

基于 DirectShow 实现局域网实时视频传输

钱小红, 李小艳

(武汉理工大学 华夏学院 信息工程系, 湖北 武汉 430223)

摘 要:随着计算机网络和多媒体技术的发展,音频、视频等多媒体信息已在网络中越来越普遍,它们深深地影响了人们的生活。MPEG-4 是时下一种流行的视频编解码模式,怎样在 Internet 上实时地传输 MPEG-4 码流已经成为当前的一个热点;同样是未来网络方向的一个研究重点。文中在目前的视频传输系统的基础上,利用新的协议标准和传输方式来实现视频的实时传输,使传输的视频实时,没有马赛克。论文研究了 DirectShow 技术,分析了其体系结构的特点与组成,研究了利用 DirectShow 实现视频采集、压缩编码、网络传输以及解码播放技术,并给出了利用 RTP/RTCP 以及 DirectShow 实现 MPEG-4 视频数据的实时直播的方法和部分核心代码,有效地保证了高质量的实时视频传输。同时在应用程序内开发源 Filter,简化了开发流程和程序的调试。测试证明系统有良好的网络适应性,视频经过传输后比较流畅,几乎没有马赛克。

关键词:视频;DirectShow;实时;MPEG-4;压缩

中图分类号:TP393.09

文献标识码:A

文章编号:1673-629X(2011)10-0218-04

Realization of LAN Real-Time Video Transmission Based on DirectShow

QIAN Xiao-hong, LI Xiao-yan

(Info. Eng. Dept. of Huaxia College of Wuhan Univ. of Techn., Wuhan 430223, China)

Abstract: With the development of computer network and multimedia technology, multimedia such as sound and video stream is becoming more and more popular in the Internet, which changed people's life greatly. MPEG-4 is used widely nowadays. How to transport real-time MPEG-4 data on the network is a hotspot, and also is very important in the future. Based on the above background, used new protocols and transmission to realize real-time video transmission. Researched DirectShow technology, analyzed the traits and organization of its architecture and researched realizing video capture, encoding, real-time transmission, decoding and playback using DirectShow. The methods and some core code of real-time video transmission of MPEG-4 using RTP/RTCP and Direct Show were given, which guaranteed the high quality of real-time video transmission. At the same time, some source filters were developed in the program, which simplified the development flow and the programming debugging. The testing results indicate that the system achieves good picture quality and fluency for MPEG-4 video under difficult network conditions.

Key words: video; DirectShow; real-time; MPEG-4; compression

0 引言

随着互联网和多媒体技术的发展,在局域网内部实时传输视频已经得到广泛应用。但由于视频数据的庞大性和时间特性,视频传输在网络上还不是非常理想,经常会出现马赛克或不能实时传输的现象。视频的传输主要研究对视频数据的抓捕、编码、传输、解码、存储和回放等操作,怎样有效地对数据进行编码以利于传输,成为网络视频应用的研究重点。在设计视频

处理程序的过程中,怎样有效地屏蔽掉硬件设备和数据格式差异呢?有效的方法是利用微软的 DirectShow 技术、MPEG-4 视频压缩标准、RTP/RTCP 协议以及组播技术,这样设计的程序具有复用性,能让程序的开发变得简单高效,并可以提高系统的通用性和可扩展性。这样的机制,能够有效地解决由于网络的不稳定因素导致的视频图像重影、抖动、花屏等问题。

1 DirectShow 技术基础

DirectX 是微软公司提供的一套基于 Windows 操作系统平台开发的高性能声音、图像、信息输入输出的编程接口,它包括 DirectDraw, DirectSound, Direct3D。DirectShow^[1]是一套基于 DirectX 之上的开发包,是一

收稿日期:2011-03-04;修回日期:2011-06-17

基金项目:教育部高等学校创新工程重大项目培育基金(705038);武汉市科技攻关计划项目(20051101013);武汉市发展计划委员会光电子信息专项项目

作者简介:钱小红(1980-),女,助教,研究方向为软件工程。



图 1 过滤器连接图

个开放性的框架,也是一套基于 COM 的多媒体编程接口^[2]。为了处理不同的硬件设备、数据源以及数据格式等等问题,DirectShow 是通过一种系统内置的或程序员自行开发的过滤器(也叫 Filter)来处理的。DirectShow 技术使 Windows 系统下的多媒体编程变得简单^[3]。

