Horus Testbed: Implementation of Real-Time Video Streaming Protocols

Ahmed Abdelhadi†, Andreas Gerstlauer† and Sriram Vishwanath‡
†University of Houston, Houston, TX
‡University of Texas at Austin, Austin, TX

Abstract—The integration of communication, computing, and control in mobile cyberphysical systems leads to many challenges. This is due to the complex interactions between these three subsystems along with the limited resources in mobile cyberphysical devices. In this paper, we design and construct mobile cyberphysical nodes to be integrated with other stationary nodes to form a wireless network. These nodes are autonomous aerial vehicles (AAVs) forming a network with each other and with ground access point nodes that we use as a testbed called Horus. In this testbed, we run our developed rate distortion optimized (RDO) and adaptive wireless video transmission protocols. In our design of these transmission protocol, we take into consideration low computation, power, and bandwidth capabilities of Horus testbed that is a good representation of mobile cyberphysical systems. We provide results of our experiments on Horus testbed. These include both spatial and temporal distortion metrics and comparison between various transmission protocols along with samples of transmitted images and corresponding distorted received images from our recorded videos.

Index Terms—Embedded Systems, Video Streaming, Real-Time Systems

I. INTRODUCTION

With the increase in real-time video transmission in wireless systems [1]–[4] accompanied with quality of service (QoS) [5]–[7] and quality of experience (QoE) [8]–[10] requirements, there is a need to implement reliable wireless real-time transmission protocols with improved QoE in real mobile cyberphysical systems with stringent wireless conditions. These systems are usually limited in resources, such as power and computation resources. This leads to many research challenges emerging from the need to integrate communication, computing, and control in mobile cyberphysical system. With the vast implementation options of mobile cyberphysical systems, it is a daunting task to integrate, cross-layer, cross-domain co-design such a system. However, for testing new protocols in real stringent setting, new versatile testbeds need to be constructed and utilized for experimentation. In this paper, we are interested in designing a mobile cyberphysical system testbed. This testbed consists of wireless autonomous aerial vehicles (AAVs) connected with each other and with ground stations. Due to the mobility of these AAVs and the fixed location of ground stations, this testbed provides a time-varying wireless channel with high frequency of channel state changes. Such system has both civilian and military applications. We implement using this testbed our developed rate distortion optimized (RDO) and adaptive video transmission protocols. In the design of these protocols, we consider the challenges that come with real-time packetized media [11], such as high data rate, real-time constraints, and dependencies between frames.

A. Prior Work

There are numerous prior work on QoS and QoE for various Open Systems Interconnection (OSI) Model layers. For example in [12]–[15], QoS is addressed for network layer, while in [16], [17] physical layer is the main focus, and in [18], [19] algorithms were developed for application layer. In this paper, we are mainly concerned with the design of the application layer. In other research work, game theory was utilized [20]–[23], or energy efficiency for LTE third generation partnership project (3GPP) in [24]–[29]. Additionally, other standards like Worldwide Interoperability for Microwave Access (WiMAX) [30]–[32], Mobile Broadband [33] in [34], and Universal Mobile Terrestrial System (UMTS) [35]–[37] were investigated. The utilization of cross-layer design is also well investigated, e.g. in [38], [39]. Various other schedulers were introduced as well, e.g. integrated services in [40], [41], differentiated services in [42]–[44], Asynchronous Transfer Mode (ATM) in [45], [46], and embedded and battery related QoS research in [47]–[51].

In this paper, we are running real-time applications, i.e. video applications. Historically, delay-tolerant applications [21], [52] were the main focus of schedulers’ algorithms [53]–[59] with optimal design shown in [60]–[63]. Currently, real-time applications, as the one considered in this paper, are of more interest with both approximate solutions [64]–[66] and optimal solutions [67]–[74]. For more improvement of QoS and QoE, carrier aggregation methods [75]–[79] for schedulers are proposed with optimal solution in [80]–[84]. Due to the need to increase the communications spectrum band, utilization of radar spectrum via carrier aggregation was recommended by the President Council of Advisers on Science and Technology (PCAST) [85], [86], Federal Communications Commission (FCC) [87]–[90], and National Telecommunications and Information Administration (NTIA) [91]–[93] as shown in [94]–[102].

Extensions of this prior work and the work presented in this paper can lead to further implementations and applications, for example to machine to machine (M2M) communications [103]–[105], multicast networks [106], ad-hoc networks [107]–[110], and various other networks [111]–[114].
B. Our Contributions

The paper has the following contributions:

- We implement a wireless testbed for testing various real-time wireless transmission techniques.
- We test the performance of our developed rate distortion optimized and adaptive video transmission techniques and measure corresponding temporal and spatial distortion metrics.
- We compare between the experimental results of these transmission techniques in a real-setting.

II. SYSTEM IMPLEMENTATION

Our network consists of a fixed number of autonomous aerial vehicles (AAVs) nodes, see section II-A for node description, that are placed in a prespecified topology, see section II-B for the proposed topologies. In this network, sources are streaming video data in real time to destinations. The routing path is given a priori and the topology of the network is fixed throughout the experiment. However the nodes are moving in a fixed circular path as described in section II-B. This continuous movement ensures the time-varying nature of wireless channel and reveals the effectiveness of the implemented algorithms. We propose to implement rate-distortion optimized algorithm and adaptive algorithms applied on MPEG2 and MJPEG compressions and measure the system performance for these networks.

