A SCALABLE AND RELIABLE HYBRID SCHEME FOR IMAGE MULTICAST APPLICATIONS

Abed Elhamid Lawabni and Ahmed H. Tewfik
Department of Electrical and Computer Engineering
University of Minnesota, Minneapolis, MN 55455
{lawabni, tewfik}@ ece.umn.edu

Abstract

In real-time multicast communication scalability, reliability, and feedback implosion are of paramount importance. In this work we have developed an efficient feedback-free, entirely receiver-driven, and reliable visual information delivery hybrid scheme (SIGMA-EC) for multicast applications over unreliable communication networks, that requires no per-receiver status at the sender. Efficient bandwidth utilization is achieved by transmitting the parity packets over separate multicast channels, which receivers dynamically join and leave, preventing unnecessary receiver processing overhead and storage requirements for those packets that have already received correctly, and giving rise to a tradeoff between bandwidth and quality. Our XOR-based scheme's encoding and decoding processing time is an order of magnitude faster than conventional RS-based loss recovery schemes. Furthermore, we obtain all these advantages using a remarkably simple structure, both conceptually as well as computationally.

I. INTRODUCTION

The rapid growth of the multicast backbone (MBONE), combined with other multicast-capable networks have paved the way for providing new Internet multimedia applications such as audio/video-on-demand, distance learning, distributed visual tracking, distributed game packages, and releasing new software updates. Such applications can have potentially thousands of concurrent receivers with heterogeneous characteristics. Hence, delivering such data in a reliable and scalable manner is of paramount importance.

IP multicast is an excellent mean of transmitting data to multiple destinations. However, it provides an unreliable datagram service, where there are no delivery guarantees.

In designing a reliable scheme for multicast applications that scales to arbitrarily large receiver sets, there are typically two challenges:

- *Feedback implosion*: As the receivers experience packets lost, their requests for retransmission can "swamp" the sender, and introduce traffic congestion in its network links.
- *State explosion*: The sender must keep state information for each receiver, such information become too large to store or manage, resulting in state explosion.

A simple solution to avoid the latter problem is by simply keeping no state at the sender, and to let the receiver handle the losses. In order to deal with the implosion problem, a wide variety of techniques have been devised to eliminate the need for retransmission requests where the back channel between the sender and the receivers has high latency and limited capacity, if it is available at all. We briefly mention a few samples of them here.

Scheduled Multicast which also known as ***Conventional Multicast Techniques:*** In the simplest scenario, service requests for the same media file arriving within a short time duration are bunched together and served in a batch according to some dynamic scheduling policy. Some examples of such approach are *First Come First Served* (FCFS), *Maximum Queue Length first* (MQL), *Maximum Factored Queue length first* (MFQ), and *Patching* [1].

Periodic Multicast: These schemes divide the server bandwidth into logical channels with equal bandwidth. The media file is partitioned into segments of equal / unequal sizes. Each segment is repeatedly broadcast on its own channel. One sample of such approach is *Pyramid Broadcasting* [2]: The server partitions the media file into unequal-sized segments, and repeatedly transmits each segment on a different multicast group as in a carousel at the same rate. The client downloads the segments in order and starts playing the content when it receives the beginning of the first segment. This is generally not feasible in a network where the client maximum download rate already limits playback quality. To this end, *Permutation-based Pyramid Broadcasting*, *Skyscraper Broadcasting*, and *Harmonic Broadcasting* are among other schemes that were proposed to resolve such a problem.

Layered Multicast: The media file is coded in multiple layers, and each layer is multicasted to a different multicast group. Each receiver joins a number of these groups based on his available bandwidth. One attractive scheme that can be categorized under this approach is the *Digital Fountain Approach* [3]: It was proposed to provide reliable distribution of bulk data. It is based on using a new class of erasure codes called Tornado codes. The basic principle behind the use of erasure codes is that the original source data in the form of a sequence of k packets, along with n-k additional redundant packets, are transmitted by the sender. The redundant packets can be used to recover lost source data at the receiver end. A receiver can reconstruct the original source data once it receives *any subset* of length k packets out of the n. One attractive feature of this approach is that different receivers can recover from different lost packets using the same redundant packets.

