Journal of Zhejiang University SCIENCE A
ISSN 1009-3095 (Print); ISSN 1862-1775 (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

Video over IP using standard-compatible multiple description coding: an IETF proposal

VITALI Andrea L.[1], BORNEO Antonio[1], FUMAGALLI Marco[2], RINALDO Roberto[3]
([1]*ST Microelectronics, Strategy and System Technology, Agrate Brianza (MI) 20041, Italy*)
([2]*Center of Excellence for Research, Innovation, Education and Industrial Labs Partnership, Milan 20133, Italy*)
([3]*Department of Electrical Management and Mechanical Engineering, University of Udine, Udine 33100, Italy*)
E-mail: andrea.vitali@st.com; antonio.borneo@st.com; marco.fumagalli@cefriel.it; roberto.rinaldo@uniud.it
Received Dec. 2, 2005; revision accepted Feb. 28, 2006

Abstract: Standard-compatible multiple description coding (MDC) and layered coding (LC) are efficient ways to ensure erasure resilient, scalable transmission of encoded multimedia sources via RTP, allowing a gradual degradation of the application quality with increasing packet loss rate and decreasing bandwidth/throughput on the network. In this paper we review the standard-compatible framework proposed to IETF. Alternative techniques such as robust source coding and channel coding techniques (ARQ: automatic repeat request, FEC: forward error correction) are presented; their integration into the proposed framework is also discussed. The performances of MDC and LC either coupled with channel coding or not, are summarized by reference to current literature. Typical cases and examples are illustrated.

Key words: Video streaming, Robust coding, Multiple description coding, Layered coding, Loss/error resilience
doi:10.1631/jzus.2006.A0668 **Document code:** A **CLC number:** TN919.8

INTRODUCTION

Communication networks are growing quickly. Throughput, delays and losses become unpredictable because of heterogeneity and congestion. This results in lack of quality of service. Today's solution is based on standard video codecs (as MPEG-2) that do not allow scalability and yield low quality at low bit-rates, complex forward error correction (FEC) codes that are not easily used with time-varying channels and must be set to face the worst case, and automatic repeat requests (ARQ) that can be used only in point to point communications and need a feedback channel. Future solution should be based on advanced video codecs that are efficient at low bit-rates (as H.264) and, at the same time, allow for scalability (as SVC); FEC should be designed jointly with codecs, exploiting the robustness that may be embedded in compressed bit-streams; ARQ should not be used to allow for broadcast communications without feedback channel.

Robust source coding

Hybrid video codecs are based on prediction (also known as motion estimation and compensation), transform, quantization and entropy coding. One or more of the following techniques can be used to embed some robustness in the compressed bit-stream: more frequent intra (not predicted) pictures, a suitable macro-block intra refresh policy or a multi-frame interleaved prediction policy in order to reduce dependency and stop error propagation due to prediction loop; more slices per picture to reset the differential coding of the DC transform coefficient and the differential coding of motion vectors; flexible macro-block order (FMO) or asynchronous slice order (ASO) to delocalize the effect of losses and ease the concealment; encoding concealment motion vectors or redundant slices; using reversible variable length codes (RVLC), error resilient entropy codes (EREC) or inserting more sync markers to reduce the portion of bit-stream affected by errors.

Multiple description coding (MDC)

MDC can be seen as another way of enhancing error resilience without using complex channel coding schemes. The goal of MDC is to create several independent descriptions that can contribute to one or more characteristics of video: spatial/temporal resolution, quality (SNR), frequency content in the transform domain. Descriptions can have the same importance (as for balanced MDC) or they can have different importance (as for unbalanced MDC). The robustness comes from the fact that it is unlikely that the same portion of the same picture is corrupted in all descriptions. The coding efficiency is reduced depending on the amount of redundancy left among descriptions; however channel coding can indeed be reduced.

Layered coding (LC)

LC is analogous to MDC. The main difference lies in the dependency among bit-streams: there is one base layer and several enhancement layers that can be used, one after another, to refine the decoded quality of the base layer. The base layer should be protected more heavily because if it is not received there is nothing to be refined by successfully received enhancement layers. Channel coding is required to protect the base layer.

