Multimedia Traffic over Wireless and Satellite Networks

Floriano De Rango, Mauro Tropea and Peppino Fazio

DEIS Dept., University of Calabria,
Italy

1. Introduction

The rapid growth of multimedia application together with advanced development in digital technology and with the increased use of mobile terminal has pushed the research toward new network technologies and standards for wireless environment. Moreover, ever-increasing computing power, memory, and high-end graphic functionalities have accelerated the development of new and exciting wireless services. Personal video recorders, video on demand, multiplication of program offerings, interactivity, mobile telephony, and media streaming have enabled viewers to personalize the content they want to watch and express their preferences to broadcasters. Viewers can now watch television at home or in a vehicle during transit using various kinds of handheld terminals, including mobile phones, laptops computers, and in-car devices. The concept of providing television-like services on a handheld device has generated much enthusiasm. Mobile telecom operators are already providing video-streaming services using their third-generation cellular networks. Simultaneous delivery of large amounts of consumer multimedia content to vast numbers of wireless devices is technically feasible over today's existing networks, such as third-generation (3G) networks. As conventional analog television services end, broadcasters will exploit the capacity and flexibility offered by digital systems. Broadcasters will provide quality improvements, such as high-definition television (HDTV), which offer many more interactive features and permit robust reception to receivers on the move in vehicles and portable handhelds. Mobile TV systems deliver a rich variety of content choice to consumers while efficiently utilizing spectrum as well as effectively managing capital and operating expenses for the service provider. Mobile TV standards support efficient and economical distribution of the same multimedia content to millions of wireless subscribers simultaneously. Mobile TV standards reduce the cost of delivering multimedia content and enhance the user experience, allowing consumers to surf channels of content on a mobile receiver. Mobile TV standards address key challenges involved in the wireless delivery of multimedia content to mass consumers and offer better performance for mobility and spectral efficiency with minimal power consumption. An important aspect of multimedia delivery contest is the possibility of make integration between different networks in order to be able of reaching users every-time every-where. Then, it is possible to use a multi-layer hierarchic platform that use satellite [1] segment together with wireless network based on 802.11 standard [2,3,4] and with cellular network in order to have an ubiquitous coverage. In

order to grapple with the continuously increasing demand on multimedia traffic over high speed wireless broadband networks it is also necessary to make network able to deal with the QoS constraints required by users. In order to provide quality of service to user applications, networks need optimal and optimized scheduling and connection admission control algorithm. These mechanisms help to manage multimedia traffic guaranteeing QoS to calls already admitted in the system and providing QoS to the new connection. In order to evaluate the quality of video traffic with the mobility it is important to examining the quality assessment techniques: Subjective and Objective quality assessment.

2. Video quality in multimedia broadcasting

In the last few years multimedia applications are grew very fast in all networks typologies and in particular in wireless networks that are acquiring a big market slice in the telecommunication field. This trend of applications has pushed the researchers to perform a lot of studies in the video applications and in particular in the compression field in order to be able of transporting this information in the network with a low impact in the system resources. In literature a lot of studies exist on video compression and new standard of compression are proposed in order to be able to transmit video traffic on different network technologies that often have resource problem in terms of bandwidth capacity and, then, a very performance compression algorithm can give a great support to network service provider in respecting the quality constrains otherwise nor achievable.

The ubiquitous nature of all multimedia services and their use in overall telecommunications networks requires the integration of a lot of technologies that aim to improve the quality of the applications received by the users. First of all, it is clear that the traditional concept of best effort paradigm in the delivery of multimedia contents is not possible to adopt because it does not match with the users requirements. This type of approach try to do its best but it is unable of guaranteeing any form of users requirements. In order to address this type of problem, recently different Quality of Services architectures have been proposed capable of guaranteeing to the multimedia streams the users constrains. The most famous architectures are Integrated Services and Differentiated Services that manage different class of services in order to allow a traffic differentiation able to discriminate also a cost differentiation for the customers. It is easy to understand that this new type of applications is based on a constantly reliable reception of information that it is possible to have only through an appropriate network management.

The new end users are always more quality aware and then they are more exigent and accordingly networks have to guarantee always more capacity in order to satisfy their users. As a consequence, there is a continuous and extensive research effort, by both industry and academia, to find solutions for improving the quality of multimedia content delivered to the users; as well, international standards bodies, such as the International Telecommunication Union (ITU), are renewing their effort on the standardization of multimedia technologies. There are very different directions in which research has attempted to find solutions in order to improve the quality of the rich media content delivered over various network types [5,6,7,8,9].

Moreover, it is very important to determine efficient quality assessment. It is really important to know how the network behaves in terms of parameters of service for end users in order to take the correct decisions during the development, evaluation, construction and operation of network services.

Over the years have developed many models, more efficient and less expensive, able to consider more types of factors involved in the telecommunication networks. An overview of models of measurement is reported in the Figure 1 below. They are divided substantially in two categories: subjective assessment and objective assessment.
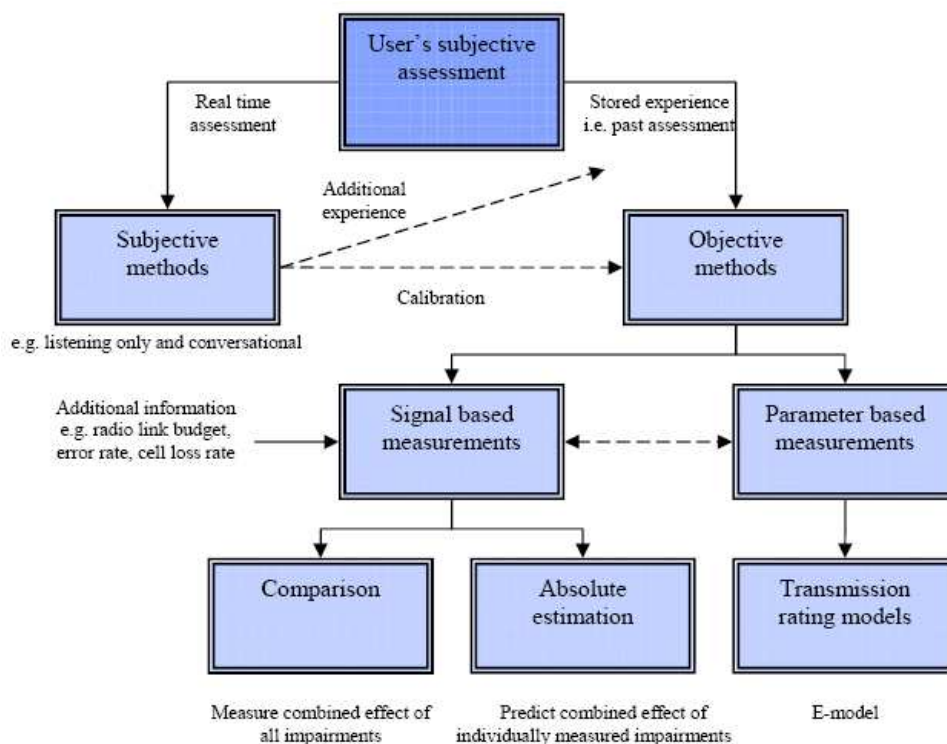