Our main goal in this work is to develop a new feedback-free, scalable, entirely receiver-driven, and reliable visual information delivery scheme for multicast applications over unreliable packet networks and wireless networks. In particular, the proposed scheme (SIGMA-EC) is a hybrid approach that integrates and interacts two major schemes:

1. An efficient Sigma filtering and interpolation-based reconstruction scheme (SIGMA) that proposed by the authors in [4] combined with packetization procedure that addresses the same issues of Multiple Description Coding (MDC) techniques.
2. A new XOR-based Erasure Code (EC) scheme that operates in conjunction with periodic multicast channels.

The novelty of this scheme relies on the fact that receivers can immediately reconstruct the visual information in incremental progressive manner, due to the robustness of the SIGMA scheme. Other conventional schemes are inapplicable to streaming media, due to the fact that receivers are unable to reconstruct the portion of the file at which playback can begin, until after receiving most or all the data needed to reconstruct the entire file. This was pointed out in [5]. Another attractive feature of this scheme is that the reconstruction quality at the receiver depends *only* on the number of packets received, but is *independent* of the place from where they were cropped. This latter feature is of paramount importance, especially in the context of visual information delivery. Furthermore, the outgoing bandwidth from the sender is independent of the number of receivers. The sender can handle a large number of receivers with heterogeneous characteristics.

The rest of this paper is structured as follows. Section II, briefly describes the relevant properties and the intuition behind the SIGMA scheme. Moreover, it provides the reader with the intuition behind the hybrid approach without going in much detail. Section III presents the new erasure code scheme, and discusses some practical aspects concerning the deployment of such codes. In section IV the hybrid approach is discussed in details. Finally, conclusions are drawn in section V.

II. THE SIGMA SCHEME

Providing quality-of-service (QoS) guarantees are critical for real-time applications such as data, voice, images, and video. The provisioning of these guarantees over wireless links is a challenging problem whose difficulty stems from the need to explicitly consider the harsh radio-channel transmission characteristics, and the underlying link-layer error-control mechanisms. Hosting mobility and its impact on the sustained bandwidth capacity, further compound this difficulty. These challenges remain as bottlenecks for different kinds of multimedia applications.

Aligned with this vision, our approach is motivated by the following question:

- How one can achieve an acceptable perceptual quality in image transmission over erroneous channels, while maintaining an acceptable level of bandwidth utilization?

A straightforward engineering sense will lead to the following notion:

- First extract the most important features of the underlying image.
- Second, transmit that extracted information in such a way that meets Multiple Description (MD) objectives.
- Finally, increase the injection of the remaining less important details if the channel situation permits that.

How to select perceptual important features of an image automatically is a wide research issue. Here we just present a prefiltering perceptual importance enhancing method-SIGMA.

In order to achieve our prefiltering perceptual important features of images, it is necessary to process the underlying image into piecewise smooth regions while preserving or even enhancing important edges. The traditional low pass filters are not able to accomplish this goal; since less pronounced edges are likely to be eliminated, resulting in many visual artifacts. Fortunately, in image smoothing and segmentation literature there are several techniques being developed to achieve such goal. Anisotropic diffusion, total variation minimization, and sigma filter are among the most efficient approaches. In this work, sigma filter is considered to achieve this goal. By preprocessing an image using Sigma filter, it would transform a natural scene image to a “cartoon-like” image. Ideally the “cartoon-like” image preserves all important information about how the scene looks like, but smoothes out some low energy detail. However, our human eyes are less sensitive to the “cartoon-like” image than image with some loss of information. The SIGMA scheme not only selects perceptual important features of the underlying image, it also preserves less pronounced edges and even enhance them.

It’s worth mentioning that this filter is motivated by the sigma probability of the Gaussian distribution. The basic idea is to replace the pixel to be processed by the average of only those neighboring pixels having their intensity within a fixed sigma range of the center pixel. Due to space constraints we refer the reader to [4] for the filter design, and for general further details.