Forward error correction (FEC)

FEC usually needs to be complex (plus it introduces substantial delay) in order to be effective and it has an all-or-nothing performance: if the correction capability is exceeded, almost nothing is delivered to the receiver. Capacity may be wasted if the worst case (worst channel conditions or farthest user) must be considered. On the opposite, when the channel is better than expected and there are no losses, FEC redundancy is useless.

While FEC is independent, MDC and LC are dependent on the nature of the data. FEC is needed by LC but can also be used with MDC. Generally speaking, it is suggested to adapt the protection level of a given description/layer to its importance, a technique commonly known as unequal error protection. It is suggested to use unequal error protection even in the case of equally important descriptions (balanced MDC). In fact, protecting only one description may be more effective than trying to protect all descriptions. If this is done, there is one description which is heavily protected. If the channel becomes really bad, this description is likely to survive losses. Then the decoder will be able to guarantee a basic quality, thanks to this description.

STATE-OF-THE-ART IN MDC

The multiple description coding (MDC) of a source consists of generating a number of data streams (2 or more) that, together, carry the input information. The objective of MDC is to encode a source into many bit-streams in such a way that a high quality reconstruction can be achieved from all the streams together at the decoder, or, if fewer bit-streams are available at the decoder end, an acceptable, but obviously poorer, quality reconstruction is attainable.

The cost of this operation is to insert a certain amount of redundancy among descriptions which are then compressed into bit-streams. The literature show how some MD approaches are more flexible than others in redundancy insertion. Varying the amount of redundancy in accordance with channel performance is crucial for the final reconstruction quality of the source: less redundancy insertion in each description is needed for error-free transmissions than for unreliable packet transfer. Many approaches have been proposed to realize MD coding: scalar quantizer (Vaishampayan, 1993), pair-wise transform coding (PTC) (Orchard *et al*., 1997), spatial and temporal down-sampling (Jiang and Ortega, 1999; Shirani *et al*., 2000; Franchi *et al*., 2005), correlating filter-bank (Yang and Ramchandran, 1998), frame expansion (Bernardini *et al*., 2004), matching pursuits algorithms (Tang and Zakhor, 2001). Such approaches differ in terms of redundancy management and complexity. Some algorithms are designed for a general source, other more specifically for a type of signal, e.g. speech, image or video.

The MDC method used in this paper is essentially based on the work in (Franchi *et al*., 2005). Since, in raw image data, the value of any given pixel can be reasonably predicted by the value of its neighbors, there is a strong correlation of inter-pixel information. By exploiting this, it is possible to create a multiple description algorithm where the source is split into N descriptions by a poly-phase down-sampler (PD) along rows and columns (Franchi *et al.*, 2002). Because of this inter-pixel correlation each generated description maintains the main features of the original image. For a better understanding of the PD multiple description (PDMD) procedure, the general scheme is shown in Fig.1 for N=4. If the transmission is error-free, the receiver over-samples the descriptions, combining them to restore the original source. In the worst case, when only one description is received, the receiver exploits the available information in order to obtain a good low-resolution image. The novelty of this scheme is that the number of decoders needed at the receiver equals the number of descriptions.

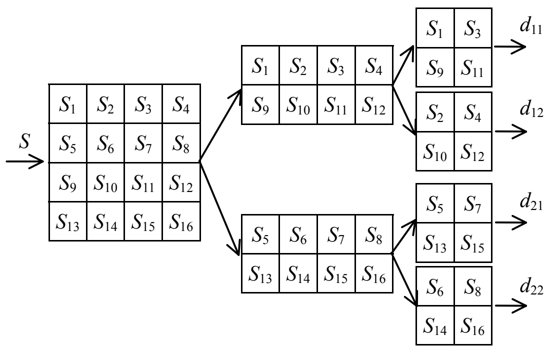

Fig.1 An example of poly-phase down-sampling system. The original image dimension is 4×4 pixels and the number *N* of multiple descriptions is 4. Each poly-phase component (description) is composed of 4 pixels that can be organized according to spatial location

Past work showed some advantages of the PDMD algorithm. One is its strong robustness in an error-prone environment (Franchi *et al.*, 2002): this is due to the exploitation of the natural correlation of the source. Moreover in the spatial domain, the algorithm provides a very simple design with consequent computational time saving. The drawbacks of the PDMD techniques are the loss of coding efficiency due to the separate coding of each description and some rigidity in redundancy control among the descriptions.

