Towards the Enhancement of UAV Video Transmission with Motion Intensity Awareness

Roger Immich
University of Coimbra
Coimbra, Portugal
immich@dei.uc.pt

Eduardo Cerqueira
Federal University of Para
Belem, Brazil
cerqueira@upfa.br

Marilia Curado
University of Coimbra
Coimbra, Portugal
marilia@dei.uc.pt

***Abstract*—The use of video-equipped Unmanned Aerial Vehicles (UAV) has been increasing recently, along with the number of available applications for military and civilian employment. This unveils the need for an adaptive video-aware mechanism capable of overcoming a number of challenges related to the scarce network resources, device movement, as well as high error rates, to ensure a good video quality delivery. Forward Error Correction (FEC) techniques can be tailored to provide adaptive protection with Quality of Experience (QoE) assurance over error-prone and high-mobility networks. Besides that, unique characteristics of each video sequence, such as the spatial complexity and the temporal intensity, strongly affect how the QoE will be impacted by the packet loss. This paper proposes an adaptive motion intensity and video-aware FEC mechanism with the aid of Fuzzy logic to safeguard UAV real-time video transmissions against packet loss, providing a better user experience, while saving resources. The advantages and drawbacks of the proposed mechanism in comparison to the related work are evidenced through experiments and assessed by using QoE metrics.**

***Index Terms*—Motion intensity, Forward Error Correction (FEC); Unmanned aerial vehicles (UAV); Fuzzy Logic; Quality of Experience (QoE); Unequal Error Protection (UEP)**

I. Introduction

In the last few years, we have seen a substantial growth of Unmanned Aerial Vehicles (UAV) use in video reconnaissance, exploitation, and surveillance [1]. The deployment of both, autonomous and nonautonomous UAVs, was formerly exclusive use of military and special operation teams. Nowadays, the easy operation and cost-effective wireless network technologies are making this type of equipment accessible for civilian use. The adoption of UAVs can be helpful in a broad range of situations, more often than not replacing fixed video cameras due to their mobility and low cost operation in contrast to manned systems. Some examples of UAVs applications are in traffic surveillance, sport events, festivals, public parades, or at any place that has the potential of gathering a large amount of people [2] [3]. It is also worth highlighting the use for monitoring and inspection of critical infrastructures, such as harbours, large industrial areas, rail-ways, long pipelines, power plants, as well as to cover large areas with lack of infrastructure, such as interior border control, countryside properties or even in natural disaster sites and rescue missions [4].

As evidenced, video-equipped UAVs can provide several benefits, however, despite the adoption of modern visualization tools, without a proper video streaming quality, which must be watchable for humans, the usability of the system will be compromised. At the same time, real-time video transmissions with ensured Quality of Experience (QoE) are resource-demanding services, especially over wireless networks. Video streaming needs a steady and continuous flow of packets, being susceptible to failure due to different factors. To start off, in wireless networks, the channel conditions can quickly fluctuate over time, particularly in high-mobility networks. Several other common communication issues also have to be taken into consideration like channel interference, multipath fading, and noise [5]. To make the matter even worse, the UAV networks often have poor connectivity quality [6]. Another problem that needs to be tackled is how to make a fair use of the available bandwidth [7]. All of these issues need to be considered to provide an efficient use of the available resources.

The QoE of video streams can be defined as the overall acceptability of end-users being related to, but differing, from the largely studied concept of Quality of Service (QoS). This means that QoE solutions assess the video quality from the end-users point-of-view, and as a result of that, it must be considered in the adaptive mechanisms. Efficient QoE-aware video stream distribution is one of the main challenges in high-dynamic wireless networks. To do that, it is necessary to define a proper adaptive control that uses QoE, video, and network characteristics to improve the use of resources and, at the same time, enhances the video quality for end-users.

A number of factors have impacted on the QoE in a UAV scenario. Besides the network-related parameters, several video characteristics are known to play an important role on the video quality, such as the image size (resolution), codec type, bitrate, the format of the Group of Pictures (GoP), as well as the spatial complexity and the motion intensity [8]. Owing to that, an adaptive mechanism is necessary to better define the amount of redundancy accordingly to the video characteristics, network conditions and, the human perception of quality, leading to an improved QoE. This is even more important in high-dynamic and error-prone wireless network, particularly if it involves mobile nodes like the UAVs. One important factor to consider in these networks is the motion intensity. It can be inferred through the Motion Vectors (MV) data, and it is used to store changes from adjacent frames in the temporal video compression. This offers a particular view of each video sequence and can be used to adaptively change the amount of redundancy to be contained in a set of live video

978-1-4799-6606-6/14/$31.00

flows. Another important video feature that needs to be tackled is the image size. Lately, a large number of video resolutions is available and can be used for different purposes. Owing to that, the adaptive mechanisms have to be flexible enough to cope with arbitrary resolutions. The analysis of this information allows us to identify the most critical pieces of information and through the use of an Unequal Error Protection (UEP) scheme it is possible to protect it accordingly.