In order to have a complete solution for our notion; we need to devise packetization and reconstruction procedures, which provide the following tasks:

- They have to meet MD objectives; if some of the original data get lost due to the channels impairments, the receiver should still be able to reconstruct that data from the survival packets. The reconstruction quality should be acceptable and increases in progressive fashion, without the need of feedback channels.
- The processing time required to reconstruct the original data and the complexity structure should be kept nearly minimal and simple, respectively.
- With a small overhead in the processing time and delay, receivers should be able to reconstruct the original data with superior quality. Furthermore, these procedures should enable multiple receivers recovering from different packet loss using the same redundant data.

These considerations underline the main intuition behind our proposed hybrid scheme. In the next section, we propose an XOR-based error recovery scheme (EC) that enable different receivers recovering from different packet loss using the same redundant data. Unlike other traditional forward error correction (FEC) schemes, which use Reed-Solomon codes (RS) as the erasure codes, our EC scheme’s encoding and decoding processing time is order of magnitude faster, since the processing time is dominated by exclusive-or operations rather than matrix inversion and multiplication over the Galois field. Moreover, its structure is much simpler. We conclude the section, by pointing out couple drawbacks of our EC scheme, and how we can easily overcome them by applying the EC scheme to a sigma prefiltered image.

III. XOR-BASED ERASURE CODE (EC) SCHEME

A. The Coding Scheme

Our scheme divides the original data into equal size segments. Each segment is packetized into a packet of n-bits length; the full description of the packetization procedure is provided in the next section. Then it arranges the data packets in an $M \times K$ array and adds to each row and each column of the array an n-bit parity packet as it is depicted in figure 1. The i-th bit of a parity packet is the modulo 2 sum (i.e., exclusive-or) of the i-th bits of the data packets in the corresponding column or row.

Let $B_{i,j,l}$ represents the l-th bit of the packet in the (i, j) position in the array, then the row-parity packets are generated by:

$$B_{i,K+1,l} = \left(\sum_{j=1}^{K} B_{i,j,l} \right) \bmod (2)$$

where $1 \le l \le n$, $1 \le i \le M$.

and the column-parity packets are similarly generated by:

$$B_{M+1,j,l} = \left(\sum_{i=1}^{M} B_{i,j,l} \right) \bmod (2)$$

where $1 \le l \le n$, $1 \le j \le K$.

Thus, if the packets are sent row by row, this EC scheme can recover any erasure of scattered packets over the array with the aid of the row-parity packets. The salient part of this scheme is its capability in recovering bursts of corrupted bits and erasure of packets of any length less than or equal to K, and that can be done with the aid of the column-parity packets. Furthermore, since the XOR operation is commutative, packets can be processed at the decoder in any arbitrary order.

B. Practical Aspects

The practical size of the packets (order of several hundred bits e.g., ATM cell, to several thousand bits e.g., IP datagram) poses a serious limitation in deployment of the RS coders due to the hardware architecture complexity. The alternative we propose to avoid this cost and to assess the encoding and decoding speed with much simple structure is by using SIGMA-EC scheme.

The reader should recognize that if only the EC scheme used as an error-recovery scheme, it would have the following drawbacks:

- Suppose that a row suffers only a single erasure packet, that packet cannot be recovered unless the last packet of the row is received.
- If two or more packets are missing in a row, they cannot fully be recovered till their respective columns are received.

However, the hybrid SIGMA-EC approach can easily overcome such drawbacks, as will be more clear through the next section, and as will be demonstrated by the experimental results.

For the time being, note that by applying the sigma filter, we convert the underlying image into piecewise smooth regions while preserving or even enhancing important edges. The combination of the smoothness process (which can also be regarded as a correlation process of the data packets) and the packetization procedure (as will be explained in the next section) is the key feature behind the success of the reconstruction procedure. The receiver doesn't need to wait until the successful receiving of the row-parity packets or the respective column-parity packets. It can start reconstructing the original data immediately with an acceptable quality, in the same spirits as the MDC schemes.
However, if the end user seeks a superior quality he has to wait until receiving any distinct number of packets which is equal to the number of the original data packets. This progressive reconstruction nature provides the end users with incremental progressive quality.

