Special Issue

An Integrated Multiple Access and Hierarchical Coding Scheme for Video Communication on Wireless Networks

RAFFAELE BOLLA, ALESSANDRO ISCRA, MARIO MARCHESE *, CARLO NOBILE, SANDRO ZAPPATORE
Dept. of Communications, Computer and Systems Science (DIST), University of Genoa, Via Opera Pia 13, Genova, Italy
{lelus,iscra,qan,zap}@dist.unige.it

Abstract. A Medium Access Control (MAC) protocol of the Reservation Random Access (RRA) family is integrated with a hierarchical video-coding scheme, to realize effective real-time video services over the up-link of a mobile wireless network. For every picture or for each group of contiguous pictures (GOP), the encoder generates an information set organized in a hierarchical layered structure. The base layer contains the information needed to a low quality video; the other layers contain data, which, if received and decoded, can enhance the video quality. The basic idea is to reserve bandwidth only to grant the transmission of base layer information, and modify the MAC scheme to manage the video enhancement data together with voice/data streams, by adaptively transmitting the incremental video information only if the performances of other traffic classes will not drop below a certain threshold. The aim of this proposal is the improvement of the video stream quality in low load conditions, without a significant reduction of other traffic classes performance during high load periods. Several simulation results show the effectiveness of the proposed method.

1 INTRODUCTION

Real-time multimedia networking services have received a lot of attention in the recent years, both in cabled (see [1-3], among others) and wireless [4-6] communication environments. In the cabled network field, for example, there is an evident effort spent in proposals and research by the scientific community to develop Quality of Service (QoS) control techniques for effective real-time communication over the Internet [7-11]. Nowadays, Networked Multimedia Applications (NMA) represent also a relevant issue for wireless and mobile environments. Among the different classes of traffic that can be considered in multimedia communications, video represents at the same time the most problematic and, potentially, the most flexible one. It is problematic because, in general, it requires a large amount of resources (e.g., bandwidth and computational power); it is flexible because video can be coded in many different ways, each of them characterized by an output bit-rate, whose values vary over a wide range. Obviously, the QoS perceived by users (P-QoS) decreases with the reduction of the output bit-rate.

In this paper a new approach for a joint source-channel coding scheme is proposed. It employs a hierarchical video coder, integrated with a MAC mechanism of the RRA family. The scenario consists of a group of mobile stations that potentially can generate three different classes of traffic (video, audio, and data) to be transmitted over the up-link channel of a wireless network.

In a wireless environment, large bandwidth requests are more difficult to manage and satisfy. As a consequence, the usage of an adaptive multi-rate video coding, often proposed for cabled networks [12-14], can be largely motivated in order to provide real-time video services.

As concerns the design of an effective digital coder, the related problem represents an issue of growing interest, specifically if the coder has to be involved in applications employing networks unable to grant the data delivery, or an upper bound to the transmission delay. Many approaches have been proposed in the literature in the past few years [15-20]. Moreover, some constraints must be taken into account while designing a packet video system: i) the bandwidth availability can significantly change over time, according to the aggregate network traffic load; ii) some video data packets may be lost during the transmission; iii) the coding scheme must be characterized by a high efficiency and a relatively low computational burden, compatible with real time requirements. In order to cope with the above constraints, a video digital coder should be designed to be scalable, i.e., capable of adaptively changing its output quality according to the available transmission bandwidth, and robust enough to guarantee good

* CNIT - Italian National Consortium for Telecommunications, Univ. of Genoa Research Unit, Via Opera Pia 13, Genova, Italy, e-mail: *mario.marchese@cnit.it*.

quality even in the presence of a moderate packet loss. The basic idea of the video system adopted in this paper on the mobile stations is an extension of the progressive coding paradigm, well known in the field of still image compression (see for instance [21, 22]). The coder is organized according to a layered architecture and generates a set of incremental information for every picture or Group of Contiguous Pictures (GOP). The base layer produces a very low bit-rate video stream. The enhancement layers provide progressive refinements, which integrate the information related to the base layer stream by adding temporal and spatial details, which, if received and decoded, enhance video quality.

Concerning the MAC and bandwidth allocation mechanisms able to take into account multimedia flows, several proposals for mobile wireless networks model video and voice as Constant Bit Rate (CBR) sources. In these cases, a fixed amount of bandwidth for each connection is reserved, without exploiting the potentially bursty characteristics of these traffic types. On the contrary, the RRA schemes were introduced to effectively transmit both data and packetized voice, by means of Speech Activity Detectors (SAD) over wireless channels, by taking into account the bursty nature of the ensuing traffic. They mix the random access and TDMA techniques, by substantially using a random access approach to manage the channel contention of data packets and of the first packets of every voice burst. Actually, the latter packets act as reservation ones: if one of them is transmitted with success, it reserves a slot per frame, until the overall voice burst has been completely transmitted. The oldest and best-known MAC protocol in this family is PRMA [23, 24], which uses Aloha as random access scheme. Several modifications of the basic mechanism have been proposed, to integrate different protocols for voice and data [25, 26], to enhance system performance under heavy traffic load conditions [27, 28], and to accommodate Variable Bit Rate (VBR) [29, 30] or Available Bit Rate (ABR) traffic [31]. We have adopted another variant, named RRA- Independent Stations Algorithm (RRA-ISA) [32], which uses an efficient and controllable MAC technique, aimed at maximizing the throughput over the next frame, by exploiting a dynamic estimation of packet presence probabilities at the mobile stations.

