ISSN: 2414-1895 DOI: 10.6919/ICJE.202207_8(7).0074

Video Transmission and Detection System based on Software Definde Radio and Object Detection

Zhouyun Wei, Feng Liu*

School of Information Engineering, Shanghai Maritime University, Shanghai 201306, China

Abstract

This paper implements real-time video acquisition transmission and detection based on software radio technology and target detection algorithms. The objective is to transmit the video captured by the camera wirelessly using software radio technology, while providing target detection processing for the received video.The system uses the GStreamer framework for video capture, encoding and multiplexing. The multiplexed TS stream is sent to GNU Radio's video transmission system and matched with a USRP device to send the TS stream to another receiving device. Finally, object detection processing is performed on the video at the receiving end to identify the type of object in the video.Object detection utilizes an improved algorithm of YOLOv4. The improved YOLOv4 achieves a 71.34% increase in detection speed and a 38.65% decrease in the number of network parameters with only a 0.27% decrease in accuracy.

Keywords

Gstreamer; Software Defined Radio; Video Ranssmission; Object Detection.

1. Introduction

With the ubiquitous deployment of wireless imaging devices in various IoT(Internet of Things) [1] applications such as intelligent video surveillance, intelligent transportation, and telemedicine, there is a high demand for efficient transmission of massive videos between different IoT entities. Depending on the use of video transmission to transmit information, higher requirements are gradually put forward for video transmission technology.

The core idea of software radio is to load different communication software on a common hardware platform to achieve the purpose of different communication methods. This design idea enables almost all communication methods to find their corresponding stations. Compared with the traditional radio station, the biggest advantage of the existing software radio is that it does not only take hardware as the core, it combines software and hardware to make it have good scalability, and at the same time, the functions that can be realized change. be more diverse. The overall structure of the transceiver of the software radio is shown in Figure 1. Its structure can be divided into four parts [2], including the communication module, that is, the software radio platform (SDR) for communication; the USRP driver module: UHD, which is used to drive the software radio external hardware devices; operating system module: including system library, system interface and kernel; and external hardware device USRP: USRP includes receiving control module, digital up/down converter (DUC/DDC), digital-to-analog converter (DAC) , analog-to-digital converters (ADCs), filters, power amplifiers. The software part transmits the data to the external device USRP of the software radio through the USB interface. The USRP will convert the digital signal into an analog signal through the internal device, and finally send the radio frequency signal through the antenna.

In recent years, with the continuous development of the field of computer vision, people have begun to process video image information, and have begun to let machines replace human eyes to complete some things, combining video transmission with real-time processing. For example, with the development of video surveillance system, traditional digital video surveillance can only collect, store, transmit, display and other functions of video [3]. The video content is judged only by the human eye, and the workload of video analysis is huge, which is a great consumption for human resources. However, at this time, combining computer vision technology with transmission can collect the received video for monitoring. Perform intelligent processing and replace manpower with machines, greatly reducing the workload for people, and at the same time, it can also complete some functions that people can't complete. The camera is used instead of the human eye to collect video information, and the video is transmitted to the technology center through wireless communication technology. The technology center uses the technology in the related computer vision field to provide real-time target tracking and detection for the transmitted video. This recognition technology for moving objects has far-reaching significance for practical life.

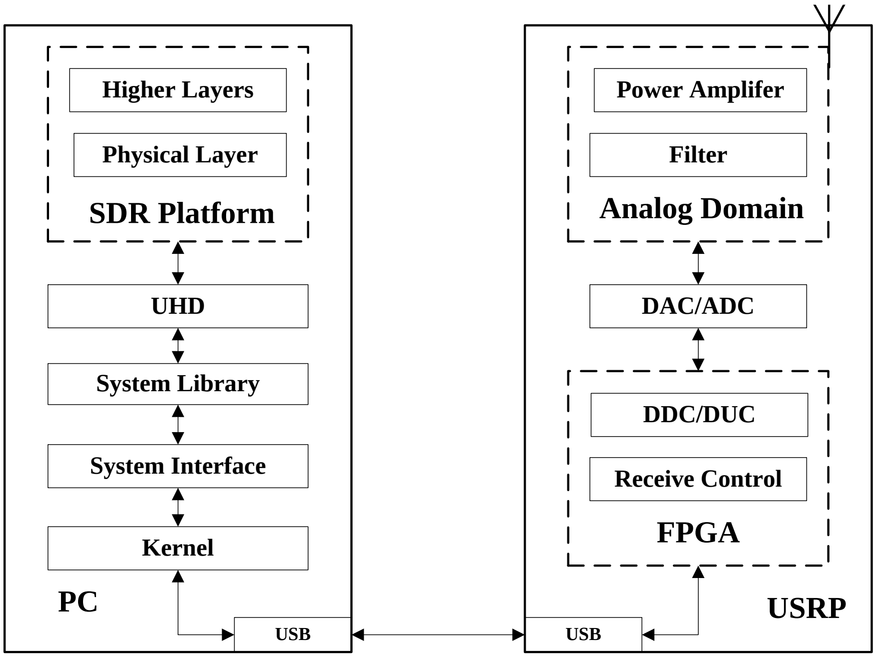


Figure 1. The overall structure of software radio transceiver [4]

2. System Design

2.1 The Overall Framework of the System

2.1.1 Sub-section Headings

The video transmission and detection system designed in this paper is mainly based on GStreamer framework, software radio technology and YOLOv4 [5] object detection algorithm. The overall framework of the system is shown in Figure 2. First, the Gstreamer [6] framework is used to build a pipeline to call the camera to collect the video, and a series of processing such as compression and multiplexing are performed on the video. Secondly, the video sending module and the video receiving module are constructed in the software platform GNU Radio[7] of software radio based on the DVB-T [8] standard. At the same time, it is equipped with a software radio external device (USRP) to complete the wireless transmission of video between the two devices. Finally, the YOLOv4 improved algorithm is used to detect the target of the video at the receiving end. The pipeline file (mkfifo videotx.ts) [9] is used to build a pipeline between video collection, transmission and detection, and streaming media technology is used to achieve the integration of collection, transmission and detection. Table 1 is the software and hardware configuration of the system, and Table 2 is the relevant parameters during system transmission.

 Volume 8 Issue 7, 2022
DOI: 10.6919/ICJE.202207_8(7).0074

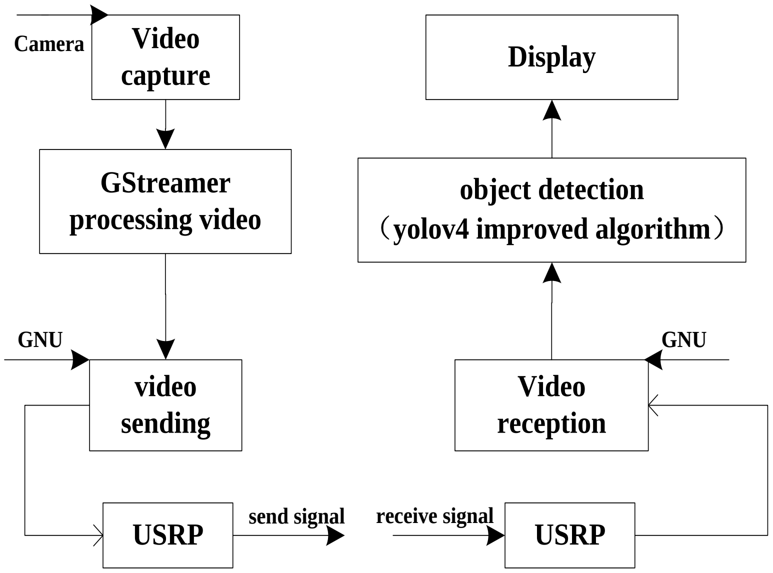


Figure 2. The overall structure of software radio transceiver

Table 1. System software and hardware platform configuration

Lab Environment	Name
Server	GTX 1080 Ti
PC operating system	Ubuntu 18.04
Basic development environment	GNU Radio3.8.2, Gstreamer 1.0, anaconda
Software Radio External Device	USRP B210
External camera	HIK 2K Camera
Detection side development framework	PyTorch