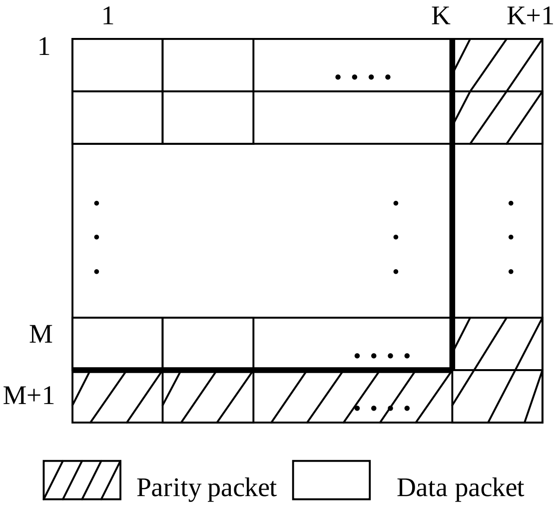

Figure 1. *Array arrangement of the data packets and the parity packets.*

IV. THE HYBRID SIGMA-EC SCHEME

The input image is first sigma filtered to produce an edge-enhanced one. Then we use wavelet transform to get one level decomposition consists of the approximation and the details coefficients. These coefficients preserve more strong edges than the ones without pre-sigma-filtering, due to the fact that vast area of a sigma filtered image is smoothed out, and certain high frequency information is kept at low-low band. Then the transformed image is scanned by raster-scan order, i.e., following the first row by the second row, and so on. Each row of the scanned transformed image is divided into non-overlapping blocks of size 8x16; each of these blocks is divided into two further blocks of size 8x8, corresponding to odd columns and even columns. Then each of these 8x8 blocks is packetized separately, and arranged into a row of a new array as illustrated earlier in figure 1. The EC scheme is applied for each row of the new array generating the rows-parity packets (two parity-packets for each row: one corresponding to the odd packets of that row, and other corresponding to the even packets) and for each column of the array generating the columns-parity packets.

Then each row of *data packets only* is set for transmission through its own multicast channel (e.g., multicast group). First row is transmitted over the first multicast channel; the second row is transmitted over the second multicast channel, and so on. Over each channel the transmission is performed in periodic procedure, that supports late-joining receivers: the source repeatedly loops through the same set of packets until all of them are exhausted. In order to reduce unnecessary receiver processing overhead and storage requirements for those packets that have already received correctly, the parity packets are periodically transmitted over two separate multicast channels: one corresponding to the different row-parity packets, and other corresponding to the different column-parity packets. These two FEC streams are only subscribed by those receivers who need them to achieve superior quality. Thus, our scheme provides a tradeoff between bandwidth / delay and quality. For example, in a live multicast of a conference, participants who want to interact with the speaker have a preference for low latency rather than superior quality. On the other hand, passive participants will sacrifice latency, network bandwidth consumption, and small processing overhead for higher image quality.
At the receiver, we assume that the arriving packets are reassembled into a resequencing buffer allowing for receiving out of order packets. For simplicity, we simulate the random loss at the packet level. We set p the probability that a certain packet gets lost.

After the initialization procedure between the server and the client, the latter can subscribe to any number of the multicast channels depending on his preferences and bandwidth availability. More importantly, the receiver can immediately start in the reconstruction procedure, by applying the same interpolation technique (*averaging*, *edge detecting*, and *filling*) that was proposed by the authors in [4]. Finally, after the interpolation procedure in the transformation domain, we apply the inverse wavelet transform to get the reconstructed image in the spatial domain.

It worth pointing out that, so far, we excluded the *difference* between the original image and the sigma-filtered one. However, when the client's preference is to achieve high quality, it is important to provide the end users with this *difference* in order to improve and enhance the visual quality of the transmitted image. Thus, the *difference* is fed to the scheme in a parallel fashion along with the sigma-filtered image. And all the previous procedures hold true, starting at the wavelet transform step. The block diagram of the hybrid scheme is given in figure 2.