The idea, which the paper is based on, is to put together the know-how concerning MAC RRA protocols and layered coding: the bandwidth of the base layer information is guaranteed, while enhancement layer information competes with best-effort data traffic to get resources. The aim of the algorithm is varying the bit-rate entering the network. The rate changing may be performed, as clearly identified in [31], directly "by the sources (typically based on some input conditions) or by the system, in response to some network/system conditions". Our paper applies this observation and proposes an operative method to manage video, audio and data sources over the up-link. As the data stations are assigned a certain transmission probability (i.e., the probability of a packet presence in the transmission queue), upper layer video sources are considered as virtual stations and the probability of having a packet to transmit acts as a threshold, becoming an important parameter to improve the performance of the mechanism. Its choice is performed by using information on the quality of service really perceived by the users (P-QoS, Perceived – Quality of Service). An off-line MOS (Mean Opinion Score) method on video flows has been adopted to get the measures of quality.

The performance analysis has been carried on by simulation. An application environment of only video and data is considered first, followed by a more complete and complex environment involving voice traffic, too. Advantages and drawbacks of both situations are envisaged and highlighted.

The paper is organized as follows. The next section contains the description of the hierarchical coding scheme, while the MAC mechanism is described in Section 3. In Section 4 some simulation results are reported and the performance of the overall system is discussed. Finally, conclusions are drawn in Section 5.

2 THE HIERARCHICAL CODING SCHEME

In this section the general features of a mobile radio channel are first examined; then a model of the physical performance of the channel is presented that enables the evaluation of the channel error rate.

2.1 BROADBAND MOBILE RADIO CHANNEL CHARACTERISTICS

The proposed hierarchical video encoder has been designed to dynamically change its output bit rate in order to meet the actual network capacity. This goal is achieved by means of a layer-structured codec, where the delivery is granted of only a part (the base layer) of the overall video traffic offered to the channel.

Due to the intrinsic real-time nature of a video flow, no automatic retransmission policy can be adopted, because of the severe delay constraints. In our network environment, a promising approach may consist of organizing the coded output in a set of streams, each of them characterized by a different level of detail. In other words, several streams at different average bit rate can be derived from the main one, by simply cutting off some amount of information related to a specific layer. In such a way, only the base layer stream is transmitted by means of groups of reserved slots; vice versa, the enhancement layers' streams are associated with "special virtual stations", a concept that will be clarified in the next section. Therefore, layers carrying image details will be transmitted according to the actual network capacity, only if resources, in terms of bandwidth, are available.

The proposed encoder is fed by a QCIF image sequence and produces five bit streams as output: the first one bears the basic, low quality/resolution information, while the subsequent bit streams carry auxiliary data, useful to progressively enhance the decoded frame resolution and the quality at the receiver end.

The encoder makes extensive use of the JPEG compression algorithm, suitably tuned to achieve a good redundancy reduction. The main reasons to base the encoder on a JPEG coding module are the following: i) the JPEG coder provides good compression ratios with a reasonable computational burden; ii) it allows to separate coding of luminance and chrominance, even with different quality factors; iii) the coding scheme can be easily adapted to efficiently process difference images rather than original ones, by simply modifying the quantization of the DCT coefficients. Other coding approaches, i.e., MPEG or M-JPEG, have been tested during the experimental phase. For the sake of completeness, a comparison of our encoder with a M-JPEG coder, characterized by a quality factor equal to the highest one of our system, is summarized. The results obtained show that the proposed coder requires a noticeably lower bit rate when only the first three layers are involved. On the other hand, the bit rate associated with the M-JPEG stream is about 20% lower than the overall load requirement of our coder when all streams are transmitted. Actually, the choice of using a custom JPEG-based scheme is due to the fact that the other tested schemes exhibited higher computational complexity, lower robustness with respect to the packet loss, so frequent in a wireless environment, and lower scalability. In more detail, concerning scalability, the layered coding requires a flexible and suited to the aim coding scheme. In this respect, layered coding has been chosen, due to the simplicity of these types of schemes to cut off parts of the traffic without any intervention on the sources. Anyway, other coding techniques, suited for a layered architecture, may also be used.

The coder works as follows. The QCIF image sequence is down-sampled, divided into GOPs, and then differentially JPEG coded, in order to produce the base layer, low-resolution, bit stream (B_1). Notice that the differential images (inter frames) of the same GOP are evaluated by subtracting the current picture from the first image of that GOP (intra frame). This allows a simple reduction of the output bit rate by means of a trivial time subsampling (i.e., packet frame dropping) of the base layer encoded frame sequence. On the other hand, only a partial benefit in terms of rate reduction can be achieved, as differences are computed by using a fixed reference picture (the first one of the GOP).

The enhancement layers' bit streams (B_2, B_3) are generated by JPEG encoding the luminance and chrominance errors, respectively, between the original pictures and the corresponding ones reconstructed by decoding the base layer bit stream. If, at the receiver, B_2 is combined with the base layer bit stream B_1, an enhancement of luminance image quality shall result. Furthermore, if B_2 and B_3 streams are added to B_1, both the luminance and chrominance quality shall be improved.

Since a non-full quality JPEG coding is applied to the error images, an amount of information still remains to be fruitfully coded and transmitted, if some bandwidth is available, with the B_4 and B_5 bit streams. The latter are obtained with the procedure used to produce the previously mentioned B_2 and B_3 streams. Obviously, as concerns the end quality of the decoded sequence, B_4 and B_5 bit streams play the same roles as B_2 and B_3, respectively. The encoder has been tested on some different image sequences, for example reproducing a speaker on a still background (see Fig. 1, reproducing three images obtained by decoding B_1, $B_1+B_2+B_3$ and all the layer streams, respectively).

