International Journal of Computer Applications Technology and Research
Volume 12–Issue 03, 09-13, 2023, ISSN:-2319–8656
DOI:10.7753/IJCATR1203.1003

KVM Terminal Software Design Based on RK3399 Processor

Xiaoqiang Xin
School of Communication Engineering
Chengdu University of Information Technology
Chengdu, China

Fengbo Wang
School of Communication Engineering
Chengdu University of Information Technology
Chengdu, China

Qinglin Huang
School of Communication Engineering
Chengdu University of Information Technology
Chengdu, China

Abstract: KVM denotes the initials of the keyboard, monitor and mouse of a computer. digital KVM systems are composed of two parts, one is the device terminal and the other is the software adapted to this device terminal. KVM systems are emerging remote centralized management technologies for multiple devices in multiple locations based on IP networks in recent years [1,2]. digital KVM systems mainly use IP networks to control remote computers to control remote computers mainly through IP networks [3]. Therefore, KVM systems are widely used in computer management and remote control of computers. In practical applications, the compression efficiency of images is relatively low and the delay of mouse movements is relatively large. To address these two shortcomings, this paper designs a software based on the RK3399 processor, which on the one hand performs fast image acquisition and high-quality encoding, and on the other hand realizes input control of the target computer through the mouse and keyboard control chip UI012.

Keywords: Digital KVM; Capture Video;Mouse and Keyboard Control; Video Codec

1. INTRODUCTION

With the rapid technological development of society's information industry, digital network systems based on embedded technology are constantly emerging [4]. With computers and networks distributed in all corners of our cities and servers of large multinational companies all over the world, digital KVM systems make it possible for network administrators not to go back and forth between servers in order to complete the maintenance and management of servers, but to access and centrally manage up to thousands of computers with just one set of I/O devices [5]. Instead, they can remotely access and centrally manage up to thousands of computers through a single set of I/O devices [5]. This can create tremendous benefits for centralized information management in enterprise server rooms or data centers [6], not only reducing energy consumption and saving space in server rooms, but also avoiding the clutter that can be caused by too many keyboards, monitors, and mice, greatly simplifying workflow and improving enterprise productivity. Among the various data center server management solutions and remote control solutions, KVM systems are a better choice because managers can manage multiple servers from a single terminal and break through distance limitations. Even if the computers are distributed in different locations, maintenance personnel do not need to travel around for maintenance and management work, but can access and centrally manage multiple computers remotely with just one set of I/O devices [7]. Traditional KVM systems are affected by the video codec technology, and the quality of the received audio and video is very poor, which is an important reason for the extremely poor user experience. Moreover, due to the limitation of network bandwidth, the problems of data transmission rate and transmission delay are also very prominent, and a series of problems reduce the efficiency of data management. In such a context, this paper uses RK3399, which supports H.265 video encoding, as the core controller to design and implement a KVM terminal software based on the characteristics of KVM terminal devices, aiming to reduce the video transmission bandwidth and overall time delay, improve the video quality, and also support mouse and keyboard control, which has good practical application value.

2. KEY TECHNOLOGIES

In our daily learning life, all kinds of audio and video can be seen everywhere, whether it is the elderly or children, whether it is on the bus or in the restaurant, people are holding their cell phones to use the network to brush a wonderful video. Video has gradually become an important way for us to know and perceive the world. Behind the wonderful video is a very serious problem, which is brought by the storage and transmission of the video. For example, a 1-hour 1080P, 30FPS lossless video requires about 660G of storage space, which obviously exceeds the storage capacity and transmission capacity of our existing media, so it is necessary for us to perform some lossy compression - that is, video codec. The goal is to represent the same video with fewer bits, thus reducing the size of this video and achieving the goal of taking up less space and being easier to transmit.

In 2013, ITU-T and ISO/IEC jointly developed a new generation of video coding standard, which is called High Efficiency Video Coding HEVC in the ISO standard, hereinafter referred to as H.265 [8]. In modern multimedia applications, the obtained video information is compressed efficiently with the help of video coding technology in order to achieve real-time playback, or storage on hard disk [9].