Table 2. Video transmission related parameter

Parameter	Value
Environment	Indoor
Center frequency	700MHz
Channel bandwidth	7MHz
Constellation map	16QAM
Transmission Mode	8K
Video pixel	640*480

2.2 Video Capture

The video capture terminal is mainly designed using the GStreamer framework. GStreamer is an open source framework for developing streaming media applications [10]. It can realize one-stop media solutions such as capture, encoding, decoding, and rendering. Its architecture is mainly Based on plugins and pipelines, the functional modules in the framework are implemented as components that can be connected, and they can easily implement various pipelines, so as to use various existing plugins to "assemble" a Full-featured streaming app. The overall design flow chart of the video capture terminal is shown in Figure 3.

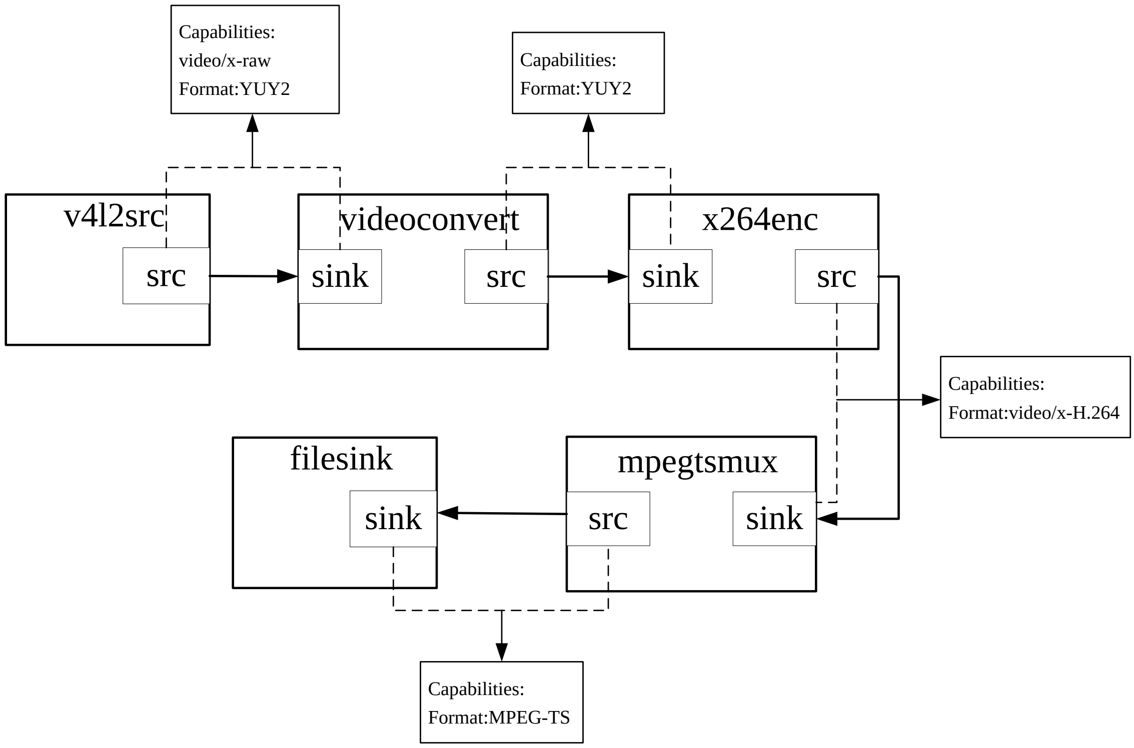


Figure 3. Design flow chart of video capture terminal

The main functions of each component are as follows:

V4l2src: Drive the camera [11] to collect video images according to a certain video format, resolution and frame rate.

Videoconvert: Ensure that the data between the various components in the pipeline are compatible with each other, so that the pipeline can be connected correctly. Automatically convert the original video source into a format that x264enc can understand, and then send the video data to x264enc.

X264enc: Encode raw video data into H.264 [12] compressed data.

Mpegtsmux: Encapsulate video data into MPEG-TS container to multiplex multiple video streams into MPEG transport stream.

Filesink: Save the received data to a certain path.

2.3 Video Transmission

The video transmission system used in this paper is based on the existing modules in the GNU Radio software platform, combined with the USRP B210 hardware equipment to build a DVB-T system.

GNU Radio is a free and open source framework for building software-defined radios, which can be combined with hardware and software to define the emission and reception of radio waves to build a radio communication system. GNU Radio is jointly developed by Python and C++. The C++ language is used to write the underlying signal processing modules. These modules are called "blocks". For example, codec modules, source modules, etc. belong to "blocks". Python is used to write the top-level module in GNU Radio, and a tool: SWIG (Simplified Wrapper Interface Generator) is used to connect the bottom-level module and the top-level module. SWIG can provide a code interface to C++, so that the Python language can connect various data processing modules into a communication system, which is generally called a flow graph [13].

USRP is a general peripheral of software radio, which is used to build SDR system together with GNU Radio. The model used in this article is USRP B210, which is an integrated, single-board,

ISSN: 2414-1895 DOI: 10.6919/ICJE.202207_8(7).0074

general-purpose software radio peripheral that covers the 70MHz-6GHz frequency band and can be used for low-cost experiments [14].

The GRC flow diagrams built in the transmission part of the system in this paper are shown in Figure 4 and Figure 5 respectively. Figure 4 is the GRC flow diagram of the transmitter, and Figure 5 is the GRC flow diagram of the receiver.