MDC/LC standard-compatible framework

The implementation of the proposed MDC/LC scheme is completely independent of the underlying multimedia codec. The creation of descriptions/layers is performed in the data domain. This is done in a pre-processing stage. Descriptions/layers can then be coded independently. At the decoder side is a post-processor stage in which decoded descriptions/layers are merged.

Specific information needs to be communicated out-of-band via SDP or MIME to specify which MDC/LC scheme is in use. Standard decoders will ignore this specific information. However, such decoders can still decode each successfully received description/layer. At the same time, decoders MDC/LC-aware will parse this information in order to properly decode and merge descriptions/layers.

Balanced MDC can even be beneficial for standard decoders. Multiplexed descriptions can be marked so that standard decoders understand they are multiple copies of the same data. Of course, decoded data will have a smaller resolution/quality. As an example, when balanced descriptions are transmitted, standard decoders will understand that the same data is transmitted multiple times in a way similar, but not equal to, repetition codes. Actually, there is no repetition but slightly different data packets are transmitted.

Decoders can be instructed to decode only the first successfully received copy.

MDC/LC implementation issues

Pre- and post-processing stages can be completely decoupled from the underlying multimedia codec. However, it must be noticed that keeping MDC/LC decoupled from the underlying codec prevents MDC/LC from giving its best. To get maximum quality for the decoded MDC/LC and to do MDC/LC encoding with the least effort, joint or coordinated encoding could be used. Also, to exploit MDC/LC redundancy and to maximize the error resilience, joint MDC/LC decoding is recommended.

As an example, video encoders can share expensive encoding decisions (motion vectors) instead of computing them; they can also coordinate encoding decisions (quantization policies) to enhance quality or enhance resilience (interleaved multi-frame prediction policies, intra refresh policies). Decoders can share decoded data to ease error concealment; they can also share critical internal variables (anchor frame buffer) to stop error propagation due to prediction.

It is worth mentioning that, if balanced descriptions are properly compressed and packed, losses/ erasures can be recovered before the decoding stage. In this case, decoders are preceded by a special processor in which lost packets are recovered by copying similar packets from other descriptions. Similar packets are those that carry the same portion of data.

MDC VS LC

In order to combat channel-induced impairments, layered coding (LC) and multiple description coding (MDC) have been proposed as source coding techniques that are robust against inevitable transmission errors (Goyal, 2001). In contrast to a conventional media coder that generates a single bit-stream, LC and MDC encode a media source into two or more sub-bit-streams. For LC, one base layer bit-stream and several enhancement layer bit-streams are generated. The base layer can be decoded to provide a basic quality of video while the enhancement layers are mainly used to refine the quality of the video that is reconstructed from the base layer. If the base layer is corrupted, the enhancement layers become useless, even if they are received perfectly. Therefore, the base layer is critically important and is usually protected using either automatic repeat request (ARQ) (Mao *et al*., 2001) or forward error correction (FEC) codes (Gallant and Kossentini, 2001).

Compared to LC, MDC has the following unique properties: the descriptions are mutually refining, equally important, and independent. Each description can be independently decoded. There is no decoding dependency between any two of the descriptions. MDC usually does not require prioritized transmission (Singh *et al*., 2000).

LC+FEC vs MDC+FEC

A few performance comparisons between LC and MDC have been reported (Reibman *et al*., 1999; 2000; Singh *et al*., 2000; Wang *et al*., 2002). Reibman *et al*.(1999) first analyzed and compared FEC-coded LC with MDC on a binary symmetric channel (BSC) and a random erasure channel (REC) (Reibman *et al*., 1999). For the BSC, cyclic redundancy check (CRC) codes and rate-compatible punctured codes (RCPC) were used together to detect and recover corrupted packets. For the REC, Hamming codes were used to recover lost packets. The encoding method used for LC was similar to the SNR-scalability method defined in the MPEG-2 standard. The channel coding was applied to both layers but the base-layer data were protected with stronger channel codes. For MDC, two descriptions were generated using the approach in (Wang *et al*., 2001), and were equally protected. They concluded that MDC was more effective than LC only in very high error probability situations. They further suggested that a multiple description source coder should use channel coding to obtain acceptable performance on memory-less channels.