Fig. 1. Overview of models of measurement

2.1 Subjective quality assessment

A method to perform quality evaluation for multimedia applications is called subjective assessment and it represents an accurate technique for obtaining quality ratings. It is basically based on experience done by a number of users. Typically, the experiments are made in a room where there are a certain number of persons (about 15-30) that they have to watch a set of video streams and, then, they have to give a rate based on own quality perception. On the basis of all rates given by all subjects included in the experiment it is formulated an average rate called Mean Opinion Score (MOS). It is clear that, being a subjective evaluation determined by the subjectivity and variability of the involved persons, this test is affected by the personal opinion that cannot be eliminated. In order to avoid this type of problem the experiments are made through precise instructions that give to the subject a set of precautions that they have to follow. Moreover, also the environment used for the test is a controlled environment. In this way it is possible to perform a set of tests and provide a quality score that is a number that results from a statistical distribution. There are

a wide variety of subjective testing methods. In literature different methods exist for giving a measure of the perceptual performance of subjects. The ITU has formalized a series of conditions that are used for quality assessment in various recommendations [10,11,12,13,14]. They suggest standard viewing conditions, criteria for the selection of observers and test material, assessment procedures, and data analysis methods. Recommended testing procedures exist as Absolute Category Rating (ACR), Degradation Category Rating (DCR), Comparison Category Rating (CCR).

In the test the presentation order of the video clips can be randomized between viewers in order to obtain a more realistic statistical score. After each video presentation, viewers are asked to judge its overall quality using the rating scale shown in Figure 2.

Fig. 2. Subjective assessment rating scale

Normally voting period was not time-limited. After choosing their quality rating, assessors had to confirm their choice using an 'OK' button. Furthermore, this eliminated the possibility of missing ratings in the test. After each video sequence and after to have gave a vote, a neutral gray background is, often, displayed on the video terminal during some second before the next sequence is presented. The test procedure and monitor selection adhered to the latest findings and recommendations for best practice from the Video Quality Experts Group (VQEG), a technical body supporting ITU standardization activities.

The use of this type of test has some disadvantages, first of all, the result of the test depend from uncontrollable attributes like experience, the mood, the attitude and culture, then, they are very expensive and impractical if you want to do frequently because of the number of subjects and tests are necessary to give reliable results. Anyway, subjective assessment are invaluable tools for evaluating multimedia quality. Their main shortcoming is the requirement for a large number of viewers, which limits the amount of video material that can be rated in a reasonable amount of time. Nonetheless, subjective experiments remain the benchmark for any objective quality metric.

2.2 Objective quality assessment

The subjective methods are not feasible during the design of a network. These are limited, impractical and very expensive. To overcome these problems have been developed, new methods that allow the calculation of values that represent the different combinations of

factors of damage that could affect the network. The primary purpose of these methods is to produce an estimate of quality, providing the results as comparable as possible to MOS values. The ITU is proposing a method of objective testing, automatic and repeatable, which takes into account the perceived quality. Different types of objective metrics exist [15]. For the analysis of decoded video, we can distinguish data metrics, which measure the fidelity of the signal without considering its content, and picture metrics, which treat the video data as the visual information that it contains. For compressed video delivery over packet networks, there are also packet- or bitstream-based metrics, which look at the packet header information and the encoded bitstream directly without fully decoding the video. Furthermore, metrics can be classified into full-reference, no-reference and reduced-reference metrics based on the amount of reference information they require.

The most known video quality metric is called Peak Signal to Noise Ratio (PSNR) that is calculated simply as mathematical difference between every pixel of the encoded video and the original video. The other popular metric is the classical Mean Square Error (MSE) is one of many ways to quantify the amount by which an estimator differs from the true value of the quantity being estimated.

3. Multimedia over wireless networks

Nowadays multimedia communication over wireless and wired packet based networks is growing up. In the past, many applications were used for video downloads, while now they take up the share of all traffic on the Internet. Most mobile devices can actively download and upload photos and videos, sometimes in real time. In addition, Voice over IP (VoIP) is heavily changing the voice telecommunications world and the enhanced television is also being delivered into the houses over IP networks by Digital Subscriber Line (DSL) technologies. Another issue in the multimedia revolution takes place inside the home environment: the electronics manufacturers, the computer industry and its partners are distributing audio and video over local-WiFi networks to monitors and speakers around the house. Now that the analog-to-digital revolution is going to be complete, the "all media over IP" revolution is taking place, with radio, television, telephony and stored media all being delivered over IP wired and wireless networks. Figure 3 shows an example of different multimedia applications in a home environment.