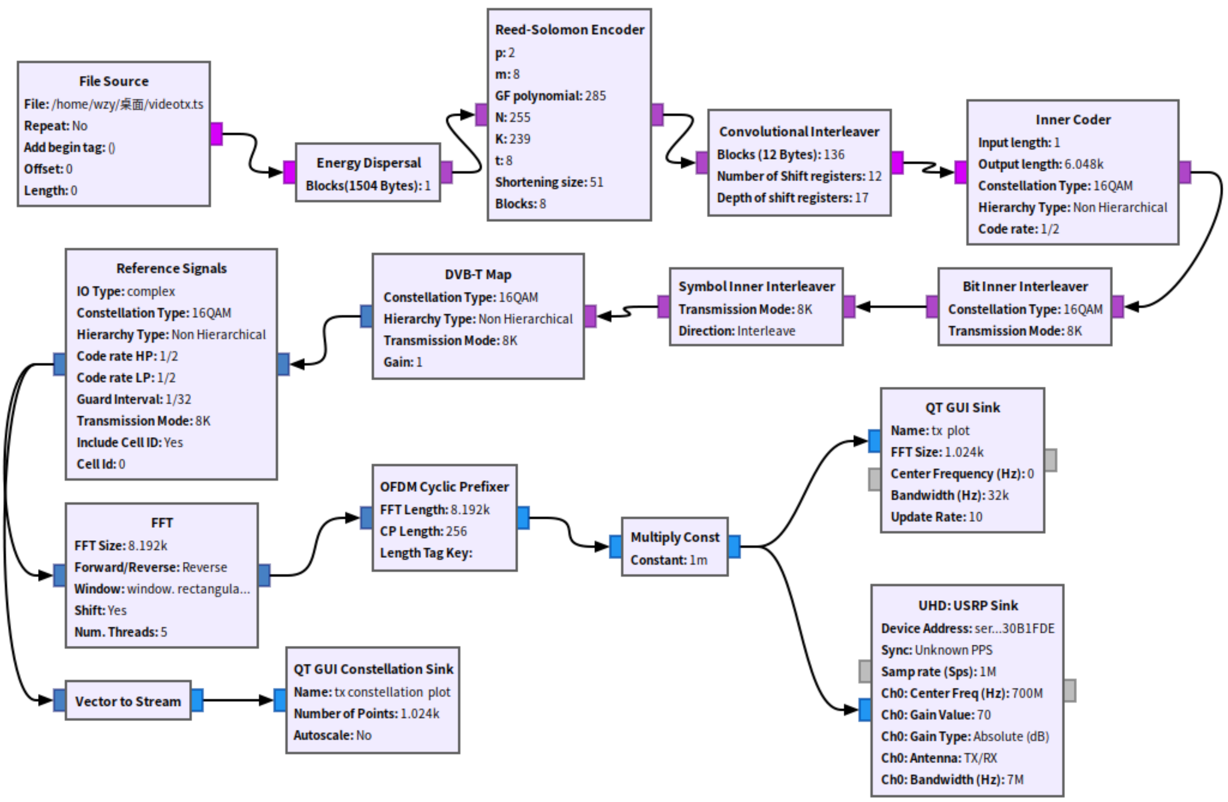


Figure 4. Transmitter GRC flow diagram

The functions of the main modules in the transmitter GRC flow diagram are as follows:

File Source: This module is mainly used to read in a file and then output it in a certain data form. Put the pipeline file with the MPEG-TS video stream in it, and input the TS stream for transmission.

Energy Dispersal: The function of this module is to add pseudo-random code to the input TS video stream, so that the probability of "1" and "0" appearing in the data stream is basically equal to prevent uneven energy distribution [15]. Through energy dispersion, it is beneficial for the receiving end to restore video stream information.

Reed-solomon Encoder: The system uses RS coding to perform outer coding on the code stream, which is used to correct some random byte errors and burst byte errors in the transmission process.

Convolutional Interleaver: It is used to prevent the burst error of the system. When there is an error that cannot be corrected by the RS encoding, through convolutional interleaving, some consecutive byte errors can be eliminated, so that the receiving end can obtain a code stream with uniform error distribution during decoding, and improve the error correction ability.

Inner Coder: It is used to solve the problem of burst bit errors during transmission. There are 5 selectable error correction code rates for inner coding: 1/2, 2/3, 3/4, 5/6 and 7/8. The higher the coding rate, the stronger the error correction ability, but the more bandwidth is used, so the appropriate code rate can be selected according to the current channel environment.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202207_8(7).0074

Bit/symbol Inner Interleaver: The bit interleaver and the symbol interleaver form the function of inner interleaving, which mainly solves the continuous bit errors that cannot be solved by the inner coding, and then maps the bit stream to the OFDM symbol.

DVB-T Map: Map the data output by the symbol interleaver into a complex number z according to a specific mapping method. There are three options: 16QAM, 64QAM, QPSK.

Reference Signals: Reference signal generator, used to generate pilot and transmission parameter signaling (TPS) to form transmission frames, paving the way for OFDM modulation.

FFT: The FFT module is used for IFFT operation to realize OFDM modulation. In the FFT module, it can be set whether to perform IFFT operation or FFT operation. If FFT operation is set, it is OFDM demodulation.

OFDM Cyclic Prefixer: In order to overcome the interference between symbols and the interference of reflected waves, a guard interval will be added to the transmission interval of the OFDM symbol. In the DVB-T system, a cyclic prefix is used to realize the guard interval.

UHD: USRP Sink: Through this module, the data stream will be transmitted to the software radio peripheral device USRP.