LC+ARQ vs MDC

Wang *et al*.(2002) considered transporting layered and multiple description coded video over wireless networks with multiple path routing and compared their performance (Wang *et al*., 2002). For LC, the SNR scalability method defined in the H.263+ standard was adopted. For MDC, they used a multiple description motion compensation coder developed by Wang and Lin (2002). They compared the performance of LC, MDC, and LC with ARQ, where ARQ was used to protect the base-layer data only. Their simulation results showed that MDC was better than LC when the underlying application had a very stringent delay constraint and that the RTT on each path was relatively long. However, LC performed better than MDC when limited retransmission of the base-layer was acceptable.

LC+FEC/ARQ vs MDC+FEQ/ARQ

Lee *et al*.(2003) carefully compared the performance between LC and MDC in multi-path environments at various packet loss rates. Two different video codes are chosen for LC and MDC: a hybrid transform and motion compensation codec and a wavelet-based video codec. Three different error protection scenarios are also considered for both LC and MDC: no error protection, ARQ-based error protection, and FEC-based error protection. According to the simulation results, MDC is more suitable for delay-sensitive applicants or for which the underlying transmission channels do not support a feedback link, while LC may be a better choice for applications that can tolerate a certain amount of delay. If FEC-based error protection can be used, MDC and LC perform similarly. However, MDC is preferred at high error rates because of its better error-resilience at those rates.

MDC vs FEC

In the most recent article (Bernardini *et al*., 2005), polyphase MDC was compared to single description whose robustness is increased by Reed-Solomon FEC at application layer (Fig.2). Two descriptions are created by separating odd and even lines. H.264 is used. Bit-streams are split into 500-bytes packets sent over a single path and over multiple paths. SD packets are protected by using a systematic (n, k) Reed-Solomon code which adds $n-k$ redundancy packets every k packets. A given number of frames (one GOP, IbbPbbPbbPbbPbb) is covered by $k1$ packets from first description, $k2$ packets from second description and k packets from single description, $n-k$ is chosen so that $k+(n-k)$ equals $k1+k2$, the aggregate bit-rate of MD is therefore equal to SD+FEC, while the decoded quality of MD in the absence of losses is lower.

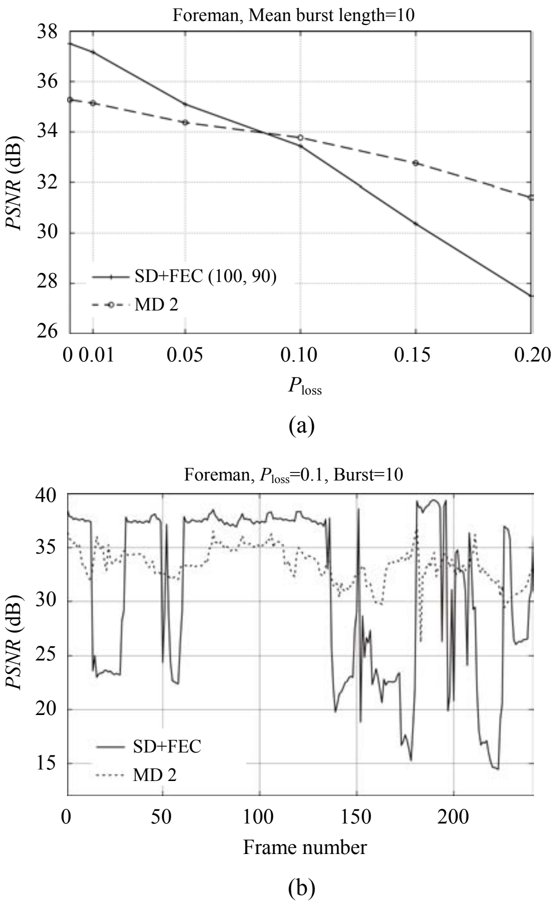

Fig.2 Two descriptions (MD 2) and single description plus (SD) Reed-Solomon FEC. (a) *PSNR* versus loss rate; (b) *PSNR* frame by frame. As can be seen the decoded quality of MD is higher for high loss rate (>5%) and the variance is lower (graphs are courtesy from (Bernardini *et al*., 2005))

In another experiment, the decoded quality in the absence of losses was made almost equal by reducing the rate k/n of the code (k is reduced by compressing more while n is kept constant).

Experiments showed that for video sequences with high motion content the inefficiency of MD can make the SD+FEC preferable. However the decoded quality of MD has a lower variance and it is higher, especially at high loss rates (>5%). MD is preferable to compensate for long burst of losses and in the presence of on-off channels.

