2010 年 2 月

碩士學位 論文

The Implementation of Video Conferencing on P2P Network Environment

朝鮮大學校 大學院

電子工學科

張明晨

The Implementation of Video Conferencing on P2P Network Environment

2010 年　2 月　25 日

朝鮮大學校 大學院

電子工學科

張明晨

The Implementation of Video Conferencing on P2P Network Environment

指導敎授　李 康 鉉

이 論文을 工學碩士學位申請 論文으로 提出함.

2010 年 2 月

朝鮮大學校 大學院

電子工學科

張明晨

張明晨의 碩士學位論文을 認准함

委員長　朝鮮大學校 教授　　金正和　(印)

委　員　朝鮮大學校 教授　　金鍾彬　(印)

委　員　朝鮮大學校 教授　　李康鉉　(印)

2009 年 11 月

朝鮮大學校 大學院

ABSTRACT

The Implementation of Video Conferencing on P2P Network Environment

Ming Chen ZHANG

Advisor: Prof. Kang Hyeon RHEE, Ph.D.

Department of Electronics Engineering,

Graduate School of Chosun University

As the development of network technology, multimedia technology and communication technology, text and voice communication cannot meet People's requirement anymore, so the development and application of video conferencing has become one of the hot network applications.

Video conference system is a communication system which uses the network to transmit sequential pictures, audio and data to many participants. Through this system, people in different places can communicate with each other freely and effectively.

At present，most video conference systems use a video server to receive and transfer video and audio data. Because of the resource limitation of video server in CPU, memory and bandwidth, the quality of video and voice

in the system gets worse rapidly as soon as the number of clients increases. So the video sever becomes the bottleneck of video conference's development. Besides, the quality and definition of video and audio are greatly affected by the compression and encoding of data. Integrating software and hardware can solve this problem, but the cost is really high. So it is of great importance to design new video conference software to suit common consumers.

Recently, the developing P2P technology can distribute the function of server to different clients, and if we can apply it to the video conferencing, it will make a great breakthrough. This paper researches modern video conferencing technologies and corresponding criteria, detail analyses network flow media and Microsoft DirectShow, then designs a framework of video conferencing system based on DirectShow. According to the design policy and the demands of project, first this paper formulates the requirement analysis, preliminary design, and then introduces how to build this system with DirectShow developed by Visual C++6.O. Finally, it formulates the details of technologies, research methods and realizing steps to build core function models. This paper aims to provide a solution for middle and small scale video conferencing system, and realizes it.

Keywords: Video conferencing, P2P, DirectShow

Table of Contents

List of Tables

List of Figures

I. Introduction

A. Research Background and Overview

With the development of network and digital communication technology in the direction of integration, digitization and intelligentization, human being lives are increasingly dependent on Internet. People have no longer satisfied with a simple voice and text communications, hope to have set of voice, text and images in multimedia communications. Therefore, in wide-area Internet platform, voice, date, images and other multimedia services have become an integrated network of development objectives and treads. The use of Internet for text, voice, video and other multimedia information, particularly the use of Internet online video conferencing has become more and more common way. Video conferencing has became a new interactive tool in this context.

1. Definition of video conference

Video conferencing system also is known as TV conferencing system, is a communication system which can send sound, images, text and other information from one place to another place. Through this system, it enables the participators who are in two or more different places to hear the sound and see the video. Participators also can communicate through the image, which means to express their views, observe each other's emotion and the relevant information or show the objects, drawings, etc. in order to enhance sense of reality. Apart from this, people can send some relevant documents, charts and so on to shorten the distance and improve the atmosphere.

2. Constitutes of video conference system

A typical multimedia video conferencing system consists of terminal equipment, communication links, multi-point control unit and the corresponding software components.

Terminal equipment: Terminal equipment is not only to complete their respective data-processing tasks, but also to complete the processing of multimedia communication protocols, receive store and playback the video and audio signal. Also record and index a large number of conference-related date and documents. The hardware configuration of terminal equipment includes audio and video signal processors, compression and decompression cards, as well as cameras, microphones, speakers and network card, etc.

Communication links: There are many choices of communication links, PSTN, LAN, WAN, N-ISDN, FrameRel or B-ISDN, ATM and so on.

Multi-point control unit: MCU is nucleus equipment of video conferencing system, it is a digital processing unit, usually located in the network nodes, office for dealing with multiple locations to communicate at the same time, its main function is sent signal by each terminal separation of extracted audio, video, data and signaling signals, sent to the corresponding processing unit, or switch audio mixing, data broadcasting and determine the routing, timing and processing of parliamentary control.

Software: Software components include protocol processing, conference services, audio and video signal processing, teamwork, management, and graphical user interfaces.

3. Application of video conference system

With the rapid development of communications networks, video conferencing systems are becoming more widespread: Provide remote-face talks, the most direct benefit is to save time and cost of various meetings and to increase its efficiency; improve and enhance business communication, customer service, product development and product display; support for distance learning, technical training, telemedicine and consultation; provide market research, information retrieval, research co-operation, engineering design, staff recruitment of new means.

B. DirectShow Technology

Audio and Video replay is the main functions of remote monitoring, video conferencing, video-on-demand and other multimedia systems. In multimedia applications, data transmission requires precise time control, in order to ensure the normal playback of multimedia content. However, the network transmission's uncertainty, making audio and video are not synchronized. And, because the frame compressed video data is usually much larger than the audio data cannot ensure that they are able to reach the receiver at the same time. In addition, the receiving end audio and video decoding rate does not match the decoded audio and video playback delay is also inconsistent. These issues will affect the audio and video media streams repetition effect.

Synchronization refers to the repetition of two or more multi-media event at a certain time of the order of the relationship between players, Synchronization refers to the repetition of two or more multi-media event at a certain time of the order to play. But also one process for the coordination of multiple media events broadcast in time domain, it exists in a series of process of multimedia information acquisition, storage, transmission and demonstration. Synchronous operation can co-ordinate and control two or more media events during playback in parallel developed by the intrinsic nature of the progress and contacts. In a distributed multimedia communications networks, multimedia synchronization includes two types: inter-stream synchronization and stream synchronization. Inter-stream synchronization is that a number of related media stream's tense relations which exist between the basic media unit; stream synchronization is the relationship that exists between the units in a single media stream. The substance of synchronization is to achieve a variety of media in the transmission remains after the original time and space constraints of the relationship.

DirectX is a Microsoft Windows-based platform for the programming interface (API); it does excellent high-speed real-time animation rendering, interactive music and environment sound, efficient multimedia data processing which general API is difficult to accomplish.

DirectShow is a member of the DirectX clan. DirectX has a lot of family members, but each have their own skills, as DirectDraw and Direct3D is responsible for two-dimensional graphic image and 3D animation speedup, DirectMusic and DirectSound is responsible for interactive music and environment sound, DirectShow deal with various broadcast media file formats, audio and video capture such as high-performance requirements of multimedia applications, providing a complete solution for the Windows platform. Using DirectShow technology can achieve the synchronization of audio and video data.

C. Streaming Media Technology of Video Conference

Streaming media refers to a continuous time-base media by using streaming technology in the Internet/Intranet, such as: video, audio or multimedia files, it does not download the entire file before playing, only put the first part of the content into memory, other data streams transmitted at any time to keep playing, but there are some delays at the beginning, the key technology is streaming.

1. Streaming Transmission

Audio and video transmitted over the network multimedia information such as the current main download and streaming two kinds of programs [1] [2]. Audio and video files are large, so need to have greater storage capacity; at the same time due to network bandwidth limitations, downloads often take minutes or even hours, so this approach is also a great delay. For audio, video or animation and other time-based media, when streaming from video/audio server to a user's computer continuous as real-time transmission, the user only takes a few seconds or ten seconds to watch after the start delay. When the video or audio and other time-based media playback on the client machine, remainder of the file will be in the background and continue to download from the server. Streaming not only significantly reduce startup delay but also never require too much cache, thus avoiding a user must wait for all of the entire files to download from the internet to view shortcomings.

Streaming has a very broad definition, and now primarily refers to the generic terms of the media transmission technology by network, and the specific meaning is to transmit the video/audio data to the destination PC by Internet. There are two methods to achieve streaming transmission:

real-time streaming and progressive streaming. In general, if the video is for real-time broadcast, or streaming media server, or the real-time protocol RTSP, namely, it is real-time streaming. If it use HTTP server, the file is sent through the progressive streaming.

a. Progressive Streaming Transmission

Progressive Streaming is downloaded in the order, user can watch re-line media while the files are downloading. People only can watch the downloaded part, cannot jump ahead to the part that are not download. Progressive Streaming is different with Real-time Streaming which can adjust speed by the user's connection when transmission. As a standard HTTP server can send file of this form do not need other special protocol, it is often called HTTP streaming. Progressive Streaming is more suitable for high-quality short segment, such as the title sequence, trailers and advertising, because the file's part is non-destructive downloaded, this approach can ensure the final quality of video playback. This means that they must experience delays before watching, especially on slower connections.

Progressive streaming files are placed in the standard HTTP or FTP server, easy to manage, basically nothing to do with the firewall. Progressive streaming is not suitable for long-fragment and a video random access requirement, such as: lectures, speeches and presentations. It does not support live broadcast and strictly speaking, it is an on-demand technology.

b. Real-Time Streaming Transmission

Real-time streaming refers to the guarantee that the media signal bandwidth and network connection with horses, so the media can be real-time watch. Real-time streaming and HTTP streaming are different, it requires a dedicated streaming media server and the transport protocol. Real-time streaming is always real-time transmission, especially for live events, and also supports random access; the user can fast-forward or rewind to view the content of the front or the back. In theory, real-time streaming cannot stop once playback, but in fact, may occur cycle pause.