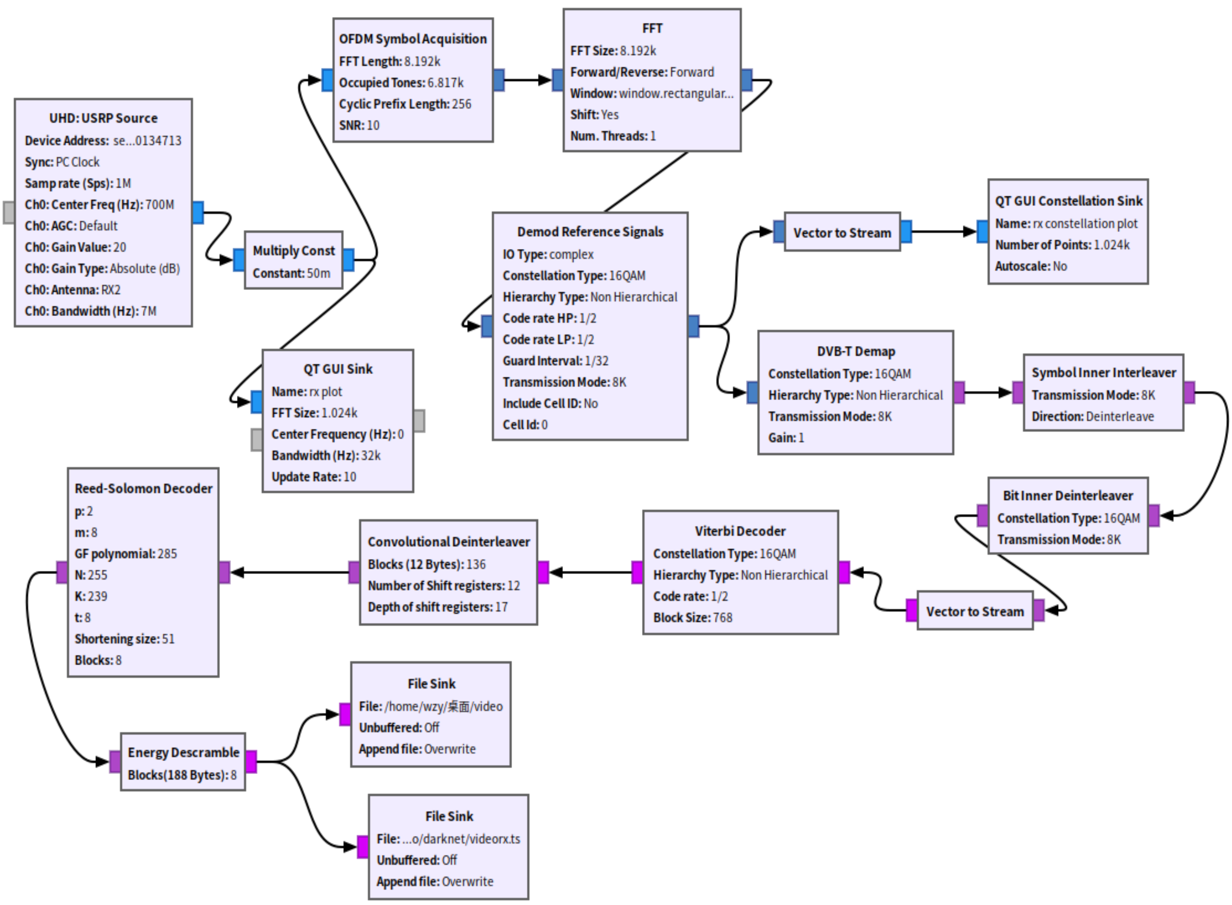


Figure 5. Receiver GRC flow diagram

The GRC flow diagram of the receiver mainly includes functions such as channel demodulation and channel decoding. First, the USRP Source module is mainly used to obtain data flow information from the hardware device USRP for subsequent processing. Then, the OFDM Symbol Acquision module realizes timing synchronization and carrier synchronization functions, and removes the cyclic prefix to obtain OFDM symbols; the FFT module is set to FFT operation for OFDM demodulation. The Demod Reference Signals module extracts the pilot signal and transmission parameter signaling in the transmission frame. The subsequent modules will perform operations such as demapping, decoding, and de-interleaving of the data information, and finally restore the original video information and save it to the File Sink. Two File Sink modules are set up, one is used to receive and save the video stream for individual playback or to analyze its performance; the other is placed as a pipeline file (mkfifo videorx.ts) to connect transmission and detection, pipeline The video stream in the file cannot be saved. After the system runs, the file is invalid, so another File Sink is needed to save the video.

2.4 Video Detection

In the video detection side, the target detection algorithm is mainly used to identify the content in the video. The algorithm used in this paper is an improved algorithm of YOLOv4. Since the system in this paper has high real-time requirements, the main purpose of the algorithm improvement is to increase the detection speed (FPS) of the model without reducing the accuracy.

2.4.1 Date Set

The data set used in target detection directly affects the results of the experiment. When training the neural network, it is necessary to select an appropriate data set. Commonly used data sets for object detection include VOC, COCO, ImageNet [16], etc. However, in order to be more in line with the exper-imental environment, we abandon the use of public data sets, use the data sets collected by ourselv-es and use the LabelImg tool for image labeling, and the labeling method is the target frame. The si-ze of each image is fixed at 416*416, in which case it is possible to increase FLOPS with minimal loss of accuracy. Wrap the category you want to detect inside the box. This dataset is mainly divid-ed into the following categories: person, computer, chair, phone, cup. The scene of the dataset is m-ainly a laboratory scene. On the one hand, the environment in the laboratory is relatively complex, with many computers, people, tables and chairs, etc., which meets the needs of the detection envir-onment. On the other hand, the video transmission is also in the laboratory scene. Video, which can correspond to the detection part. When training with your own labeled dataset, the YOLO format is used, that is, one image corresponds to one label file, and the format of the label file is xml. During the training process, the xml file will be implemented in txt format through code, which contains the corresponding image. Each target box on , where the value represents the category and the target b-ox location. During the training process, the data set is divided into training set and test set, 80% of the data set is used as the training set and 20% is used as the test set.

2.4.2 YOLOv4 Improved Algorithm Design

The convolutional layer and the fully connected layer in the neural network involve a lot of parameters. Although the backbone network of YOLOv4 uses CSPDarknet53 to effectively extract and extract feature information and ensure a certain accuracy, the amount of parameters is too large. Problem, this paper chooses to use the lightweight network GhostNet [18] to replace the backbone network of YOLOv4.

GhostNet is a brand-new lightweight network proposed by Huawei researchers, which surpasses the current state-of-the-art methods in terms of speed and calculation, and can also guarantee comparable accuracy. GhostNet is based on a new neural network model proposed by the Ghost module, and it is different from traditional convolution. It first uses normal convolution to calculate some feature maps with fewer channels. Then use a small amount of computation to get more feature maps, and then combine these results together to form a new feature map. This method can greatly reduce the amount of calculation, and the plug-and-play feature of the Ghost module is also very convenient. Therefore, this paper uses the GhostNet network to replace the backbone network CSPDarknet53 of YOLOv4.

The improved network structure is shown in Figure 6. First, the input 416*416 image is extracted through the backbone network GhostNet, and the features of three scales are output, which are 52*52, 26*26, and 13*13 in turn. Then the 13*13 features are added to the receptive field through the SPP module, and after output, they are input into PANet together with the other two feature layers for the fusion of detail features and deep features, which will undergo upsampling, downsampling,

ISSN: 2414-1895 DOI: 10.6919/ICJE.202207_8(7).0074

convolution and Stitching operation. Finally, the feature layer after feature fusion is sent to the detection network, and three targets of different scales are predicted respectively, and the prediction results are output.