As concerns the output bit rate, the performance of the system is summarized in Table 1. The latter presents the bit rates associated with the five layers at different frame rates (first row). Regarding the quality evaluation of the decoded sequence, some preliminary results of the Mean Opinion Score (MOS) are reported in Section 4.

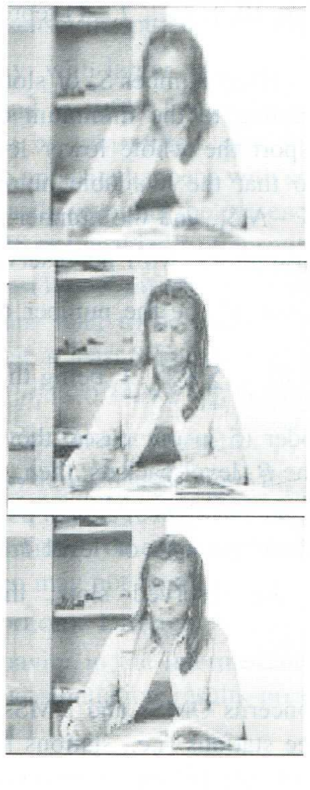

Figure 1: A frame of the test sequence used to evaluate the performance of the coder. The top one is the reconstructed image by using the B_1 flow, while the center and the bottom images are obtained by decoding B_1, B_2, B_3 and all the flows, respectively.

Table 1: Bit rate in kbits/s for the different level streams versus frames/s.

Flow	12fps	8fps	4fps	3fps	2fps	1fps
B_1	56	38	17	15	10	5
B_2	134	88	49	34	20	10
B_3	16	14	3	3	3	2
B_4	371	242	123	90	63	30
B_5	18	14	6	7	4	3

3 The Multiple Access Protocol

We consider a cellular system, where the Base Station (BS) and the Mobile Stations (MS) within a cell share a common channel; the latter is structured according to a TDMA time frame of S slots. Each MS is supposed to generate only one traffic type of the three considered: data, voice and video; so, we suppose to have N_v Video MSs (VMS), N_o vOice MSs (OMS) and N_d Data MSs (DMS).

Let $t = 1, 2, \ldots$ be the discrete time unit corresponding to a TDM-frame. Each VMS i generates information every T^i time units; the i-th VMS generates $R_k^i = R(t_k^i)$ total video packets (whose length is supposed to fit a single slot) at the time instants $t_k^i = T^i + kT^i$, k=1, 2, …, where $\tilde{T}^i \in [0, 1, , \ldots, T^i\text{-}1]$ represents a time offset.

We reserve a fixed number S_v of slots/frame for each VMS (corresponding to the minimum capacity in slots needed to transport the whole lower level coded video information), so that the available number of slots per frame is $S_a = S - N_v S_v$, and the number of video packets for VMS i at instant t_k^i with no reserved bandwidth is $\tilde{R}_k^i = R_k^i - S_v$. Let $R_k^{i,l}$ be the number of packets of bit stream B_l, with $R_k^i = \sum_{l=1}^{L} R_k^{i,l}$, L being the number of layers of the encoder (5 in our case); the S_v slots are first filled with all the B_1 level packets, then, if space remains, with the B_2 packets and so on. Moreover, we indicate with $\tilde{R}_k^{i,l}$ the "residual" packets of level l and VMS i, not transmitted in the reserved slots; this implies that $\tilde{R}_k^i = \sum_{l=1}^{L} \tilde{R}_k^{i,l}$.

For what concerns OMSs and DMSs, all packets are generated at free stations (i.e., stations not waiting for a retransmission or not having a reserved slot during a burst) according to Bernoulli distributions with parameters σ_v and σ_d for voice and data, respectively. OMSs produce bursts of several packets (the model used for voice is a two-state Markov chain as in [32]) so that only the first packet of a burst is generated according to the above distribution, while the other packets in the burst occur deterministically, one per frame. Data sources always generate single packet bursts.

To introduce the basic algorithm, let us consider first only voice and data stations. In the RRA-ISA scheme, at the beginning of a frame, after observing the collision feedback for all slots available at the end of the previous one, the base station computes the access rights for all slots of the new frame and distributes them to the stations within the cell.

The aim of the algorithm is to maximize the one-step conditional throughput over the frame, by assuming the independence of the presence of packets at the stations and by enforcing a station to be able to transmit in no more than one slot in a frame. In the following, the generic frame t is considered, dropping, for simplicity, the time index from the notation. Let $u_i^s \in \{0, 1\}$ be the access right (1 enabled, 0 not enabled), for station i in the s-th slot of the frame, $u^s = \mathrm{col}[u_1^s, \ldots, u_N^s]$, $u = \mathrm{col}[u^1, \ldots, u^{S_a}]$, where N is the total number of MSs, p_i the packet presence probability for station i, conditioned to the past channel history, and

$$\bar{J}^s(u^s) = \sum_{j=1}^{N} p_j u_j^s \prod_{r=1, r\neq j}^{N} \left(1 - p_r u_r^s\right) \tag{1}$$

Expression (1) represents an approximation of the conditional probability of success (expected one-step throughput) for slot s, which assumes the independence of the presence of packets at the stations. By further enforcing a station to be able to transmit in no more than one slot in a frame, i. e.

$$\sum_{s=1}^{S_a} u_i^s \leq 1, \ \forall i \tag{2}$$

we can write the throughput over the entire frame as

$$\bar{J}(u) = \sum_{s=1}^{S_a} \bar{J}^s(u^s) \tag{3}$$

It can be proved [32] that, given a control vector $\breve{u}$ satisfying (2), not with N components equal to 1, we can find $\hat{u} = \breve{u} + \tilde{u}$, with $\breve{u}^T \bullet \tilde{u} = 0$, such that $\bar{J}(\hat{u}) > \bar{J}(\breve{u})$, if and only if $\exists \bar{s}$, such that

$$\prod_{i=1}^{N} \left(1 - p_i \breve{u}_i^{\bar{s}}\right) - \bar{J}^{\bar{s}}(\breve{u}^{\bar{s}}) > 0 \tag{4}$$

Since it is easily shown (see [33, 34]) that the optimal control vector must allow the station m with highest marginal probability to transmit, a recursion can be obtained [32], which constructs the maximum throughput step by step.