The growing and the emergence of communication infrastructures, like the Internet and wireless networks, enabled the proliferation of the above mentioned multimedia applications (music download to a portable device, watching TV through the Internet on a laptop, viewing movie trailers posted on the web through a wireless link). New applications are surely revolutionary, like sending VoIP to an apparently conventional telephone, sending television over IP to an apparently conventional set top box or sending music over WiFi to an apparently conventional stereo amplifier. The exposed applications include a big variety of new multimedia related services but, unfortunately, the Internet and the wireless networks do not provide full support for multimedia applications. The Internet and wireless networks have stochastic and variable conditions influenced by many factors, however variations in network conditions can have heavy consequences for real-time multimedia applications and can lead to unacceptable user experience, because multimedia transmissions are usually delay sensitive, bandwidth intense and loss tolerant. The theory that has traditionally been taught in information theory, communication and signal processing may not be directly applied to highly time-varying channel conditions and, as a

consequence, in recent years the area of multimedia communication and networking has emerged not only as a very active and challenging research topic, but also as an area that requires the definition of new fundamental concepts and algorithms that differ from those taught in conventional signal processing and communication theory.
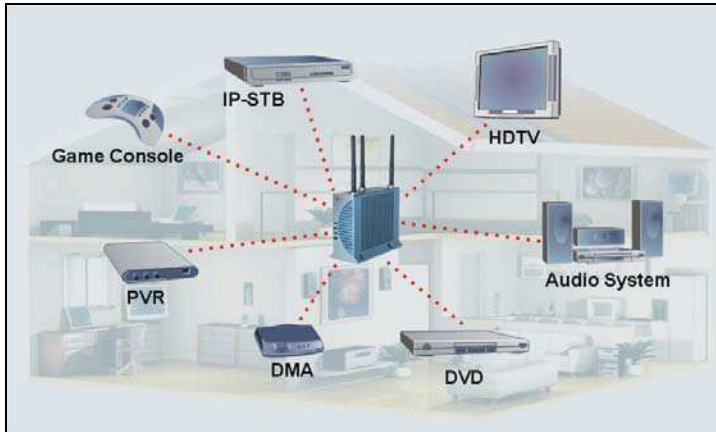

Fig. 3. Different Wireless IP-based multimedia applications (STP – Set Top Box for on-demand video, HDTV – High Definition TV, DMA – Digital Media Adapter for multimedia home extensions, PVR – Personal Video Recorder, etc.).

It is clear that Best-Effort (BE) IP networks are unreliable and unpredictable, expecially in wireless networks, where there can be many factors that affect the Quality of Service (QoS), that measures the performance of a transmission system via parameters that reflect its transmission quality, such as delay, loss and jitter. In addition, congested network conditions result in lost video packets, which, as a consequence, produces poor quality video. Further, there are strict delay constraints imposed by streamed multimedia traffic. If a video packet does not arrive before its "playout time", the packet is effectively lost. Packet losses have a particularly devastating effect on the smooth continuous playout of a video sequence, due to inter-frame dependencies. A slightly degraded quality but uncorrupted video stream is less irritating to the user than a corrupted stream. Controlled video quality adaptation is needed to reduce the negative effects of congestion on the stream whilst providing the highest possible level of service and quality. The applications based on streaming operations are able to split the media into separate packets, which are transmitted independently in the network, so that the receiver is able to decode and play back the parts of the bit stream that are already received. The transmitter continues to send multimedia data packets while the receiver decodes and simultaneously plays back other already received parts of the bit stream.

The philosophy of playing back received packets allows the reduction of the delay between the transmission time instant and the moment at which the user views the multimedia content. Having a low delay is of primary importance in such kind of systems, where interactive applications are dominant (for example, a video conference or a video on-demand architecture).

The transmission of multimedia content can be categorized into three main classes: unicast, multicast and broadcast, depending on the number of senders and receivers. Unicast

transmission connects one sender to one receiver (point-to-point connection, p2p), as downloading, on-demand streaming media and p2p telephony. The main advantage of unicast is that a feedback channel can be established between the receiver and the transmitter, so the receiver can return information to the sender about the channel conditions which can be used accordingly by the transmitter to change transmission parameters. In a multicast communication the sender is connected to multiple receivers that decide to join the multicast group; multicast is more efficient than multiple unicast flows in terms of network resource utilization, since the information must not be replicated in the middle nodes (it is obvious that in a multicast communication the sender cannot open a session toward a specific receiver). In a broadcast transmission, the sender is connected to all receivers that it can reach through the network (an example is a satellite transmission); the communication channel may be different for every receiver.

One possible approach to the problem of network congestion and resulting packet loss and delay is to use feedback mechanisms to adapt the output bit rate of the encoders, which, in turn adapts the video quality, based on implicit or explicit information received about the state of the network. Several bit rate control mechanisms based on feedback have been presented in the last few years. As the Real-Time Control Protocol (RTCP) provides network-level QoS monitoring and congestion control information such as packet loss, round trip delay, and jitter. Many applications use RTCP to provide control mechanisms for transmission of video over IP networks. However, the network-level QoS parameters provided by RTCP are not video content-based and it is difficult to gauge the quality of the received video stream from this feedback.

Now, in the following paragraphs, multimedia transmission techniques will be deeply introduced, as well as different policies dedicated to evaluate the quality (and the distortion) of the perceived content.

3.1 Background and coding

The transmission of video over wireless media is becoming very popular for the different variety of applications and networks. The advantages of these transmissions over wireless media are evident, but the transmission rate will always be limited due to the limitations introduced by physical layer. So, having the possibility of compressing the video before transmission is crucial in such kind of environments, especially for real-time traffic, where some constraints are required. For these reasons, video decoder must tolerate delay and packet losses: the standards in video coding (like MPEG-4 and H.264/AVC) [16,17,18,19] are today very popular, because of their capability to adapt to these environments.

These standards, like the previous well known ones, use a hybrid coding approach. The Motion Compensated Prediction (MCP) is combined with transform coding of the residual components in order to obtain a codec that can be very useful nowadays. As described in previous paragraph, video communications can be categorized into unicast, multicast and broadcast services, with different peculiarities, depending on the desired service, like the on-line generated or pre-encoded content, video telephony and downloading.