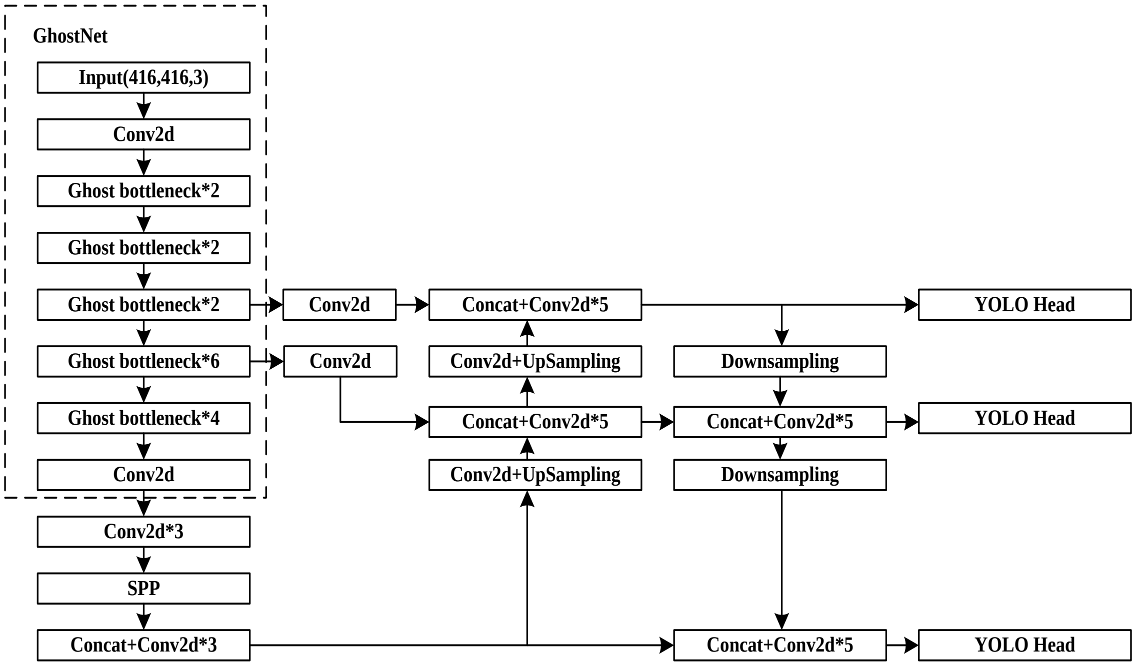


Figure 6. YOLOv4 improved algorithm structure diagram

The idea of transfer learning [19] is used during the experiment, that is, instead of training the backbone network with its own dataset, use the backbone network that has been pre-trained on ImageNet. In addition, in order to make the model converge faster, the warmup strategy is used[20], and a small learning rate is set at the beginning, which can make the model gradually stabilize, and then use the preset learning rate for training, which not only can The convergence speed is faster, and the trained model works better. The training in the experiment can be divided into two stages: the freezing stage and the thawing stage. The freezing stage refers to freezing the backbone network for training. At this time, the weights pre-trained by GhostNet on imagenet will be loaded, the rest of the network will be trained, and a small value will be set. The learning rate is 0.0001; after that, the backbone network will be released during the thawing phase and continue training. At this time, the warmup learning rate is set to 0.00005. The parameter quantity, average precision (MAP) and detection speed (FPS) of the final model are shown in Table 3.

Table 3. The values of various indicators of the detection model

Index	YOLOv4 model	YOLOv4 improved model
Parameter quantity	64040001	39285534
MAP	97.57%	97.28%
FPS	28.6659	49.1159

Experiments show that this improvement is effective and can greatly improve the speed of the network without affecting the accuracy. It can be seen from the experimental data that the improved accuracy has only dropped by 0.27%, which is within the acceptable range, but the detection speed has increased by 71.34%, while the amount of network parameters has dropped by 38.65%.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202207_8(7).0074

3. System Performance Analysis

After the experimental deployment is completed, the GNU Radio on the receiving terminal is usually run for a while, and it is allowed to receive for a while, so as to avoid the loss of video data on the receiving terminal. Then the GNU Radio on the transmitter side will run, and although the transmission part has started running, there is no data flow in it, so the oscilloscope in the system will not have a waveform. After that, the sending end Gstreamer will be run to start collecting video. When the constellation diagram appears on the receiving end, the Gstreamer code of the receiving end will be run to decode and play and display. Finally, run the detection code to observe its target detection. Since the evaluation of the detection model has been analyzed in Chapter 2, this chapter will first display the various video result graphs of transmission and detection, and then analyze the constellation diagram during the transmission process, the image quality of the video, and the data packets in the TS stream. Analyze the experimental results.

3.1 Experimental Result

First, the system will fix the distance between the two USRPs at 0.5 meters. Since the absolute maximum value (dB) of the USRP gain is unknown, the normalized value is used to represent the gain of the system. Set the transmitter gain to 1, which maps to the maximum gain value of the USRP, and set the receiver gain to 0.2. The FFT spectrum diagram of the transmitter and receiver of the system is shown in Figure 7, and the constellation diagram of the transmitter and receiver is shown in Figure 8.

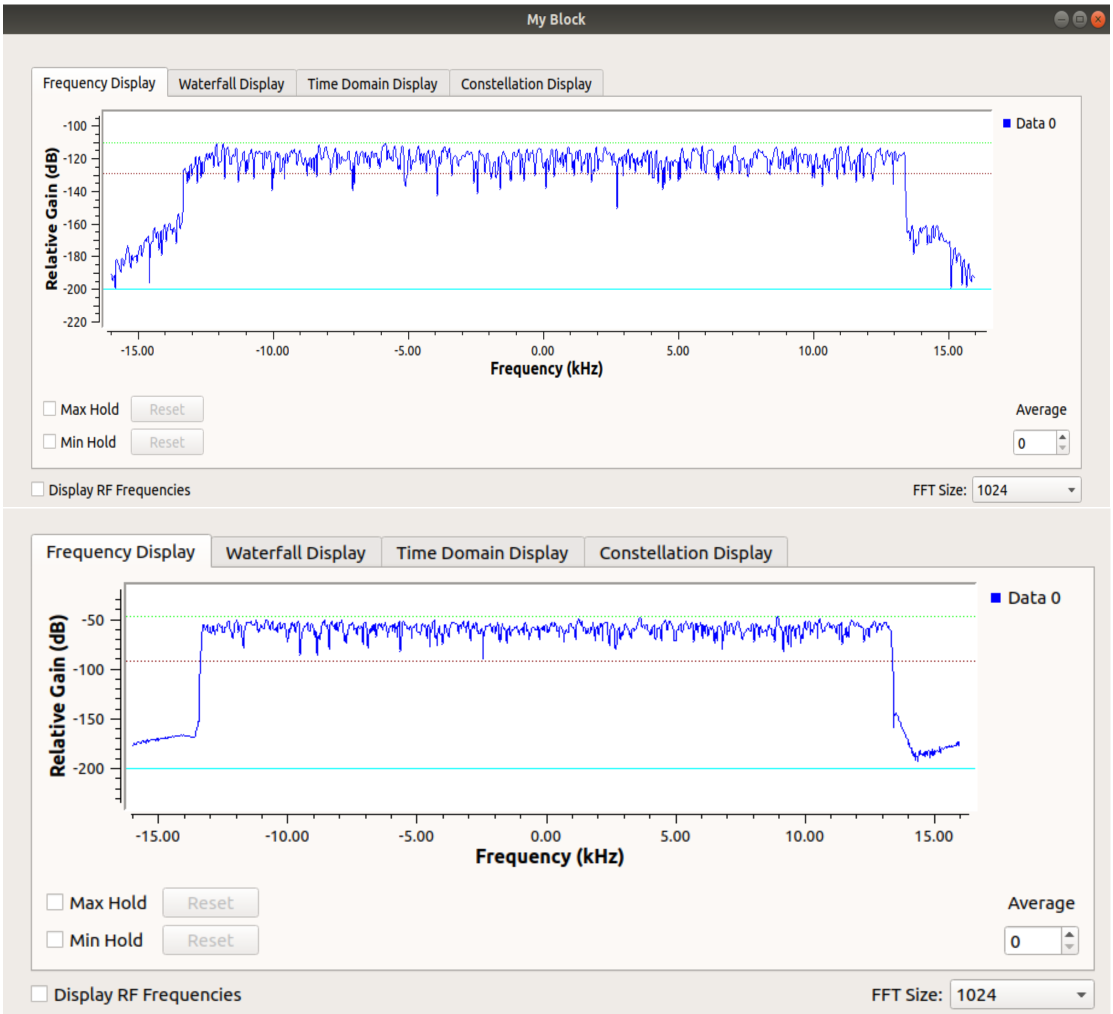


Figure 7. FFT spectrogram of the receiver (top). FFT spectrogram of the lower transmitter (bottom).