Real-time streaming should be supported by adequate network bandwidth, when the network congestion or the problems arising, due to missing information of an error is ignored, the video quality will decline. For ensuring video quality, progressive streaming is better. Real-time streaming requires a specific server, such as QuickTime streaming server, real server and the windows media server. So the system has more complex settings and management than the standard HTTP servers. Real-time streaming is also in need of special network protocols, such as: RTSP (Real Time Streaming Protocol) or MMS (Microsoft MediaServer).

2. Real-Time Transport Protocol of Streaming Media

At present, network is moving in the direction of multi-media business, especially Internet. Many interactive multimedia applications, include video-on-demand, video conferencing and distance education, etc. which are quickly started in the TCP/IP networks. In order to meet the requirements of these applications, multimedia streaming real-time data transmission technology has been tremendous growth, with the corresponding agreement also has been improved and enhanced. To deliver high quality multimedia information in Internet depends on the following three factors: First, the network speed increase; second compression technology development; third is new protocol and technologies for real-time transmission. At present these aspects have significant improvement.

There are several kinds of Real-time transmission of multimedia data streaming network protocols: Real-time Transport Protocol (RTP), Real-time Transport Control Protocol (RTCP), Resource Reservation Protocol (RSVP), and Real-time Streaming of protocol (RTSP), IPv6 protocol. They cooperate with each other which support the new streaming media transmission protocol, shown in Fig. 1-1.

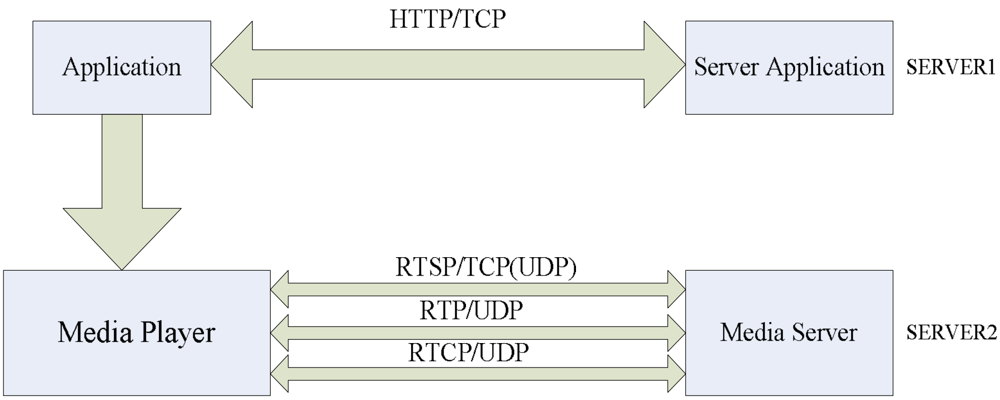

Figure 1- 1 New protocol relation of new stream media

RTP is a real-time transmission protocol which provides peer to peer transport service. It is used to support for real-time data transmission in the single-objective broadcasting and multi-objective broadcasting network service.

Applications using the RTP protocol running on top of the RTP, while the implementation of the RTP program running on the UDP-top, the purpose is to use the UDP port number. Table 1-1 shows, RTP can be seen as the transport layer sub layer. Multimedia applications generated by the voice and television data blocks are encapsulated in RTP packets; each RTP packet is encapsulated in the UDP message segment, and then encapsulated in IP packets.

In 1889, information packet structure contains a number of widely used multimedia fields, including audio-on-demand, VOD (video on demand), Internet telephony and videoconferencing. RTP specification does not make standard for sound and video compression format, it can be used to transmit common format. For example, audio of WAV or GSM (Global System for Mobile communications) formats, MPEG-1 and MPEG-2 television, can also be stored in a proprietary format used to transmit sound and television files.

Table 1- 1: RTP is on the transfer layer

	TCP/IP model
	Application
Transfer Layer	RTP
	UDP
	IP
	Data Link
	Physical

From the application developer's point of view, the RTP implementation of the program can be regarded as part of the application, because the developers need to put RTP into applications. Developers need to write RTP protocol program into the application program of RTP packets, and then the application sent RTP packets to the UDP-socket interface (socket interface), as shown in Table. 1-2; Similarly, at the receiving end, RTP packets via UDP sockets interface to input to the application, so developers need to write procedures for the implementation of RTP protocol to applications which is withdrawn from the RTP packet media data.

Table 1- 2 Socket between RTP and UDP

TCP/IP Model	
Application	
RTP	
—	Socket
UDP	
IP	
Data Link	
Physical	

RTP is allowed to assign a separate RTP packet stream for each media source, such as video cameras or microphones. There are two groups in video conferences, which could open four packet streams: two cameras and two microphones, which carry television streaming transmission of voice streaming. However, many of the popular encoding techniques, MPEG-1 and MPEG-2 encoding, they regard sound and television images tied together to form a single data stream in process; build an RTP packet stream in one direction.

RTP packets are not restricted only to application in a single target radio; they can also be transmitted in one-to-many of the multi-objective broadcast trees or in many-to-many of the multi-objective broadcasting tree. For example, for many-to-many multi-objective broadcasting, in such applications, the entire sender usually sent their RTP packet stream to multi-objective broadcasting tree with the same the broadcast address.

3. Streaming Media Transmission Mode

a. Unicast

It needs to establish a separate data channel between the client and media server, each packet can only be sent from a server to a client, this type of transmission known as unicast. Each user must send a separate query separately to the media server, and the media server must copy the packet which forms each user. Server has heavy burden caused by this large redundancy, need long time to respond, or even stop playing.

b. Broadcast

Broadcasting refers to the user passively receive data streams. In the broadcasting process, the client receives flows, but cannot control the flow. For example, the user can not pause, fast forward or rewind the stream. In broadcast, a single copy of the data packet will be sent to all users on the network. Use unicast to send, you need to copy multiple copies of the data packets in order to send to those users by many point to point way, while use the broadcast to send data packets, only need to sent a single copy to all users on the network, it does not care whether the user needs. While the one-to-many network transmission can be achieved, but the broadcast communications have their own drawbacks, the main drawbacks are: easily lead to network congestion; broadcast data cannot cross the router; as a peer communication, it does not have permission to set up; within any one local area network users can receive broadcast data, while the sender does not know number of receiver.

Taken together, the network broadcast is not suitable for stream media transmission, because in video transmission application, not only need to be transmitted in LAN, but also in WAN; also need to consider the security of data, and it easily lead to network obstruction.

c. Multicast

IP Multicast builds a network with multicast capability, allowing a router to copy the data packets to multiple channels. By using multicast mode, a single server sends a continuous data stream to hundreds of thousands of clients on the same time without delay.

Media server needs to only send a packet instead of many; all requesting clients share the same packet. Information can be sent to any address of the client on the network, reduce the amount of transmitted packets. The network utilization efficiency greatly increased, the cost dropped significantly.

Multicast absorbs the strengths of these two methods and overcomes the weaknesses of these two methods; send the data packets of a single copy to the clients that needs. Multicast will not copy many packets and transmit the copies to those clients do not need, ensure the network bandwidth.

D. Peer-to-Peer (P2P) Technology

Peer-to-Peer network (P2P) belongs to one type of network, it is through a large number of peer nodes to connect any network, computers can communicate with each other directly without going through the central server in this network, these peers run the task of other client and server.

1. Definition of P2P

P2P is a peer-to-peer acronym; P2P can be understood as "partner to partner" or "point to point". In the P2P network, concept of client and server is gone, replaced by Servents (server + clients). The earliest P2P applications are resource-sharing. Napster was the first commercial P2P software (an MP3 sharing program), users around the world in one night, and eventually to stop using due to legal reasons, but it opens the door to P2P applications. P2P software, the corresponding agreement and more in-depth study also will be developed. As the features of P2P technology, in the file exchange, peer computing, collaborative work, instant messaging, search engines, online games, Internet-based file storage system, particularly in the application of streaming media, it is more and more important.

2. Characteristics of P2P

While P2P networks have different models, but it has some of their commonalities compared with C/S or B/S network. As follows:

a. Non-central

Network resources and services spread across all nodes, the transmission of information and services are achieved directly between the nodes, without the involvement of intermediate links and servers, avoid a possible bottleneck. Characteristics of the non-center bring its scalability, robustness, and other areas.

b. Extendibility

In P2P networks, along with the user to join, not only the demand for services increased, overall system resources and service capabilities are also synchronized to expand, and always be able to more easily meet user needs. The distribution of the whole system is full, there is no bottleneck. In theory its scalability can be considered almost infinite.

c. Tuneability

P2P architecture is inherently resistant to attack, has high fault-tolerance. As the services are scattered among the various nodes, and part of the destruction of the node or network has little effect on other parts. P2P network can automatically adjust the overall topology of other nodes to maintain connectivity when some nodes are failure. P2P networks are usually self-organization approach to building up, and allow nodes to join and leave freely. P2P networks also can be based on network bandwidth, number of node, load and other changes to make adaptive adjustments.

d. High Performance/Cost Ratio

The performance advantage of P2P is an important reason of widespread concern. With the development of hardware technology, personal computers computing, storage capacity and network bandwidth, performance is rapid growth in accordance with Moore's theorem. Using P2P architecture can effectively use the Internet to spread a large number of ordinary nodes, distribute the computing tasks or stored data to all nodes. Use one of the idle computing power or storage space to achieve high-performance computing and mass storage purposes. Through the use of a large number of idle resources in the network can be used to provide more lower cost computing and storage capacity.

e. Load Balancing

In P2P network environment, each node is both a server is the client, it can reduce the need for the traditional C/S structure of the server computing power, storage capacity requirements, as well as the resources are distributed across multiple nodes, better realization of the entire network load equilibrium. In centralized video conferencing, because of the bottleneck of server's CPU, storage capacity, network bandwidth, when the user increases, the image quality was experiencing deterioration. We are anxious to find a new idea to solve this problem. So P2P distribute the task of server to the client, it can not only reduce the burden on the server, but also make full use of client resources.