Several adaptive mechanisms have been proposed with the aid of Forward Error Correction (FEC) techniques. This technique is known to be successfully used in real-time video transmission services [9], by sending redundant data along with the original set, thus improving the QoE, without increasing the delay. It is known that the wireless channel resources are limited and can be unfairly distributed among the users, thus an adaptive FEC-based mechanism is required. To further improve the adaptive mechanism, it should also be UEP- and QoE-aware, assuring the redundancy of only the most important video sequences from the user point-of-view. This will produce less network overhead, while increasing the human perception when watching live video flows.

This paper describes a novel adaptive **M**otion **INT**ensity and video-aware mechanism (MINT-FEC) to enhance the resilience of UAV real-time video transmissions. One of the major weaknesses in the mechanisms found in the literature is the use of unnecessary redundancy, where they generate and send sequences of all video frames and not only the most important ones from the human perspective, such as I- and P-Frames. These are the most important frames in the MPEG standard. The loss of one I- or P-Frame will be more noticeable by the end-user because the error will only be corrected when another I-Frame arrives, in other words, in the beginning of the next GoP. Thus, those frames need to be protected with redundant information; To tackle this issue, the MINT-FEC dynamically adapts itself by using fuzzy logic, to add a precise amount of redundancy to only the most QoE-sensitive data, while ensuring high-quality video and downsizing the usage of scarce wireless resources. The video quality is important to users and authorities (e.g., paramedics and firefighters) to do an accurate assessment of each situation, reducing human reaction times. Additionally, by sending less redundant data we are automatically using less power, which saves energy. The MINT-FEC was assessed using objective QoE metrics and real videos obtained from UAVs.

The remainder of this paper is structured as follows. Section II describes the related work. The novel MINT-FEC mechanism is shown in Section III, and its assessment is presented in Section IV. Conclusions and future work are given in Section V.

II. Related Work

A range of different solutions has been proposed to enhance the video quality over wireless networks. To the best of our knowledge, none of those proposals improve the real-time video transmission of UAVs using the motion intensity. This is a key characteristic to know how the video quality will be impacted when packets are lost in dynamic wireless networks. The Adaptive Cross-Layer FEC (ACFEC) provides packet-level error correction [10]. A MAC layer loss counter determines the amount of FEC redundancy. Because of that, in good wireless conditions, this counter will be low, generating a small amount of redundant traffic. While this is true, a proper network overhead assessment was not conducted and thereby it is difficult to ascertain the proposal efficiency. On top of that, the video characteristics are neglected. This information is known to impact on the video resilience to packet loss and, therefore, on the QoE levels.

Another approach defines a dynamic FEC block length [11]. This block is configured using the packet loss rate and also the number of continuous losses to enhance the video transmissions. The weakness of this mechanism is that it only uses network parameters disregarding key multimedia information, such as video characteristics and QoE results. These details are very important to set a precise amount of redundancy in any adaptive QoE-aware mechanism.

The Cross-Layer Mapping Unequal Error Protection (CLM-UEP) [12] assigns a custom amount of redundancy by analysing the packet loss rate and the frame type. Nevertheless, the CLM-UEP does not utilize an important video detail, namely the motion intensity. As previously mentioned, this characteristic can have a significant weight to determine a precise amount of redundancy, which allows saving significant network resources and energy.

III. Adaptive Motion Intensity Awareness Mechanism

In the light of the open issues aforementioned, especially the lack of adaptive QoE-aware mechanisms that take into consideration efficient indicators of motion intensity, this study describes and evaluates a novel cross-layer adaptive FEC-based mechanism with motion intensity awareness (MINT-FEC). The main goal is to ensure a good QoE in real-time video transmissions of UAVs, while saving wireless resources and energy. This proposal improves our previous work [13] and the main enhancements are described next.

A novel approach was taken in the MINT-FEC, as the motion intensity is now given by combining the spatial complexity and temporal intensity. Spatial complexity is how distinct is one frame from another, as well as the colour and luminance saturation. The temporal intensity can be defined as how fast and how much the image is changing frame-by-frame. The joint use of them provides a more accurate motion classification to be used for adding a precise amount of redundancy. Another improvement from the previous work is the video resolution independence. By normalizing the values of all the video characteristics, as well as using the motion vector distance and macroblock size, it is possible to add, on-the-fly, an adaptive amount of redundancy to videos with arbitrary resolution. A detailed description of all the novel components is given further.

The use of fuzzy logic allows the design of a comprehensive and dynamic mechanism [13]. This is possible because it can take into consideration a large number of videos and network details and still be fast enough to operate in real-time schemes, as expected in highly dynamic UAV networks.

Additionally, fuzzy logic can be considered a problem-solving methodology that aims to define what the system should do, rather than attempting to fully understand its operation. It adopts a simple approach to provide definitive conclusions relying on imprecise, ambiguous, or vague information.