Volume 8 Issue 7, 2022
DOI: 10.6919/ICJE.202207_8(7).0074

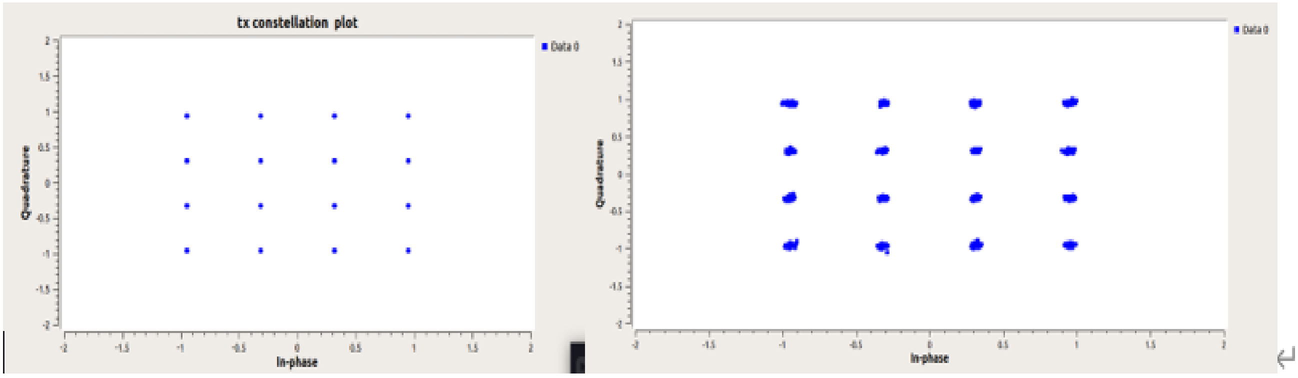


Figure 8. Constellation plot. The left picture is the constellation plot of the transmitter, and the right picture is the constellation plot of the receiver.

According to Figure 7, it can be seen that the FFT spectrogram of the transmitting end and the receiving end are similar, which means that the receiving end can receive data, but the spectrogram of the receiving end oscillates strongly, because there is noise in the actual channel. According to the constellation diagram, it can be shown that the receiving end successfully receives the video stream. According to the shape of the constellation diagram, it can be seen that the constellation mapping mode is 16QAM. At the same time, the constellation diagram of the receiving end can indicate the size of the bit error rate. If the constellation diagram is too scattered, the bit error If the rate is too high, the received video quality will be poor.

Figure 9 is the transmission diagram of the video at the experimental site. A stopwatch timer is used on the display, and the stopwatch value is at 6.57 seconds to shoot, and observe the delay of sending and receiving the video. Although there may be a little delay in the video, but Observed from the human eye's point of view, it does not affect the human eye's reception of video information.

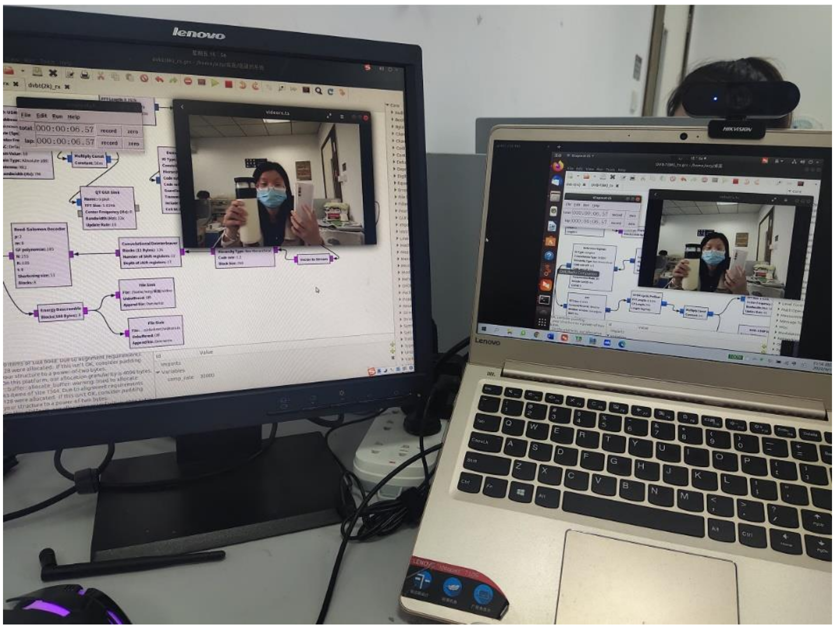


Figure 9. Experimental site transmission diagram

Figure 10 is a screenshot of the video of the transmitter and receiver. It can be found that the receiver can receive the video normally, and the video resolution is acceptable. Figure 11 is the target detection

ISSN: 2414-1895 DOI: 10.6919/ICJE.202207_8(7).0074

diagram obtained by using the improved yolov4 algorithm. It can be seen that people, computers, cups, etc. are correctly identified, and even some small targets can be correctly detected.

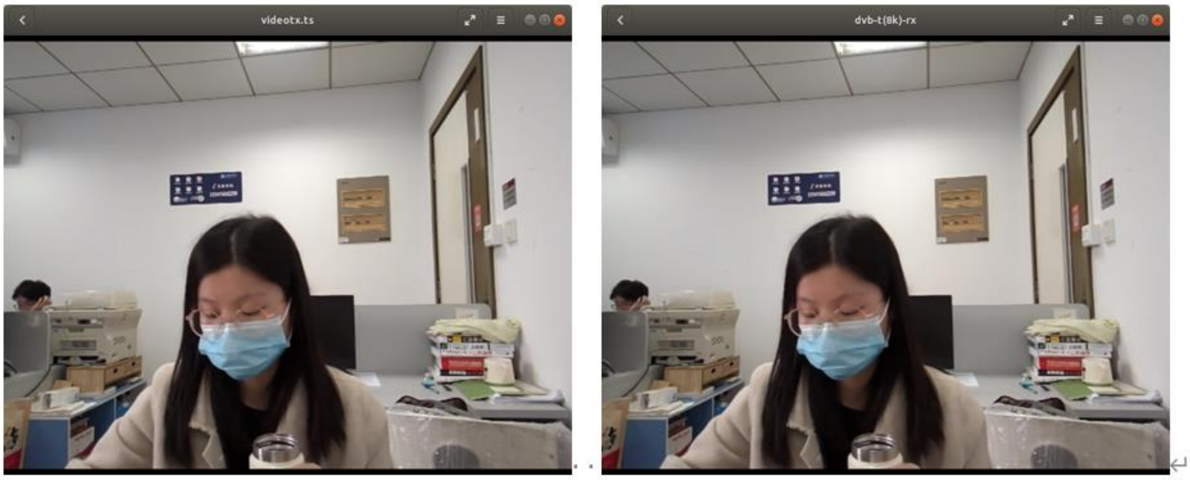


Figure 10. Video screenshots during transfer. The picture on the left is a screenshot of the video from the transmitter, and the picture on the right is a screenshot of the video from the receiver.