3. Classifications of P2P

P2P research can be divided into a central directory type, flooding type, document routing model by topological relations.

(1) Central directory type: it stores node address information and the stored data of all the nodes. It can be able to return quickly to find the most suitable one or more of the destination node. Such as Napster.

(2) Flooding type: also known as the full distribution of unstructured networks, the network there is no central server, peer nodes and never take the initiative to publish share information. In the overlay network, using a random graph organization, the node degree subject to "Power-law" rule, find the destination node faster, have better fault tolerance in the face of dynamic changes of the network reflects, and therefore have a better availability. At the same time can support complex queries, such as multiple keywords with a regular expression queries, fuzzy queries, etc. The most typical case is like Gnutella.

(3) Document Routing: also known as the fully distributed structured topology networks, each of the nodes is assigned a random ID value. And have a certain number of other values such as node ID, issued a new document based on content and name of the hashing algorithm to generate a file ID value. These are the basic structural model; these structures can also be mixed to generate new models, such as semi-distributed structure. The most typical cases, such as, Kazaa JXTA.

II. Video Conferencing System Scheme

Video conferencing system [3-4] is a communication technology, computer technology, microelectronic technology in one of the remote off-site means of communication, the system is a typical voice and video communication. The transmitter in the communications, image and audio signals are transferred into digital signals, at the receiving end then it can reproduce the visual and auditory access to information. Compared with the conference call, video conferencing possesses strong intuitive and informative characteristics.

A. DirectShow

1. Summarize

DirectShow is not only an open application framework, but also a set of COM-based programming interface. DirectShow system functions are as shown in Fig. 2-1.

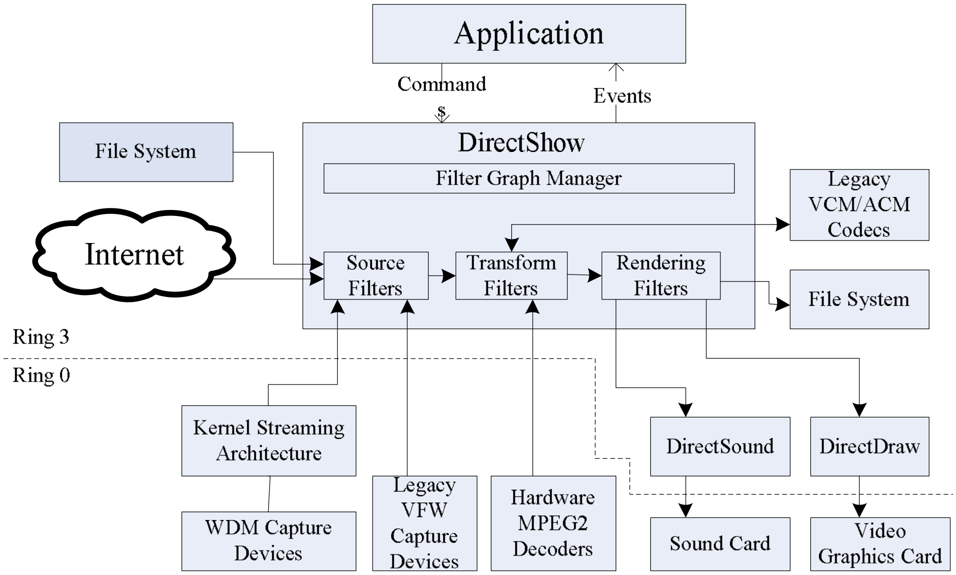

Figure 2- 1 The DirectShow.

From above figure to be seen, which is the biggest DirectShow system, the basic working principle is the "pipeline": Filter the cell components in series together, a unified control by the Filter Graph Manager. The system input can be a local file system, the hardware card, Internet, etc. The system output can be a sound card (sound reproduction), graphics card (video display), the local file system, of course, could eventually send the data to the network [5].

In fact, computer applications in the field of many modules can interact with DirectShow system. In other words, scope of application of DirectShow is very wide. Purely from the local systems that DirectShow can be achieved in different formats of media files decode and play, or convert between formats, you can collect data from the local machine collection devices and save it as audio and video files, you can receive, watch analog TV and so on. From the perspective of network applications, DirectShow can also be used for video on-demand, video conferencing, video surveillance and other fields. In fact, broadly speaking, DirectShow system is suitable to all streaming data processing, these data can be audio, video, multimedia data, but not be limited to multimedia data.

DirectShow is based on COM (Component Object Model) system by a number of modular software components. In this system, the most basic building blocks are known as filter of the software components. The processing of multimedia data is divided into a number of steps by filters, each step by a filter is in order to complete the implementation of the multimedia data stream with a simple operation. Filter has input and output, it accepts input and produces output [6]. For example, a decoder filter, its input is based on some kinds of format, encoded multimedia data streams, which filter the output after decoding the data stream.

2. Filter

Filters are composed of the most basic components of DirectShow. DirectShow for data stream processing can be roughly divided into several separate processes; each process is complete different tasks. The filter is the basic unit of the completion of these processes. In fact, the user's application program is a function of several different filters combined filter map (Filter Graph).

DirectShow Filter can be divided into the following categories:

a. Source Filter

The source filter is the filter map to deal with the input data filter. It is to obtain the raw data from external devices and make a simple processing, and then send the data down to a filter [7].

b. Transform Filter

Transform filter is the filter map at the core of a filter from the data acquisition and process it into raw data stream into other forms of multimedia data streams and compression coding or decoding; a data flow broken down into multiple data streams, such as putting an audio and video mixing flow broken down into separate audio stream and a separate video stream; the combination of multiple data streams are broken into a single data stream, etc. [8].

c. Render Filter

Presented filter is in the final level of filter layouts; its role is to processed data stream presented to the external device. Here that the external equipment includes file system, display cards, sound cards, network cards, etc. [9].

3. Filter Graph

Any applications developed using DirectShow, must create multiple filters and make proper connections, so the data stream can be transmitted from the source filter through to the Render Filter output by users. Collection of these filters is called filter map (Filter Graph). Fig. 2-2 is an example of a Filter Graph.

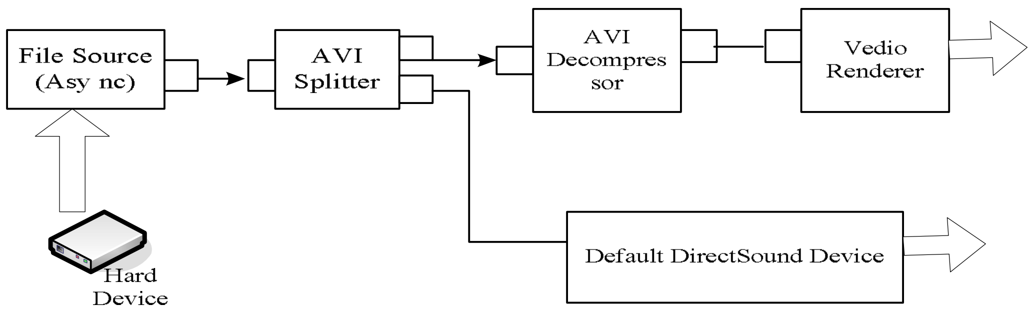

Figure 2- 2 The example of Filter Graph

4. Multimedia Data Sample and Multimedia Data type

Two filters connected, they must use the same data type. This will ensure the next one filter can handle a filter from the data obtained. The data is transferred between filters after a COM package, known as the multimedia data sample (MediaType), and the use of MediaSample or IMediaSample2 interface. In the actual data, also contains a time stamp in order to achieve synchronization.

B. Real-Time Transport Protocol (RTP)

Real-time streaming transport [10] denotes that ensure that the signal bandwidth and network connection with horses, so the media can be real-time watch. Real-time streaming and HTTP streaming transmit differently and require a dedicated streaming media server and transfer protocol. Real-time streaming is always real-time transmission, especially for live events, and also supports random access; the user can fast-forward or rewind to view the content of the front or the back. In theory, real-time streaming playback cannot stop once, but in fact, may occur in cycle pause.

Streaming media technology not be achieved without the support of a variety of new network protocols, TCP need more overhead, it is not suitable for real-time data transmission. In the streaming implementation of the program, transmit control information with HTTP/TCP in common way, while transmit real-time multimedia data t with the RTP/UDP.

RTP (Real-Time Transport Protocol) is used for Internet, multimedia streaming applications a real-time transport protocol, which is based on multicast or unicast network services, providing end to end network transport functions suitable for transmission, such as audio, video real-time data, or simulation data. RTP is defined as one-to-many transmission situations or work, the purpose is to provide time information and the achievement of flow synchronization. RTP typically use UDP to transfer data, but RTP can also be other protocols such as TCP or ATM on top of work. When an application to start a RTP session will use two ports: one is for RTP, the other is for RTCP. RTP itself can not transmit data packets in order to provide a reliable delivery mechanism does not provide flow control or congestion control, it relies on RTCP to expand in large multicast networks monitor and control data transmission and provide minimal control and identification functions. RTP is usually not as a separate algorithm for network layer to achieve, but as part of the application code. Real-Time Transport Control RTCP (Real-Time Transport Control Protocol) and RTP protocols provided with flow control and congestion control services. In the RTP session, the participants periodically send RTCP packets. RTCP packet contains the packet has been sent.