Figure 4 illustrates a generic scheme of a video transmission system, including the main components that go from the transmitter device to the receiver device. Data compression is feasible also because video content presents significant redundancy (in time and in space) that reduces the amount of transmitted packets significantly. According to Figure 4, the video encoder generates data units containing the compressed video stream, which is stored in the encoder buffer before the transmission. In general, the transmission system damages

(quantitatively or qualitatively) individual data units, also introducing some delay; then, the encoder/decoder buffer are used in order to balance the bit-rate fluctuations produced by the encoder and by the wireless channel. In general, a coded video stream can be considered as composed by a sequence of data units, called Access Units (AU) in MPEG-4 or Network Abstraction Layer Units (NALU) in H.264. The AUs/NALUs can be labeled as data unit specific information (as in MPEG, where they are labeled on the basis of relative importance for video reconstruction). On the other hand due to spatial and temporal prediction the independent compression of data units cannot be guaranteed without significantly losing compression efficiency.
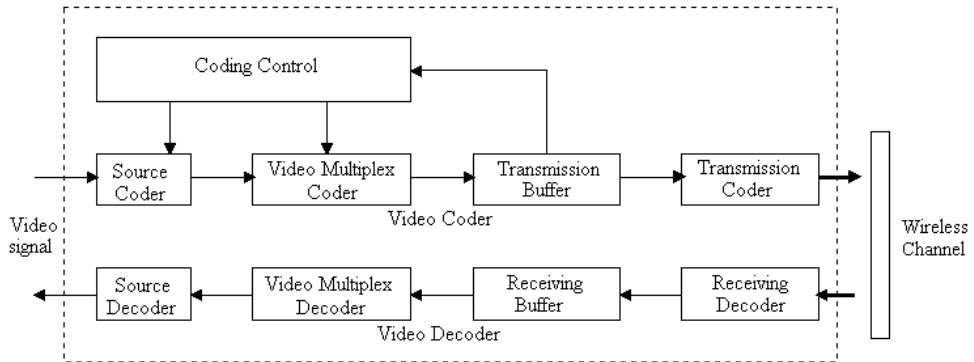

Fig. 4. Blocks diagram of a generic audio/video transmission system.

Errors introduction and its effects are considerably different in wired or in wireless networks, because of the different phenomena that impact on the medium. For wireless networks, fading and interference cause burst errors in form of multiple lost packets. Moreover congestion can result in lost packets in a wired IP network. Nowadays, even for wireless networks, systems are able to detect the presence of errors in a packet on physical layer and the losses are reported to higher layers. These techniques usually use Cyclic Redundancy Check (CRC) mechanisms. By consequence the video decoder will not receive the entire bit-stream. Intermediate protocol layers such as User Datagram Protocol (UDP) might decide to completely drop erroneous packets so deleting all the encapsulated data units. In fact, video data packets are considered lost if their delay overcomes a fixed and tolerable threshold, defined by the user video application.

Figure 5 illustrates a typical simplified version of an end to end video system when the MCP compressed video is transmitted over non error-free channels. In this environment, t represents the time, S_t is a single video frame composing a network packet P_t, C_t (with values 1 or 0) indicates if P_t is correctly received or discarded, while $S_t(C)$ represents the decoded frame as a function of the channel error pattern. Let assume that a packet is transmitted over a channel that forwards correct packets to the decoder, in case of successful transmission, the packet is forwarded to the normal decoder operation such as Entropy Decoding and Motion compensated Prediction.

The prediction information and transform coefficients are reloaded from the coded bit-stream to reconstruct the current frame S_{t-1}. After that, the frame is forwarded to the display buffer and also to the reference frame buffer to be used in the MCP process to reconstruct
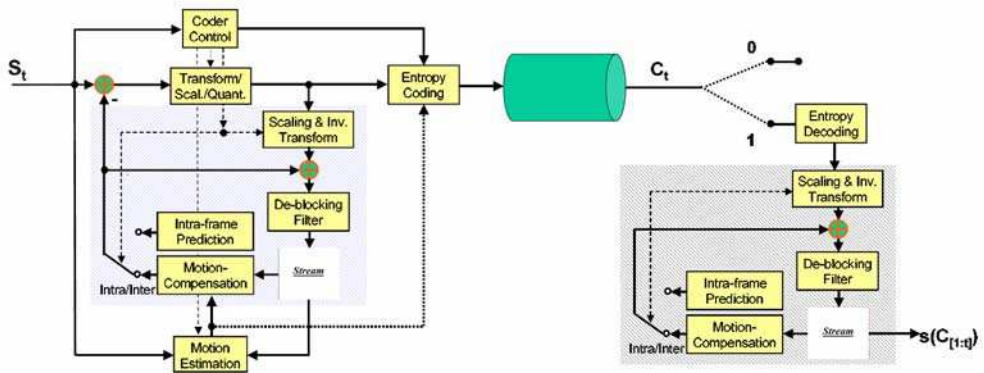
Fig. 5. End to end video transmission system.

the following inter-coded frames (i.e. the frame st frame). When the packet P_t is lost, i.e. at the reference time t, $C_t = 0$, the Error Concealment (EC) is necessary to be enabled, so the decoder just avoids the decoding operation and the display buffer is not updated, so the displayed frame is still S_{t-1}. In this case the viewer will understand that there has been a loss of motion, since the continuous display update is not maintained. Also the reference frame buffer is not updated as a result of this data loss.

In case of successful reception of packet P_{t+1}, the inter-coded frame S_{t+1}, reconstructed at the decoder, will in general not be identical to the reconstructed frame S_{t+1} at the encoder side, because as the encoder and the decoder refer to a different reference signal in the MCP process, there will be a reconstruction mismatch in reference signal when decoding S_{t+2}. For this reason it is obvious that the loss of a single packet P_t affects the quality of all the inter-coded frames: S_{t+1}, S_{t+2}, S_{t+3}, etc. This phenomenon is present in any predictive coding scheme and is called error propagation. If predictive coding is applied in the spatial and temporal domains of a sequence of frames, it is referred to as spatio-temporal error propagation.