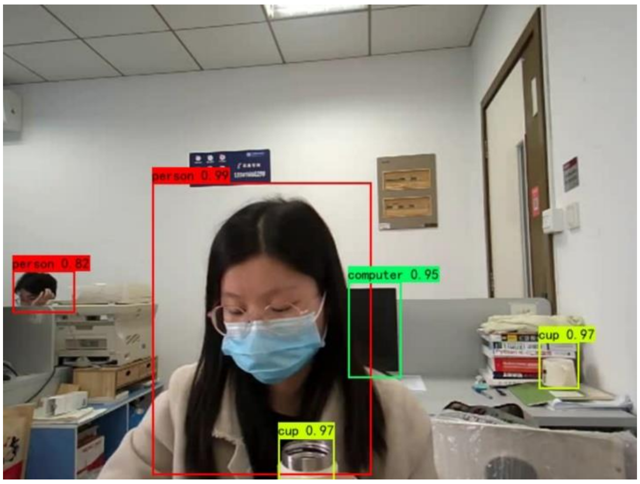


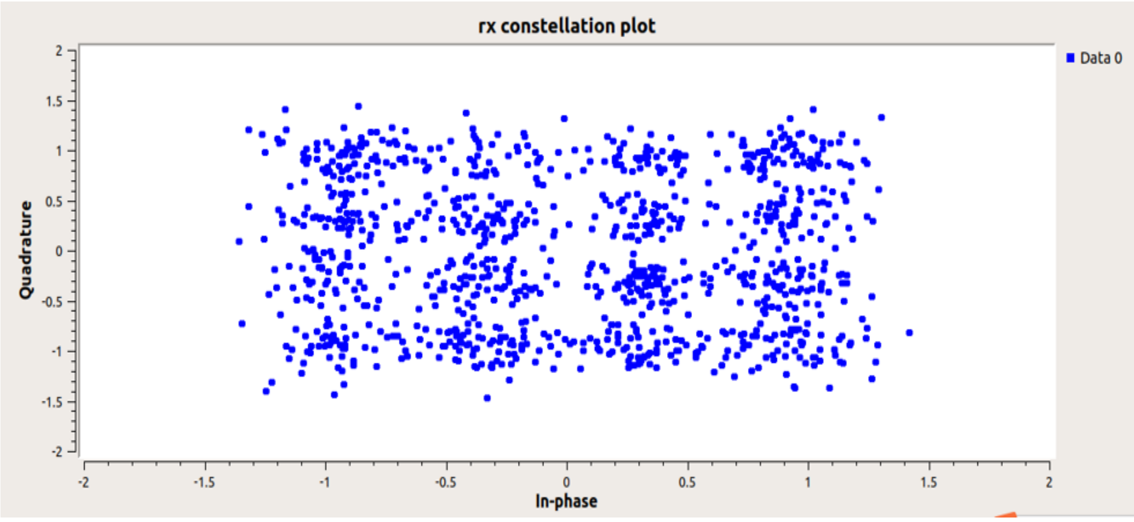
Figure 11. Video object detection diagram

3.2 Video Transmission Performance Analysis.

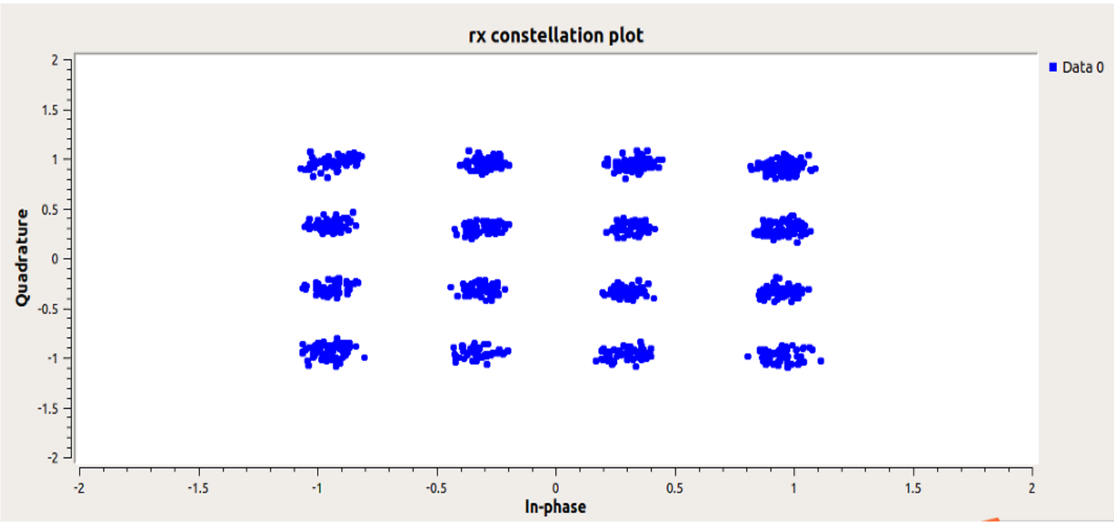
In this paper, by changing the antenna gain of the transmitter and receiver, the constellation diagram of the receiver, the video image quality and the packet loss rate of the TS data packets in the video transmission process are analyzed. When running GNU Radio, the receiver constellation diagram will be displayed directly by the QT GUI Constellation Sink. Regarding the video image quality, the TS data will be first transcoded into YUV format files, so as to analyze the SSIM (Structural Similarity, structural similarity) and PSNR (Peak Signal-to-Noise Ratio) between the received video and the transmitted video through the MSU software. peak signal-to-noise ratio). Since the MSU will calculate the value of each frame, we analyze the video quality by its average value (MSSIM, MPSNR), and the TS packet is analyzed by using TS Expert.

The transmission modes and parameters remain unchanged. The video is collected for about five minutes, and the receiving gain of the receiving end is controlled to be 0.2, and the gain of the transmitting end is changed, that is, the signal-to-noise ratio is changed, so as to observe the receiving situation of the receiving end. Figure 12 is the constellation diagram of the receiving end under different receiving gains, Table 4 is the average SSIM and PSNR values obtained by using the MSU video quality detection software, and Table 5 is the packet loss rate of the TS packet.

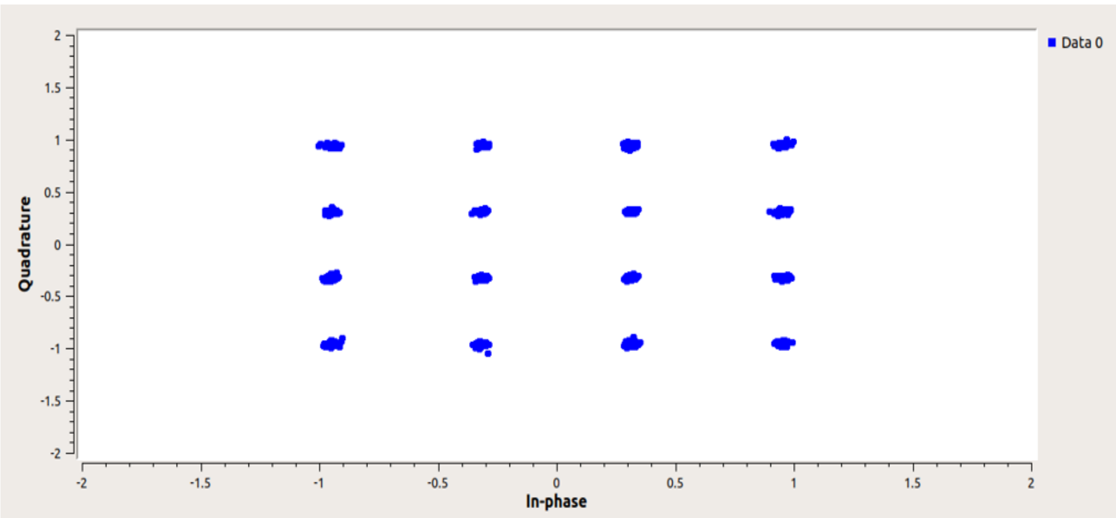
 Volume 8 Issue 7, 2022
DOI: 10.6919/ICJE.202207_8(7).0074



(a) The transmit gain is 0.2



(b) The transmit gain is 0.6



(c) The transmit gain is 1

Figure 12. Constellation plot of receiver under different transmit gains

 Volume 8 Issue 7, 2022
DOI: 10.6919/ICJE.202207_8(7).0074

Table 4. Video image quality under different transmit gains

Transmit gain	MSSIM	MPSNR
0.2	0.88472	37.35696
0.6	0.92337	41.27569
1	0.98541	54.15836

Table 5. Packet loss rate of TS stream under different transmit gains

Transmit gain	Number of TS packets at the transmitter	Number of TS packets at the receiver	The number of lost TS packets	Packet loss rate
0.2	45385	45385	0	0
0.6	45385	45385	0	0
1	45385	45385	0	0

According to the above chart, when the control transmission distance is 0.5 meters and the receiving gain is 0.2, as the transmit gain increases, that is, the signal-to-noise ratio gradually increases, the constellation diagram at the receiving end is closer to the center point, and the smaller the bit error rate, the better the receiving end. The video effect will be better; and when the transmit gain is smaller, or even smaller than the receive gain, it may not be possible to judge the way of constellation mapping according to the constellation diagram.