H.265 is an improvement on H.264, the technique of efficient prediction in the structure is retained and the framework used is still the hybrid coding framework. In order to optimize the compression performance of H.265, numerous coding tools such as flexible CTU quadtree image division, inter-frame prediction asymmetric motion division and multi-directional intra-frame prediction patterns are then used. Its newly adopted coding techniques are as follows.

International Journal of Computer Applications Technology and Research
Volume 12–Issue 03, 09-13, 2023, ISSN:-2319–8656
DOI:10.7753/IJCATR1203.1003

(1) image chunking technology. H.264 standard uses macroblock to divide a frame image, the maximum size of macroblock is 16×16. H.265 uses three concepts of coding unit CU (Coding Unit), prediction unit PU (Prediction Unit) and transform unit TU (Transform Unit) to describe the whole coding process. Among them, CU is the basic unit of H.265 coding, and the size of CU can range from 64×64 to 8×8, which can effectively segment images of various resolutions.

(2) Predictive coding technology. H.265 still uses time and space correlation to perform intra-frame or inter-frame coding to eliminate redundant information, but thanks to the use of multi-angle prediction, advanced displacement vector prediction technology, high-precision motion compensation technology and other technologies, the prediction accuracy and coding efficiency are greatly improved.

(3) Transformation and quantization. H.265 reduces the correlation between image blocks by transforming the frequency domain and compresses the prediction residual information to achieve further compression. Both intra-frame prediction and inter-frame prediction, the obtained prediction residuals are transformed by the transform matrix to achieve the concentration of signal energy. The quantization and inverse quantization processes are integrated into the transform matrix operations, and the size of the quantization step (Qstep) depends on the quantization parameter (Qp). When the quantization step is larger, the quantization error is larger and the image quality suffers a more significant degradation, but the compression effect is better. Conversely, the decrease of quantization step size reduces the quantization error and the loss of image quality, but the compression effect is poorer.

(4) The in-loop filter filters the reconstructed image obtained after inverse transform and inverse quantization to remove the block boundary effect and reduce the distortion of the reconstructed image to the original image in the coding process. The loop filtering process of H.265 adds adaptive sampling point compensation after the deblocking filter, and compared with H.264, H.265 increases the minimum block of the deblocking filter from 4×4 to 8×8 to reduce the number of filtering times, and for each 8×8 For each 8×8 grid, the first judgment is whether it is the boundary of TU or PU, if it is not the boundary, no filtering is needed, if it is the boundary, the corresponding boundary intensity is calculated, and the boundary intensity value is 0, 1 or 2. For the luminance component, the boundary with intensity 1 or 2 will be filtered; for the chrominance component, only the boundary with intensity 2 will be filtered. For the luminance boundary with intensity 1 or 2, and the chrominance boundary with intensity 2, it is also necessary for the quantization parameters to meet certain conditions before filtering.

The advantage of H.265 is that it adopts advanced coding and technology, and the compression rate is doubled compared with H.264. The disadvantage is that the complexity of coding computation is increased several times, especially the computation of CTU division method selection, intra-frame prediction and inter-frame prediction mode in pattern decision is greatly increased. Therefore, H.265 video coding puts higher requirements on the hardware implementation platform.

3. DESIGN SOLUTIONS

3.1 Overall software framework

The software design of this paper revolves around the structure of the RK3399 processor, relying on its rich interfaces, running the Linux real-time operating system on its ARM Cortex-A53 and Cortex-A72 [10], managing the entire software system, including the driver support for each device and the various requirements of the business, as follows: configuring the TC358743XBG through I2C working mode through I2C; audio and video acquisition and encoding through MPP; debugging service content through UARTO port and network port connected to the host; controlling UI012 through UART1 port to complete the keyboard and mouse control of the host. The overall framework of the system is shown in Figure 1.

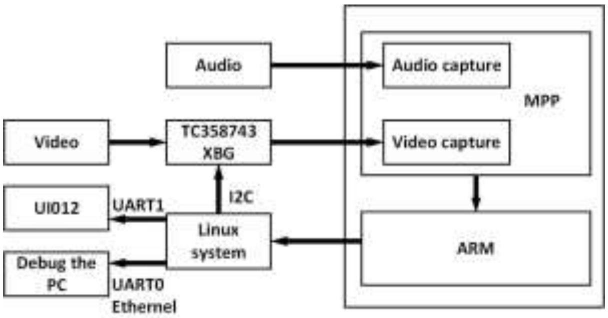

Figure 1. Overall system framework

3.2 Video module design

Rockchip Micro provides a media processing software platform MPP, which shields the application software from the complex underlying chip-related processing, and directly provides multimedia processing interface MPI interface to complete the corresponding module functions. Developers directly through the combination of calls MPI to achieve different system functions, which framework greatly reduces the development cycle, cost and development difficulties, but also makes the system maintenance becomes more convenient.