The design process of MINT-FEC starts with the definition of several fuzzy components, such as the sets, membership functions, and rules. Fuzzy sets are different from classical sets since an element can have a degree of membership instead of just belonging or not belonging to a set. Membership functions are used to represent the degree of relationship between the elements of the sets. At last, the rules are responsible for defining how the system will behave. This is a complex process that has to be executed offline only once. After that, all the generated information can be loaded into the fuzzy interface engine to be used in real-time. This offline process is very important as it allows a faster and more accurate mechanism to perform its procedures on-the-fly, since simpler tasks are done.

The first step is to quantify the spatial complexity, which is how much spatial information a frame is carrying compared to the previous one. The most common way to compute this difference is using the Sum of Absolute Differences (SAD) [14]. It is not a complex operation, however, it is very time-consuming because it compares each pixel from both frames, making this impractical in real-time. Another way to find this information is through the frame sizes. The problem of using the frame size is that several video characteristics can impact on it, such as different resolutions (picture size), content, as well as temporal intensity. To be able to compare the frame sizes among different videos, it is necessary to normalize all the information. Using Eq. (1) the average frame size is calculated, and the same operation is also executed for P- and B-frames. After that, through Eq. (2), all frame sizes are normalized, as before, this is also done for P- and B-frames. This process is performed for each video sequence separately. Table I shows the adopted notation.

TABLE I
ADOPTED NOTATION

NOTATION	MEANING
$\mu I_s, \mu P_s, \mu B_s$	Frame size average
$\hat{\mu} I_s, \hat{\mu} P_s, \hat{\mu} B_s$	Normalized frame size average
$I_{s(i)}, P_{s(i)}, B_{s(i)}$	Frame size of the i_{th} frame
nF	Number of frames in the video sequence
$\lvert MV \rvert$	Euclidean distance of a motion vector
$\lvert MV_{(i)} \rvert$	Euclidean distance of the i_{th} motion vector
MB_h	Macroblock height
MB_w	Macroblock width
aMB	Macroblock area
$aMB_{(i)}$	Area of the i_{th} macroblock
nMB	Number of macroblock in the frame
$TI_{\Delta t}$	Temporal intensity

$$\mu I_s = \frac{1}{nF} \sum_{i=0}^{nF-1} I_{s(i)} \tag{1}$$

$$\hat{\mu} I_s = \frac{\mu I_s}{\mu I_s + \mu P_s + \mu B_s} \tag{2}$$

Once all the frame sizes are normalized, it is possible to perform an exploratory analysis to cluster all frames of all video sequences together according to their sizes. The technique used was hierarchical clustering, because it allows the partition of the data, as much as possible, into the most homogeneous groups [15]. Based on the linkage distance of the clusters it was possible to divide them into three distinct groups, namely “small”, “medium”, and “large”. After defining the clusters, a boxplot was used to summarize and display the distribution of the data. This is an important tool in the exploratory analysis because it displays the shape of the distribution of each cluster along with the central value and the variability. Fig. 1 shows the boxplot for the spatial complexity. The fuzzy sets for spatial complexity were defined using the information displayed by the boxplot, as shown by Fig. 2.

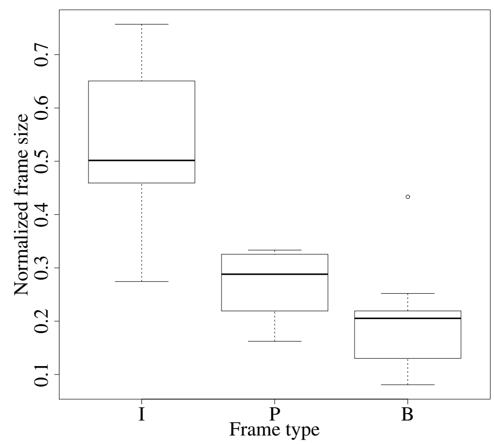

Fig. 1. Spatial Complexity

```
FuzzyOperator& op = FuzzyOperator::DefaultFuzzyOperator();
FuzzyEngine engine("complex-mamdani", op);

InputLVar* Isz = new InputLVar("I-size");
 Isz->addTerm(ShoulderTerm("SMALL", 0.274, 0.459, true));
 Isz->addTerm(TriangularTerm("MEDIUM", 0.274, 0.651));
 Isz->addTerm(ShoulderTerm("LARGE", 0.502, 0.757, true));
engine.addInputLVar(Isz);

InputLVar* Psz = new InputLVar("P-size");
 Psz->addTerm(ShoulderTerm("SMALL", 0.162, 0.219, true));
 Psz->addTerm(TriangularTerm("MEDIUM", 0.162, 0.325));
 Psz->addTerm(ShoulderTerm("LARGE", 0.288, 0.333, true));
engine.addInputLVar(Psz);

InputLVar* Bsz = new InputLVar("B-size");
 Bsz->addTerm(ShoulderTerm("SMALL", 0.081, 0.13, true));
 Bsz->addTerm(TriangularTerm("MEDIUM", 0.081, 0.219));
 Bsz->addTerm(ShoulderTerm("LARGE", 0.205, 0.252, true));
engine.addInputLVar(Bsz);
```

Fig. 2. Spatial complexity (Frame size sets)

After defining the set, the membership functions need to be outlined. This is problem-dependent, as well as a complex task being difficult to find an optimal solution [16]. Considering that, it is better to use piecewise linear functions (formed of straight-line sections). These functions are both simpler and more efficient regarding to computability, leading to lesser resource requirements. Fig. 3 shows the graphical representation of our chosen membership functions for the frame sizes.