A. Node Description

In this section we describe the nodes used to build the network. We are considering AAVs as aerial nodes that form the network under test, see Figure 2. Due to the nature of these nodes we have constraints on the weight and dimensions of the components and also the power consumption which directly affect the transmission range. After elaborate search for the suitable components the following list represents the components that are used to construct our system nodes and their corresponding functions, see Figure 1\(^1\) for the block diagram:

1) **Via EPIA Nano-ITX** [115]: is an x86 computer which is the central unit for managing and operating the communication between nodes. It includes an Atheros [116] wireless card with IEEE 802.11 2.4 GHz frequency band which is used for wireless video transmission

2) **Zigbee** [117]: includes IEEE 802.15 900 MHz frequency band which is used for sending way-points for AAV navigation

3) **IMU\(^2\)** unit [118]: controls the AAV movement during flight and sustains the required circular path that is shown in Figure 4 and Figure 5.

4) **GPS\(^3\)** unit [119]: measures the location of the AAV which is important in topology management.

5) **Remote Control Unit (R/C module)** [120]: uses 72 MHz frequency band and is the main manual ground control on the AAV. In our project it is used to control take-off and landing of AAV.

6) **Servo motors** [121]: control flaps and rudder and motor speed for controlling direction and speed of AAV.

7) **DC motor** [122]: is the main moving force of AAV and is connected to propeller.

8) **Video Camera**: it capture real-time video.

We divide the block diagram into three section: propulsion which consists of motor, speed control, and motor battery, control that includes IMU, GPS, R/C module, servos, control battery, and communication that represents the main payload of x86 computer, wireless card, camera, and communication battery.

Our choice of the components and the sections of the block diagram is for the purpose of achieving:

1) **Reliability**: To ensure reliability, we choice the components with different frequency bands to avoid any interference between the different transmissions. The following

\[^1\]UART, PWM, and USB stands for Universal Asynchronous Receiver/Transmitter, Pulse-Width Modulation, and Universal Serial Bus, respectively

\[^2\]IMU stands for Inertial Measurement Unit

\[^3\]GPS stands for Global Positioning System
are the frequencies used:

a) R/C module: has a carrier frequency of 72 MHz where each node is on a different channel to avoid interference

b) Zigbee network: operates on the 900 MHz frequency band for sending feedback data between nodes
c) WiFi network: uses 2.4 GHz for sending and receiving compressed video data

2) Safety: is achieved by separating the power supply of each section of AAV (propulsion, control, communication). So, for example, if the propulsion battery is drained, we can still have control over the AAV for safe landing and avoiding a possible crash

3) Availability: we are using off the shelf components

B. Implementation Network Topology

In Horus testbed project, we consider two main network topologies:

1) Unicast network topology: we have in this case one AAV moving in a circular pattern as shown in Figure 4. The AAV is the source that records video of the landscape and transmits this video to the destination. The destination is a regular laptop running Linux.

2) Multiple unicast network topology: we have in this case two AAV moving in a two different circular paths as shown in Figure 5. The AAVs are both sources that records video of the landscape and transmit these videos to one destination as shown in Figure 3. The destination is also a laptop running Linux.

The AAVs are controlled manually during take-off and landing only using R/C module running at 72MHz. In the air, the AAVs are controlled automatically by the on-board IMU module. The IMU module uses stored way-points and GPS to follow the desired path. The path can be changed while the AAVs are in the air by changing the way-points on a ground
laptop and sending these new way-points from laptop to the on-board IMU by zigbee at 900MHz. The AAV actual path and the relative viewing area on the location used to run the experiments are shown in Figure 6.

C. Implementation Distortion Measure

As we experience drops in video transmission, we experience different video duration between video stored in the transmitter and the receiver. Hence, we use two distortion metrics for our measurements in this experiment. First, we measure the number of drops in the transmission which represents the temporal loss in information. Second, we measure the spatial loss in the video by slicing the video into images and comparing between the best and worst received images with respect to the source images.

1) Temporal Distortion: We measure the temporal distortion by measuring the size of the stored video files. The video file is compressed by MPEG2 encoder and stored at the source while being transmitted to the destination. At the destination, a copy of the video received is stored in MPEG2 compression while being viewed in real-time. Therefore calculating the percentage of the stored video at the destination to the stored video at the source is proportional to measuring the packet drops in the wireless network. These compressed videos are then decompressed into raw video. We calculate the percentage of the raw video size at the receiver to the raw video size at the transmitter. This percentage gives an approximate measure of the frames reconstruction at the receiver. MPEG2 decoder performs inter-frame estimation for the missing packets in the frame. The estimation process causes some of the frames to be distorted. For the distorted frames, we perform another measure which is the spacial measure in section II-C2. For final temporal measure, we compare difference in duration between the video viewed at the receiver and the original at the transmitter. These three temporal distortion measures are used for measuring the difference in performance between different algorithms.

