TRANSMISSION OF LOW-MOTION JPEG2000 IMAGE SEQUENCES USING CLIENT-DRIVEN CONDITIONAL REPLENISHMENT*

J. J. Sánchez-Hernández[1], J. P. García-Ortiz[1], V. González-Ruiz[1], I. García[1] and D. Müller[2]
[1]University of Almería, Comp. Architecture and Electronics Dept., Almería, Spain
[2]European Space Agency, ESTEC, Noordwijk, Netherlands

Keywords: JPEG 2000, Conditional replenishment, Video transmission.

Abstract: This work proposes a strategy for browsing interactively sequences of high resolution JPEG 2000 remote images. These sequences can be displayed in any order (forward and backward) and following any play/timing pattern. In order to increase the quality of the reconstructions where the retrieved images are only known at the moment of the visualization, this work has proposed and evaluated a novel technique based on conditional replenishment. This solution profits from the SNR/Spatial scalability of JPEG 2000 to determine which regions of the next image should be transmitted and what regions should be reused from the previously reconstructed image. Experimental results demonstrate that, even without motion compensation and with a transmission exclusively controlled by the client, the reconstructions are consistently better, both visually and from a rate-distortion point of view, than those that only remove the spatial redundancy (such as Motion JPEG 2000). Other advantages of our approach are that no data overhead is generated, the computational complexity is very small compared to similar techniques, and the fact that it can be used with any JPIP server.

1 INTRODUCTION

Some of the powerful features of the new JPEG 2000 multi-part standard (International Organization for Standardization, 2004) are very efficient lossless/lossy compression, random access to the compressed data streams, incremental decoding and high scalability. These characteristics make JPEG 2000 a state-of-the-art solution for remote browsing of high-resolution images. Using the JPIP protocol, defined in Part 9 (International Organization for Standardization, 2005) of the JPEG 2000 standard, clients can interactively explore remote image data by specifying a window of interest (WOI). This data exchange uses the available bandwidth efficiently and does not require any recoding or additional processes. The server extracts only the required data from the images and transmits it to the clients.

JPEG 2000 has already been successfully used in many scientific areas; e.g., tele-microscopy or tele-medicine. A promising application in space sciences is the JHelioviewer project (Müller et al., 2009), developed by the European Space Agency (ESA) in collaboration with the National Aeronautics and Space Administration (NASA). Its main goal is to provide an interactive data browsing, visualization and access platform to accommodate the staggering data volume of 1.4 TB of images per day that are returned by the Solar Dynamics Observatory (Pesnell, 2008). Among other data products, SDO is providing full-disk images of the Sun taken every 12 seconds in ten different ultraviolet spectral bands with a resolution of 4096×4096 pixels. As of today, the combination of JPIP and JPEG 2000 seems to offer the best solution in order to efficiently browse image data sets of this magnitude.

The basic functionality of JHelioviewer allows users to explore the available data for a given point in time. An interesting extension of this functionality is that it enables users to move smoothly through a sequence of time-coded solar images given a specific time range while displaying the changes in the WOI. This work therefore focuses on JPIP applications, such as JHelioviewer, that are designed to in-

*This work has been funded by grants from the Spanish Ministry of Science and Innovation (TIN2008-01117) and Junta de Andalucía (P08-TIC-3518), in part financed by the European Regional Development Fund (ERDF).

J. Sánchez-Hernández J., P. García-Ortiz J., González-Ruiz V., García I. and Müller D..
TRANSMISSION OF LOW-MOTION JPEG2000 IMAGE SEQUENCES USING CLIENT-DRIVEN CONDITIONAL REPLENISHMENT.
DOI: 10.5220/0003518600110016
In *Proceedings of the International Conference on Signal Processing and Multimedia Applications* (SIGMAP-2011), pages 11-16
ISBN: 978-989-8425-72-0
Copyright © 2011 SCITEPRESS (Science and Technology Publications, Lda.)