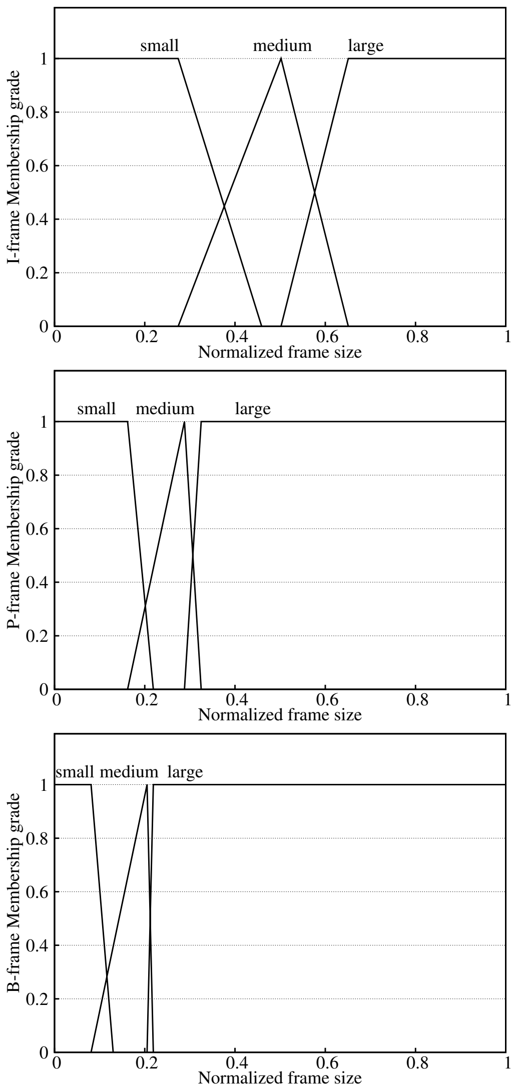

Fig. 3. Frame size membership function

Apart from the spatial complexity, the fuzzy components for the motion intensity also need to be created. The analysis of this intensity is performed through the Motion Vectors (MV) information. The basic idea of MV is to describe the change in place or position of objects as a sequence of small translations on a plane. In order to better represent this movement, instead of counting the number of MV, it is computed how far each vector is pointing using Euclidean distance, which is given by Eq. 3. This is necessary because it is possible to have one frame with several vectors pointing to a close distance meanwhile, another frame with fewer vectors, pointing farther away although, and thereby having higher motion intensity.

$$|MV| = \sqrt{(x - x')^2 + (y - y')^2} \tag{3}$$

As defined in the MPEG standard, the MV describes the movement of macroblocks (MB) from some position in one frame to another position in another frame. It is important to note that not all MB have the same size, as well as videos with higher resolution will have more MB than videos with lower resolution. To be able to compare video sequences with different MB sizes and resolutions, we decided to use the MB area, given by Eq. (4), together with the MV. Additionally, using Eq. (5) it is possible to calculate for each macroblock how many pixels have been moved and how far away, which can be translated as temporal intensity.

$$aMB = MB_h \times MB_w \tag{4}$$

$$TI_{\Delta t} = \frac{1}{nMB} \sum_{i=0}^{nMB-1} aMB_{(i)} \times \left|MV_{(i)}\right| \tag{5}$$

Using the aforementioned details, another exploratory analysis was performed to classify the video sequences in terms of temporal intensity. This time, instead of breaking the video sequences in frames, the whole video was analysed. The values found through (5) were used to cluster the videos into three distinct groups, namely "low", "medium", and "high" temporal intensity. Additionally, in the same way as before, a boxplot was used to summarize and display the data distribution, as well as to create the sets, as presented in Fig. 4.

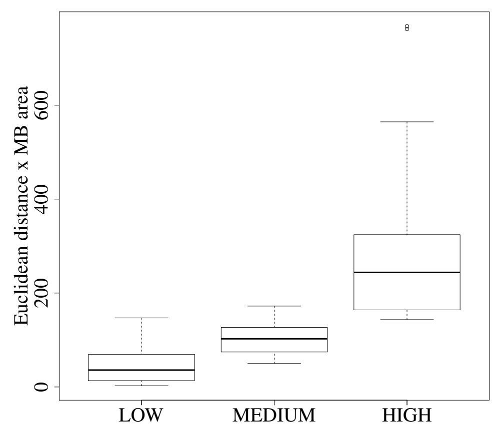

Fig. 4. Temporal Intensity

Another important step is to define the packet loss rate (PLR). The primary objective is to find out the impact of different PLR in the QoE for a set of videos with the same characteristics. This is necessary because video sequences have a certain tolerance for losses. For example, due to the natural video resiliency to packet loss, in our approach a loss rate of 10% can be considered low, in other services such as voice over IP, it might be considered unacceptable. To find the PLR that best represents the video quality, a number of network simulations using a broad collection of UAV video sequences were carried out. For PLR between 0% and 10% the video quality was good. A tolerable video quality was observed for majority of the videos between 5% and 20% of PLR. Over 15% of PLR the quality quickly decreased in videos with higher motion intensity, and over 25% it became unacceptable. Based on the results, the PLR set was defined, as shown in Fig. 5.