Performance summary

Summarizing the results from these studies, we might conclude that MDC and LC have some aspects in common, such as the sub-streams representation and the intrinsic scalability, but they differ in error-resiliency capability and the importance of having a feed-back channel. In general, about the performance, we can state that MDC has advantages over LC for networks with no feedback or a long RTT, or for those applications that have very stringent delay constraints (Singh *et al*., 2000; Wang *et al*., 2002). However, if the networks or the applications support prioritized transmission or error control, LC might be better than MDC (Reibman *et al*., 1999; 2000; Wang *et al*., 2002).

IETF DRAFT

The common stack for video streaming is composed by RTP/RTCP over UDP over IP. These protocols take care of sequence numbering, stream time stamping and session state control. These allow for stream synchronization and loss detection. Other protocols are needed to communicate stream type and session parameters (SDP via RTSP or SAP).

Related RFCs and drafts

Robust source coding does not need IETF standardization.

FEC coding is covered in many RFCs: RFC2733 for generic FEC based on packet XOR operations; RFC3009 for MIME types registrations related to FEC parity; RFC3452 for FEC building block; RFC3453 for FEC used in Reliable Multicast; RFC3695 for various FEC schemes. Drafts are discussed in the MMUSIC group (fec-grouping), the AVT group (ulp, unequal level protection, updates RFC2733) and RMT group (bb-fec-basic-schemes, bb-fec-ldpc and bb-fec-raptor-object, fec-bb-revised) for the definition of various building blocks, LDPC codes, Raptor codes, etc.

LC combined with unequal error protection is also covered extensively in many RFCs: RFC3450 for Asynchronous Layered Coding (ALC) protocol instantiation; RFC3451 for Layered Coding Transport (LCT) building block. Drafts are discussed in the RMT (bb-lct-revised and pi-alc-revised).

MDC is not yet standardized. The proposed draft (mdc-lc) was submitted recently in July 2005 to the AVT group (Vitali and Fumagalli, 2005).

SDP media-level attributes for MDC/LC

Descriptions/layers are sent in different, separate and independent RTP streams. Specific information must be communicated out-of-band to use an MDC/LC scheme. SDP (Session Description Protocol, described in RFC 2327) is used to send the required information.

Descriptions/layers are identified by a given payload type, e.g., "m=video 49170 RTP/AVP 97 98" indicates that two descriptions/layers are used for media flow on port 49170, the transport protocol is RTP/AVP (IETF Realtime Transport Protocol using the Audio/Video profile carried over UDP), their payload type is 97 and 98. Existing media attributes can be used, e.g., "a=rtpmap: 97 H264/90000" indicates that H.264 has been used to encode the video with 90 kHz sampling clock; "a=fmtp: 97 …" indicates the format specific parameters used by the encoder; etc.

In the MDC/LC framework there are two sets of parameters (more detailed information can be found in (Vitali and Fumagalli, 2005)). One of these two sets of parameters must be sent via SDP using media-level attributes ("a=").

The first set of parameters is related to the pre-processor: it describes the creation of each description/layer so that smart receivers can compute how to merge them depending on the loss pattern.

In a way analogous to RFC3388, payload types are tagged, e.g., "a=X-mdclc-tag: 97 D1" indicates that payload type 97 is tagged as D1 (first description).

The tag "S" has a special meaning: it is reserved to indicate the original multimedia data which is the default starting point for descriptions/layers generation. "E" is used to indicate the ending point, the reconstructed data. The suffix "_Q" is used to indicate decoded (and quantized) descriptions/layers. The prefix "T_" is used to indicate temporary descriptions/layers used for computations in pre- and post-processors. "D" should be included as part of the tag to indicate that it is an independent description. "L" should be included as part of the tag to indicate that it is a dependent layer. Lower numbers should be used to indicate more important descriptions/layers. "_" is used as separator to improve readability.

Then, each tag is associated with a mathematical expression to specify how the corresponding video sub-sequence is computed. Common pixel-wise operations are specified by mathematical and logical operators. Three functions are used for up-sampling, down-sampling and finite impulse response filtering, e.g., "a=X-mdclc-pre: D1=dn(2, 0, Y, fir([0, 0.75; 0.25], S));" indicates that description D1 is obtained in the pre-processor by filtering the original sequence S with a finite impulse response filter (FIR) and then down-sampling the output of the filter. Filter coefficients (0, 0.75 and 0.25), down-sampling factor (2:1) and phase (0) are specified.