2) Spacial Distortion: We measure the spacial distortion by measuring the Structural SIMilarity (SSIM) [123] index of the frames successfully received/reconstructed at the receiver. The SSIM index is a method for measuring the similarity between two images. The SSIM index is a reference metric where the measuring of image quality based on an initial distortion-free image as reference. We choose SSIM because it outperform traditional methods like peak signal-to-noise ratio (PSNR) and mean squared error (MSE), which have proved to be inconsistent with human eye perception.

The SSIM metric is calculated on various windows of an image. The measure between two windows \( x \) and \( y \) of common size \( N \times N \) is:

$$SSIM(x, y) = \frac{(2\mu_x \mu_y + c_1)(2\sigma_{xy} + c_2)}{\mu_x^2 + \mu_y^2 + c_1(\sigma_x^2 + \sigma_y^2 + c_2)}$$  (II.1)

where \( \mu_x \) is the average of \( x \), \( \mu_y \) is the average of \( y \), \( \sigma_x^2 \) is the variance of \( x \), \( \sigma_y^2 \) is the variance of \( y \), \( \sigma_{xy} \) is the covariance of \( x \) and \( y \), \( c_1 = (k_1L)^2 \), \( c_2 = (k_2L)^2 \) are two variables to stabilize the division with weak denominator, \( L \) is the dynamic range of the pixel-values (typically this is \( 2^\text{#bits per pixel} - 1 \)), and \( k_1 = 0.01 \) and \( k_2 = 0.03 \) by default.

In order to evaluate the image quality this formula is applied only on \( luma \). The resultant SSIM index is a decimal value between -1 and 1, and value 1 is only reachable in the case of two identical sets of data.

D. Implemented Video Transmission Protocols

For implementing our transmission protocols, Linux operating system is used. Additionally, for video compression, GStreamer [124] is used, and for running our transmission protocols, Click Modular Router network stack [125] is used.

1) RDO with MPEG2 Codec: In this technique, the output of the MPEG2 compression block is connected to a software switch to select between two different MPEG2 compression. The selection is decided based on the received Received Signal Strength Indicator (RSSI) from receiver.

2) Adaptive with MPEG2 Codec: This technique is similar to RDO in switching between two transmission schemes based on the received RSSI. But the switching is between unlimited transmission rate and limited transmission rate of 100 kbps.

3) Adaptive with MJPEG Codec: This technique uses MJPEG for video transmission which requires very low complexity compared to MPEG used in the other two methods. This technique switches between two encoders, one with low quality image of 30% and another with high quality image of 80%.

In the above transmission protocols, the decision to switch from one compression to another is done via pre-specified thresholds based on the received Received Signal Strength Indicator (RSSI). The RSSI is measured at the source, i.e. AAVs, from the beacons inherently sent by the , i.e. laptop.

E. Implementation Results

Using the AAV nodes described in Section II-A, we run flight experiments using the topologies described in Section II-B at Lester Field [126] in Austin, Texas, shown in Figure 6 (source Google maps [127]). In these flight experiments, the AAVs capture live video of the landscape and transmit the recorded video in real-time to the ground station (laptop). We run the three mentioned techniques in Section II-D for both unicast and multiple unicast topologies and measure the video distortion using the metrics described in Section II-C.

MJPEG codec outperforms MPEG2 codec in SSIM index. This is due to the inter-frame estimation that is performed in MPEG2 decoder and causes partially received frames to be viewed as distorted frames. While these partially received frames are dropped in MJPEG decoder. Hence, we conclude that SSIM index is compression dependent. The calculated

<table>
<thead>
<tr>
<th>Video Compression</th>
<th>MPEG2</th>
<th>MPEG2</th>
<th>MJPEG</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSIM index</td>
<td>0.77</td>
<td>0.78</td>
<td>0.96</td>
</tr>
</tbody>
</table>

### TABLE I

AVERAGE SSIM INDEX FOR DIFFERENT ALGORITHMS
values for each experiment is shown in Table I and Figure 7. A summary of the comparison between the three proposed algorithms is shown in Table II. This summary includes temporal distortion metric differences between various algorithms by comparing the different duration of received videos. We can see that MPEG2 outperform MJPEG in this metric. Some samples of observed frame distortions are shown in Figure 8.

### III. Conclusion

This paper constructs a reliable wireless network testbed for testing new wireless networks protocols called Horus. In this testbed, we implement the problem of streaming packetized media over a wireless network using a rate distortion optimized algorithm. We provide a comparative study to illustrate some practical trade-offs for reliable and optimized video transmission. The main conclusion that emerges from Horus is that in order to provide a good real-time video transmission performance, one should consider both the computation power and the bandwidth limitations. For low computation power, we have MPEG2 as the more suitable solution. For limited bandwidth but high computation power, MPEG4-H264 could potentially be the more suitable solution because it has better utilization of the available bandwidth compared to MJPEG and MPEG2. This is the trade-off between bandwidth and computation limitation and the video quality that can be achieved.

### References

Fig. 8. Samples of Distortion (image from source video on the right and image from destination video on the left)