Before using the MPP development platform for software development, the initialization of the MPP system needs to be completed. After the application completes the MPP service, it also needs to de-initialize the MPP system to release the system resources. the initialization of MPP includes the initial processing of video cache pool, binding system interfaces, etc. MPP provides MPI including video input, video processing, video output, video processing subsystem, video coding, video detection and analysis, area management, audio and video graphics subsystem. The rich multimedia processing interfaces cover all aspects of audio and video multimedia processing for users to develop specific functions.

The video module is designed to perform a series of tasks related to video. The purpose of this module is to implement multi-resolution video acquisition for the KVM market, which can be used to adapt to various resolutions in the KVM market; to perform corresponding processing such as scaling, frame rate control or video encoding quality selection according to user requirements; to feed the acquired video data to the H.265 encoder for encoding and compression of the data according to the MPP programming framework. The captured video data is encoded and compressed according to the MPP programming framework and sent to the H.265 encoder to obtain a high compression rate video stream for

International Journal of Computer Applications Technology and Research
Volume 12–Issue 03, 09-13, 2023, ISSN:-2319–8656
DOI:10.7753/IJCATR1203.1003

network transmission. The video processing flow is designed in accordance with the design purpose of this module.

In order to perform adaptive capture of the current actual video resolution, the TC358743XBG decoder chip needs to be accessed via the I2C interface to get the actual resolution information of the current video, and then the multimedia module is designed and connected to the VI module according to the MPP programming framework, and the video channel and cache pool are initialized according to the current resolution. After initialization, the video image is captured, and it can be cropped, color space conversion and other processing, and output multiple image data of different resolutions; the VPSS module receives the image sent from VI, and can scale the image, frame rate control and other processing, and realize the same source output multiple image data of different resolutions for encoding; the VENC module receives the image data captured by VI and output after processing by VPSS, and supports multi-channel real-time encoding, and each encoding independent.

Video input (VI) unit to achieve the function: through I2C to obtain the current video resolution, and according to the resolution information to complete the initialization work, VI and VPSS unit binding, will be outside the processor video data received through the ITU-R BT1120 interface, directly to the VPSS. In this process, VI can crop, color space conversion and other processing of the received original video image data, and realize the original video image input all the way, and output all the way video image function. The overall design of the video processing flow is shown in Figure 2.

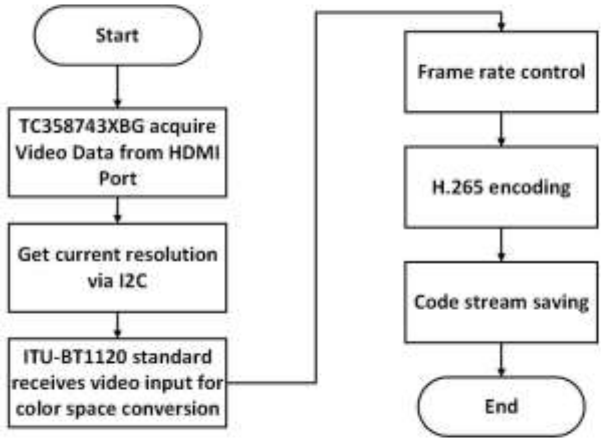

Figure 2. Video processing flow design flowchart

The video processing (VPSS) unit implements the functions:Uniform pre-processing of a pair of input images, such as scaling, frame rate control, etc. The frame rate is controlled in the VPSS by channel CHANNEL, which binds the physical channel and takes the physical channel output as its own input, and when the image data of the physical channel passes through CHANNEL, it can choose to discard some of the data to achieve When the image data of the physical channel passes through CHANNEL, it can choose to discard some data to achieve frame rate control, and it can crop the data in the channel to achieve the effect of cropping the image.