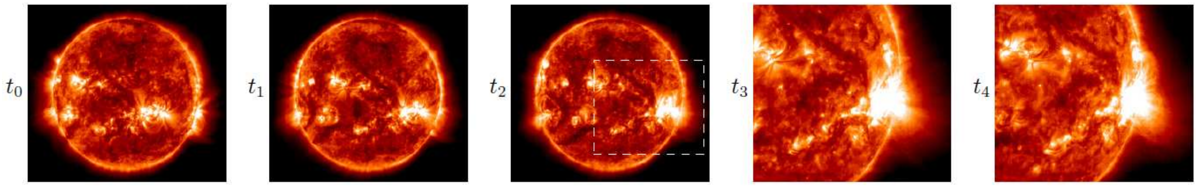

Figure 1: An example of different time instances from a remote browsing session.

teractively explore remote image sequences. Fig. 1 shows an example of five sequential time instances during a remote browsing session using JHelioviewer. Once the user has selected a time range, the server builds a virtual JPEG 2000 file which only contains links to those solar images whose date/time stamp belongs to that time range. The client starts a JPIP session for that file and requests the first image, displayed at time t_0. The user can then watch the sequence of images belonging to the time range, one by one, jumping to any image, and changing direction at any moment. In this example, at time t_2 the user zooms in on a certain region, thus changing the current WOI. From this new WOI, the user continues moving forward through the sequence, to instances t_3 and t_4. In the "video mode", where the images are visualized without pause for a given frame-rate that the user can control, the amount of data that it is retrieved from the JPIP server depends on the available bandwidth and the frame-rate. In any case, images are transmitted by quality in order to increase the quality of the reconstructions as much as possible.

The fact that SDO takes images of the entire Sun at high temporal resolution results in a high degree of spatial correlation between consecutive images (about 33dB of PSNR on average between images that have been taken with a cadence of 432 seconds[2]) and therefore, they show a certain degree of temporal redundancy. For this reason, we focus this paper on the removal of this redundancy by means of Conditional Replenishment (CR).

2 RELATED WORK

CR has been previously used with success to encode low-motion image sequences. In (Mounts, 1969), Mounts proposed a method for the lossy (8:1) compression of head-and-shoulder digital TV signals. In (McCanne et al., 1997), McCanne et al. explored the possibility of CR for the MBone (multicast backbone) thanks to its reduced complexity, the option to generate scalable video and its data loss resilience. More recently, three major solutions that offer good compression results and decoding flexibility have been studied. In (Cheung and Ortega, 2007), Cheung and Ortega have investigated the performance of a I14P motion-compensated predictive video codec that is based on Distributed Source Coding and support forward and backward playback. In (Devaux et al., 2009), Devaux et al. have proposed a solution based on CR to exploit the temporal redundancy of video surveillance sequences that have previously been encoded with JPEG 2000. Although this approach does not rely on motion compensation, the sender (or server) has to determine, and to transmit to the receiver, the optimal points at which it is better to use the JPEG 2000 data than the prediction (generated with the last reconstructed frame and an estimation of the background). Finally, the work of Naman et al. is similar in essence, but they propose the use of CR with multiple reference frames and motion compensation to encode those image sequences with a large amount of movement (Naman and Taubman, 2009) and the same architecture but without motion compensation for more static sequences (Naman and Taubman, 2010).

After an analysis of all these approaches our work is motivated by the following main reasons: (1) all the proposed approaches need to process the sequence of images on the server-side in some way and therefore the sequence (and in some cases even the order of the images) must be known a priori, (2) the computational resources required by the sender to perform real-time transmission are extremely high due to the amount of data to be managed (4Kx4K images, and at least 20 images/second) and (3) none of them are fully JPIP compliant because additional data has to be transmitted from the sender to the receiver. Therefore, none of the existing solutions can be easily applied to an interactive browsing system like JHelioviewer. Moreover, in (Ortiz et al., 2010) a prefetching strategy has been proposed that would further enhance the user experience and that is fully compatible with the proposal of this paper, but would be difficult to implement with the approaches described above.

The remainder of this paper is structured as fol-

[2]This result has been determined by computing the average PSNR of one image respect to the next (in time) one. SDO offers the possibility of retrieve images at a higher cadence, up to 12 seconds.