After delineating the packet loss rate, the amount of the redundancy set must be defined. In the same way as done

```
InputLVar* PLR = new InputLVar("PacketLossRate");
 PLR->addTerm(TriangularTerm("LOW", 0, 10));
 PLR->addTerm(TriangularTerm("MEDIUM", 5, 20));
 PLR->addTerm(TriangularTerm("HIGH", 15, 100));
engine.addInputLVar(PLR);
```

Fig. 5. Packet loss rate input set

before, the definition of this set also enfolds several experiments. With the help of human knowledge in the field, it was specified what would be considered a "small", "medium", and "large" amount of redundancy. Fig. 6 displays the result.

```
OutputLVar* redundancy = new OutputLVar("RedundancyAmount");
 redundancy->addTerm(ShoulderTerm("SMALL", 0.55, 0.70, true));
 redundancy->addTerm(TriangularTerm("MEDIUM", 0.60, 0.80));
 redundancy->addTerm(TriangularTerm("LARGE", 0.75, 1));
engine.addOutputLVar(redundancy);
```

Fig. 6. Motion activity output set

With all sets defined, it is necessary to create the rules. This activity also involves human knowledge about the video characteristics, namely spatial complexity and temporal intensity, as well as the frame type, and the PLR. For example, in videos with high spatial complexity the I-Frame needs a greater amount of protection, because it holds a lot of information. On the other hand, in videos with high temporal intensity, the I-Frame also needs to be protected, but the P-Frame plays an important role because it holds the temporal information about that sequence, and needs to have almost the same protection as the I-Frame. Fig. 7 shows two rules that represent this case.

```
RuleBlock* block = new RuleBlock();

block->addRule(new MamdaniRule("
 if ( SpatialComplexity is HIGH and
      PacketLossRate is HIGH and
      FrameType is I )
   then RedundancyAmount is HIGH", engine));

block->addRule(new MamdaniRule("
 if ( TemporalIntensity is HIGH and
      PacketLossRate is HIGH and
      FrameType is I or P )
   then RedundancyAmount is HIGH", engine));
```

Fig. 7. Packet loss x video characteristics rules

Once all the fuzzy rules and sets are defined, they are employed in real-time in the Fuzzy Logic Controller (FLC). The offline process needs to be performed just once, after that the FLC will be able to compute a precise amount of QoE-aware redundancy on-the-fly.

IV. Performance Evaluation and Results

The MINT-FEC goal is to ensure a high perceived QoE for end-users, while avoiding unnecessary network overhead, saving the already scarce wireless resources. The experiments were conducted in the Network Simulator 3 (NS-3). The assessment scenario consists of up to four UAVs operating in autonomous mode, with 4G LTE radio at 800MHz. To better reflect a UAV scenario, the Gauss-Markov distribution mobility model was used. This model provides a uniform spatial distribution of the nodes and also simulates inertia in the movements, which is a characteristic of UAVs in autonomous mode. The ground control station is equipped with a portable base station and antenna. All UAVs are in line-of-sight and communicating in ad-hoc mode. Only real UAV video sequences were used in the experiments. To be more precise, twenty of each video resolution (1080p, 720p, and SVGA), giving a total of sixty video sequences. All of them were encoded with both same GoP length of 19:2 and same H.264 codec. Considering the portable base station power, the ad-hoc communication, and the very demanding high definition videos, the flying range was limited to a radius of 2000 meters from the base station. Due to the harsh environment and the low-gain antenna, the PLR can range from 0% to 35%. Fig. 8 shows the packet loss distribution in our experiments. At the receiver side, a Frame-Copy error concealment was used, meaning that when a frame is lost, it will be replaced with the last good one received. Table II shows the simulation parameters.