DirectShow 规定了该怎样利用过滤器(Filter)来处理多媒体数据,过滤器是(Filter)DirectShow 的基本组成部分,是过滤器图中最小的功能模块,过滤器按照基本功能大致分为三类^[4]:源过滤器 SourceFilter、变换过滤器 TransformFilter 和提交过滤器 RenderFilter。为了完成应用程序要实现的功能,需要将各个过滤器按照其要实现的功能串联起来,形成一个流水线。每个过滤器通过输入针(In Pin)、输出针(Out Pin)连接,这些所有通过针(Pin)连接起来的过滤器就形成了过滤器图表(Filter Graph),如图 1 所示。要控制各个过滤器的连接和信息流在过滤器图标中的流动,DirectShow 提供了更高层次的 Filter Graph Manager(过滤器图表管理器),因此不需要对过滤器图表中的单个过滤器进行单独的处理。DirectShow 能在过滤器图表(Filter Graph)工作的时候接收到系统传过来的各种事件,然后通过 message(消息)的方式发送给所在的应用程序,这就实现了应用程序和 DirectShow 之间的交互。

2 视频直播系统分析与设计

本系统利用 DirectShow 作为视频处理通用平台,它支持宽松的格式变化,提供高品质的流媒体回放功能。MPEG-4 视频压缩标准不仅可提高压缩率,同时也可实现多媒体内容互动性以及拥有较高的压缩率和很强的容错能力,具有灵活的编码和解码特点。MPEG-4 代表了基于模型/对象的第二代压缩编码技术,充分利用了人眼视觉特性,抓住了图像信息传输的本质,适合于作为传输的图像压缩标准。论文采用 MPEG-4 的 XVID 编解码器(XVID Codec)对视频进行编解码,XVID 继承并发展了曾经的 DIVX^[5],并且使其性能、效率得到了具大的提升,它被认为是目前最快的 MPEG-4 编码,符合多媒体流实时传输的应用要求。系统发送端先用 USB 摄像头采集视频,再用 XVID 对采集的数据进行压缩编码,然后依据实时传输协议(RTP)对编码后的视频数据进行封装,最后一步用组播的方式传输出去。系统结合 RTP 还采用了实时传输控制协议(RTCP)和用户数据报协议(UDP)来共同

完成视频的传输,因为 RTP 自身没有独立地传输数据、提供可靠传输机制、流量控制及拥塞控制的能力。

2.1 系统总体结构设计

为了达到视频传输的实时性,总的思想是最少地发送冗余信息,最大程度上发送最新的视频。局域网实时视频传输采用服务器/客户机模式,利用 VC++ 实现。服务器(Server)端的主要功能有:获取视频、实时预览视频、编码视频、存储视频、发送视频;客户(Client)端的主要功能有:接收视频、解压视频、存储视频和播放视频。系统功能流程结构见图 2。

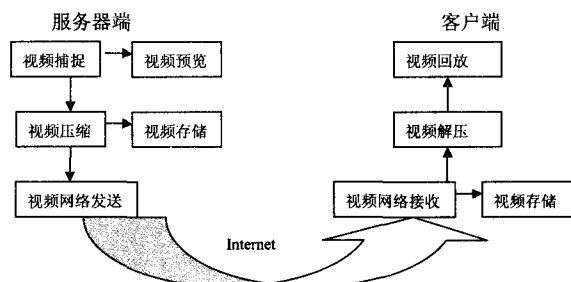


图 2 系统功能模块结构图

视频通过摄像头捕捉后实时地进行预览,然后通过 MPEG-4 编码过滤器(XVID)将摄像头采集到的视频数据进行压缩编码,形成满足某个标准的、适应于在网络上传输的视频流,再把经过编码后的视频数据备份到硬盘上,然后将编码过的视频数据打包,最后在 UDP/RTP/RTCP 协议下通过组播(multicast)的方式发送出去;视频接收模块需要指定组播(multicast)地址,它接收来自服务器端的视频数据,并把视频数据存储在硬盘上,再通过解码过滤器(XVID)将接收到的视频流进行解压并播放。

2.2 视频直播系统对应的 DirectShow Filter 过滤器图表

系统的软件结构图如图 3 所示。

DirectShow 作为一种 COM 组件^[6],内置了许多有用的 Filters,但它们主要是针对怎样处理本地的多媒体数据流的,对多媒体流的网络传输缺乏有效的支持,文中的难点是设计开发视频发送过滤器 Sender Filter 和接收过滤器 Receiver Filter。一般来说,通过 DirectShow 开发网络应用系统,必须用到 Winsock 接口编程、multicast(组播编程)、实时传输协议 RTP/实时传输控制协议 RTCP 等来开发 DirectShow 中没有内置的而程序又需要的 Filter 以实现应用需求。发送端和接收端的 Filter Graph 分别如图 4 和图 5 所示。该系统的发送端实现思路如下:用 USB 摄像头采集数据,同