SCITEPRESS SCIENCE AND TECHNOLOGY PUBLICATIONS

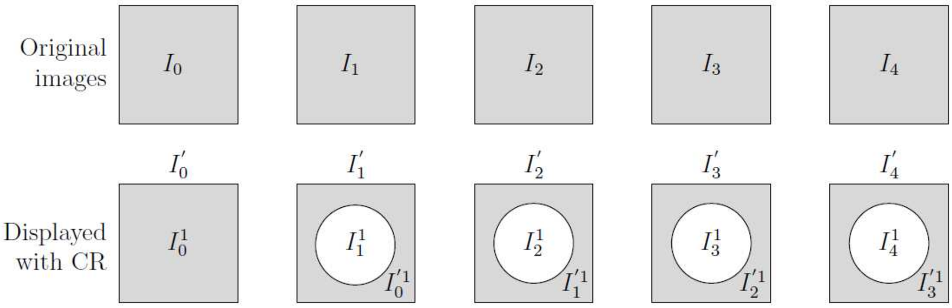

Figure 2: Example with the first four iterations of our algorithm.

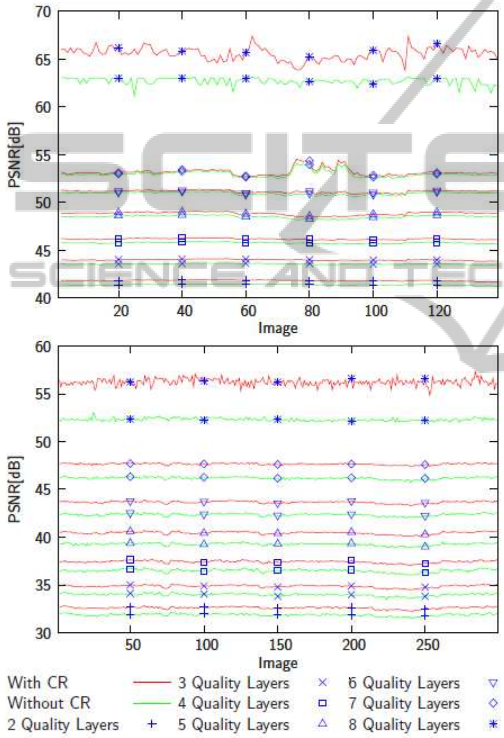

Figure 3: A comparison between our proposal ("with CR") and the standard method ("without CR"). On the top for the SDO/AIA sequence. Below, the *Akiyo* sequence.

lows. Section 3 introduces our client-driven retrieving algorithm. Section 4 provides experimental results and comments about the visual increment of the quality of our reconstructions. Section 5 summarizes our work and describes the main future research lines that could extend it.

3 PROPOSAL

In large-scale solar image sequences, such as in many other videos where significant portions of the image do not change much for a long period of time, there is room for the use of CR, even though the predictions are not motion compensated. In this case, it is possible to perform a real-time reconstruction of a 4Kx4K-resolution video using a frame-rate of 25 images/second in machines with modest computational resources and limited band-widths.

In our discussion it will be supposed that the band-width of the communication channel between the JPIP server and a client as well as the frame-rate selected by the user remain constant over time. This fact does not affect the description nor the efficiency of our approach.

We denote the image sequence with I and the n-th image of this sequence with I_n, where $0 \leq n < N$. The images that are finally displayed with CR are denoted as I'. Because in the context of JPEG 2000 the images can be decompressed from the same code-stream with several spatial resolutions, we define that $I_n = I_n^0$ is the image at the highest resolution and I_n^r is the spatially scaled version of I_n, that has a resolution of $S/2^r$, where S is the size of the images and $0 \leq r < R+1$. The encoding parameters for the image sequence are P, Q and R, where P denotes precinct dimensions, Q the number of quality layers and R the number of wavelet decomposition levels, or stages.

Basically, Our algorithm works as follows: the first image of the sequence is progressively transmitted for a given period of time. This is the image that is visualized. After, a thumbnail of the second image is transmitted with the aim of knowing which precincts of the second image should be updated by the CR. Notice that the JPEG 2000 packets that are necessary to reconstruct this thumbnail are also needed to reconstruct the second image at high resolution. This process is repeated with the following images. It also takes care to compute the distortion between images with the last update of each precinct that are not neccesary in the previous image. With these ideas in mind, we propose the following client-driven transmission scheme:

1. Let $q \leftarrow 1$ be the quality layer used for the current

SCITEPRESS SCIENCE AND TECHNOLOGY PUBLICATIONS

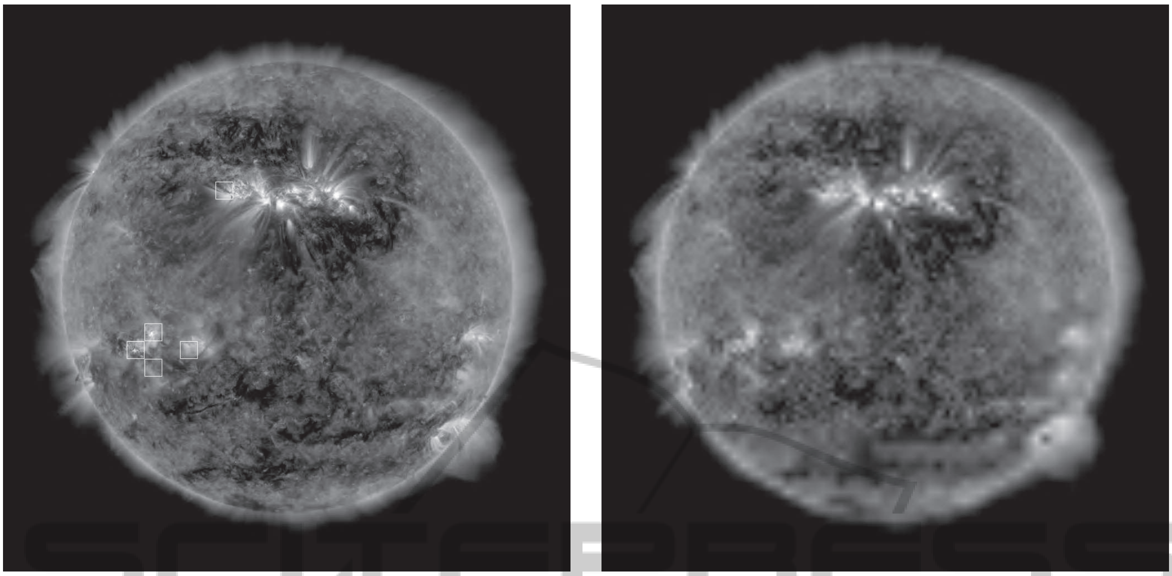

Figure 4: Two SDO/AIA reconstructed images, left with and right without CR. The refreshed precincts have been highlighted with a white square.

SCITEPRESS SCIENCE AND TECHNOLOGY PUBLICATIONS

Figure 5: Two *Akiyo* reconstructed images, left with and right without CR. The refreshed precincts have been highlighted with a white square.

reconstructions.

2. Let $n \leftarrow 1$ be the index of the reconstructed image.
3. Retrieve the next quality layer of I_{n-1} and initialize I'_{n-1} with this value.
4. Retrieve only the next quality layer of I_n^{R+1} (the thumbnail image).
5. Store in list L and sort in descending order, the precincts whose MSE (Mean Square Error) between I_{n-1}^{R+1} and I_n^{R+1} are larger than a given threshold λ.
6. $I'_n \leftarrow I_n(L) \cup (\overline{I'_{n-1} \cap I_n(L)})$, i.e., copy from I'_{n-1} to I'_n those precincts that remain "constant" and update from I_n those precincts that are in list L with quality q.
7. $n \leftarrow n+1$.
8. If $n < N-1$, go to step 4.
9. $q \leftarrow q+1$.
10. If $q \leq Q$, go to step 2.

With the objective of describing the operation of the algorithm, Figure 2 shows a schematic example of the transmission of the first five images of a sequence where the significant changes only happen in the center of the pictures. The data that are finally displayed are labeled with “Displayed with CR”. As it can be seen, with CR only the center of the images is retrieved with a higher quality (in this case, with 1 quality layer) and therefore, for a given bit-rate, CR reconstructions should be better than without CR from a SNR and from a visual point of view.

4 EVALUATION

This proposal has been tested with two different low-motion image sequences. The first one is composed by a set of 140 4Kx4K-resolution SDO/AIA images taken in the 17.1 nm channel (to show the finest structure and the one that is most dominated by Fe IX/X emission from 10^6K hot plasma), which has been log10-valued-scaled and corrected for exposure time variations. The second sequence[3] is composed by the CIF (352x288 pixels) color (although only the luma component has been used in our experiments) images of the *Akiyo* sequence. This sequence includes a static background and foreground with very little motion, only a head-and-shoulder video of an almost static announcer. The encoding parameters for SDO/AIA sequence has been: $P = 128$, $Q = 8$ and $R = 8$; and $P = 32$, $Q = 8$ and $R = 3$ for the *Akiyo* sequence. Experimentally, a good value for the threshold parameter λ was found to be 1. The reconstruction of the SDO images has been performed with a resolution of 1024x1024 pixels because this is the largest size that most actual displays can show of a power-of-two-resolution-scaled 4Kx4K image. In the case of *Akiyo*, no scaling has been applied.