The second set of parameters is related to the post-processor: it describes explicitly the merge of descriptions/layers for basic receivers that cannot compute how to merge them. Smart receivers can compute a second set for the merge given the first set.

Again, in a way analogous to RFC3388, tags are grouped. There is one group for each loss pattern. All descriptions/layers listed in the indicated group must be received in order for the merging to occur, e.g., "a=X-mdclc-group: lp0 D1 D2" indicates that the loss pattern "lp0" uses description "D1" and "D2" for reconstruction.

Each group is then associated with a mathematical expression to specify how the video sequence is to be computed from the corresponding set of sub-sequences, e.g., "a=X-mdclc-post: lp0 E= up(2,0,Y,D1)+up(2,1,Y,D2);" indicates that the reconstructed data "E" is obtained by up-sampling descriptions in group "lp0" (D1 and D2) and summing the result. Again, up-sampling factors (1:2) and phases (0 for D1 and 1 for D2) are specified.

RFC3388 cannot be used for tagging and grouping: its use for simultaneous encoding (MDC) or layered coding (LC) is explicitly forbidden.

Mathematical expressions allow the management of several MDC and LC schemes. MDC schemes: Basic poly-phase down-sampling (PDMD); poly-phase down-sampling applied to filter bank to get unbalanced multiple descriptions (UMD) or frame expanded MDC. LC schemes: classical layered coding (like MPEG-2 spatial scalability); wavelet layered coding.

Typical cases and examples for MDC

2MD: two descriptions can be generated by down-sampling vertically 2:1 the video sequence. This is equivalent to separating odd and even lines. 4MD: four descriptions can be generated by down-sampling vertically and horizontally 2:1. This is equivalent to separating odd and even lines and taking every other pixel. Note that in both cases the quantity of data to be encoded is not changed: 2×½ for 2MD and 4×¼ for 4MD (Fig.3).

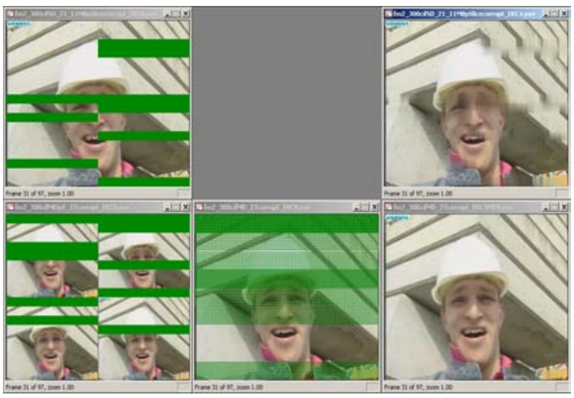

Fig.3 Single description (SD) and four multiple descriptions (4MD): same aggregate bit-rate, same number of packets and same packet size, 25% packet loss rate. The complexity and quality of the concealment can be clearly evaluated

The frequency content of each description can be controlled by applying a suitable filter prior to down-sampling. As an example: the 4th description of 4MD can be made alias free by applying a simple 3×3 low-pass FIR filter.

Frame expansion is a way to expand the original data so that some controlled redundancy is added. In literature, frames expansion was used with quantized filter banks. Here frames expansion is used with down-sampled filter banks. 3MD: 2 descriptions can be generated by separating odd and even lines as done in 2MD; the 3rd description is simply the average of odd and even lines. 3MD can be seen as equivalent to an FEC code with rate 2/3: one single erasure can be recovered. The redundancy can be controlled easily by quantizing more heavily the third description. Another example: 4 descriptions can be generated by separating odd and even lines and taking every other pixel as done for 4MD; the 5th description can be generated by averaging separated pixels (i.e. averaging descriptions). Note that the quantity of data is 3×½ for 3MD and 5×¼ for 5MD.

Unbalanced Multiple Description (UMD): one description corresponds to the original sequence; the other description is simply a down-sampled version with 1/4th the size of the original. The second description is more important and should be more protected than the first and is to be used in case of losses to enhance the concealment. The redundancy can be controlled easily by quantizing more heavily the second down-sampled description, whose protection level can be increased by simply increasing the intra refresh rate. This is useful because the second description is used only for the concealment of the first. Note that the quantity of data to be encoded is 1+¼.