In the online mode of the RK3399, the VPSS supports sending images to the encoding unit for encoding according to the line unit, sending while capturing, which can reduce the delay time generated during the process of processing the complete frame image by the VPSS and then sending it to the encoding unit.

Video encoding (VENC) unit functions: This unit supports multiple video encoding in real time, and each encoding is independent of each other, and the encoding protocol and encoding profile can be different. A typical encoding process includes the reception of the input image, the encoding of the image and the output of the code stream.

The coding specifications supported by RK3399 processor include H.264 coding algorithm supporting BP, MP and HP coding levels, JPEG coding algorithm, MOTION JPEG coding algorithm, and most importantly, H.265 coding algorithm supporting MP coding level, which makes the efficiency of video compression reach the first line level and is the basis for the design of this paper to achieve high quality video transmission with low latency and bandwidth. The basis of compressed video transmission under low latency and bandwidth conditions.

The VENC unit consists of the receiving channel subunit and the encoding channel subunit, and the flow chart of encoded data in VENC is shown in Figure 3. If the input image is larger than the encoding channel size, VENC will reduce the source image to the size of the encoding channel and then encode the image; if the input image is smaller than the encoding channel size, VENC will discard the original image because VENC does not support enlarging the input image for encoding; if the input image is comparable to the encoding channel size VENC accepts the source image directly and encodes it.

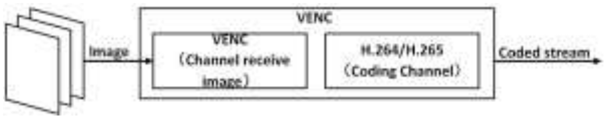

Figure 3. Data flow diagram of VENC

3.3 Web Media Data Interaction Design

The TCP/IP (Transmission Control Protocol/Internet Protocol) protocol is the result of protocol research and development work carried out on the packet-switched network ARPANET (Advanced Research Project Net). Today TCP/IP protocol is widely used by developers and has become the industry standard for communication on the Internet. The four-layer TCP/IP model is used more often in the development process, which mainly includes network layer, transport layer, application layer, and data link layer, and each layer is responsible for a different function.

The process of establishing the server and client structure in this paper is as follows, and the schematic diagram is shown in Figure 4, where the client actively connects to the server to communicate and interact with data.

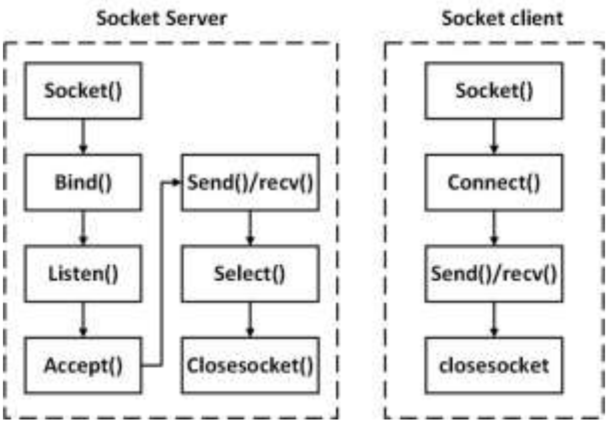

Figure 4. Schematic diagram of the server client

The data sending interaction unit requires the RK3399 system side to send the encoded audio and video data over the TCP network, and after the data reaches the client, a series of processing such as decoding the audio and video data is performed.

In this paper, a thread is created for the data sending interaction unit, and the thread completes a series of operations to obtain video frame data, the process of which is shown in Figure 5.