Let ρ^s represent the set of stations enabled to transmit in slot s (i.e., all the stations i with $u_i^s=1$), and let $\breve{u}^s$ be the corresponding control vector. We can now state the following

Algorithm:

Step 1: (initialize): let $\rho^s = \emptyset, \forall s$;

Step 2: search for the slot $\bar{s}$ in which (4) is satisfied and the l.h.s. is the largest;

Step 3: if slot $\bar{s}$ exists, and not all stations have been enabled

- then enable the station $\bar{i}$ that has the largest p_i and has not been enabled in another slot, that is
$\bar{i} = \arg \max_{w \notin \rho^s, \forall s} \{p_w\}$
$\rho^{\bar{s}} \leftarrow \rho^{\bar{s}} + \{\bar{i}\}$
and go to 2
- else stop

The presence probabilities p_i are updated at the end of each frame, on the basis of collision feedback and of the knowledge of the packet generation rates σ_d and σ_v (see [32] for the complete algorithm and a comparison with PRMA, which shows a significant performance improvement in favor of RRA-ISA).

For what concerns the presence of video, one "Virtual" VMS (VVMS) is associated to each video packet with an assigned "presence probability" p_i^l different for each coding level l. Each VVMS is treated as a normal station, with the exception that only one VVMS can be enabled to transmit in a slot and, after it has been enabled, it exits from the system. It is important to note that $\tilde{R}_k^i$ packets of VMS i, not transmitted in the interval $[t_k^i, t_{k+1}^i)$, are dropped. By temporarily resuming the time index t, we indicate with $\tilde{N}_{vv,t}$ the number of VVMSs still active in the system at frame t. Having defined the previous quantity, the index t can be omitted again in the following. On the basis of the description reported above, the algorithm is changed as follows.

Algorithm:

Step 1: (initialize): let $\rho^s = \emptyset, \forall$ s;

consider the total number of stations $N = N_d + N_o + N_{vv}$. Enable the first S_a stations (real or virtual) with larger presence probabilities (p_i and p_i^l) values, one each available slot; i.e., if we consider the MSs ordered in decreasing order of presence probability and the available slots numbered from 1 to S_a, let

$\rho^j \leftarrow \rho^j + \{j\}$ for $j = 1,\ldots, S_a$

Step 2 consider slots in which VVMSs have been enabled (if any), no longer available, and a total number of stations $N = N_d + N_o$ (only voice and data stations).

Step 3: search for the slot $\bar{s}$ in which (4) is satisfied and the l.h.s. is the largest;

Step 4: if slot $\bar{s}$ exists, and not all stations have been enabled

- then enable the station $\bar{i}$ that has the largest p_i and has not been enabled in another slot, that is
$\bar{i} = \arg \max_{w \notin \rho^s, \forall s} \{p_w\}$
$\rho^{\bar{s}} \leftarrow \rho^{\bar{s}} + \{\bar{i}\}$
and go to 3
- else stop

In this way, p_i^l i=1,...N; l=1,..L can represent a sort of stochastic thresholds, by means of which the base station dynamically identifies the congestion period of the system, during which the high quality video information should be sacrificed in favor of congested data and voice traffic. In this sense, the parameters p_i^l act like an auto-adaptive priority: with larger values of p_i^l the voice and data packets tend to suffer longer delays (and the video packets tend to be transmitted, also in quite high load conditions); the opposite happens with smaller values. For example, it is easily seen that the presence probability of a packet in a DMS not enabled to transmit since r frames is given by $p = 1-(1-\sigma)^r$; if $\sigma_d = 0.05$ and $r = 8$, then $p \approx 0.34$. This means that, to be enabled before a VMS, a DMS has to wait at least 8 frames from its last successful transmission if $p_i^l > 0.34$ $\forall i$, $\forall l$. We have used the term "adaptive priority" for the quantities p_i^l because the behavior of the system changes with the load. When data and voice offered load is really low, the video information is always transmitted; otherwise, when the same load is very high, video may not be transmitted at all (except for the granted portion). Different values of p_i^l permit to control the behavior in the intermediate load conditions, as shown in the simulation results presented in the following Section.

4 SIMULATION RESULTS

Several simulation tests have been performed to evaluate the behavior of the proposed scheme in different situations and to compare this dynamic approach with a static strategy.