Typical cases and examples for LC

Two layers are created as follows: the original data is down-sampled to 1/4th the original and encoded. This is the base layer. The base layer is decoded, up-sampled and subtracted from the original data, generating what can be seen as a prediction error to be encoded. Note that the quantity of data to be encoded is 1+¼.

Three layers are created in an analogous way: the original data is down-sampled to 1/16th the original and encoded, this is the base layer. The base layer is decoded, up-sampled and subtracted from the original down-sampled to 1/4th to get the first enhancement layer. The first enhancement layer is decoded, up-sampled and subtracted from the original to get the second enhancement layer. Note that the quantity of data to be encoded is 1+¼+1/16.

Classical layered encoding suffers from overhead. Wavelet coding does not suffer from overhead because enhancement data is down-sampled critically. Two layers are created using the 1D vertical Haar wavelet. The base layer is the average of odd and even lines. The enhancement layer is the difference between odd and even lines. Note that the quantity of data to be encoded is 2×½.

Four layers are created using the 2D Haar wavelet. The base layer is the average of odd and even lines and columns. The other three enhancement layers are the horizontal, vertical and diagonal difference with respect to the base layer. Note that the quantity of data to be encoded is 4×¼.

Synchronization issues

Descriptions/layers are sent in different, separate and independent RTP streams. Synchronization of descriptions/layers is critical in order to produce the correct result and can be accomplished in one of the following ways.

Timestamp synchronization: data to be merged can be sent in packets having the same timestamp. This is possible if the sampling clock is the same.

Sequence number synchronization: data to be merged can be sent in packets having the same sequence number. This is possible if each packet contains the same portion of the data. This means also that packets may have variable length.

Payload synchronization: data to be merged can be identified parsing the payload.

The post-processor can then identify packets to be decoded and merged by looking at their timestamp, sequence number or payload specific synchronization information.

Multiplexing and interleaving issues

Descriptions should be offset as much as possible when streams are multiplexed. In this way a burst of losses does not cause the loss of the same portion of data in all descriptions at the same time.

If interleaving is used, the same criterion is to be used: descriptions should be spaced as much as possible. In this way a burst of losses does not cause the loss of the same portion of data in all descriptions at the same time.

CONCLUSION

Multiple description coding (MDC) schemes have been presented in Section 3. The proposed approach, down-sampling in the pixel domain, is preferred because it is independent of the underlying video codec and it is relatively simple to implement in pre- and post-processors.

The performance of MDC has been compared to layered coding (LC) in Section 4. LC needs prioritization while MDC does not but can benefit unequal error protection. Experimental results show that MDC is preferable when packet loss rate is relatively high or there is no time for retransmission.

The proposed standard-compatible framework is discussed in Section 5. It mimics the grouping semantic of RFC3388. The complex syntax is used to support classical LC but it is not strictly needed. It can indeed be simplified: one filtering operation followed by down-sampling can be used to support MDC and wavelet LC. Related parameters can be grouped and sent as a single media-level SDP attribute.

Foreseen applications are summarized in the following list:

(1) Divide-and-conquer approach for HDTV distribution: HDTV sequences can be split into SDTV descriptions; no custom high-bandwidth h/w is required.

(2) Easy picture-in-picture: with the classical solution, a second full-decoding is needed plus downsizing; with MDC/LC, it is sufficient to decode one description or the base layer and paste it on the display.

(3) Adaptation to low resolution/memory/power: mobiles decode as many descriptions/layers as they can based on their display size, available memory, processor speed, and battery level.

(4) Pay-per-quality services: user can decide at which quality level to enjoy a service, from low-cost low-resolution (base layer or one description only) to higher cost high-resolution (by paying for enhancement layers/more descriptions).

(5) Easy cell hand-over in wireless networks: different descriptions can be streamed from different base stations exploiting multi-paths on a cell boundary.

(6) Adaptation to varying bandwidth: the base station can simply drop descriptions/layers; more

users can be easily served, and no transcoding process is needed.

(7) Multi-standard support (simulcast without simulcast): descriptions can be encoded with different codecs (MPEG-2, H.263, H.264); there is no waste of capacity as descriptions carry different information.