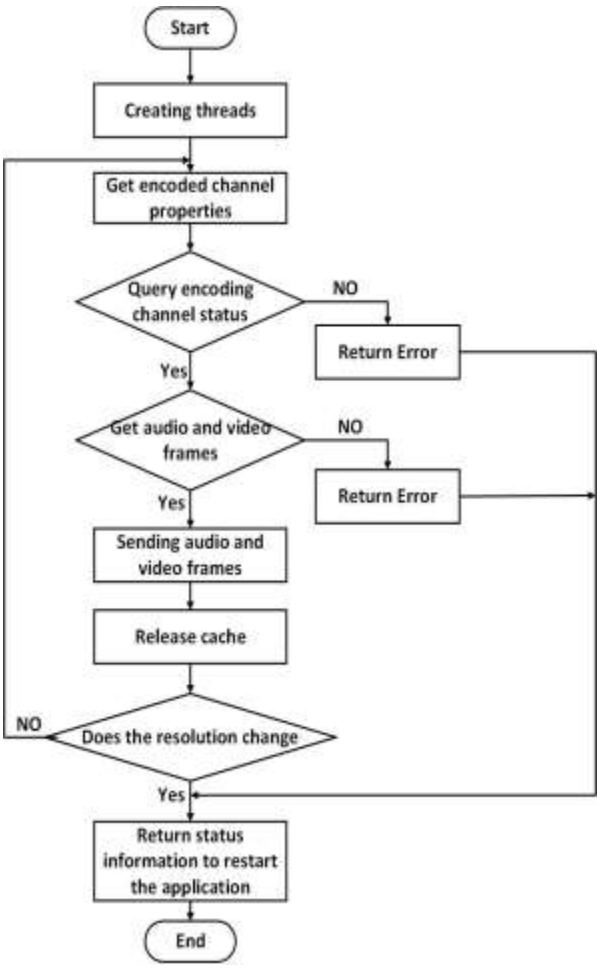

Figure 5. Flow chart of sending data interaction

The data reception interaction unit requires the RK3399 system side to receive the keyboard and mouse data from the client via TCP network and encode the data according to the data interaction format of the UI012 chip.

When a packet is received in the thread, the packet is parsed and if a control signal is received for the keyboard and mouse, the corresponding keyboard and mouse control interaction is executed. In this paper, a thread is created for the data reception interaction unit, in which the socket interface is used to complete the reception of keyboard and mouse signals and control commands sent from the client, and the encoder parameters are adjusted according to the control commands through MPI in this thread, and the flow is shown in Figure 6.

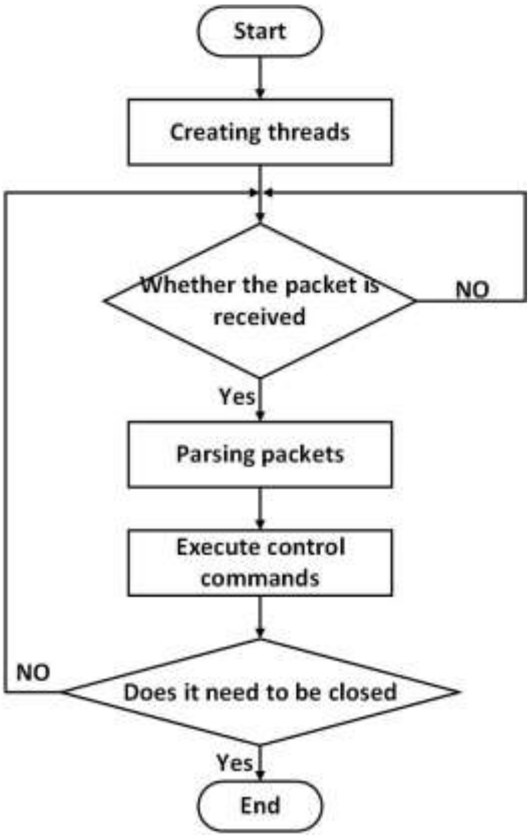

Figure 6. Receiving data interaction flow chart

3.4 Keyboard and mouse interaction design

The keyboard and mouse signal data is received by the remote client through the data reception interaction unit, and after parsing and unpacking, the actual control signal is obtained, and then the RK3399 processor controls the UART1 serial port to output the corresponding data according to the interaction protocol of UI012 to achieve the actual effect of controlling the controlled machine by UI012.

The UART1 pin of RK3399 is a multiplexed pin, and the default initialization is not enabled, so you need to set the pin multiplexing in core.c file, The code of the intercepted fragment is shown in Figure 7 below.

```
#define MUXCTRL_REG_BASE 0x200F0000
#define UART1_RXD_REG
IO_ADDRESS( MUXCTRL_REG_BASE + 0x7c )
#define UART1_TXD_REG
IO_ADDRESS( MUXCTRL_REG_BASE + 0x84 )

writel ( 1 , UART1_RXD_REG ) ;
writel ( 1 , UART1_TXD_REG ) ;
```

Figure 7 Pin reuse fragment code

When the UART1 multiplexing configuration is completed, you can directly open the UART1 device and write the data received from the client by the data reception interaction unit and encapsulated according to the UI012 data protocol.