RTP/RTCP protocol is the core of streaming media technology, in the streaming media technology has an important role. RTP/RTCP is designed to support real-time multimedia communication and transport-layer protocol designed by the IETF as an RFC an 889 release. RTP is located in UDP on top, is responsible for the transmission of multimedia data. Although the RTP/UDP not TCP is less reliable, and cannot RSVP and ensure the quality of service real-time services and the need for RTCP real-time monitoring of data transmission and guarantee the quality of service, but because UDP transmission is more delay than TCP, and audio streaming is a good match with video streaming, so in actual use, RTP/RTCP with the UDP is for the audio/video media, while the TCP is for data and control signaling transmission, thus making the network resources have been put to good use .

RTP in the multimedia data (UDP packets) in the head use time-stamp and serial number issued, if coupled with an appropriate buffer, then the receiver can be recorded out of order packets, synchronized video, audio, and data according to time-stamp and serial number information "regeneration, recovery of" data package, and improve the playback results. RTCP is responsible for monitoring network quality of service, communication bandwidth and transmission over the Internet related news, and notify the sender of these messages. Once the network available bandwidth narrowing, RTCP will immediately notify the sender that information, sending end and then adjusted according to the information sent, thus making multi-media communications to continue.

RTP is used to send data one-way and non-confirmed manner. Each RTP packet header is as shown in Table 2-1 contains a number so that the receiver can recover the original time stamp indicative timing marks, and enable the receiver to handle lost duplicate or erroneous packet sequence number. RTP is not only suited to one receiver (unicast), but also suited to multiple recipients (multicast) to send audio and video streaming.

Table 2- 1 RTP Header Format

V=2	P	CC	M	PT	Sequence Number
Timestamp					
Synchronization Source SSRC Identifier					
Contributing Source CSRC Identifier					

Where V is protocol version, P is adjustment switch (O to O are, if the switch placed on, then the corresponding blocks of the tail as being filled bytes, CC is CSRC counters (CSRC chunks contained in the number of domains), M is tag (the relevant configuration files), PT is load type (in the corresponding definition ie C,)

RTP can better handle multimedia applications in real-time characteristics. Streaming media applications and traditional data applications, different from their sender, receiver and network requirements are different. When transmitting audio or video stream, there is nothing wrong with the overall loss of some of the data, as long as the audio or video to avoid the big interval. The traditional data communication requires the accurate transfer process.

RTP of Real-time Transport Control Protocol RTCP makes the receiver can feedback to the RTP sender (or vice versa). For example, a receiving application can send the application to send video streams slow down the speed. Sending a video stream at a lower speed can still be seen, but may come about jitter or resolution is not high. RTCP specifications of the guidelines are to help control the flow of programmers to avoid the consumption of too much network bandwidth.

C. P2P Video Conferencing system Architecture

P2P technology enables video conferencing system in the reliability and stability has further improved, therefore, video conferencing systems and P2P technology together, which would ensure a reliable video conferencing system, has been achieved. P2P-based video conferencing technology, the structure of the system there are many, summed up mainly in the following three ways:

1. Server based P2P

The system still follows H.323 standard, the system exists in the server and client, and server is divided into the main server and from the server. Master server is responsible for the task are: user management, conference reservations, conference configuration, notices of meetings, the meeting was terminated, the meeting be extended from the server is responsible for the task are: audio and video reception, storage, processing, forwarding, and audio and video synchronization control; text chat and file transfer; support simultaneous operation of more than a whiteboard; desktop sharing, application sharing. The client is the general participants, can be desktop computers, mobile devices.

A single server performance and capacity is limited, can only support a certain number of users simultaneously online, and the system stability and performance cannot be effectively guaranteed. In the large-scale application environments need to use multiple servers to work together to support large-scale simultaneous users online, by configuring the intelligent connection module can be connected online users evenly distributed to various servers, the full and effective use of server clusters. The user first connected to the main server, asked to join the meeting, the main server, the type of the user from the server's load will be assigned to the appropriate user from the server. P2P is only used between the server modes (with an index of multi-server).

For large-scale conferencing systems require multiple servers to work together in order to support large-scale simultaneous users online, due to large amounts of data between servers need to transfer, the link can easily become a bottleneck to improve the quality of audio and video, the system uses P2P technology to improve the transmission efficiency. Between master and slave servers using P2P connection is working the way, each server maintains one other server resources index.

2. No server P2P approach

Architecture schematic is shown in Fig. 2-4, in this architecture is not the server. Client is not only the client but also the server, When the Client receives the information from other client when it is the client, when the client to send messages to other Client when it is the server.

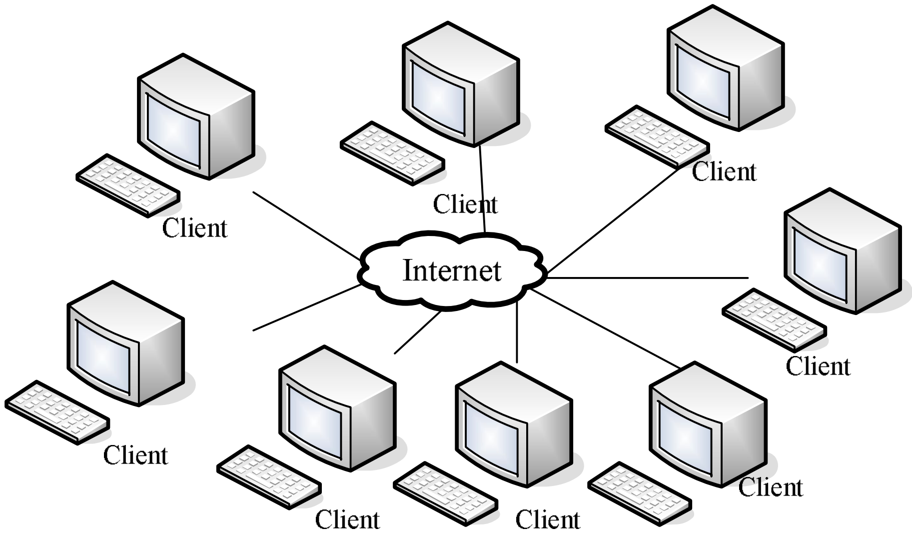

Figure 2- 3 No server-based video conferencing architecture

When the client A needs another client B's information, clients search resources using the forms of unicast, multicast, broadcast and other forms of adaptive to form a dynamic network. If find the resources needed in the client A and client B to establish a connection between the transmission of information; if not found then saved to the local radio the other Client's own needs, the client received from other connection information relating to client B, and then establish to connect and send information, in order to prevent broadcast storms, generally broadcast hop limit is 3; if it did not receive connection information relating to client B, then as a client B has been to leave the conference room.

3. With a central server in P2P mode

This architecture exists in the server and client, but the server does not handle and store client information (video, audio, documents, etc.), server is only connected to the preservation of video conferencing systems available client list and their list of the information, server initializes communication between the two client, and then the two connected client to establish communication channels to stay connected and send messages. Serve: responsible for the tasks are: user management, conference reservations, conference configuration, notices of meetings, the meeting was terminated, to extend the meeting, conference statistics, conference inquiries. Transmission of information between the clients through the Server is no longer the storage, handling and forwarding, but direct transmission between the clients. This originally from the Server is responsible for bandwidth and system resources, demanding part of the task will be delegated to client, client is the client is the server. As the level of the continuous improvement of computer hardware and software, Client has been able to undertake such tasks. Architecture schematic diagram shown in Fig. 2-5:

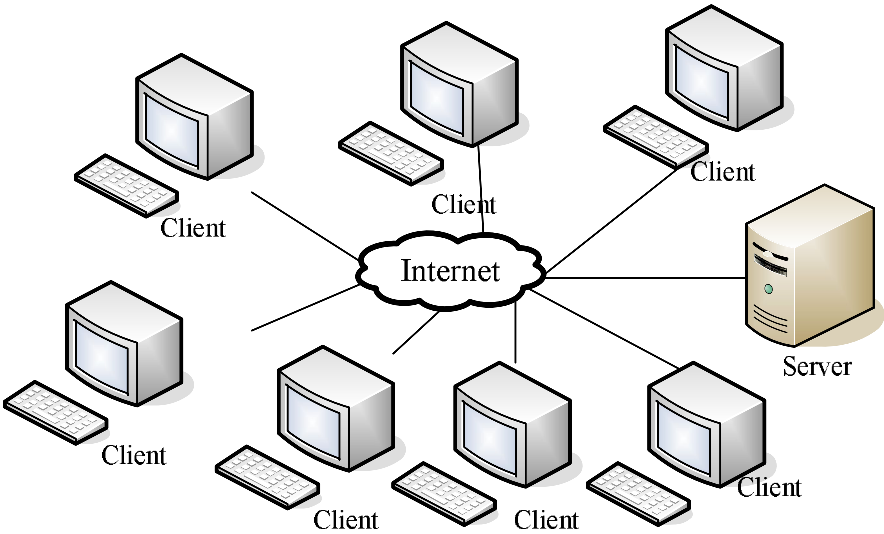

Figure 2- 4 With a central server-based video conferencing architecture

In such a system, structure, Server is no longer the bottleneck of video conferencing systems, due to adopt P2P manner, to a certain extent reduce the network congestion.

D. Comparison of three kinds of architecture

An architecture based on the first meeting of system owners, have higher demand from the server, this server is not flexible enough for users to be expected, and accordingly arranged the server, the server requirements are also relatively high. Therefore, the cost of such meetings is also very high. With the traditional centralized video conferencing, as also there may be congestion.