The MCP technique makes the reconstructed frame S_t not only depending on the actual channel behavior C_t, but also on the previous channel evolution $C[1..t]$.

After these considerations, remarking that the error propagation has direct consequences on the perceived video content, it is preferable that a video coding system provides some of the following features:

- a reliable communication mean, in order to avoid transmission errors;
- an algorithm or device dedicated to the reduction of the visual effects of errors in the received frames;
- an algorithm or device for the minimization of the error propagation effect.

Generally, in MCP coding, the video content that belongs to a single frame is not encoded as a single entity: it is composed by a macro-block and individual data units are syntactically accessible and independent. The features that are employed in these systems to correct the introduced errors are the Forward Error Correction (FEC) and Backward Error Correction (BEC) or any combinations of those. It is very important, at the receiver size, that the erroneous and missing video content is observed, localized and, if possible, eliminated or minimized.

The modern video coding systems also have the capability to inform the video encoder about the loss of the video content, so the encoder can adapt itself to enhance the

transmission quality. The macro-blocks assignments, error control methods and feedbacks transmission, for example, must be used if we desire a robust, efficient and 'error-free' application.

As known in literature, if Shannon's separation principle [20] is observed, the main goal in video transmission (that is low error or error-free propagation) can be reached: the compression (or coding) features and the transmission (or link layer) ones can be optimized separately in order to avoid losses in video transmissions.

However, in several applications and environments (low delay situations), error-free transport may be impossible and the features like channel loss correction, detection and localization of unavoidable errors, their minimization, reduction of distorted frames impact become essentials. When we are dealing with 'no-wired' systems and QoS guarantees must be given, error control such as FEC and retransmission protocols are the primary features that should be provided.

The rules that are provided by the video coding standards, such as H.263 [21], MPEG-4 [22] and H.264, beyond the syntax and the semantics of the transmitted data, specify the decoder behavior in case of reception of an error-free bitstream, so the deployment of video coding standards still provides a significant amount of freedom for decoders that have to process erroneous bitstreams. In this way, a processing algorithm may be better than another one, in terms of computational complexity or quality of the received content. In the last years, video compression tools have evolved significantly over time: previous standards, like H.261, MPEG, MPEG-2 [23, 24], are very limited in error resilience capabilities, while the latter ones (starting from the H.263 [25]), heavily influenced the applications, with a lot of improvements and tools for error resilience. In the meantime, the new emerging standard MPEG-4 Advanced Simple Profile (ASP) introduced a radically different approach, based on some resilience tools, like Reversible Variable Length Coding (RVLC) and Re-synchronization Markers (RM) were introduced [26]. Figure 6 shows the evolution in time of the video compression standards.
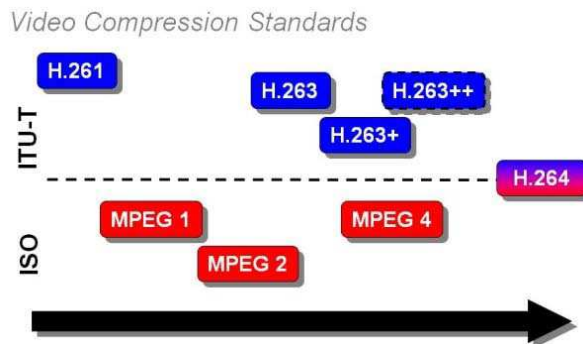

Fig. 6. The evolution of video compression standards.

As exposed in previous paragraphs, it is evident that the main aim of a communication system is to transfer the data generated by an information source efficiently and reliably over a non-ideal channel, respecting some QoS constraints. Among the different employed components (encoder, modulator, demodulator, source decoder) a channel encoder is used to insert additional and redundant data to the information sequence, so that channel errors can be detected or corrected. Shannon's channel coding theorem affirms that if there is

channel capcity (that is to say it is larger than the data rate), a coding scheme can be useful to reduce error probabilities. In general, the Bose-Chadhuri-Hocquenghem (BCH) codes are used and they represent an Error Control scheme based on block coding. They are a generalization of the Hamming codes and they add redundancy bits to the payload, in order to form code words; they also introduce the capacity of error correction (limited in number) [27]. The non–binary BCH-family codes that are massively used are the Reed Solomon (RS) codes. RS codes groups the bits into symbols, achieving good burst error suppression capability.

In order to achieve low-error (or error-free) communication, channel coding schemes must be implemented for the worst case channel characteristics. Obviously, when the channel is currently good, the channel protection, dimensioned for the worst case conditions, results in inefficient utilization of the wireless link. In addition, complex hardware or software structures may be required to realize the powerful long codes, required to defeat the worst case error patterns.

These drawbacks can be overcome by using the *Adaptive Coding* (AC) [28,29], by adding the redundancy on the basis of channel conditions and characteristics. Different adaptive algorithms have been proposed for video streaming applications, as the one proposed in [28]: in that work the fading level is estimated and, then, the algorithm evaluates the proper coding ratio in order to actively protect packets from losses. In this way, channel utilization can be immediately and efficiently improved.

3.2 Error perception

As discussed earlier, packet losses affect the quality of multimedia communications over wireless packet networks and the amount of quality degradation strongly vary on the basis of the meaning of the lost data. For the designing of an efficient loss protection mechanism, a reliable estimation method for multimedia data is needed. Providing an accurate estimation algorithm is required and in [30] the importance of a video coding element is outlined: for example, it is very useful to introduce the macro-block or packet concept as a value directly related to the distortion that would be introduced at the decoder by the loss of that specific element.

In Figure 7 the main phases of the "Analysis By Synthesis" (ABS) algorithm are illustrated; it is useful to compute the possible distortion of each element. It consists of the following steps, applied for each packet:

-   decoding, with concealment, of the bitstream simulating the loss of the packet being analyzed (synthesis stage), by adding some errors;
-   quality evaluation, in order to compute the distortion caused by the loss of the packet. The source and the reconstructed pictures are compared using Mean Square Error (MSE);
-   storage of the obtained value as an indication of the perceptual importance of the analyzed video packet.