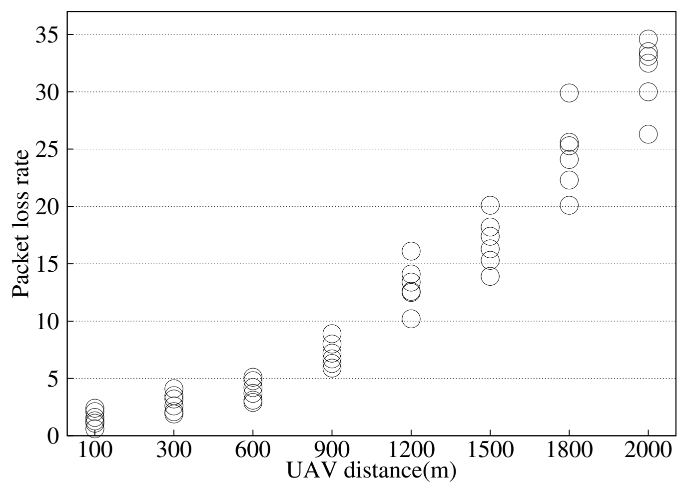

Fig. 8. PLR distribution

TABLE II
SIMULATION PARAMETERS

PARAMETERS	VALUE
Display sizes	1920x1080, 1280x720, and 800x600
Frame rate mode	Constant
Frame rate	29.970 fps
GoP	19:2
Codec	H.264
Container	MP4
Propagation model	FriisPropagationLossModel
Mobility model	Gauss-Markov
UAV velocity	45-65 km/h (28-40 mph)
LTE Frequency band	800MHz
LTE Mode	FDD
LTE Bandwidth	5 MHz
eNodeB Operating Power	22 dBm
Antenna Gain	16 dBi

Five different schemes were simulated as follows: (1) without any FEC mechanism. This is only to serve as a baseline for comparison with the others; (2) a non-adaptive video-aware FEC (I- and P-Frames are equally protected) using a pre-set value of 75% of redundancy (Video-aware FEC). This value was chosen because it showed a good tradeoff between QoE and network overhead in several PLR; (3) our previous adaptive FEC-based mechanism (uavFEC) [13]; (4) a related work implementation of the Cross-Layer Mapping Unequal Error Protection (CLM-UEP) [12]. At last, (5) adopts our novel MINT-FEC mechanism.

Fig. 9 shows the average Structural Similarity Metric (SSIM) results. Values closer to one mean better video quality. Regardless of the fact that this is an objective metric, it gives good results, which are consonant to the human visual system [17]. A foreseen situation can be clearly noticed, the farther away the UAVs are from the ground control station the worst is the video quality. In the first case (without FEC), a good video quality is noticed up to 600m. This is expected on the grounds that some video sequences tend to have a natural resiliency to packet loss. In particular, this situation is true in videos with low motion intensity, which usually scores higher results in QoE-aware assessment. Between 600m and 900m the video quality is already affected and a sharp decline is perceived after that. At the same time, in the FEC-based schemes, such as Video-aware FEC, uavFEC, CLM-UEP and, MINT-FEC, the video quality was kept good for a long distance, until 1200m. Notwithstanding, our proposed mechanism outperforms all its competitors in terms of video quality, providing even better results over higher distances. A comprehensive comparison analysis is given further.

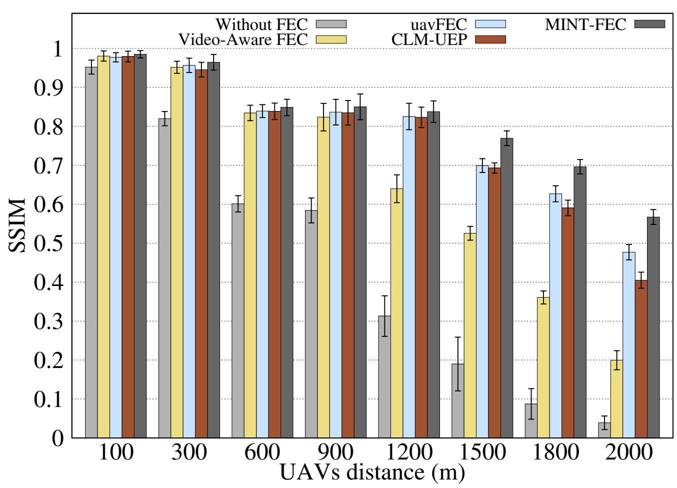

Fig. 9. Average SSIM QoE for all scenarios

The MINT-FEC already shows that it can provide enhanced video quality, especially over higher distances, however it is equally important to do so with lower network overhead. A mechanism with a low network footprint is essential considering the scarce wireless channel resources and the uneven bandwidth distribution. In our experiments the network overhead can be found by summing the size of all video frames transmitted, which includes the redundant data, and after that subtracting the original frame size. The Video-aware FEC scheme is non-adaptive and due to that it has a constant network overhead, as shown in Fig. 10. This is not suitable for UAVs, because even when they are close to the base station, with a low PLR, a large amount of redundancy is added, wasting resources. On the other hand, the adaptive mechanisms (uavFEC, CLM-UEP, and MINT-FEC) allow a better use of the network resources, as also shown in Fig. 10. In all three mechanisms, the initial amount of redundancy is small and starts to become larger as the UAVs move away from the ground station. Both our previous uavFEC mechanism and CLM-UEP perform close to each other up to 1200m. After that uavFEC starts to add more redundancy to provide better video quality. Here again, the MINT-FEC performs better than the others. Up to 1500m, it induces less network overhead, while providing higher video quality. Subsequently, it adds a slightly higher redundancy and still smaller than the other schemes, in favour of a considerably better video quality. This proves that we were able to identify the most important video portions and protect them accordingly.