We have considered a TDMA channel of available capacity C = 1440 kbits/s, a frame duration of 16 ms, 40 slots of 576 bits each; 64 bits/slot are considered reserved for FEC codes and other overhead. To better understand the behavior of the proposed method, only one active video source has been used (so, in the following, we drop the index i over the VMS parameters). The video stream is coded with the algorithm proposed in Section 2 with a frame rate of 8 frames/s. We have fixed T = 1.5 s, i.e., we collect 12 video frames before starting a new video transmission (the inter-frame coding covers a set of 12 frames in the algorithm used). The video set, produced every 1.5 s, is divided into packets of fixed length (512 bits, one packet each slot). We consider two slots every frame reserved for video to grant the transmission of the B_l flow without loss. We have assigned the same value of "virtual presence probability" to the packets of all levels, i.e., $p^l = p$, $\forall$l, to allow a simpler interpretation of the results. The DMSs generate packets 512 bit long, by following a Bernoulli distribution with σ_d = 0.05 for all data stations. OMSs utilize a Speech Activity Detector (SAD) with 32 kbits/s voice encoding, providing, in this way, no more than one packet per frame for each single OMS. For the voice stations, we have used a simplified version of the priority scheme presented in [32]: the presence probability in a OMS of the first packet in a burst has been multiplied by a fixed factor x. We have chosen x = 14, because it has resulted a suitable value for the tested situations. As described in [32], voice, due to its sensitivity to delay, requires a higher priority than the other traffic types in order to obtain the best performance for the whole system. Each simulation covers a real network time of about 25 minutes.

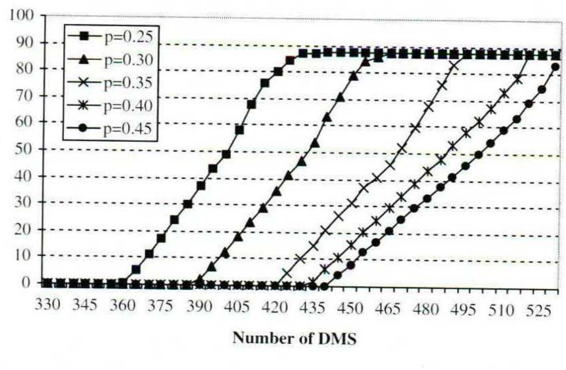

Figure 2: Lost video packets percentage versus the number of DMSs for five values of p.

The first group of results concerns a mix of data and video traffic without voice. Fig. 2 shows the percentage of lost video packets with respect to the overall video stream versus the number of DMSs, for five different values of "virtual presence probability" p: 0.25, 0.30, 0.35, 0.4, 0.45. The behavior of the system, concerning video performance, can be controlled by changing the value of the parameter p. Video packet discarding starts with lower data load if a small value of p is set (e.g., it happens with about 360 DMSs if p = 0.25), while discarding begins at higher load with larger values of p (about 440 DMSs if p = 0.45).

Figs. 3, 4 and 5 show the average delay in slots suffered by data packets versus the number of DMSs, for p = 0.25, p = 0.35 and p = 0.45, respectively. The results have been obtained by using three different schemes: i) the "adaptive" scheme, which corresponds to the application of the proposed algorithm; ii) the "fixed" scheme, which corresponds to reserve a fixed number of slots/frame (9 slots in this case) for the video stream, to allow the deterministic transmission of all packets; iii) the "only data" scheme, which is not a real access scheme, but reports the delay of data packets without the presence of the VMS and represents a lower bound for the measure of the delay. In this last case we have considered only 38 slots available to correctly compare the results with the other two cases. It can be easily seen that the proposed method tends to maintain the delay under a certain threshold, which changes with the value of p. In particular, the average delay is maintained less than 10 slots (160 ms) for p = 0.25 (Fig. 3), 16 slots (256 ms) for p = 0.35 (Fig. 4) and 40 slots (640 ms) for p = 0.45 (Fig.5). It is important to observe that the higher priority to the video traffic, which, as evidenced in Fig. 2, allows to accept a larger number of DMSs (after fixing a threshold on the video packet loss), imposes a higher delay to the same DMSs, as is clear in Fig. 5. Similar considerations could be reported for the other values of p.

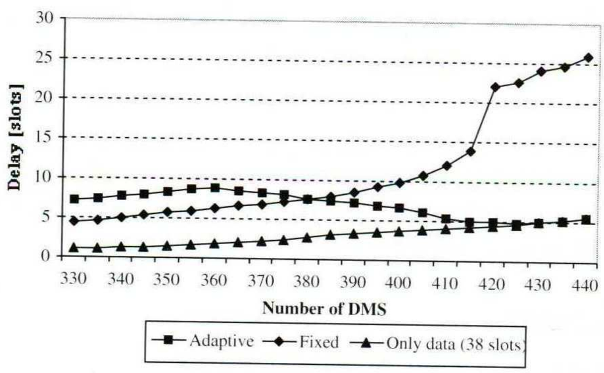

Figure 3: Data packets delay [slots] versus the number of DMSs for p = 0.25 .

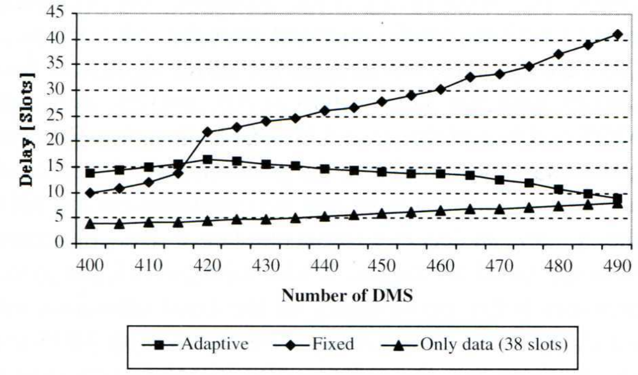

Figure 4: Data packets delay [slots] versus the number of DMSs for p = 0.35.

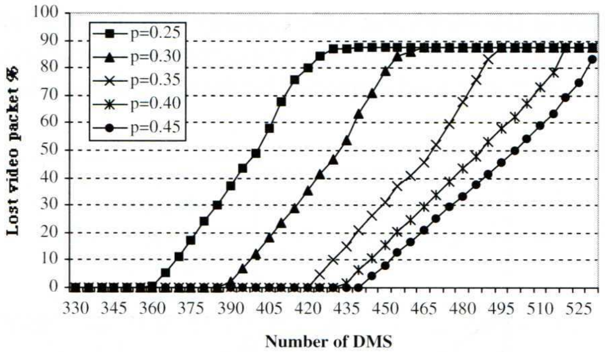

Figure 5: Data packets delay [slots] versus the number of DMSs for p = 0.45 .