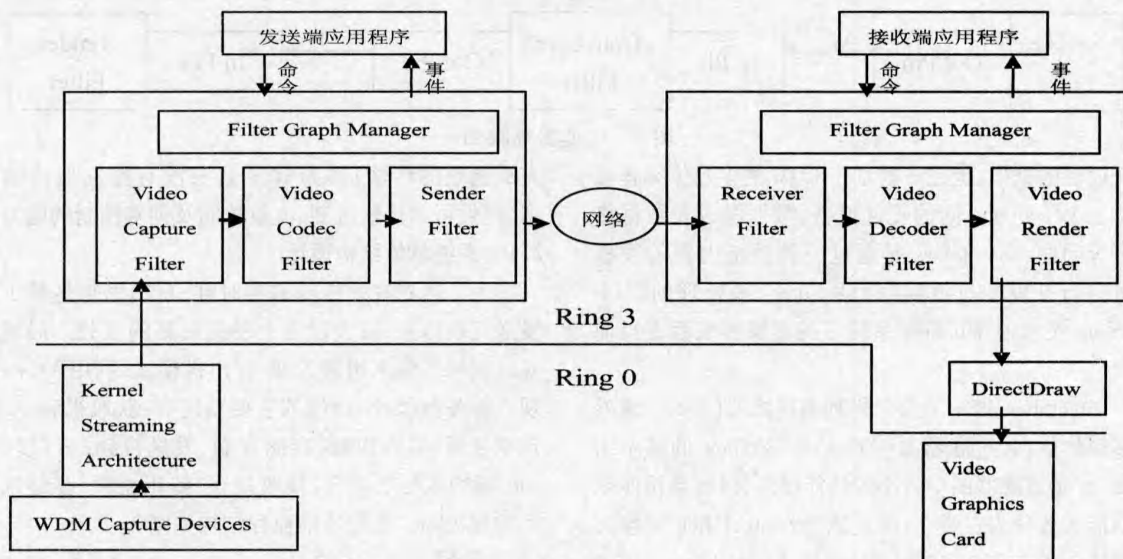


图 3 系统软件总体结构

时实现视频的预览和存储,然后用 XviD MPEG-4 Codec 对采集到的数据进行 MPEG-4 的编码,然后连到一个发送 Filter 把编码后的数据发送出去。客户端接收到的是 MPEG-4 视频流,要想播放接收到的视频,必须先通过视频解码过滤器(XviD MPEG-4 Video Decoder)进行解码,然后调用视频渲染过滤器渲染。

2.3 系统开发环境和开发工具

本系统开发硬件要求较低,只要一台微机和一个摄像头即可,全部设计任务集中在如何实现软件上,因此本系统可以说是个纯软件系统。此系统是在 Windows XP 平台上开发的,编程环境使用的是 Microsoft DirectX 9.0 SDK^[7] 和 Microsoft Visual C++ 6.0, Microsoft DirectX 9.0 SDK 为系统提供运行的时候需要的库程序。

3 基于 DirectShow 的实时视频传输实现

系统的客户端和服务端都是在 PC 机上实现的,在程序中,需要将 DirectX 9.0 SDK 的 LIB 库文件

和头文件加入工程。下面介绍一下系统实现的关键步骤。要完成视频的实时捕捉,首先建立视频捕捉程序,必须首先获取并初始化 ICaptureGraphBuilder2 接口,然后在此基础上调用 ICaptureGraphBuilder2::RenderStream,将各个针(Pin)连接起来组成 Filter Graph;视频压缩编码部分,Microsoft 为 MPEG-4 标准提供了 ISO Microsoft MPEG-4 Video Codec 编码器,只要将它们安装在开发 PC 机上,编程时 DirectShow 会自动把两个新的过滤器添加进来,程序中可以像调用别的过滤器一样调用它们,即可将它们加入到过滤器图形管理器中,实现 MPEG-4 视频编解码;视频解码部分通过 XviD MPEG-4 Video Decoder(MPEG-4 解码过滤器)来实现,XviD MPEG-4 Video Decoder 包括一个输入 Pin 和一个输出 Pin,存放在 DirectShow Filters 目录下,在接收 Filter 后面连上该 Filter 的输入 Pin 即可再将其输出 Pin 连上 Video Renderer Filter 即可把接收到的视频数据播放出来。

