态的互动服务, 如远程医疗、诊断、教学和家电控制等一系列军用、民用和工用场合. 由于 Web 访问已成为获取信息和服务的主要方式, 所以一般把机器人构建在 Internet 的一个 Web 站点上, 网络用户通过 Web 浏览器对其进行远程控制, 并得到信息反馈. 这种基于 Web 的机器人最大的特点就是开放性, 只要能连上 Internet 就能对其控制, 而且普通用户不必了解机器人复杂的操作原理也能进行控制.

由于有不确定性的大时延存在, 所以监督控制是实现 Web 机器人的最佳的控制方案<sup>[1]</sup>. 此控制方

案将操作人员置于控制结构闭环之外, 从而减小了传输时延对整个系统的影响. 操作人员只是发送目标任务给远端, 任务由远端回路自己完成, 而操作人员根据远端返回的信息再发送下一个任务命令. 由此可见, 在监督控制中, 视觉信息的获取、处理及传输是人机交互的重要环节, 它直接影响了交互性能的优劣和操作者指令的决断.

在国外已有一些公开的 Web 机器人<sup>[2]</sup>, 本地用户通过 Web 浏览器, 访问远端连有被控机器人的 Web 服务器, 通过在本地显示的操作界面发送指令. 服务器收到指令后交由机器人的控制器执行, 执行结果和最新环境信息利用摄像机采集图片返回到服务器, 并将其转化为 HTML 格式, 发送至用户端显示(图 1).

系统视觉反馈放在整个控制环中, Web 用户发出命令后要等到画面更新后才可以根据最新的图片信息进行下一步操作. 不仅时延大, 而且没有良好的人机界面, 人机交互环境既不生动也不友好. 为了解决

\* 基金项目: 国家 863 计划资助项目(2001 AA422260).

收稿日期: 2003 - 08 - 01

这个问题,也有一些系统通过提供虚拟环境的人机交互界面增加现场沉浸感和减少延迟对系统的影响,但是这样对远端系统的环境不变性要求很高,且只适合固定的机器人系统,对环境可变的移动机器人就不适合。

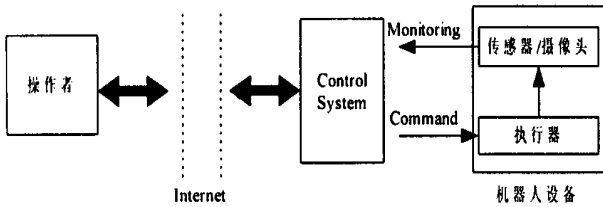


图 1 监督控制

Fig.1 Supervisory control

计算机网络通讯技术和多媒体技术的发展使得实时传输图像已成为可能,如果把视觉反馈独立出控制环(图 2),实时传输图像,操作者可以及时地得到远端的环境信息并及时地根据信息发送指令.视觉反馈与控制是并行的,比传统的反馈—控制—再反馈—再控制的串行操作相比极大地减少了系统的延时,并且提供了良好的人机界面,提高了系统的交互性能。

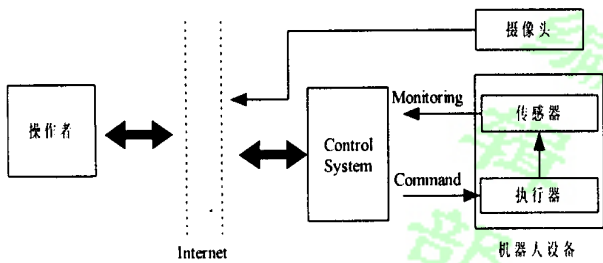


图 2 改进后的控制

Fig.2 Improved control

## 2 视频传输关键问题 (Key problems of video transmission)

与传统的单帧图片传输相比,视频的实时传输有以下关键问题需要解决:首先,由于视频图像信息的数据量非常大,所以对采集的原始图像进行压缩;其次,要在网络负载变化和时延不确定性的情况下保持图像的连续性;另外,对于 Internet 上多变的网络结构,多种类型的操作系统,如何使得视频流顺利地传送到客户端,并且被客户端成功地接收播放等都是系统的关键问题。