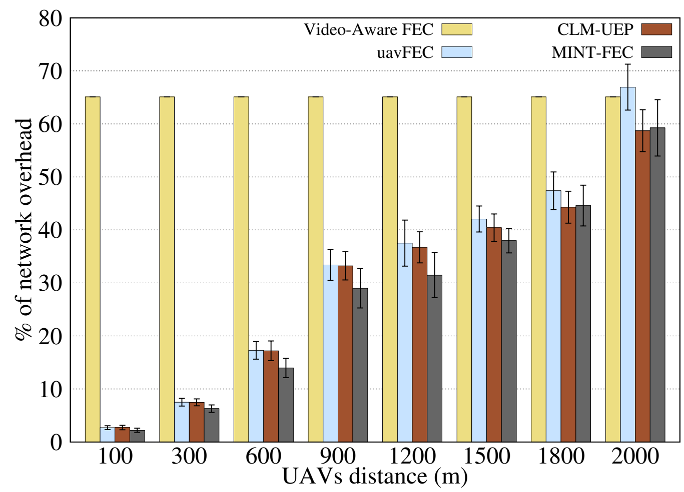

Fig. 10. Average Network overhead

To further understand the MINT-FEC achievements, a comparison against CLM-UEP and uavFEC is given in Fig. 11. The first case is the comparison between CLM-UEP and MINT-FEC, and the second one, uavFEC against MINT-FEC. The graph shows the average percentage of QoE and network overhead improvement. In the QoE assessment, a positive percentage means that our proposed mechanism achieved higher video quality, which is desirable. On the other hand, in the network evaluation, a negative percentage means that the MINT-FEC generated less overhead, which is also advantageous.

In both cases, MINT-FEC presented a slightly better video quality until 1200m, which was between 0.59% and 2.01%, and between 0.69% and 1.63%, respectively. While outperforming the other mechanisms in terms of quality, our mechanism was also able to considerably reduce the network overhead. It added on average 16.20% less redundancy than CLM-UEP and 16.65% less redundancy than uavFEC, up to 1200m. This proves that our mechanism is capable of better identifying the most QoE-sensitive data and adds a precise amount of QoE-aware redundancy to it, resulting in higher video quality and less network overhead. After this threshold, the MINT-FEC starts to increase the amount of redundancy

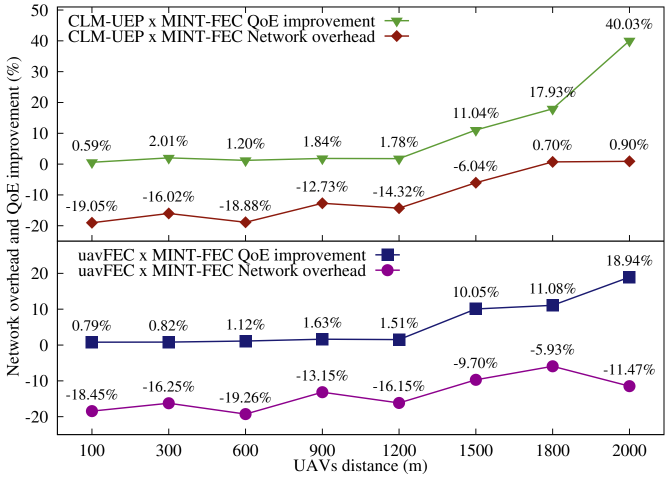

Fig. 11. QoE and Redundancy comparison

to improve the video quality. This happens for the reason that our mechanism was designed to sustain a higher video quality over long distances, when the connection is more susceptible to errors. In doing that, the videos are received with up to 40% better quality in comparison to the CLM-UEP mechanism and up to 18% higher quality than our previous uavFEC mechanism. Another important advance of the MINT-FEC over the uavFEC was the network overhead reduction beyond 1200m. In this case, the MINT-FEC mechanism managed to reduce the overhead by up to 11.74%. This is an additional proof that our mechanism is doing a better work to infer the motion intensity using spatial complexity and temporal intensity, allowing the definition of a precise amount of redundant information to the most sensitive data. Thanks to that, we are able to deliver higher video quality leading to a better user perception.

V. Conclusion and Future Works

The ever-growing use of UAVs to deliver video underlines the need for an adaptive QoE-aware mechanism to improve the resiliency to packet loss, while keeping the user perception high. The MINT-FEC enhances the video quality over long-range transmissions, providing both higher QoE and lower network overhead. This allows a thorough use of the already scarce wireless resources. The advantages and disadvantages of the proposed mechanism were highlighted through a set of experiments, proving that we were able to precisely identify the motion intensity in arbitrary video sequences and better protect the most QoE-sensitive data, which is translated in higher video quality.