Table 2: MOS values obtained by decoding one or more layers (first column) for different levels of packet loss (first row).

	0%	5%	10%	15%	20%	30%
B_1	1.7	1.7	1.6	1.5	1.5	1.5
B_1+B_2	2.5	2.5	2.2	2	1.7	1.7
B_1+..+B_3	3	2.8	2.1	2.4	1.6	1.5
B_1+..+B_4	4.3	3.9	3.5	3.3	2.9	2.8
B_1+..+B_5	4.4	4.4	4.3	3.9	3.7	3.6

The next results are related with the PQoS (Perceived Quality of Service). The evaluation of the quality in the decoded sequences at the receiver has been approached by using the Mean Opinion Score (MOS). The results of the evaluation have been reported in Table 2. Each column contains the MOS values measured by using a fixed number of layers and a certain packet loss value. These results should be considered as indicative since they have been obtained with a limited number of measures. Fig. 6 shows the maximum number of DMSs that can be accepted in the system versus the value of p. Two different constraints on the minimum value of acceptable MOS have been used. These last results show the effect of the MAC scheme and of the parameter p on the PQoS.

Fig. 6 may also be used to select the best value of p in the situation presented, namely, the value of p guaranteeing the largest number of acceptable DMSs with a guaranteed level of PQoS. No data delay bound is taken into account in this case.

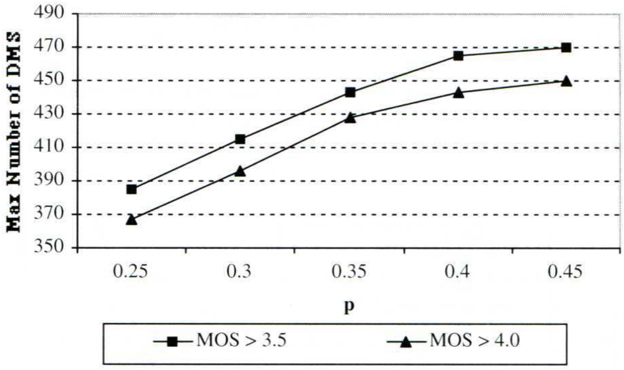

Figure 6: Maximum number of DMSs that can be accepted in the system with two constraints on the PQoS versus p.

The three successive graphs have been obtained with OMSs and one VMS only. Fig. 7 shows the percentage of lost video packets versus the number of OMSs, for three different values of virtual presence probability p: 0.25, 0.35, 0.45. As in the previous situation, the behavior of the system with respect to the video performance can be controlled by tuning the value of the parameter p. The effect of modifying p is less effective in this case, because voice traffic maintains a higher priority than the other types of traffic. The voice packet drop percentage versus the number of OMSs is reported in Fig. 8. The results obtained with three values of the parameter p (0.45, 0.35 and 0.25) are compared with the values obtained without video (with 38 slots available) and with a fixed allocation of 9 slots to the video stream (i.e., 29 slots available for OMSs). Fig. 8 shows that the use of the proposed mechanism provides better performance with respect to a fixed allocation, and that the performance of video slightly decreases for higher values of p (p = 0.45) and is substantially equivalent to the "only voice" situation for lower values of p (p = 0.25). Fig. 9 contains the maximum number of acceptable OMSs with a voice packet drop percentage lower than 1%. A voice packet is dropped if it experiences a delay over 32 ms (2 frames). The graph shows that the parameter p allows to tune the system behavior: if a low value of p is chosen, the system behaves as in the presence of only voice; on the other hand, a higher value of p makes the system similar to a fixed allocation.

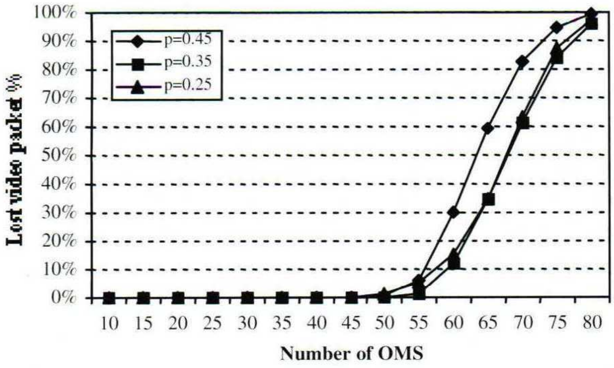

Figure 7: Lost video packets percentage versus the number of OMSs for three values of p.

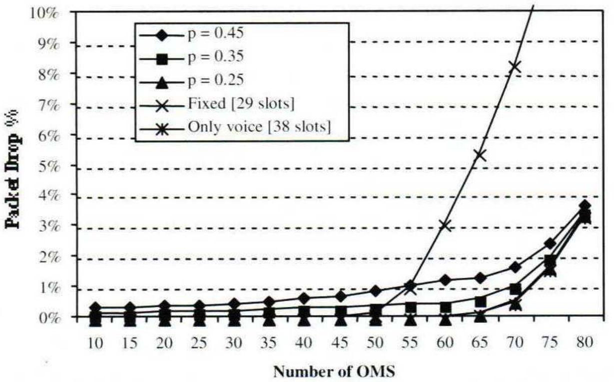

Figure 8: Packet drop percentage for voice traffic versus the number of OMSs for three values of p, without video (only voice), and with a fixed capacity reservation to the video (fixed).