Little modifications of the standards are necessary in order to implement the modified encoding process. The algorithm reconstructs the encoded frames by the simulation of the decoder operations. It can be surely used for video coding, but the obtained values depend on the adopted encoding (if the video will be compressed with a different encoder or if a different packetization policy is used, values will be very different. The main disadvantage of the exposed technique consists of the interdependencies usually present between data
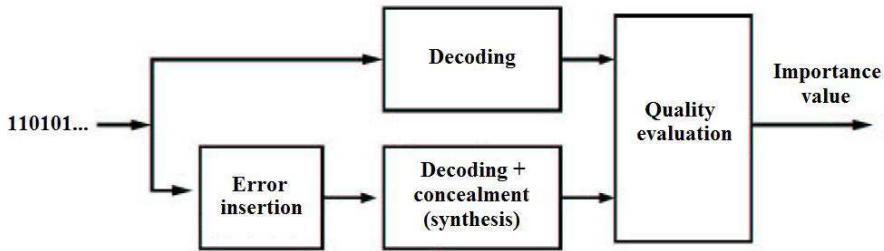
Fig. 7. Analysis by synthesis blocks diagram.

units, so the simulation of the loss of an isolated data unit is not completely realistic, particularly for high packet loss rates: all possible combinations of the events should be considered, weighted by its probability, and its distortion computed by the ABS technique, obtaining the expected importance value. The application of the ABS algorithm is easier when considering elements of the video stream which do not influence next frames (that is to say, no error propagation is present). If propagation is present, the distortion introduced in next frames should be evaluated until it is negligible (for example until a *I* frame is reached in a MPEG stream). In this case, the application of the ABS scheme is harder, due to the need of pre-computation of the perceptual distortion.

Another scheme for distortion evaluation is represented by the Distortion Matrix (DM) [31]: it allows to compute the distortion introduced when some frames are dropped in a sequence of Group Of Pictures (GOP) (as in an MPEG-2 stream). It is calculated under the assumption that once a specific P frame or I frame is lost, all the depending frames in the current GOP are replaced with the latest successfully decoded frame. This assumption makes the (DM) model unsuitable, because today all video decoders tend to mitigate the error propagation, reducing the distortion after a single loss. The GOP structure of video k is described by the GOP length L^k and the number of B-frames B^k between two I or P frames. For example, with $L^k = 9$ and
$B^k = 2$, the GOP structure will be I B B P B B P B B. The obtained distortion matrix as in [31] is shown in Figure 8.

$$
\begin{array}{c}
R: \\
I: \\
P_1: \\
P_2: \\
B_1: \\
B_3: \\
B_5:
\end{array}
\left[
\begin{array}{ccccccccc}
D_I^R & D_{B_1}^R & D_{B_2}^R & D_{P_1}^R & D_{B_3}^R & D_{B_4}^R & D_{P_2}^R & D_{B_5}^R & D_{B_6}^R \\
/ & D_{B_1}^I & D_{B_2}^I & D_{P_1}^I & D_{B_3}^I & D_{B_4}^I & D_{P_2}^I & D_{B_5}^I & D_{B_6}^I \\
/ & / & / & / & D_{B_3}^{P_1} & D_{B_4}^{P_1} & D_{P_2}^{P_1} & D_{B_5}^{P_1} & D_{B_6}^{P_1} \\
/ & / & / & / & / & / & / & D_{B_5}^{P_2} & D_{B_6}^{P_2} \\
/ & / & D_{B_2}^{B_1} & / & / & / & / & / & / \\
/ & / & / & / & / & D_{B_4}^{B_3} & / & / & / \\
/ & / & / & / & / & / & / & / & D_{B_6}^{B_5}
\end{array}
\right]
$$

Fig. 8. DM for Lk = 9 and Bk = 2.

The entries in the distortion matrix are the MSE values observed when replacing frame F_{loss} as part of the concealment strategy. The column left to the distortion matrix shows the

replacement frame for every row of the matrix. R is a frame from the previous GOP that is used as a replacement for all frames in the current GOP if the *I-frame* of the current GOP is lost. From this matrix, the resulting distortion for any possible loss

pattern can be determined. The total distortion for the GOP is computed as the sum of the individual frame loss distortions. This matrix can be determined during the encoding of the video. The number of columns of the distortion matrix corresponds to the GOP length L^k. For more details see [31].

The ABS and DM end to end distortion estimation techniques can be categorized as "frame-based" methods, because they concern the distortion introduced at the frame level; other distortion evaluation techniques are well-known in the literature and they can be classified as "pixel-based" methods. In particular, the block-based approach generates and recursively updates a block-level distortion map for each frame [32,33,34]. Nevertheless, since inter-frame displacements influence sub-block motion vectors, a motion compensated block may inherit errors propagated from prior frames. In contrast, pixel-based policies estimate the distortion on a "pixel-basis", so they have the advantage of providing high accuracy. Obviously, on the other hand, the complexity goes increasing. An innovative approach has been proposed in [35], where the distortion on pixel-basis is calculated by exhaustive simulation of the decoding procedure and averaging over many packet loss patterns. Another scheme is illustrated in [36], where only the two most likely loss events are considered.

However, in [37] it has been demonstrated that, by using the ROPE scheme, low complexity can be preserved, without losing the optimality of the distortion estimation algorithm. ROPE recursively calculates the first and second moments of the decoder reconstruction of each pixel, while accurately taking into account all relevant factors, including error propagation and concealment.
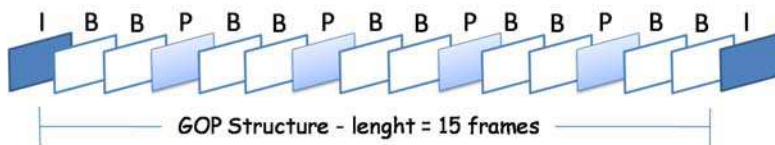
4. MPEG Standards

Fig. 9. GOP Structure

MPEG is the acronym of Moving Picture Experts Group, a working group which has the role of developing video and audio encoding standards [38].