The MINT-FEC outperforms the others, adaptive and non-adaptive, mechanisms in the experiments. In scenarios with up to 1200m, it was able to deliver videos with slightly higher quality and at the same time generating substantially less network overhead. This implies that an enhanced video quality was perceived without wasting wireless resources. On the other hand, over 1200m, due to harsher conditions, our mechanism starts to increase the redundancy providing a considerably higher QoE than the other mechanisms. This is definitely a desired tradeoff between network overhead and video quality.

As future work, other mobility scenarios are going to be evaluated, an analysis about the delay and buffer size impact will be conducted, as well as subjective QoE assessments.

Acknowledgment

This work was funded by the Brazilian National Counsel of Technological and Scientific Development (CNPq) and also supported by the Intelligent Computing in the Internet of Services (iCIS) project (CENTRO-07-ST24-FEDER-002003), co-financed by QREN, in the scope of the Mais Centro Program.

References

[1] R. Kumar, H. Sawhney, S. Samarasekera, S. Hsu, H. Tao, Y. Guo, K. Hanna, A. Pope, R. Wildes, D. Hirvonen *et al.*, “Aerial video surveillance and exploitation,” *Proceedings of the IEEE*, vol. 89, no. 10, pp. 1518–1539, 2001.

[2] A. Puri, “A survey of unmanned aerial vehicles (uav) for traffic surveillance,” *Department of computer science and engineering, University of South Florida*, 2005.

[3] İ. Bekmezci, O. K. Sahingoz, and Ş. Temel, “Flying ad-hoc networks (fanets): A survey,” *Ad Hoc Networks*, vol. 11, no. 3, pp. 1254 – 1270, 2013.

[4] M. Bernard, K. Kondak, I. Maza, and A. Ollero, “Autonomous transportation and deployment with aerial robots for search and rescue missions,” *Journal of Field Robotics*, vol. 28, no. 6, pp. 914–931, 2011.

[5] M. Lindeberg, S. Kristiansen, T. Plagemann, and V. Goebel, “Challenges and techniques for video streaming over mobile ad hoc networks,” *Multimedia Systems*, vol. 17, pp. 51–82, 2011.

[6] E. Frew and T. Brown, “Networking issues for small unmanned aircraft systems,” *Journal of Intelligent and Robotic Systems*, vol. 54, no. 1-3, pp. 21–37, 2009. [Online]. Available: http://dx.doi.org/10.1007/s10846-008-9253-2

[7] T. Liu and W. Liao, “Interference-aware qos routing for multi-rate multi-radio multi-channel ieee 802.11 wireless mesh networks,” *Wireless Communications, IEEE Transactions on*, vol. 8, no. 1, pp. 166 –175, jan. 2009.

[8] E. Aguiar, A. Riker, E. Cerqueira, A. Abelm, M. Mu, T. Braun, M. Curado, and S. Zeadally, “A real-time video quality estimator for emerging wireless multimedia systems,” *Wireless Networks*, pp. 1–18, 2014. [Online]. Available: http://dx.doi.org/10.1007/s11276-014-0709-y

[9] A. Nafaa, T. Taleb, and L. Murphy, “Forward error correction strategies for media streaming over wireless networks,” *IEEE Communications Magazine*, vol. 46, no. 1, pp. 72–79, 2008.

[10] L. Han, S. Park, S.-S. Kang, and H. P. In, “An adaptive fec mechanism using cross-layer approach to enhance quality of video transmission over 802.11 wlans.” *TIIS*, pp. 341–357, 2010.

[11] M.-F. Tsai, N. K. Chilamkurti, S. Zeadally, and A. Vinel, “Concurrent multipath transmission combining forward error correction and path interleaving for video streaming,” *Comput. Commun.*, vol. 34, pp. 1125–1136, June 2011.

[12] C.-H. Lin, Y.-C. Wang, C.-K. Shieh, and W.-S. Hwang, “An unequal error protection mechanism for video streaming over ieee 802.11e wlans,” *Computer Networks*, vol. 56, no. 11, pp. 2590 – 2599, 2012.

[13] R. Immich, E. Cerqueira, and M. Curado, “Improving video qoe in unmanned aerial vehicles using an adaptive fec mechanism,” in *Wireless Networking for Moving Objects*, ser. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2014 (to appear).

[14] J. Vanne, E. Aho, T. Hamalainen, and K. Kuusilinna, “A high-performance sum of absolute difference implementation for motion estimation,” *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 16, no. 7, pp. 876–883, July 2006.

[15] R. Tibshirani and G. Walther, “Cluster validation by prediction strength,” *Journal of Computational and Graphical Statistics*, vol. 14, no. 3, pp. 511–528, 2005.

[16] K.-W. Wong, D. Tikk, T. Gedeon, and L. Koczy, “Fuzzy rule interpolation for multidimensional input spaces with applications: a case study,” *Fuzzy Systems, IEEE Transactions on*, vol. 13, no. 6, pp. 809–819, Dec 2005.

[17] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: from error visibility to structural similarity,” *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, 2004.