The second architecture based conferencing system to users, they are not too high, as long as the user can access the Internet, there is audio and video input and output devices can enter the conference room and meeting room according to their ability and meeting rooms demand for a certain task. But the users to search resources in a very long time, the user's information list will not automatically update in a timely manner, if a user left the conference room, other users connected to this node will become isolated point, re-connected to the conference to take a very a long period of time, to send out search requests were often not had any response, therefore, such structures still needs to improve.

Architecture based on a third set of conferencing system for more than two kinds of advantages of the architecture designed to avoid its shortcomings. The server's mission has been to reduce a lot of demands on the server is not very high. And because there is a server, so the meeting of the management is very convenient, user-selected meetings, join the meeting, the allocation of the terminal number, informed the meeting that the other participants in the other participants in the Conference, a list of the user out of session, and a variety of permissions management, etc. are all done by the server.

III. The implementation of video conference system based on P2P technology

Based on the idea P2P built a roommate central server-based video conferencing system, the system is consist of the central server and various peer client component. Central server provides public service and conference rooms of the management functions, but does not participate in data transfer between peer. The wind was distributed among the nodes in the structure of message passing to each other through the dissemination of information and search resources, through the application process to deliver multicast data.

A. The solution of NAT through problem

With the popularity of the Internet to connect computers on the Internet, more and more people began to realize that lack routed IP address resource issues. A large number of enterprises adopt the program of private IP addresses to form their own networks to form relatively independent and closed networks, when LAN in the formation of all sizes this networking mode has been widely used. It can be said that in today's online world, use the private network IP address of the network is much larger than the number of devices using the public network IP address of the number of network equipment. As the private network IP address cannot be routed between the public networks, so the private network equipment can not directly access the public network resources. As the public network IP address of lack of resources, it is impossible for each private network device a public network IP address. In order for the private network devices can access the private network of external resources, NAT [11] [12] (Network Address Translation technology) have emerged.

But we are now based on the idea to construct our P2P video conference system, it is often the machine will cause the public network and private network equipment for data transmission, such as video streaming, audio streaming. Therefore, NAT has become the biggest obstacle, so we have a pass through NAT technology. In the STUN approach, the NAT technology typically consists of the following four categories: FullCone type, RestrictedCone type, Port Restricted type and Symmetric model.

1. UDP through NAT

We have to conduct network transmission, the main use of the TCP [17] transport layer protocol and UDP protocols. TCP protocol is reliable, connection-oriented transport protocol. UDP is unreliable, connectionless protocol. According to TCP and UDP protocols implementation principle for the NAT to carry out penetration, mainly referring to UDP protocols. Let us look at the use of UDP protocol to penetrate What the NAPT principle is: one external network IP address of the computer want to NAPT behind the computer communication within the network condition is to ask the internal network behind the NAT external network IP address of the computer initiative the computer to launch a UDP packet. External network IP address of the computer to use the data received in UDP packets get to the NAT's external network IP address and mapped port, the future can be and intranet IP-based computer to communicate in a transparent.

B. Central server implementation

Center Server is mainly responsible for clients, conference room management; it holds the entire P2P network node connection information for all nodes to provide directory services. If one node wants to connect with another nodes, you must first get through the server's IP address of other nodes and the nodes provide the services or to share information ID. As the actual network topology may be very complex, some computer may be behind a firewall, while the firewall allows only HTTP protocol usually adopted, so we use HTTP as the communication protocol between the nodes and servers.

The server software architecture as shown in Fig. 3-1: Mainly a user management module, conference room management module, data security and encryption modules composed of three modules. Its user management module by the user registration, online user management, user search modules, primarily to complete user management; conference room management module from the conference registration, conference booking, conference room monitoring, conference room search, Council member management module composed of the completion of the management of the main conference room. The Fig.3-2 shows the test of the main server.

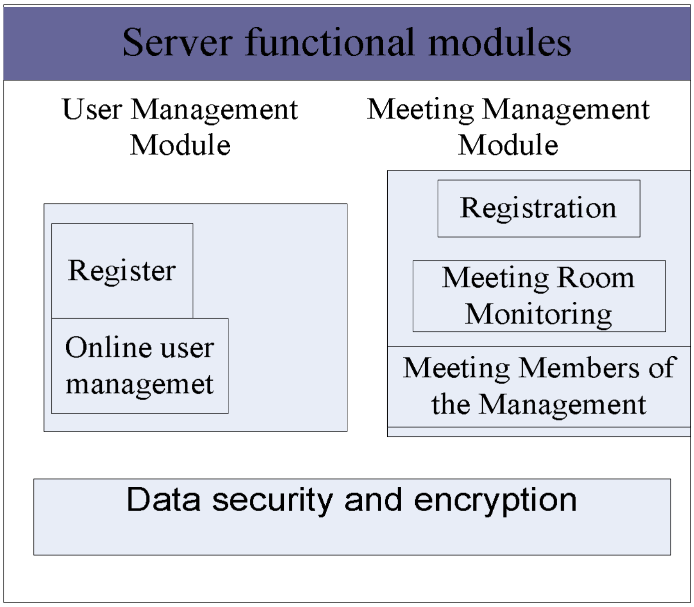

Figure 3 - 1 Core server functional module

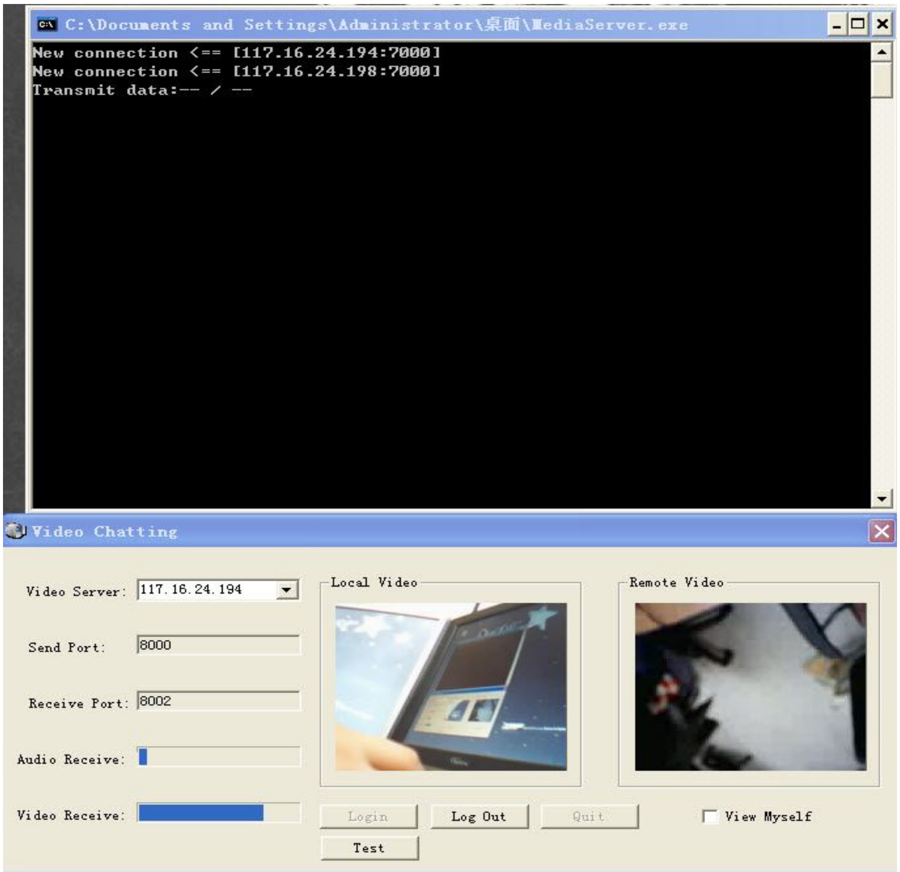

Figure 3 - 2 Test of main server

C. Peer-side implementation

In the system for each user is a Peer, the system topology is a node. Each Peer is equal; it can receive information from other Peer there, but also provide services for other Peer. Each user according to their ability to undertake the task of the meeting, in order not to allow users to suffer a particular task, the right to influence its access to services and the provision of other services, capabilities, we require a user to receive the media stream, for the Unicast child nodes no more than 5, if the completion of its lack of capacity can also refuse to provide services. Users can also according to their own preferences, asked to select a streaming media format, the source. User login system will be every 3 seconds will be sent to the central server heartbeat factor. After the user into the conference room, members in the conference room of the activities based on the time the conference room; server in addition to monitor its status, if necessary, according to the conference room of the Council members request the administrator to manage things, not between the various users in the right The information, data transmission services.