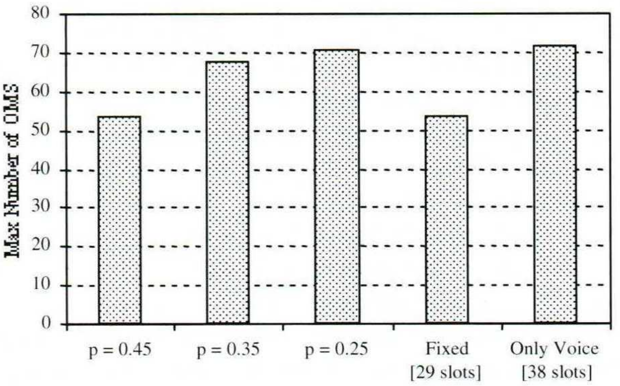

Figure 9: Maximum number of OMSs acceptable in the system with the constraint of less than 1% of voice packet drop, for p=0.45, p=0.35, p=0.45, fixed assignment and only voice.

Fig. 10 shows the behavior of the system with both OMSs and DMSs. It reports the maximum number of acceptable DMSs versus the number of OMSs. A constraint on the voice packet drop (less than 1%) is taken into account. In more detail, four situations are compared: p=0.45 with no MOS constraint over the video traffic (in this case the video traffic does not suffer significant loss), p=0.25 with the video MOS > 3.5, p=0.25 with video MOS > 4.0, and the case with fixed allocation. The conclusions may be similar to the previous case. It can be observed that, with p=0.25 and no constraint on the MOS, the system can accept a larger number of data connections than the fixed allocation, while, with p=0.45, the system provides better performance of the fixed allocation only for a high number of OMSs. The results with MOS constraint show that the video quality progressively deteriorates when the data load increases. As a consequence, the mechanism proposed allows to choose a reference quality for the video traffic and the corresponding acceptable number of DMSs (data threshold); if the input load is over such a threshold, the video quality deteriorates, but the voice constraint is always respected until the maximum overall traffic load supported by the system is reached.

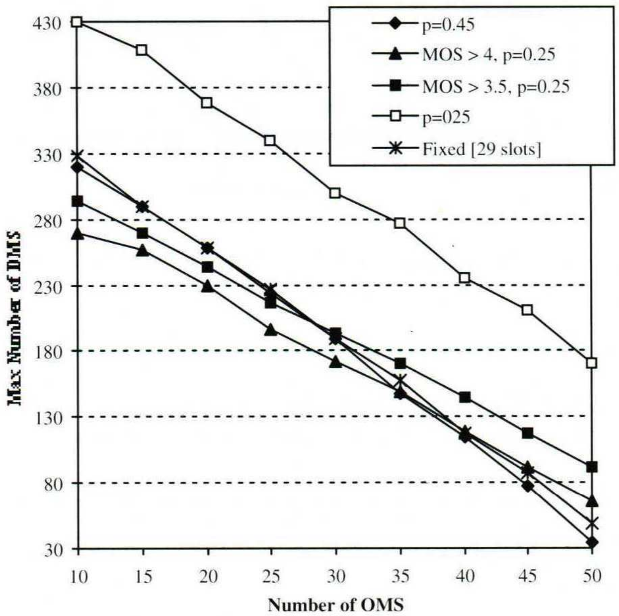

Figure 10: Maximum number of DMSs versus the number of OMSs, which respects the constraint on the packet drop probability, with p = 0.45 and no constraint on the MOS, with p = 0.25 and MOS > 3.5, p=0.25 and MOS > 4.0 and with fixed bandwidth allocation ("fixed (29 slots)").

5 Conclusions

The paper has proposed an integrated coding and multiple access scheme for the transport of video streams in multi-service mobile radio networks. Each video source is coded with a hierarchical scheme, which generates, for every frame or number of contiguous frames, different sets of hierarchical layered information. The lowest layer contains the minimum information to guarantee a minimum level of video quality; the other layers, if received, improve the quality of the service. The basic idea is to design a MAC scheme that grants resources only for the lowest level video, while the packets of the other levels contend for the available resources with the other types of traffic and are partially or completely discarded if the total offered load is too high. Some parameters give the opportunity of controlling the level of the load at which the video packet discarding begins and, then, to decide the quality of service the other types of traffic should receive.

The proposed scheme has been tested by simulation to verify its effectiveness. Performance in terms of delay and packet loss and PQoS in terms of MOS have been considered for video, voice and data. The results confirm that this technique can effectively exploit the capability of hierarchical coding to dynamically adapt the video source generation to the network load, in order to maintain a certain grade of quality.

Manuscript received on July 23, 1999.

REFERENCES

[1] Real time video services in multimedia networks. *IEEE Journal on Selected Areas in Communications*, Vol. 14, No. 6, August 1997.

[2] Traffic management and switching for multimedia. *IEEE Communications Magazine*, Vol. 37, No. 1, January 1999.

[3] Special Issue on Transmission and Distribution of Digital Video. *IEEE Network,* Vol. 12, No. 6, 1998.

[4] T. Ojanpera and M. Ohyama. An overview of third-generation wireless personal communications: a European perspective. *IEEE Personal Communications Magazine*, Vol. 5, No. 6, pages 59-65, 1998.

[5] T. Murase and M. Ohyama. Evolution of personal multimedia communications services in Japan *IEEE Personal Communications Magazine*, Vol. 5, No. 6, pages 66-74, 1998.

[6] Wireless access broadband networks. *IEEE Journal on Selected Areas in Communications*, Vol. 16, No. 6, 1998.