虽然目前图像压缩在方法上已取得许多惊人的进展,但就其成熟性而言,目前流行应用的仍然是

JPEG 或 MPEG 技术.因此,在视频采集及图像帧编码环节的设计中也主要考虑这两种编码方式.在原理上这两种编码方式基本相同.两者的不同之处在于:JPEG 针对单独的图像图片进行处理(帧内压缩),而 MPEG 则利用了图像序列间的相似性的技巧(帧间压缩).因此 MPEG 的整体压缩倍率要优于 JPEG.但是考虑到传输信道链路特点, JPEG 对于一次误码最多损失图像序列中的一幅,而 MPEG 将导致后续的多幅帧间编码序列无法复现.所以,根据系统工作的可靠性需求决定了 JPEG 成为优选方案。

为了对远程机器人现场进行监控,系统的图像传输要求实时进行,需要图像的实时传输和播放,因此要求服务器方发送持续稳定的数据流,客户方一边接受数据一边播放.为此采用实时图像传输协议 RTP(Real Time Transport Protocol)和实时传输控制协议 RTCP(Real Time Control Protocol). RTP 是一层独立的传输协议,属于网络结构的应用层,可以通过传输层的 TCP 和 UDP 提供一种与底层网络无关的数据传输. TCP 是一种可靠的有连接的通讯方式,能保证信息正确到达目的地,但网络延时比较长.此方式可用于控制命令的传送. UDP 是一种可靠性不高的无连接方式,但是它的网络延时短.因此,选择 UDP 作为图像传输的传输层协议。

在 Internet 上视频图像要经过各种路由才能送到客户端,其中有些特殊的网络会配置网络地址转换(NAT)组件共享有限的公共 IP 地址,并阻止外部用户直接访问内部计算机. NAT 是 Network Address Translation 的缩写,采用网络地址转换技术,局域网内部的非法互联网 IP 地址通过 NAT 可以转化成合法互联网 IP 地址,实现对外界网络如 Internet 的合法访问.对内, NAT 将内部客户机发出的每一个 IP 数据包地址进行检查和翻译,记录包内的请求端 IP 地址数据并重新打包成合法的外部 IP 地址发送到互联网.对外, NAT 把从互联网获得的数据包根据请求端记录把目的 IP 地址在数据包内部进行重组,使其转换为局域网客户端的 IP 地址后发送到客户端.由于有这样的路由存在,所以采用一般的端到端的视频“Push”推送技术是不能把视频信号传送到在 NAT 后的客户端的.需要采用“请求—传送”机制,即服务器端有个端口监听客户请求,当客户请求的数据包到达后得到目标端的地址和端口,并用此信息构建底层的数据包,发送视频流。

针对以上问题选择了 JMF 作为开发工具. JMF (Java Media Framework) 是由 Sun 等公司推出的一

种开发流媒体应用的应用程序接口(API),它支持多媒体数据的回放和实时传输等媒体操作.在数据回放上,JMF支持大多数的媒体格式如(MPEG,JPEG等).在实时传输上,JMF支持实时传输协议RTP(Real Time Transport Protocol)和实时传输控制协议RTCP(Real Time Control Protocol).它采用统一的结构和消息传递协议,包括JMF API和RTP API两个部分,前者的主要功能是捕捉、处理、存储和播放媒体,后者主要是在网络上传输和接收媒体流.而且JMF提供了一个RTP与传输层的接口,可以实现基于任何传输层(TCP/UDP)的网络流媒体应用.另外它是一种采用Java语言开发流式媒体应用的API,通过利用Java平台的优势,为访问底层的媒体框架提供了一个通用的跨平台Java API.虽然Java经常被批评为使用效率低下,但是由于JMF使用本地库播放媒体,所以可以克服Java潜在的效率问题.对整个系统而言效率瓶颈还是在网络上,尤其是当系统运行在WAN上<sup>[3]</sup>.

### 3 系统实现 (System Realization)

我们通过实验验证视频传输的可行性和合理性.整个系统如图3所示:一台PC机和一个摄像头通过USB口相连,构成视频服务器,负责视频数据的采集、压缩和传输.为了防止外部IP直接访问控制机器人和视频服务器,Web服务器、Pioneer 2自主移动机器人和视频服务器构成一个局域网,因此服务器也同时起到网关的作用.另外用了几台PC机器构建了一个模拟网络,其中一台机器提供NAT服务.