传输过滤器的实现:

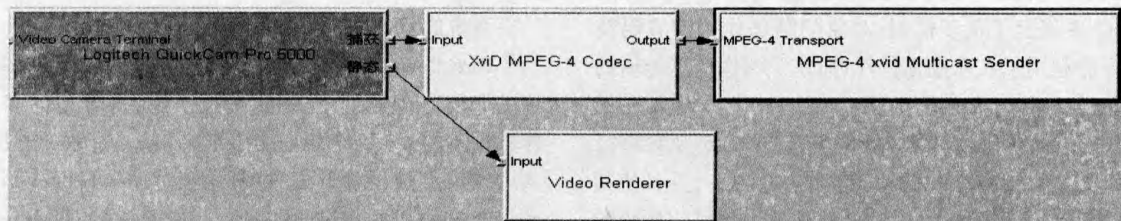


图 4 服务器端 Filter Graph

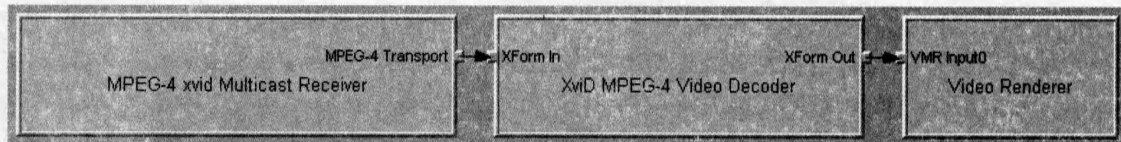


图 5 客户端 Filter Graph

文中的难点在于怎样实现传输过滤器,包括实现网络发送过滤器(文中命名为 MPEG-4 xvid Multicast Sender)和网络接收过滤器(文中命名为 MPEG-4 xvid Multicast Receiver^[8]),实时传输协议 RTP/RTCP 协议功能封装在这两个过滤器中,从编码过滤器输出的数据会流到下一个过滤器(MPEG-4 xvid Multicast Sender),它把传来的视频流封装成 RTP 包,包中的源数据是 MPEG-4 视频数据,RTP 包头包括数据包的时间戳以及发送顺序号等。将 RTP 封装好后,再按照尽快传输的原则根据视频帧发送时间间隔要求将包发送到网络中去。MPEG-4 xvid Multicast Receiver 负责接收来自网络的视频数据包,将视频数据包解开封装可以获得包的序列号,对包进行分析排序并且存入接收端缓冲区中。传输过滤器以类的形式来定义,Filter 作为一个基于 COM 的组件,它的 Com 特性实现都是从基类中派生出来的,在实现 Filter 的 Com 接口时需要选择合适的基类进行派生^[9],发送 Filter 负责将上一级的 Filter 传过来的数据发送到网络上,实现时可继承自 CBaseFilter 和 ISpecifyPropertyPages,接收 Filter 作为源 Filter,可选择继承自 ISpecifyPropertyPages 和 CSource,两个传输过滤器的实现过程如下^[10]:

第一步:声明新的传输过滤器类;

第二步:一个新的 Filter 必须有入口函数,这里通过注册函数和注销函数完成,需要在新类中定义注册函数 DllRegisterServer 和注销函数 DllUnregisterServer;

第三步:在新类中需要定义实现过滤器实例的函数和过滤器引脚的实例函数;

第四步:为了提供向外的接口,必须实现 NonDelegatingQueryInterface 函数;

第五步:为了获取过滤器上各个 Pin 的对象指针,必须实现纯虚函数 GetPin;

第六步:要想获得过滤器上各个针(Pin)的数量,需要实现纯虚函数 GetPinCount;

第七步:对 Sample 数据进行处理。

Filter 完成数据的处理、发送和接收都是在端口(pin)中实现的,DirectShow 也为 pin 端口提供了基类进行派生,在实现传输过滤器时还要定义输入针类 CVideoInputPin,它继承自 CBaseInputPin(是一个用于实现输入 Pin 的抽象类),其实现发送过滤器(Filter)上的输入针(Pin),发送过滤器通过这个 Pin 接收来自压缩编码过滤器传过来的数据,然后以某种规则发送。

CBaseInputPin 支持 IMemInputPin 接口(支持 Push Model),CJRTPSenderFilter 主要实现下面的功能:

(1) BeginFlush 和 EndFlush 两个函数的实现;

(2) 为了获得 Pin 上推荐使用的媒体类型,必须重写方法 GetMediaType,文中首选的媒体类型为 MEDI-

ATYPE_Video,子类型为 CLSID_XVID,XviD 对应的是视频数据的格式;

(3) 为判断输入 Pin 连接的媒体数据类型,需重写方法 CheckMediaType;

(4) 为接收上一级 Filter 传入的 Media Sample,需重写方法 Receive()。其中有一个指向 IMediaSample 的指针参数;

(5) 定义 InitJRtpParam() 函数以便可以使用 JRTPLIB^[11] 库来对网络进行初始化。

接收过滤器是第一层过滤器,只有输出针,它的输出针(OutputPin)端口从 CSourceStream 基类派生而来。要实现端口,需按照 DirectShow 关于 Filter 连接规则,Inputpin 和 Outputpin 都要实现向外提供媒体格式检查功能,以便进行媒体协商,即实现接收 Filter 的 GetMediaType,调用这个函数可以获得引脚连接时推荐给用户使用的媒体类型。接收过滤器中推荐首先选用的媒体类型是 MEDIATYPE_Video, CheckMediaType, 这个函数检查 Pin 连接时所用的媒体类型;DecideBufferSize,这个函数可以设置缓存(存放 Sample 数据)的大小,另外还要定义将接收的视频数据写入包的函数 FillBuffer、处理媒体数据的函数 DoBufferProcessingLoop 以及 InitJRtpParam() 函数,它可以使用 JRTPLIB 库来对网络进行初始化。

完成了传输过滤器的实现以后,必须把相应的过滤器.ax 文件复制到 C 盘的 system32 目录下^[12],再对它们进行注册,运行 regsvr32.exe 文件就可以注册传输过滤器,经过注册后的传输过滤器就如同 DirectShow 内置的 Filter,用户可以像用 DirectShow 自带的过滤器一样直接使用它们。

4 结束语

本系统结合使用 VC++、MPEG4、RTP/RTCP 和 DirectShow 技术,使视频直播系统实现变得简单而高效。在这个系统的设计和实现过程中,作者用 RTP 和 RTCP 协议对 MPEG-4 视频数据进行了组包算法设计,因为 DirectShow 没有内置多媒体数据传输过滤器,所以为了实现系统作者本人自行开发了基于 DirectShow 和 RTP 协议的发送 Filter 和接收 Filter。此系统充分利用了 DirectShow 处理多媒体流的功能,开发完的程序结构简单、清晰。经过实验测试表明通过该系统的视频马赛克小,连续性好,系统具有比较好的实时性、比较强的容错性和扩展性,代码重用率高。

参考文献:

- [1] 陆其明. DirectShow 实务精选[M]. 北京:科学出版社, (下转第 225 页)

的,并模拟出群体运动的心理状态,能够真实地再现群体逃逸等行为。

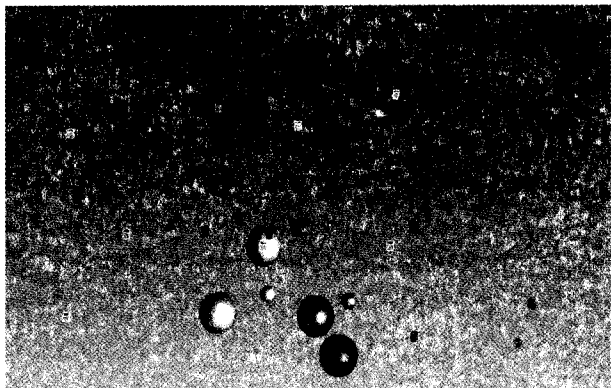


图 2 蜂群的初始状态

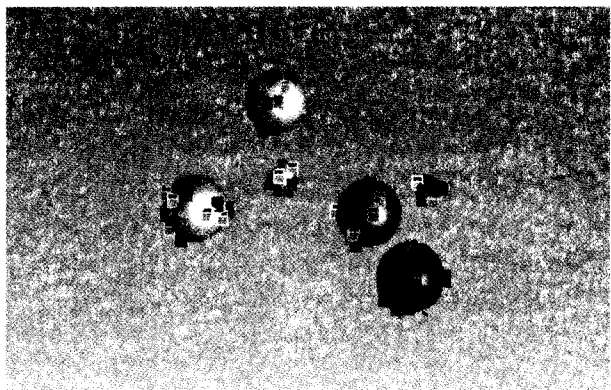


图 3 蜂群分散后状态

4 结束语

人工蜂群算法是集群智能思想的一个具体应用,文中将人工蜂群算法应用到三维群体动画中。本算法具有较好的收敛性,同时使用本算法可以模拟出个体在运动过程中的心理状态,从而使得制作出的群体动画更加的真实,但人工蜂群算法应用在群体动画的领