[7] X. Xiao, L. M. Li. Internet QoS: a big picture. *IEEE Network*, Vol.13, No. 2, pages 8-18, 1999.

[8] Integrated Services Working Group. http://www.ietf.org/html.charters/intserv-charter.html, http://search.ietf.org/rfc/rfc1889.txt.

[9] R. Braden, L. Zhang, S. Berson, S. Herzog, S. Jamin. Resource ReserVation Protocol (RSVP) - Version 1 Functional Specification. *RFC 2205*, Sep. 1997, http://www.ietf.org/rfc/rfc2205.txt.

[10] Integrated Services over Specific Link Layer, http://www.ietf.org/html.charters/issll-charter.html.

[11] Differentiated Services Working Group, http://www.ietf.org/html.charters/diffserv-charter.html.

[12] H. Kanakia, P.P. Mishra, Amy R. Reibman. An adaptive congestion control scheme for real time packet video transport. *IEEE/ACM Trans. Networking*, Vol. 3, No. 6, Dec. 1995.

[13] D. Reininger, D. Raychaudhuri, M. Ott. Dynamic QoS for video in broadband networks. *IEEE Network*, Vol. 12, No. 6, pages 22-34, Nov./Dec. 1998.

[14] R. Bolla, A. Iscra, M. Marchese, S. Zappatore. A perceived quality of service optimization for video communication in best-effort networks. In. *Multimedia Appl., Services and Tech. (ECMAST `98)*, Berlin, Germany, May 1998, pages 366-379.

[15] A.R. Reibman. DCT-based embedded coding for packet video. *Signal Processing: Image Communication*, Vol. 3 No. 2-3, pages 231-237, June 1991.

[16] M.R. Macedonia and D.P. Brutzma. Mbone provides audio and video across the Internet. *IEEE Computer*, Vol. 27, No. 4, pages 30-36, April 1994.

[17] S. McCanne, M. Vetterli, and V. Jacobson. Low-complexity video coding for receiver-driven layered multicast. *IEEE J. Select. Areas Commun.*, Vol. 15, No. 6, pages 983-1001, Aug. 1997.

[18] T. Turletti, C. Huitema. Videoconferencing in the Internet. *IEEE/ACM Trans. Networking*, Vol. 4, No. 3, pages 340-351, June 1996.

[19] D. Hoffman, M. Speer. Hierarchical video distribution over internet-style networks. In *IEEE Int. Conf. on Signal Processing*, Vol. 1, pages 5-8, Sept. 1996.

[20] G. Karlsson. Asynchronous transfer of video. *IEEE Commun. Mag.*, Vol. 34, No. 8, pages 118-126, Aug. 1996.

[21] A.N. Netravali and B.G. Haskell. *Digital Pictures, Representation and Compression*. Plenum Press, New York and London, 1988.

[22] M. Rabbani and P.W. Jones. *Digital Image Compression Techniques*, SPIE Optical Engineering Press, Bellingham, Washington, USA, 1991.

[23] D.J. Goodman, S.X. Wei. Efficiency of packet reservation multiple access. *IEEE Trans. Vehic. Technology,* Vol. 40, No. 1, pages 170-176, 1991.

[24] S. Nanda. Stability evaluation and design of the PRMA joint voice data system. *IEEE Trans. Commun.*, Vol. 42, No. 5, pages. 2092-2104, 1994.

[25] M. Paterakis, A.C. Cleary. On the voice-data integration in third generation wireless access communication networks. *Europ. Trans. Telecommun.*, Vol. 5, No. 1, pages 11-18, Jan./Feb. 1994.

[26] S. Jangi, L.F. Merakos. Performance analysis of reservation random access protocols for wireless access networks. *IEEE Trans. Commun.*, Vol. 42, No. 2/3/4, pages 1223-1234, 1994.

[27] R. Cappella, E. Del Re, R. Fantacci. A modified PRMA protool for joint voice data packet wireless networks. *European Trans. Telecommun.*, Vol. 10, No. 1, pages. 23-31, 1999.

[28] R. Fantacci, F. Innocenti. Performance evaluation of a modified PRMA protocol for joint data packet wireless networks. *IEEE Trans. Commun.*, Vol. 47, No. 12, 1999.

[29] Z.J. Haas, D. A. Dyson. A dynamic packet reservation multiple access scheme for wireless ATM. *Mobile Networks and Application*, Vol. 4, No. 2, pages 87-99, 1999.

[30] X. Qiu, V. O. K. Li. Dynamic reservation multiple access (DRMA): a new multiple access scheme for personal communication system (PCS). *Wireless Networks*, No. 2, pages 117-128, 1996.

[31] M. Thomas, D. Goodman. Multi-rate PRMA: a time division protocol for adjiustable bit-rate sources. In *IEEE 47th Vehic. Technol. Conf.* (VTC'97), Vol. 3, pages 1360-1363, 1997.

[32] R. Bolla, F. Davoli, C. Nobile. The RRA-ISA multiple access protocol with and without simple priority schemes for real-time and data traffic in wireless cellular systems. *Mobile Networks and Applications*, No. 2, pages 45-53, 1997.

[33] M. Aicardi, G. Casalino, F. Davoli, R. Minciardi. A multiaccess scheme for the maximization of one-step throughput. *Internat. J. of Mini and Micro Computers*, Vol. 8, No. 3, pages 65-69, 1986.

[34] M. Aicardi, G. Casalino, F. Davoli. Independent stations algorithm for the maximization of one-step throughput in a multiaccess channel. *IEEE Trans. Commun.*, Vol. COM-35, No. 8, pages 795-800, Aug. 1987.