Peer-end user software architecturea and its test as shown in Fig. 3-3 and 3-4

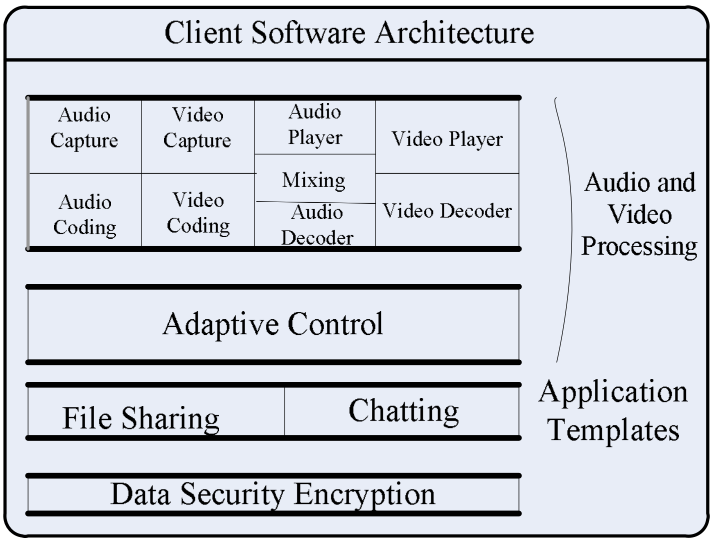

Figure 3 - 3 Software structure

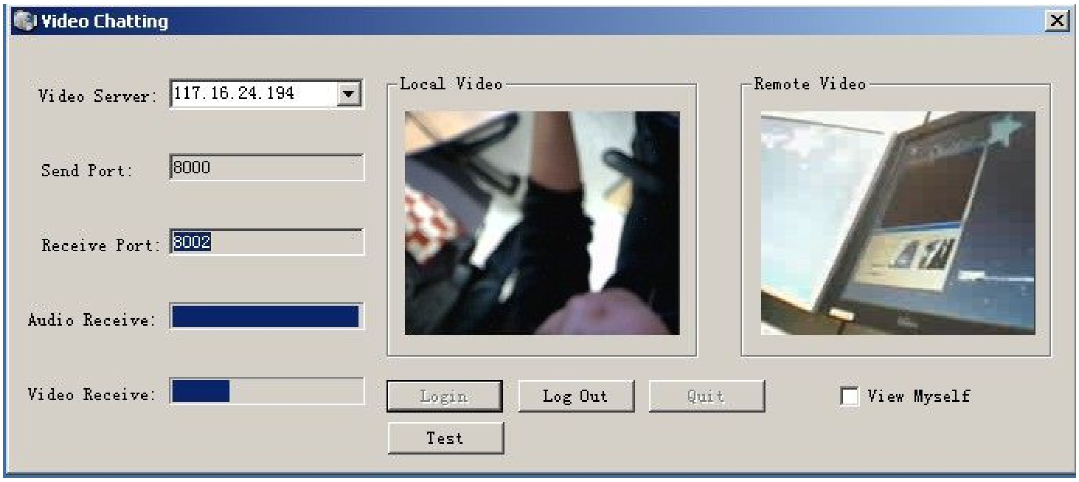

Figure 3 - 4 Test of client

D. The design of video and audio module

In the DirectShow, VCM has been AVI Compressor filter encapsulates. Similarly, ACM encoder is also encapsulated the ACM Wrapper filter. DMOs are DMO Wrapper filter package. System Device Enumerator provides a unified approach to enumerate, and create these compressors, we do not consider the low-level operation.

1. Capture device enumeration

In short, all the directshow filters are classified by category and each classification is identified by the GUID. For the video compression device, GUID type is CLSID_VideoCompressorCategory, the audio is CLSID_AudioCompressorCategory. In order to enumerate a specific type, the system device enumerator establishes an enumerator object to support the IEnumMoniker interface. APP uses this interface to return device monikers, each device moniker perform a directshow filter example. Using this moniker to establish filter, or have the friendly name of this device without having to build filter.

In order to enumerate video or audio compressor in the user's system:

(1) CoCreateInstance create a system device enumerator, which has a class ID: CLSID_SystemDeviceEnum.

(2) Classify GUID with the filter called ICreateDevEnum:: CreateClassEnumerator. Back to IEnuMoniker interface pointer.

(3) Use IEnumMoniker::Next enumerated devices moniker. Returns an IMoniker interface, and demonstrate that moniker.

```
void OnInitDialog(HWND hDlg)
{
    HRESULT hr;
    ICreateDevEnum *pSysDevEnum = NULL;
    IEnumMoniker *pEnum = NULL;
    IMoniker *pMoniker = NULL;

    hr = CoCreateInstance(CLSID_SystemDeviceEnum, NULL,
        CLSCTX_INPROC_SERVER, IID_ICreateDevEnum,
        (void**)&pSysDevEnum);
    hr = pSysDevEnum->CreateClassEnumerator(
            CLSID_VideoCompressorCategory, &pEnum, 0);

    while (S_OK == pEnum->Next(1, &pMoniker, NULL))
    {
        IPropertyBag *pPropBag = NULL;
        pMoniker->BindToStorage(0, 0, IID_IPropertyBag,
                        (void **)&pPropBag);
        VARIANT var;
```

```
        VariantInit(&var);
        hr = pPropBag->Read(L"FriendlyName", &var, 0);
        if (SUCCEEDED(hr))
        {
            LRESULT iSel = AddString(GetDlgItem(hDlg,
                        IDC_CODEC_LIST), var.bstrVal);
        }
        VariantClear(&var);
        pPropBag->Release();
        pMoniker->Release();
    }
SendDlgItemMessage(hDlg, IDC_CODEC_LIST,
                    LB_SETCURSEL, 0, 0);
    pSysDevEnum->Release();
    pEnum->Release();
}
```

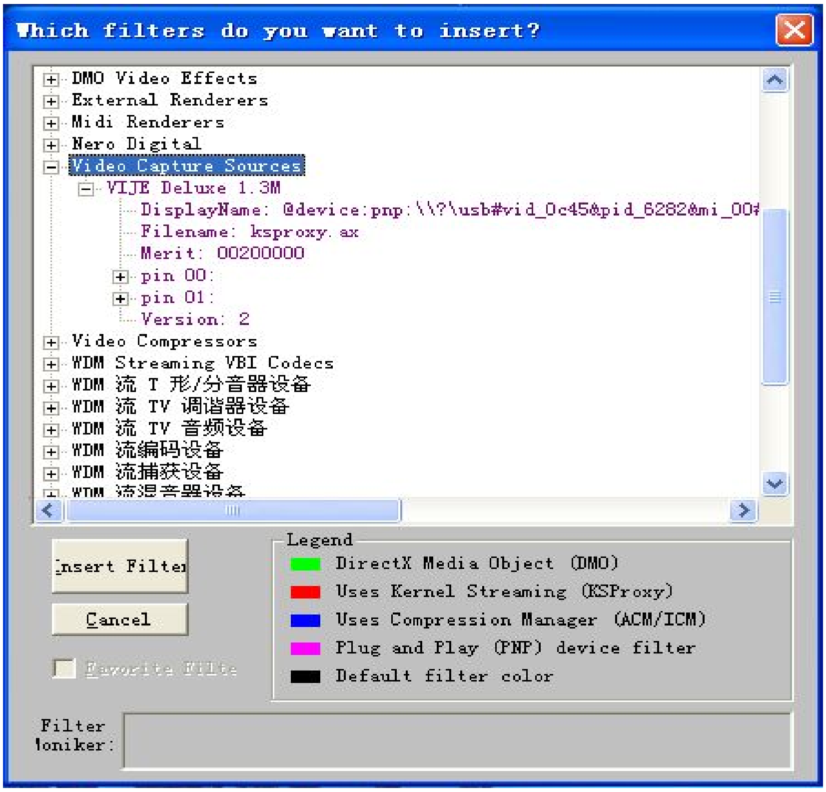

Figure 3 - 5 System Device Enumerator

Fig. 3-5 denotes enumeration for the use of the system components of a system to enumerate all the filter examples, video capture sources directory in which the filter is a video capture device.

Video capture device enumeration process as shown in Fig. 3-6:

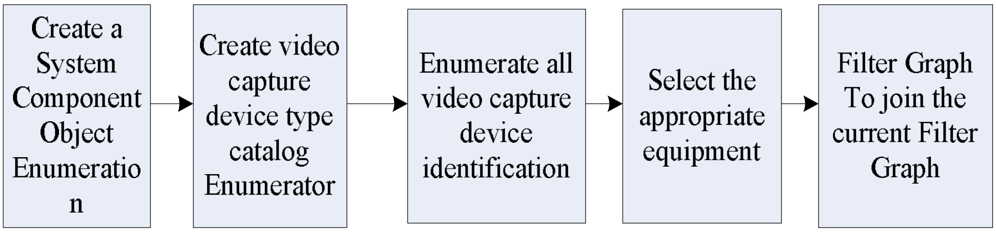

Figure 3 - 6 Flowchart of Video Capture Device enumeration process

For audio capture device enumeration methods and selection methods, processes, and video capture device enumeration methods and processes are consistent, but when using the interface method ICreateDevEnum::CreateClassEnumeratord the parameters should be replaced by CLSID-AudioInputDevicecategory type, shown in Fig. 3-7.