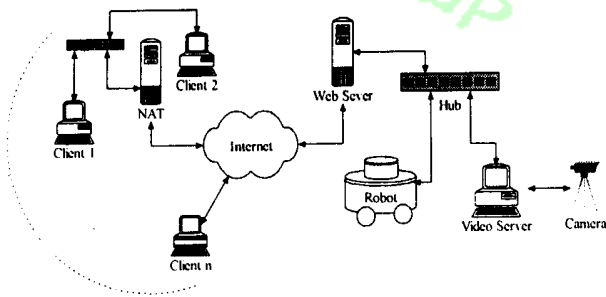


图3 系统结构图

Fig.3 Diagram of system architecture

一方面,用户在本地发送控制命令给远端的机器人,由远端的机器人系统对控制命令加以解释、执行,从而控制机器人完成相应的动作;另一方面,由远端的摄像机采集现场的图像,并实时传输到控制方.摄像头完成图像的数字化后,首先要对图像数据

进行预处理,如进行图像辨识与数据压缩,然后通过网络传送到用户站点.用户站点把接收到的已压缩的图像信息进行解压缩,还原成原来的图像,显示在Web窗口,Web操作者可以根据实时的视觉反馈信息控制远端机器人.

图4为视频系统API的总体结构.当视频服务器收到请求后,先调用设备管理器(Capture Device Manager)捕获所有可以得到的媒体设备,并进一步查找支持Video的媒体设备,得到此设备媒体对象(Capture Device Info).从设备媒体对象中得到媒体定位器(Media Locator),创建原始数据源(Data Source).考虑到网络传输,所以需要为所用信道控制器(Track Control)设置输出图像的特定RTP格式,由于网络机器人控制要求较高的图像清晰度,故选用JPEG-RTP格式.处理器(Processor)完成格式转换构件新的数据源(Data Source).最后创建会话管理器(Session Manager),构建输出流,开始输出.本系统使用Java小应用程序Applet负责实时媒体的接收和播放.在Applet初始化(init)时创建会话管理器,并注册一个接收流监听器(Receive Stream Listener).Applet的更新(Update)方法负责监听所有的接收流(Receive Stream Event)事件,当它检测到有数据流到达时,构建一个播放器(Player).在播放器中注册控制监听器(Controller Listener),当播放器构件完毕时触发播放器构建完毕事件(Realize Complete Event).Controller Listener的Controller Update()方法负责监听所有的控制事件,当它检测到组件构建完毕事件(Realize Complete Event)发生时返回播放器,即在监视器上显示视频图像.

通过JMF提供的RTPConnector接口可以实现基于任何底层网络的流媒体应用.JMF没有提供RTPConnector的缺省实现,软件中RTPConnector的实现依赖于用户所选的传输协议,在本实验中我们选择UDP协议.为了实现UDP与RTP之间的通信,需要设计一个类来实现PushSourceStream和OutputStream接口中的方法.在PushSourceStream中主要实现read()和setTransferHandler()两个方法.read()用于将底层网络中的UDP数据包(DataGram Packet)中的数据传输给Session Manager,setTransferHandler()注册一个SourceTransferHandler对象,PushSourceStream回调SourceTransferHandler对象的transferData()函数.同样在OutputStream中需要实现write()方法,它用于将Session Manager的数据输出到底层网络.为了使



UDP 包成功地穿越 NAT,在客户端 applet 发出请求前,服务器端的 DatagramSocket 必须先循环监听端口,当客户端的 applet 启动后,底层的 Datagram-