域还有待进一步研究。

参考文献:

- [1] Reynolds C. Flocks, Herds, and Schools: A Distributed Behavioral Model [C]//Proceedings of SIGGRAPH' 87. [s. l.]: IEEE Press, 1987: 25-34.
 - [2] Tu Xing. Artificial Animals for Computer Animation: Biomechanics, Locomotion, Perception, and Behavior [M]. [s. l.]: Springer-Verlag, 1999.
 - [3] 董刚,赵龙,田尊华. 基于智能体的群体动画创作方法的研究与实现[J]. 计算机仿真, 2005, 22(12): 143-146.
 - [4] Danniel K, Ilenkovic V J M. Multiple translational containment part I: An approximate algorithm [J]. Algorithmica, 1997, 19: 148-182.
 - [5] 冯琳,生绪博,孙宇哲. 基于遗传算法的移动平台动画图片存储的研究[J]. 计算机技术与发展, 2008, 18(7): 31-34.
 - [6] Karaboga D. An Idea Based On Honey Bee Swarm for Numerical Optimization [R]. [s. l.]: Erciyes University, 2005.
 - [7] 暴励,曾建潮. 自适应搜索空间的混沌蜂群算法[J]. 计算机应用研究, 2010, 27(4): 1330-1334.
 - [8] 胡中华,赵敏,撒鹏飞. 基于人工蜂群算法的 JSP 的仿真与研究[J]. 机械科学与技术, 2009, 28(7): 851-856.
 - [9] 丁海军,冯庆娴. 基于 boltzmann 选择策略的人工蜂群算法[J]. 计算机工程与应用, 2009, 45(31): 53-55.
 - [10] Wedde H F, Farooq M, Zhang Yue. BeeHive: An efficient fault-tolerant routing algorithm inspired by honey bee behavior [C] // Lecture Notes in Computer Science. Berlin / Heidelberg: Springer, 2004: 83-94.
 - [11] 樊小毛,马良. 0-1 背包问题的蜂群优化算法[J]. 数学的实践与认识, 2010, 40(6): 155-160.
 - [12] 李慧贤,程春田. 网络环境下的高效动态任务调度算法[J]. 华南理工大学学报(自然科学版), 2006, 34(1): 82-95.
-
- (上接第 221 页)
- 2004.
 - [2] Oberg R J. 深入学习: COM + 高级编程[M]. 刘谦,译. 北京:电子工业出版社, 2000.
 - [3] 张铮,张元. 基于 DirectShow 的无线移动视频监控系统[J]. 微计算机信息, 2006, 22(11): 44-47.
 - [4] 许长桥,黄明和. DirectShow 中构建过滤器图表技术[J]. 计算机应用研究, 2002(2): 78-80.
 - [5] 陆其明. DirectShow 开发指南[M]. 北京:清华大学出版社, 2003.
 - [6] 潘爱民. COM 原理与应用[M]. 北京:清华大学出版社, 1999.
 - [7] Microsoft DirectX C++ SDK Document [EB/OL]. 2003. [http://msdn.microsoft.com/en-us/library/ee418759\(v=VS.85\).aspx](http://msdn.microsoft.com/en-us/library/ee418759(v=VS.85).aspx).
 - [8] Steger C. An Unbiased Detector of Curvilinear Structures[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(2): 113-116.
 - [9] 程泽凯,陆小艺. 文本分类中的特征选择方法[J]. 安徽工业大学学报, 2003(7): 220-224.
 - [10] Maintz J B A, van den Elsen P A, Viergever M A. Evaluation of Ridge Seeking Operators for Multimodality Medical Image Matching [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 1996, 18(4): 353-356.
 - [11] Liesenborgs J. JRTPLIB 3. 1. 0 [EB/OL]. 2006. <http://wenku.baidu.com/view/e02f4064783e0912a2162a02.html>.
 - [12] 雷刚勇,解梅. 基于 DirectShow 的 MPEG-4 远程视频监控系统[J]. 电视技术, 2005(5): 154-158.