MPEG video traffic is characterized by constant transmission rate of two groups of picture (GOP) (see Figure 9) per second and 15 frames per GOP. Since the number of bytes in a frame is dependent upon the content of the video, the actual bit rate is variable over time. However, the MPEG video supports also the constant bit rate (CBR) mode. There are three types of frames:

- I-Frames (intraframes)—encoded independently of all other frames;
- P-Frames (predictive frames)—encoded based on immediately previous I or P frames;
- B-Frames (bidirectionally predictive)—encoded based on previous and subsequent frames.

Many works in literature faced the problem of analyzing and describing the structure of MPEG-2 traffic streams, trying to find a way to emulate them through statistic and stochastic streams generators, preserving the proper and intrinsic nature of the original stream. In [39] the input process model is viewed as a compromise between the Long Range Dependent (LRD) and short range dependence (SRD) models. Simulation results were found to be better than those of a self-similar process when the switch buffer is relatively small. The MPEG video model presented in [40] is a Markov chain model based on the 'Group of Pictures' (GOP) level process rather than the frame level process. This has the advantage of eliminating the cyclical variation in the MPEG video pattern, but at the expense of decreasing the resolution of the time scale. Typically a GOP has duration of a half second, which is considered long for high speed networks. Of particular interests in video traffic modeling are the frame-size distribution and the traffic correlation. The frame size distribution has been studied in many existing works. Krunz [41] proposed a model for MPEG video, in which, a scene related component is introduced in the modeling of I frames, but ignoring scene effects in P and B frames. The scene length is i.i.d. with common geometric distribution. I frames are characterized by a modulated process in which the local variations are modulated by an Auto-Regressive (AR) process that varies around a scene related random process with log-normal distribution over different scenes; i.e., two random processes were needed to characterize I frames. The sizes of P and B frames were modulated by two i.i.d. random processes with log-normal marginal. This model uses several random process and need to detect scene changes, thus complicating the modeling process. In [42,43] the adaptive source is modeled by means of a discrete-time queuing system representing a virtual buffer, loaded by the video source when its quantizer scale parameter is changed according to the feedback law implemented in the encoder system. The whole paper is based on the Switched Batch Bernoulli Process (SBBP) that has been demonstrated to be suitable to model an MPEG video source; in fact, being a Markov modulated model, an SBBP is able to capture not only the first-order statistics but also the second-order ones which characterize the evolution of the movie scene.

In this paper we introduce a new concept of GOP-rate modeling, based on the discretisation method originally proposed in [44] for the wireless channel study. After the GOP-rate trend has been analyzed for the whole duration of the stream, it has been discretised in a certain number of states. Then the associated Markov Chain parameters have been evaluated.

MPEG algorithms compress data to form small bits that can be easily transmitted and then decompressed. It achieves its high compression rate by storing only the changes from one frame to another, instead of each entire frame. The video information is then encoded using a technique called Discrete Cosine Transform (DCT). MPEG uses a type of lossy compression, since some data is removed. But the diminishment of data is generally imperceptible to the human eye. The major MPEG standards include the following:

- MPEG-1: The most common implementations of the MPEG-1 standard provide a video resolution of 352-by-240 at 30 frames per second (fps). This produces video quality slightly below the quality of conventional VCR videos.
- MPEG-2: Offers resolutions of 720x480 and 1280x720 at 60 fps, with full CD-quality audio. This is sufficient for all the major TV standards, including NTSC, and even HDTV. MPEG-2 is used by DVD-ROMs. MPEG-2 can compress a 2 hour video into a few gigabytes. While decompressing an MPEG-2 data stream requires only modest computing power, encoding video in MPEG-2 format requires significantly more processing power.

- MPEG-3: Was designed for HDTV but was abandoned in place of using MPEG-2 for HDTV.
- MPEG-4: A graphics and video compression algorithm standard that is based on MPEG-1 and MPEG-2 and Apple QuickTime technology. Wavelet-based MPEG-4 files are smaller than JPEG or QuickTime files, so they are designed to transmit video and images over a narrower bandwidth and can mix video with text, graphics and 2-D and 3-D animation layers. MPEG-4 was standardized in October 1998 in the ISO/IEC document 14496.
- MPEG-7: Formally called the Multimedia Content Description Interface, MPEG-7 provides a tool set for completely describing multimedia content. MPEG-7 is designed to be generic and not targeted to a specific application.
- MPEG-21: Includes a Rights Expression Language (REL) and a Rights Data Dictionary. Unlike other MPEG standards that describe compression coding methods, MPEG-21 describes a standard that defines the description of content and also processes for accessing, searching, storing and protecting the copyrights of content.
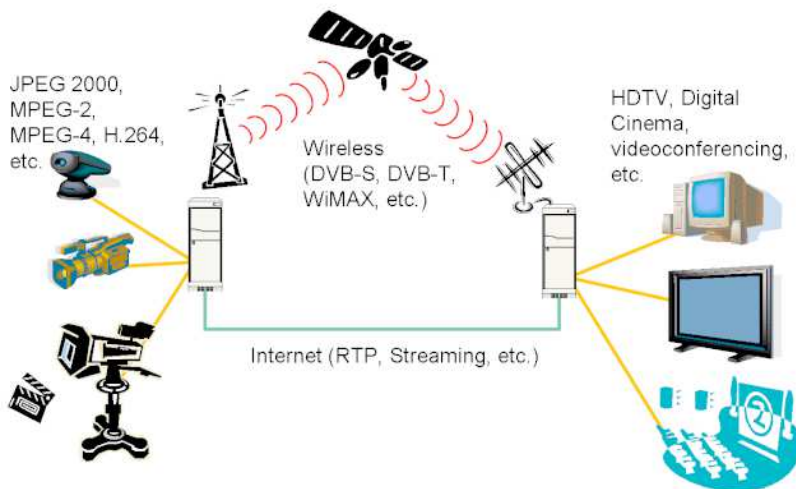
5. Multimedia over satellite