Results for evaluating the reconstruction performance for the hybrid scheme are given in figures 3-5. Figures 3 and 4 show the robustness and the progressive nature of the proposed scheme. Depending only on the SIGMA part, the receiver can achieve an acceptable quality. However, by subscribing to the two FEC streams and sacrificing latency, the reconstruction quality had improved dramatically, as it is clear in figure 5.

V. CONCLUSION

In this work, we presented the design and implementation of a new feedback-free, scalable, entirely receiver-driven, requires no per-receiver status at the sender, and reliable visual information delivery scheme for multicast applications over unreliable communication networks. The novelty of this hybrid approach relies on the following facts:

- Receivers can immediately reconstruct the visual information in incremental progressive manner.
- This integrated approach provides a tradeoff between bandwidth / latency and quality.
- The outgoing bandwidth from the sender is independent of the number of receivers.
- A remarkably simple structure, both conceptually as well as computationally.
- This approach not only reduces the unnecessary receiver processing overhead and storage requirements for those packets that have already received correctly; it also provides a congestion control mechanism.

REFERENCES

[1] K.A. Hua, Y. Cai, and S. Sheu, " Patching: A Multicast Technique for True Video-on-Demand Services", *Proc. ACM Multimedia '98,* Bristol, U.K., Sept. 1998.

[2] S.Viswanathan and T. Imielinski, " Metropolitan Area Video-on-Demand Services Using Pyramid Broadcasting", *Multimedia Systems*, Aug. 1996, pp.197-208.

[3] J. W. Byers, M. Luby, M. Mitzenmacher, and A. Rege, " A Digital Fountain Approach to Reliable Distribution of Bulk Data," *ACM SIGCOMM '98*.

[4] Abed Elhamid Lawabni and Ahmed H. Tewfik, "An Efficient Scheme for Image Transmission over Error-Prone Channels: Sigma Filtering and Image Interpolation", *2001 IEEE International Conference on Multimedia and Expo. (ICME),* Tokyo-Japan, Aug 22-25, 2001.

[5] A. Mahanti, D. L. Eager, M. K. Vernon, and D. S-Stukel, "Scalable on-Demand Media Streaming with Packet Loss Recovery", *ACM SIGCOMM'01,* Aug. 27, pp. 97-108, 2001.

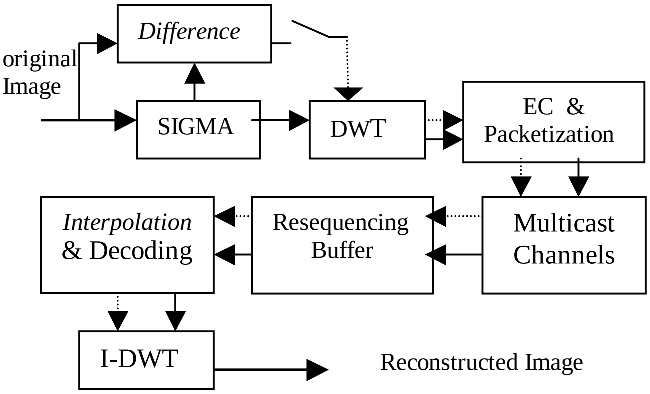

Figure 2. *Block Diagram of the Proposed Hybrid Scheme.*

Figure 3. Sigma filtered version plus the *difference* of the reconstructed Lena using the SIGMA scheme only*;* with *odd* packets *only*, and *even* packets *only*. PSNR values are 25.2190 dB, and 25.3325 dB, respectively.

Figure 4. Sigma filtered version plus the *difference* of the reconstructed Lena using the SIGMA scheme only; with p=0.1 and p=0.25. PSNR values are 30.6094 dB, and 28.7077dB, respectively.

Figure 5. Sigma filtered version plus the *difference* of the reconstructed Lena using the hybrid SIGMA-EC scheme with p=0.25, PSNR value is 35.2248 dB.