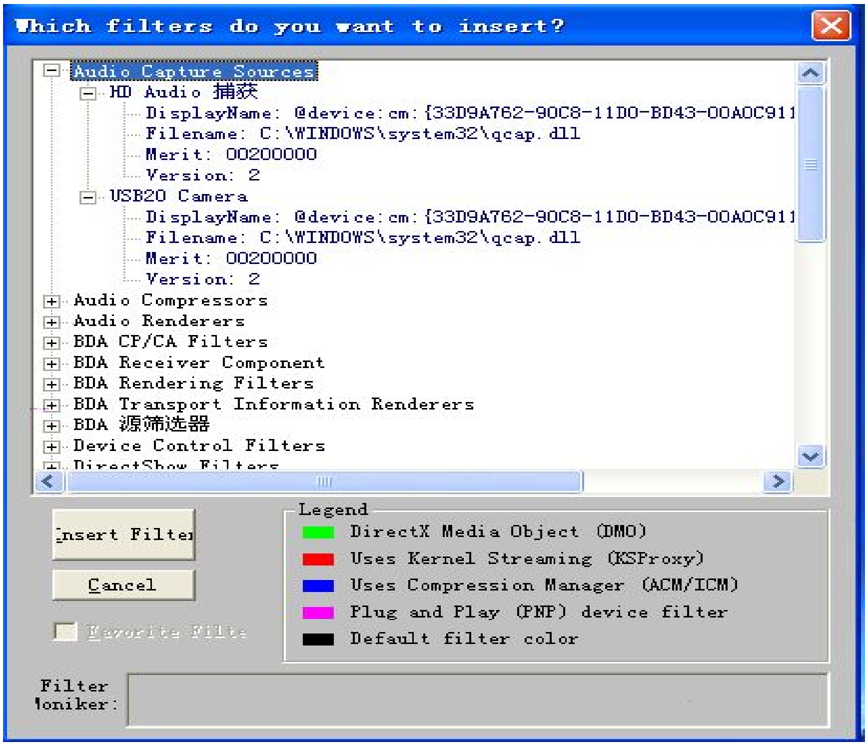

Figure 3 - 7 System Audio Capture Device Enumerator

2. Create a filter

Call IMoniker:: BindToObject method. This method returns an IBaseFilter interface pointer. Like the following way:

```
IBaseFilter *pFilter = NULL;
hr = pMoniker->BindToObject(NULL, NULL, IID_IBaseFilter,
                                    (void**)&pFilter);
if (SUCCEEDED(hr))
{
  LRESULT iSel = AddString(GetDlgItem(hDlg,
                                    IDC_CODEC_LIST), var.bstrVal);}
```

Then using RenderStream, connect the source filter and render filter, choose to add a number of intermediate filters.

To obtain the current Filter Graph

Because in the call SetOutputFileName in, capture graph builder object establishes a filter graph, all you have to add filter which need into a same filter graph. By ICaptureGraphBuilder:: GetFiltergraph get the newly created filter graph. The returned pointer is parameter gcap.pFg.

```
// The graph builder created a filter graph to do that.  Find out what it is,
// and put the video capture filter in the graph too.

    hr = gcap.pBuilder->GetFiltergraph(&gcap.pFg);
    if (hr != NOERROR)
    {
        ErrMsg("Error %x: Cannot get filtergraph", hr);
        goto SetupCaptureFail;
    }
```

Add audio / video filters to the current Filter Graph

```
    hr = gcap.pFg->AddFilter(gcap.pVCap, NULL);
    if (hr != NOERROR)
    {
        ErrMsg("Error %x: Cannot add vidcap to filtergraph", hr);
        goto SetupPreviewFail;
    }

hr = gcap.pFg->AddFilter(gcap.pACap, NULL);
    if (hr != NOERROR)
    {
```

```
        ErrMsg("Error %x: Cannot add audcap to filtergraph", hr);
        goto SetupCaptureFail;
    }
```

3. Rendering capture filter and video preview

a. Capture Pin for rendering video capture filter and audio capture

ICaptureGraphBuilder::RenderStream connects the source filter pin to the renderer filter.

Pin type is optional: capture pin (PIN_CATEGORY_CAPTURE) or preview pin (PIN_CATEGORY_PREVIEW). The following example shows the connection video capture filter (gcap.pVCap) to render the capture pin in gcap.pRender.

```
// Render the video capture and preview pins

    hr = gcap.pBuilder->RenderStream(&PIN_CATEGORY_CAPTURE,
NULL, gcap.pVCap, NULL, gcap.pRender);
    // Error checking
```

Again ICaptureGraphBuilder::RenderStream connecting audio capture filter (gcap.pACap) to render the audio renderer.

```
    if (gcap.fCapAudio)
    {
        hr = gcap.pBuilder->RenderStream(&PIN_CATEGORY_CAPTURE,
NULL, gcap.pACap, NULL, gcap.pRender);
    // Error checking
```

b. Rendering preview pin to video capture filter

Called again ICaptureGraphBuilder:: RenderStream, from the capture filter's preview pin to the video renderer. Code is as follows:

```
hr = gcap.pBuilder->RenderStream(&PIN_CATEGORY_PREVIEW, NULL, gcap.pVCap, NULL, NULL);
```

When you get IVideoWindow interface, you can call the method IVideoWindow like put_Owner, put_WindowStyle, or SetWindowPosition to get the handle of video preview window, set the window properties.

The interface in the video capture device Filter can video capture the device output image display parameters (brightness, hue, saturation, contrast, etc.).

Image display parameter adjustment process shown in Fig. 3-8:

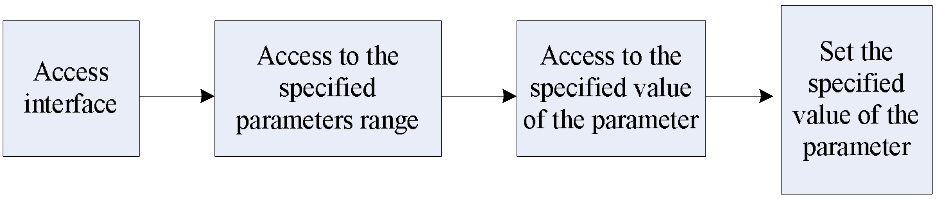

Figure 3 - 8 Flowchart of Image display parameter adjustment

Establishment of a complete capture filter graph, you can preview audio, video, or capture the data. The Filter Graph to build video capture as shown in Fig. 3-9:

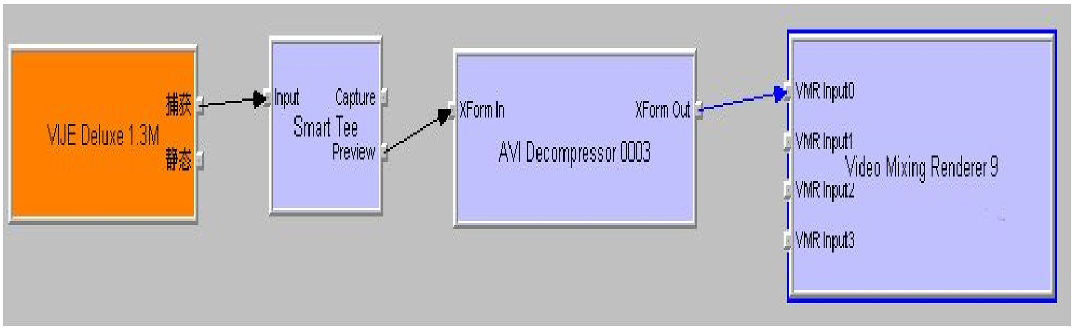

Figure 3 - 9 Equipment used for video capture preview Filter Graph

Because there is no preview of some capture device output pins, created to capture charts, a Smart Tee received the request for additional capture output pin back, so that will be able to fit two outputs, one for preview, one for output to Filter media encoder input pins. In the smart Tee output pins, preview pin and capture the actual output pins of the sample data is the same, set a good image display parameters, as long as the sample data directly by pushing modalities, pushed to the next filter, and video encoding filter can be.

c. Access to the captured information

Interface obtained through the IAMDroppedFrames. Testing the number of missing frames (IAMDroppedFrames::GetNumDropped)， the number of capture (IAMDroppedFrames::GetNumNotDropped)。IAMDroppedFrames:: GetAverageFrameSize approach provides a frame to capture the average size (unit: byte). Use this information to know the total number of capture bytes and frames per second (rate).

E. The design of audio video player

The main function of audio video player is to get audio streaming and video streaming from the Pin of output of audio, video decoder filter and send to audio render and video render to play out.

Fig. 3-10 and Fig. 3-11 show the flow chart of the video streaming and audio streaming processing.