(8) Enhanced carousel: instead of repeating the same data over and over again, different descriptions are transmitted one after another; the decoder can store and combine them to get a higher quality.

References

Bernardini, R., Durigon, M., Rinaldo, R., Zontone, P., Vitali, A., 2004. Multiple Description Coding of Video Streams Using Frames and the H.264/AVC Standard. Proc. ICIP 2004. Singapore.

Bernardini, R., Durigon, M., Rinaldo, R., Vitali, A., 2005. Comparison between Multiple Description and Single Description Video Coding with Forward Error Correction. MMSP.

Franchi, N., Fumagalli, M., Lancini, R., 2002. Flexible Redundancy Insertion in a Polyphase Down-Sampling Multiple Description Image Coding. Proc. ICME 2002. Lausanne, Switzerland, p.26-29.

Franchi, N., Fumagalli, M., Lancini, R., Tubaro, S., 2005. Multiple description video coding for scalable and robust transmission over IP. *IEEE Trans. on CSVT*, **15**(3): 321-334.

Gallant, M., Kossentini, F., 2001. Rate-distortion optimized layered coding with unequal error protection for robust internet video. *IEEE Trans. CSVT*, **11**(3):357-372.

Goyal, V.K., 2001. Multiple description coding: compression meets the network. *IEEE Signal Proc. Mag.*, **18**(5):74-93.

Jiang, W., Ortega, A., 1999. Multiple Description Coding via Polyphase Transform and Selective Quantization. Proc. VCIP 99.

Lee, Y.C., Kim, J., Altunbasak, Y., Mersereau, R.M., 2003. Layered coded vs. multiple description coded video over error-prone networks. *Signal Processing: Image Communication*, **18**(5):337-356. [doi:10.1016/S0923-5965(02)00138-8]

Mao, S., Lin, S., Wang, Y., Panwar, S., 2001. Reliable Transmission of Video over Ad-hoc Networks Using Automatic Repeat Request and Multi-Path Transport. Proc. of VTC.

Orchard, M., Wang, Y., Vaishampayan, V., Reibman, A., 1997. Redundancy Rate Distortion Analysis of Multiple Description Image Coding Using Pairwise Correlating Transforms. Proc. of IEEE ICIP 1997. Santa Barbara, CA.

Reibman, A.R., Jafarkhani, H., Orchard, M.T., Wang, Y., 1999. Performance of Multiple Description Coders on a Real Channel. Proc. of the ICASSP.

Reibman, A.R., Wang, Y., Qiu, X., Jiang, Z., Chawla, K., 2000. Transmission of Multiple Description and Layered Video over an EGPRS Wireless Network. Proc. of ICIP.

Shirani, S., Gallant, M., Kossentini, F., 2000. Multiple Description Image Coding Using Pre- and Post-Processing. International TCC. Las Vegas, Nevada, USA.

Singh, R., Ortega, A., Perret, L., Jiang, W., 2000. Comparison of Multiple Description Coding and Layered Coding Based on Network Simulations. Proc. of the SPIE VCIP.

Tang, X., Zakhor, X.A., 2001. Matching Pursuits Multiple Description Coding for Wireless Video. Proc. ICIP 2001. Thessaloniki, Greece.

Vaishampayan, V.A., 1993. Design of multiple description scalar quantizer. *IEEE Trans. Inform. Theory*, **39**(3):821-834. [doi:10.1109/18.256491]

Vitali, A., Fumagalli, M., 2005. Standard-compatible Multiple-Description Coding (MDC) and Layered Coding (LC) of Audio/Video Streams. Http://www.ietf.org, draft-vitali-ietf-avt-mdc-lc-00.txt.

Wang, Y., Lin, S., 2002. Error-resilient video coding using multiple description motion compensation. *Trans. on CSVT*, **12**(6):438-452.

Wang, Y., Orchard, M.T., Vaishampayan, V.A., Reibman, A.R., 2001. Multiple description coding using pairwise correlating transforms. *IEEE Trans. on Image Processing*, **10**(3):351-366. [doi:10.1109/83.908500]

Wang, Y., Panwar, S., Lin, S., Mao, S., 2002. Wireless Video Transport Using Path Diversity: Multiple Description vs. Layered Coding. Proc. ICIP.

Yang, X., Ramchandran, K., 1998. Optimal Multiple Description Subband Coding. Proc. of IEEE ICIP 1998. Chicago, IL.