When analyzing the video image quality, SSIM is generally between 0 and 1. The closer it is to 1, the closer the received image quality is to the original image. If the PSNR is lower than 30dB, the received image quality is poor and there is strong distortion. Greater than 40dB, the received image quality is very good. Therefore, according to Table 4, as the transmit gain increases, the structural similarity and the peak signal-to-noise ratio increase. When the transmit gain is at the maximum, the video received image is very similar to the original image.

According to the packet loss rate in Table 5, it is found that the transmit gain has no effect on the number of received TS packets. As long as the receiving end generates a constellation diagram, regardless of the bit error rate, the receiving end can completely receive TS packets without packet loss. Therefore, the transmit gain only affects the video reception quality.

When the transmite gain is guaranteed to remain unchanged and only the receive gain is changed, the conclusion is the same as above. When the receiving gain is reduced, that is, the signal-to-noise ratio gradually increases. At this time, the closer the constellation map is to the center point, the more obvious the constellation mapping, the better the received video quality, and no packet loss occurs.

4. Conclusion

Based on software radio technology and target detection algorithm, this paper constructs and implements a video transmission and detection system. Firstly, the overall framework of the system is introduced, and then the design of video capture, video transmission and video detection is described in detail, and the pipeline construction method of video capture, the GRC flow diagram of the transmitter and receiver in GNU Radio, and the detection end are given respectively. The method of the model. Finally, the camera, two PCs and two sets of USRP equipment are used to build the transmitter and receiver respectively in the actual environment. At the receiving end, use GStreamer to play and display, and perform target detection processing at the same time, observe the receiving video effect and detection results at the receiving end, and observe the impact on the video transmission system by adjusting the antenna gain. The experimental results show that the video

Volume 8 Issue 7, 2022
DOI: 10.6919/ICJE.202207_8(7).0074

transmission system built on the software radio platform can successfully transmit TS video files, and the received video can be processed in real time by means of pipeline files. Through experiments in an indoor environment, we can observe the effects of noise and antenna gain on video transmission. In addition, the system can further optimize the video transmission and detection system by modifying the signal processing module in GNU Radio, or changing the real-time processing method.

Acknowledgments

This research was funded by the Innovation Program of Shanghai Municipal Education Commission of China under Grant 2021-01-07-00-10-E00121 and the Natural Science Foundation of Shanghai under Grant 20ZR1423200.

References

[1] M.N. Sadat, E. Vargas-Alfonso, R. Dai, et al. QoE-VS: A Cross-Layer QoE-Aware Video Streaming Platform Using Software-Defined Radio[C]. 2020 IEEE 92nd Vehicular Technology Conference (VTC2020-Fall). IEEE, 2020: 1-6.

[2] B. Bloessl, F. Klingler, F. Missbrenner, et al. A Systematic Study on the Impact of Noise and OFDM Interference on IEEE 802.11 p[C]. 2017 IEEE Vehicular Networking Conference (VNC). IEEE, 2017: 287-290.

[3] Hua Liang. Detection and Tracking of Moving Objects in Multi-Camera Video Surveillance [J]. Doctoral Dissertation, National University of Defense Technology, Changsha, 2009.

[4] X.W. Tang, X. L. Huang. A design of SDR-based pseudo-analog wireless video transmission system[J]. Mobile Networks and Applications, 2020, 25(6): 2495-2505.

[5] A. Bochkovskiy, C.Y. Wang, H.Y.M. Liao. Yolov4: Optimal speed and accuracy of object detection[J]. arXiv preprint arXiv:2004.10934, 2020.

[6] Information on https://gstreamer.freedesktop.org.

[7] Information on https://wiki.gnuradio.org.

[8] L. Shavanthi, K. Nagamani, H.B. Sandya. Implementation of DVB Standards using Software-Defined Radio[C]. 2018 3rd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT). IEEE, 2018: 860-864.

[9] S. Nimmi, V. Saranya, R. Gandhiraj. Real-time video streaming using GStreamer in GNU Radio platform[C]. 2014 International Conference on Green Computing Communication and Electrical Engineering (ICGCCEE). IEEE, 2014: 1-6.

[10] M. Lanoë, E. Senn. Consumption analysis and estimation in the design of GStreamer based multimedia applications[C]. Proceedings of the 2012 Conference on Design and Architectures for Signal and Image Processing. IEEE, 2012: 1-7.

[11] L. Yinli, Y. Hongli, Z. Pengpeng. The implementation of embedded image acquisition based on V4L2[C]. 2011 International Conference on Electronics, Communications and Control (ICECC). IEEE, 2011: 549-552.

[12] D. Roy, T. Mukherjee, M. Chatterjee, et al. Adaptive Video Encoding and Dynamic Channel Access for Real-time Streaming over SDRs[C]. 2018 IEEE 37th International Performance Computing and Communications Conference (IPCCC). IEEE, 2018: 1-9.

[13] Xin Su, Xibin Xu. The principle and technology of software radio [M]. Beijing: People's Posts and Telecommunications Press, 2010. 52-61.

[14] O.N. Samijayani, P. Gitomojati, D. Astharini, et al. Implementation of SDR for video transmission using GNU radio and USRP B200[C]. 2017 5th International Conference on Cyber and IT Service Management (CITSM). IEEE, 2017: 1-4.

[15] Miao Yu. DVB-T algorithm research and FPGA implementation applied to UAV image transmission [D]. Harbin Institute of Technology, 2019.

[16] B. Hariharan, P. Arbeláez, L. Bourdev, et al. Semantic contours from inverse detectors[C]. 2011 international conference on computer vision. IEEE, 2011: 991-998.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202207_8(7).0074

[17] J. Deng, W. Dong, R. Socher, et al. ImageNet: A large-scale hierarchical image database[C]. 2009 IEEE conference on computer vision and pattern recognition. IEEE, 2009: 248-255.

[18] K. Han, Y. Wang, Q. Tian, et al. Ghostnet: More features from cheap operations[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020: 1580-1589.

[19] K. Bousmalis, G. Trigeorgis, N. Silberman, et al. Domain separation networks[J]. Advances in neural information processing systems, 2016, 29.

[20] K. He, X. Zhang, S. Ren, et al. Deep residual learning for image recognition[C]. Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.