4. EXPERIMENTAL TESTING

4.1 Video delay test

Place Tables/Figures/Images in text as close to the reference Once the data transmission interaction and video encoding module is designed, the encoded video stream can be sent directly to the remote client for decoding and playback through the data transmission interface. The test method is to use the laptop as the video input source, and use the HDMI interface to copy the video output all the way to 108OP60 to the KVM device, which encodes the video stream at 1920×1080 resolution, 30fps frame rate, and 2Mbps bit rate, and sends the encoded video stream to the client of the master. At the same time a stopwatch timer is run on the display desktop of the laptop, and the laptop (i.e., the controlled machine) is placed together with the monitor of the master PC.

Operations such as Word browsing and video viewing are performed on the controlled PC to increase the complexity of the video image. Under normal operation of the KVM system, a photo is taken of the controlled PC and the local PC screen to calculate the time difference between the two stopwatch software displays. The time difference value is the delay time between the video capture source and the remote client interface display.The video latency test results are shown in Table 1, and l0 latency test results were recorded for statistical analysis. The minimum video latency is 80ms, the maximum is 101ms, and the average latency is 89ms, which is the leading level of KVM terminal equipment in terms of latency and can bring users a good experience.

Table 1. Video delay test results

Number of times	Extension time
1	92ms
2	84ms
3	80ms
4	97ms
5	87ms
6	101ms
7	83ms
8	93ms
Average value	89ms

4.2 Keyboard and mouse delay test

When the keypad and mouse delay test, the master PC prints the current system time when it sends out the control keyboard and mouse signals, and runs an auxiliary software for testing the keypad and mouse delay on the controlled PC server. When the controlled PC receives the keyboard signal or mouse signal, it sends UDP packets to the master PC. After receiving the UDP packet, the master PC prints the current system time, then the difference between the two system printing times is the keyboard and mouse signal delay time.Keystroke delay test results are shown in Table 2.

Table 2. Keystroke delay test results

Number of times	Extension time
1	15ms
2	17ms
3	15ms
4	15ms
5	14ms
6	15ms
7	14ms
8	15ms
Average value	15ms

5. CONCLUSIONS

This paper designs and implements KVM terminal software based on the RK3399 processor in order to achieve low latency, high quality KVM video encoding transmission and keyboard and mouse control with low bandwidth. Under the MPP framework of the RK3399, the audio and video capture and encoding modules are designed, and each module in the MPP framework is connected together according to the multimedia processing interface MPI to complete the overall audio and video capture and encoding work. Then the control of the mouse and keyboard is completed.

6. REFERENCES

[1] Huang YF. IP network multimedia communication technology [M]. Beijing: People's Post and Telecommunications Publishing House, 2003.Ding, W. and Marchionini, G. 1997 A Study on Video Browsing Strategies. Technical Report. University of Maryland at College Park.

[2] Wang Da. Computer network remote control [M]. Beijing: Tsinghua University Press, 2003.Tavel, P. 2007 Modeling and Simulation Design. AK Peters Ltd.

[3] Bovberg J. KVM Over IP [J]. Windows It Pro, 2011,15(2):1322-1323.

[4] Zhang P. Software design of embedded network multi-channel audio and video server[D]. Jilin University, 2007.

[5] Thomas C L, Anderson R L, Gilgen R L, et al. Network based KVM SWIening systen. uo,US7555567 [P]. 2009.

[6] Hao Liu J. Design and implementation of a virtualized server management system [D]. Xi'an University of Electronic Science and Technology, 2011.

[7] Thomas C L. Anderson R L,Gilgen R L, et al. Network based KVM switching system: Us,US7555567[P]. 2009.

[8] Jaja E, Omar Z,Rahman A H A, et al. Efficient Motion Estimation Algorithms for HEVC/H.265 Video Coding [M].Berlin: Springer Berlin Heidelberg, 2015: 287-294.

[9] He Junfeng. Implementation and application of motion detection technology in digital surveillance [J. China Security, 2004,(4):47-49.