The transmission of the SDO/AIA sequence has been simulated using a bit-rate of 27×10^6 bits/second (see the top plot of Figure 3). Although lower band-widths have been tested, the refreshing-rate (25 frames/second) of the precincts generated by the CR procedure is too slow to represent all the motion that the rotation movement of the Sun generates in the images. Nevertheless, a bit-rate of 8×10^5 bits/second has been enough to generate good reconstructions for the *Akiyo* sequence (see the bottom plot of Figure 3).

A visual comparison can be done using Figure 4 and 5 that shows the reconstruction of the second image of the SDO/AIA sequence and the fourth image of the Akiyo sequence, respectively. In these experiments, much smaller bit-rates (1163400 bits/second for SDO/AIA and 343800 bits/second for Akiyo) have been used in order to generate large visual differences between using CR and not. A frame-rate of 25 images/seconds has been considered for both sequences.

[3]Downloadable from http://trace.eas.asu.edu/yuv/akiyo/akiyo_qcif.7z.

5 CONCLUSIONS

This work shows how to improve the remote visualization of JPEG 2000 image sequences by means of client-driven conditional replenishment. The proposed system is fully compatible with the JPIP standard because we only have changed the order in which the packets are retrieved from the server. With the idea of improving the quality of the reconstruction of sequences with a higher degree of movement, motion compensated predictions could be generated by the clients. In any respect, this modification of the tested algorithm does not affect the applicability of our proposal.

SCITEPRESS SCIENCE AND TECHNOLOGY PUBLICATIONS

REFERENCES

Cheung, N.-M. and Ortega, A. (2007). Flexible video decoding: A distributed source coding approach. In *IEEE Workshop on Multimedia Signal Processing*, pages 103 – 106.

Devaux, F.-O., Meessen, J., Parisot, C., Delaigle, J.-F., Macq, B., and Vleeschouwer, C. D. (2009). Remote interactive browsing of video surveillance content based on jpeg 2000. *IEEE Transactionf on Circuits and Systems for Video Technology*, 19(8):1143 – 1157.

International Organization for Standardization (2004). Information Technology - JPEG 2000 Image Coding System - Core Coding System. ISO/IEC 15444-1:2004.

International Organization for Standardization (2005). Information Technology - JPEG 2000 Image Coding System - Interactivity Tools, APIs and Protocols. ISO/IEC 15444-9:2005.

McCanne, S., Vetterli, M., and Jacobson, V. (1997). Low-complexity video coding for receiver-driven layered multicast. *IEEE Journal on Selected Areas in Communications*, 15(6):983 – 1001.

Mounts, F. (1969). A video encoding system with conditional picture-element replenishment. *Bell System Technical Journal*, 48:2545 – 2554.

Müller, D., Fleck, B., Dimitoglou, G., Caplins, B. W., Amadigwe, D. E., Ortiz, J. P. G., B. Wamsler, A. A., Hughitt, V. K., and Ireland, J. (2009). JHelioviewer: Visualizing large sets of solar images using JPEG 2000. *Computing in Science and Engineering*, 11(5):38–47.

Naman, A. and Taubman, D. (2009). Rate-distortion optimized jpeg2000-based scalable interactive video (jsiv) with motion and quantization bin side-information. In *IEEE International Conference on Image Processing*, page 3081 3084.

Naman, A. and Taubman, D. (2010). Predictor selection using quantization intervals in jpeg2000-based scalable interactive video (jsiv). In *IEEE International Conference on Image Processing*, pages 2897 – 2900.

Ortiz, J. G., Ruiz, V. G., García, I., Müller, D., and Dimitoglou, G. (2010). Interactive browsing of remote jpeg 2000 image sequences. In *IEEE International Conference on Pattern Recognition*, pages 3179 – 3182, Istambul, Turkey. IEEE press.

Pesnell, W. (2008). The Solar Dynamics Observatory: Your eye on the Sun. In *37th COSPAR Scientific Assembly*, volume 37 of *COSPAR, Plenary Meeting*, pages 2412–+.

SCITEPRESS
SCIENCE AND TECHNOLOGY PUBLICATIONS