Socket 向服务器端发送请求,这时服务器得到目标端的地址和端口,并用此信息构建底层的 Datagram-Packet,发送视频流。

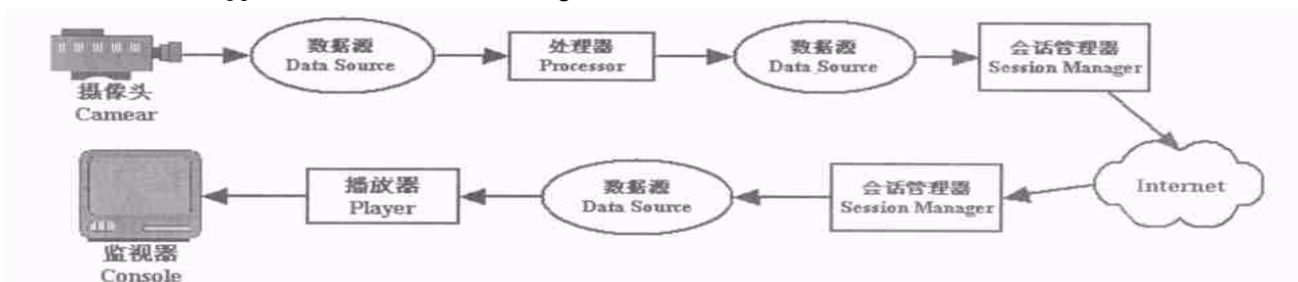


图 4 视频系统 API 的总体结构

Fig.4 API architecture of video system

### 4 实验 (Experiments)

应用本系统,实现了上海交通大学与新加坡国立大学 ECE 学生实验室之间的网络机器人视频流传输.图 5 为实际的人机交互界面。

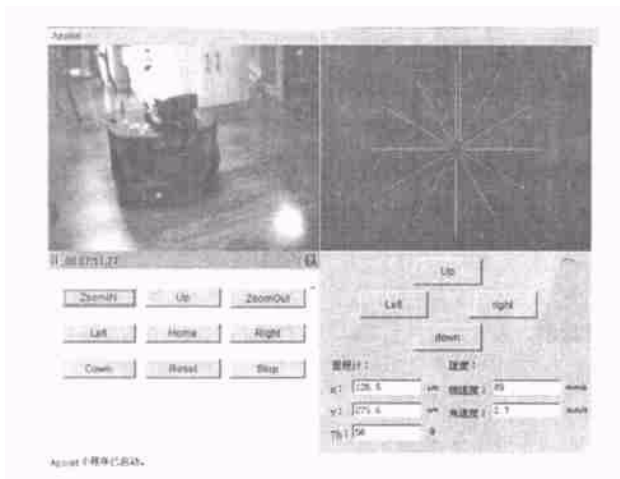


图 5 实际交互界面

Fig.5 Real interactive interface

图的左上角为实时视频传输窗口,左下角为一个两自由度的摄像头的控制窗口.它可以控制摄像头上下左右、放大缩小等动作.图的右上角是机器人实验环境的地图,左下角是 P2 机器人控制和数据反馈窗口.本视频服务器以每秒 30 帧的速度传送 320 × 240 的 JPEG 图片,实验表明系统接受视频数据流时图像连续稳定,实现了视频数据的实时通讯。

### 5 结论 (Conclusion)

本文介绍了基于 JMF 的 Web 机器人实时视频的应用的开发.根据对系统的分析,JMF 技术满足

Web 机器人实时视频传输中各方面的要求,并且它提供了大量的有用类库,使得程序开发方便.同时本系统对视频捕获设备无特殊要求,适用范围广,有良好的平台无关性,有较好的适应性和可扩展性.采用了 JMF 中的 RTPConnector 设计,它可以在任何底层网络中实现视频数据的实时传输.实验表明,系统接受视频数据流时,图像连续稳定,实现了视频数据的实时通讯.相信随着网络带宽的增加,JMF 技术的改进,系统功能的完善,本系统的性能也会有很大的改善。

### 参考文献 (References)

- [1] Taylor K, Dalton B. Internet robots: a new robotics niche[J]. IEEE Robotics & Automation Magazine, 2000, 7(1):27 - 34.
- [2] 庄严,王伟,恽为民.基于网络的机器人控制技术研究现状与发展[J].机器人,2002,24(3):276 - 282.
- [3] Markousis Th, Tsirikos D, Vazirgiannis M, et al. WWW-enable delivery of interactive multimedia documents[J]. Computer Communications, 2000, 23(3):242 - 252.
- [4] Goldberg K, Gentner S, Sutter C, et al. The Mercury project: a feasibility study for Internet robots[J]. IEEE Robotics & Automation Magazine, 2000, 7(1):35 - 40.
- [5] Stein M R. Interactive Internet artistry[J]. IEEE Robotics & Automation Magazine, 2000, 7(2):28 - 32.
- [6] Simmons R, Fernandez J, Goodwin R, et al. Lessons learned from Xavier[J]. IEEE Robotics & Automation Magazine, 2000, 7(2):33 - 39.
- [7] 王政,常一志,张建光.基于 Internet 的机器人遥控操作系统的初级构筑[J].应用科技,2000,27(8):11 - 14.
- [8] 李小海,王旭永.基于 WWW 的机器人远程控制的关键技术及典型实现[J].工业控制计算机,2000,13(2):51 - 53.

### 作者简介:

高玉华 (1779-),女,硕士生.研究领域:网络机器人.  
苏剑波 (1969-),男,博士生导师,教授.研究领域:机器人视觉,多机器人协调。