Fig. 10. Multimedia Broadcast Satellite Scenario

The new tendency of the multimedia applications on telecommunication networks is a full interactivity between users and network (see Figure 10). Thanks to this interactivity, users are able to manipulate what they receive on their terminal requiring to the network what they want. The interactivity has changed the way to exploit the network that pass from an asymmetry to a symmetry network. Then it is possible to use a return channel in order to make the users choices. This evolution in telecommunication also have consequences for satellite communications that, of sure, is the most oriented broadcast medium. Originally, the return channel of the satellite networks was designed as a terrestrial channel through different technologies such as PSTN, ISDN, GSM. Nowadays the most promising technology used for the satellite interactivity is a satellite one, in fact, the new standards for

satellite communications propose a return channel via satellite (RCS) like the DVB-RCS standard. The reasons for this choice are a lot, firstly customers prefer have a single technology for simplicity of management. It is very simple to have a single box in which all technical equipment are concentrates without have the necessity of interact with different objects. Another important reason regards the increased traffic in terrestrial networks that can produce problem in services providing such as service blocks and consequently reduction in quality of service. Moreover, the use of return satellite channel guarantees a greater bit rate available for the applications reaching a 2 Mbps against the few hundred of kbps of the terrestrial solutions. An example of this greater available bit rate can be seen on a file transfer application that for a 100 Mbyte file it will need about 7 minutes against the about 3 ½ hours of the terrestrial lines that operates a 64 kbps. There is also an advantage both for the users and the operators, that is to have both channels on the same medium. This produces a better control in QoS applications and in network management by operator, in fact, the terrestrial infrastructure is often not controlled by the same operator as for satellite, and this is certainly true when national boarders are crossed.

Until 60 years ago, each individual communicated with about 100 other persons, of which 80–90 percent lived in closed vicinity. Twenty years ago this picture changed: each individual communicates with about 500 other individuals, of which 80–90 percent do not live in close vicinity. This was made possible by the emergence of advanced communications systems and by the integration of different technologies of communication that have allowed to reach users in very isolated areas. The emergence of new technologies has allowed the integration of different devices that originally worked separately as computers, TVs, telephones. Thanks to the Internet and to communication technologies always more sophisticates all consumers devices are able to work in a merged manner. The rapid technological advances will allow in a next future to have a exchange of information in every places and in every time.

Fig. 11. DVB Satellite Terminal

Satellite plays a key role in the telecommunications networks. It is able to resolve the problem of last mile providing connections to those areas where no investment return will be possible because a large investment is required to bridge services between the local exchange and the customers. Satellite, thanks to its broadband nature, is able to provide

connection in the rural areas and isolated areas with the same investment of other areas. Users has only to install a satellite terminal (see Figure 11) and subscribe to the service and they are able to receive satellite information. Moreover, thanks to the new RCS standard the customers can also use the network interactivity exploiting the return satellite channel that is faster than terrestrial one which is based on telephony network that are limited in providing high bit rates to subscribers. It will be necessary to perform a fiber cabling also between local exchange and subscribers in order to make faster the connections. Big investments are required by telephone companies to perform a fiber cabling in order to guarantee high bandwidth to each subscriber.

It is clear that the increase of multimedia services poses the problem of provide high bandwidth to the users by operators and in this context the use of satellite platform can represents an optimal solution in terms of costs. Moreover the satellite segment guarantees also an overall integration with all communication technologies.

5.1 Digital Video Broadcasting (DVB) system

In the mid-eighties, the ability to transmit digital images was still remote and it was thought that it was neither technically nor economically feasible, at least in the short term. The main reason was the high bit rate required, especially for the transmission of digital images in motion (from 108 to 270 Mbps). The most important issue was to improve the quality of the TV, so huge amounts of capital were invested in research and development of IDTV (improved Definition Television) and HDTV (High Definition Television). The situation in the early nineties has completely changed: the creation of efficient compression algorithms has resulted in immediate result, the birth of the standard JPEG image for fixed and then the MPEG standard for moving images; thanks to MPEG compression the amount of data required for transmission of digital images up to decrease a bit rate of between 1.5 and 30 Mbps has been reduced drastically, depending on the resolution chosen and the type of content of the transmitted images. Soon we saw that digital television would have satisfied the requirements of suppliers of services, but only if they had adopted a common standard, which in 1993 gave rise to the DVB Project [45]. DVB, short for Digital Video Broadcasting, is used with reference to digital television services in accordance with the standards developed by a consortium of 300 organizations (both public and private) of more than 50 countries, operating in different sectors: from production to broadcast TV of television sets by the rules of the frequency spectrum to the study of protocols for access to a network. The members of this organization work with the DVB project to develop a set of standards, technical recommendations and guidelines available to the various manufacturers. Once given the specifications, these standards are published by ETSI (European Telecommunication Standards Institute), and then made accessible to all. So thanks to these open standards, manufacturers can create interoperable DVB systems and can also easily adapt to different transmission channels (satellite, terrestrial, cable, etc.). Although it was born for the European landscape, the DVB platform is beginning to be accepted as a world standard solution, a number of radio and television programs based on DVB standards are currently operating in North and South America, Africa, Asia and Australia.

DVB means is based on the standard ISO 13818 encoded MPEG-2 and multiplexing specifications. It defines how to transmit the signals using MPEG-2 satellite, cable and terrestrial repeaters, and as transmitting the information system, the program guide, etc.. DVB allows the broadcasting of "data containers" that can include digital data of any type, a

Thank You for previewing this eBook

You can read the full version of this eBook in different formats:

- ➢ HTML (Free /Available to everyone)

- ➢ PDF / TXT (Available to V.I.P. members. Free Standard members can access up to 5 PDF/TXT eBooks per month each month)

- ➢ Epub & Mobipocket (Exclusive to V.I.P. members)

To download this full book, simply select the format you desire below