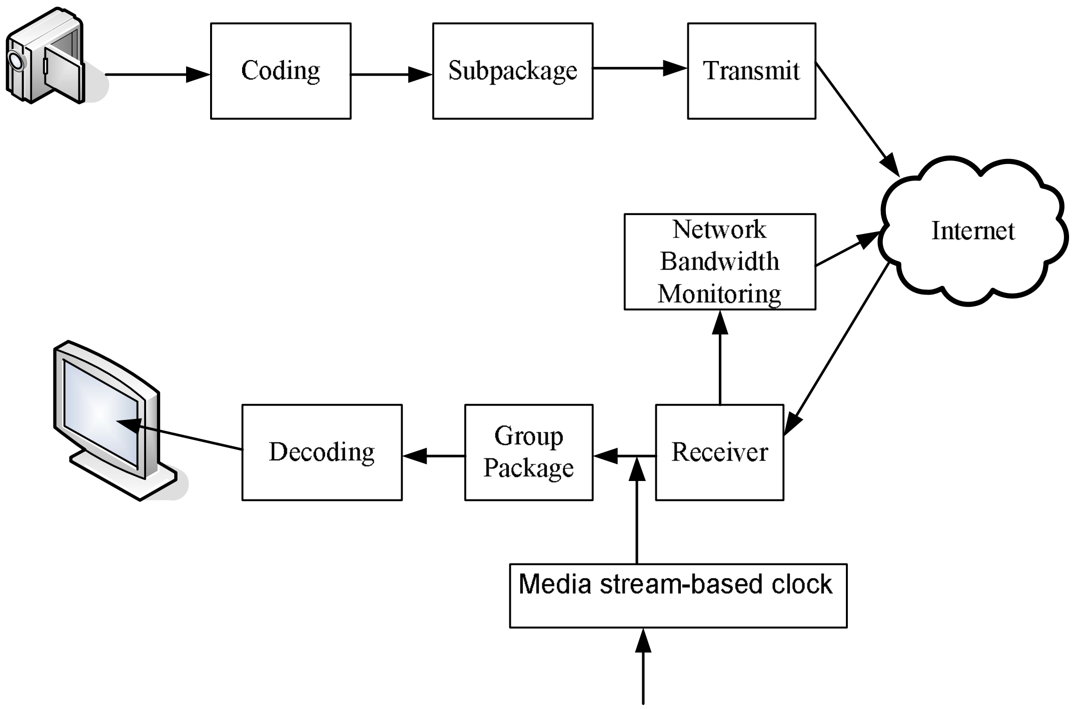

Figure 3 - 10 Video process flowchart

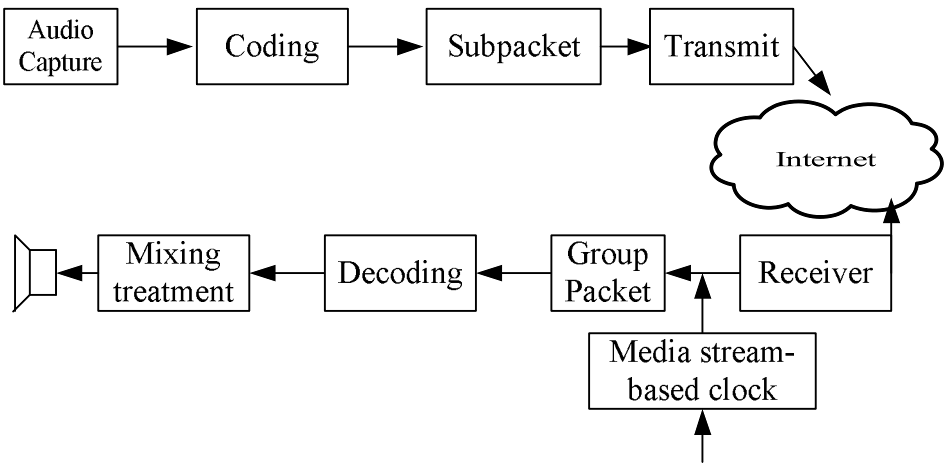

Figure 3 - 11 Audio process flowchart

F. Conferencing Login and Management

1. Conferencing Login

Safety and reliability of the network has been the focus of attention. In order to ensure the safety and reliability of the conferencing must be authenticated, transmitted data should be encrypted. Thus, the system participate in the conferencing on the server side the default user name and password, be stored in the database, and a user name and password, and conferencing -related information to encrypted email notified the conference attendees.

System in dealing with the "Login" button click event, the program first checks whether the user name and password blank, if it is empty, to prompt and exit operations. Otherwise, the user name and password for the conditions in the query data from the data table, the data returned, that the user name and password are correct, or else that the user name and password is not correct.

```
void CLogin::OnConfirm()
{
	CString c_password;
	m_password.GetWindowText(c_password);
	if (m_username.IsEmpty()||c_password.IsEmpty())
	{
		MessageBox("ID and PW is empty!","Attention ",64);
		return;
	}
	CString sql = "select * from tb_login where m_name = ? and
m_password = ?";
	dataManage.p_Com->ActiveConnection =
dataManage.p_Con.GetInterfacePtr();
		dataManage.p_Com->CommandText =(_bstr_t) sql;
	_ParameterPtr m_param1,m_param2;
	  m_param1 =
dataManage.p_Com->CreateParameter("a",adVarChar,adParamInput,30);
	m_param1->Value = (_bstr_t)m_username;
	 dataManage.p_Com->Parameters->Append(m_param1);
	m_param2 =
dataManage.p_Com->CreateParameter("b",adVarChar,adParamInput,30);
	m_param2->Value = (_bstr_t)c_password;
	dataManage.p_Com->Parameters->Append(m_param2);
	try
	{
		dataManage.p_Record =
dataManage.p_Com->Execute(0,NULL,adCmdText);
		if (dataManage.p_Record->BOF &&
dataManage.p_Record->ADOEOF)
		{
			MessageBox("ID and PW are wrong");
```

```
        }
        else
        {
    AnimateWindow(m_hWnd,2500,AW_SLIDE|AW_HIDE|AW_BLEND);
            ::FreeLibrary(h_prohandle);
            EndDialog(0);
            }
    }
    catch(_com_error &e)
    {
        MessageBox(e.Description());
    }
    dataManage.p_Com->Parameters->Delete("a");
    dataManage.p_Com->Parameters->Delete("b");
    }
```

2. Conferencing Management

When the user to enter the conference room, its identity authentication, user experience certificates to enter the conference room to the user after the allocation of certain resources to get him the information (including the user logged on Peer's IP) to send the conference room monitor module. When the module was found after a member left the conferencing room or a member take the initiative to leave the room, remove the member list of the information displayed in the member to recover the user's resources, and to inform the other members of the conference room of a member has left. Right and exit the relationship between the members of the service or services to the users accordingly.

3. The realization of other features

Text chat and file transfers are extra performance which audio and video cannot be implemented in order to complete the system.

Text chat module is the important complementary function modules of video conferencing system, which enables users to carry out a simple text chat. Text chat can be divided into public chat and one to one chat: public chat, chat information through the application layer multicast to disseminate to the meeting each member of the user interface; one on one chat takes two to start a conversation after the before they can chat between the two, the message is only transmitted between the two, other users cannot see.

Text chat control using UDP protocol to be implemented chatting for the individual, the organization and all the people. It can be direct discussion within the group.Its interface functions are as follows:

(1)BOOL InitChat(LPCTSTR Name,long Port)

Initialized chat control, set the port number used by chat, and in the specified port to monitor the arrival of the data. Name used for their-own user name, Port is the port number used.

(2)BOOL AddIp(LPCTSTR Name,LPCTSTR Ip)

To add a chat user, given by the application. Name is the newly added user name, Ip is their IP addresses. Add someone to participate in a meeting.

(3)BOOL OnKey(long nChar)

Handle shortcut keys messages, through the main application, the incoming key messages. It handles F2, F3, F4 (broadcast, multicast, and the last one whisper of) shortcuts.

(4)BOOL DeleteIp(LPCTSTR Name)

Delete a user.

(5)CloseChat()

Close chat control, close the chat socket.

File transfer module and text chat achieve essentially the same way, the difference is just a small amount of data which has been different. Users only need to open the file to be transmitted and to be read a certain size of the buffer, and then call the underlying transport module provides a function to send can be. Fig.3-12 shows all the test of project.

Figure 3 - 12 All the test of project

IV. Conclusion

In order to achieve a small-scale multi-point video conferencing system, this thesis propose our scheme as the target and does the research on P2P [13-16] network technology, application-layer multicast, streaming media technology, as well as audio and video processing technology which based on DirectShow framework. Recalling the work, mainly including the following aspects:

(1) Do a more in-depth study of P2P technology theory; include characteristics of P2P technology, P2P classification, and three kinds of P2P-based video conferencing system architecture.

(2) To understand streaming media, we do systematic analysis and research for streaming media development framework DirectShow and design a series of audio video processing scheme based on DirectShow framework.

(3) We research about the existent problem of multicast in depth, analyze and compare the advantages and disadvantages of application-layer multicast with multicast, we put forward subject for small-scale P2P video conferencing to solve main problem.

A complete video conferencing system is a very big project, at present, we have completed the basic design goal, but the system there are a number of key issues need to be resolved in future.

Bibliography

[1] Li Rui, Yu Zhangwu, Wang Hao, Research on Multi-zone and Hierarchical Routing for Video Gatekeeper Based on Agent. IEEE,2005:701-719.

[2] Tan Shi Zhe, Liu Wei Dong. UPnP 基础结构以及行为规范的描述. 组合机床与自动化加工技术，2004，6：38~41

[3] 尹传箭，陈哲，张永林. 基于 IE 浏览器的 IP 组播视频会议系统. 计算机工程与设计，2006,27(19): 3598-3600

[4] T. Dagiuklas, P. Galiotos. Architecture and Design of an enhanced H.323 VoIP Gateway, 2002:1209-1213

[5] MICHAEL S. Probe Based Dynamic Server Selection for Multimedia Quality for Service [J]. PELGarduate, 2004,7(2):76-80.

[6] 雷刚勇，解梅. 基于 DirectShow 的 MPEG-4 远程视频监控系统[J]. 电视技术， 2005，4(5):91-92.

[7] HUA J, YANG S T, LI J H. Method of Video Captures and Encoding based on DirectShow Technology [M]. Shanghai Computer Society, 2004:56-58.

[8] 张勇，罗静. 基于 DirectShow 的多媒体文件音视频的重新压缩[J]. 现代电视技术，2005，6(5):34-39.

[9] CHEN Yue. Design of a Video Capture System of Mobile Robot based on DirectShow [M]. Huanzhong University of Science and Technology, 2004:105-106.

[10] LIU W Y. Network MPEG-4 Stream Processing System based on DirectShow [J]. Atomic Energy Press, 2007, 22(7):84-85.

[11] 朱婧，常会友. 一种基于 UDP 穿越 NAT 方案的设计与实现。计算机工程与应用， 2006，35：152~155.

[12] 张兵. 基于多缓冲区实时音频采集和传输实现. 计算机工程， 2006，32(12): 272~275.

[13] Zhang Jie, Zhang Yao, Huang Tao-ying. An improved selection of clips control strategies in P2P network, [J]. Micro-computer information, 2007. (21).

[14] Wen-Jie Luo, Peer-to-Peer (P2P) summarize [EB/OL], 2008-03-22.

[15] Beverly Yang, Hector Garcia-Molina. Improving Search in Peer-to-Peer Networks [J]. IEEE Computer Society, 2002, 5:1-10.

[16] Li Gong. Guest Editor’s Introduction: Peer-to-Peer Networks in Action [J]. IEEE Internet Computing, 2002, 6(1):37-39.

[17] Jun-Zhou Luo, Bo-Tao LI, Ming Yang, Jun Wu, Jian Huang. TCP/IP Protocol and Network Programming. Tsinghua University Press. 2004, 10.