A Coding Scheme for Storage Surveillance in Timeline using SPIHT and ROI Coding

Roger Fredy Larico Chávez, Yuzo Iano, Osamu Saotome & Rogério Seiji Higa

***Abstract*—This work is a SPIHT coding scheme to generate a stream, with the option to prune it. The regions of interest ensure that the data which will be erased first is of low importance. A simple algorithm that prunes the SPIHT stream of a frame is scheduled in a given timeline.**

Some results are shown and conclude that the proposed scheme can be successfully used.

***Index Terms*— image compression, surveillance video, SPIHT, ROI coding, wavelet transform, storage video.**

I. Introduction

In many surveillance application systems, the captured video needs to be stored for analysis or transmitted to operators at a remote location [1]. The major problem in the case of stored videos is that the physical space of memory is limited. Because of that, the number of frames stored is restricted by the space available, and so is the timeline. This paper presents a proposal to increase the timeline of a stored video with the least quality impact on a region of interest (ROI), considering the same amount of physical storage. This work uses the schema of hierarchical coding and SPIHT (Set Partitioning in Hierarchical Trees) to assume that there is a defined region of interest. The main goal is to show how they influence the hierarchical coding, improving the storage specifically in the case of expanding the Timeline.

II. Video Surveillance Storage

Advanced surveillance systems use tags or labels embedded in the video in order to show the characteristics (or information) present in the footage. This speeds up the search process for motion, intrusion, or some event of interest in the surveilled region. For example, monitor a specific place in the parking lot. The moment a car gets out or comes in is quickly found using only logic tags (coded with some multimedia compression), it is not necessary to decode the video.

In this text, the conventional time of the clock is called the 'Timeline' (measured in years, months, weeks, days, hours, etc.).

The Fig. 1. shows the size of a random sequence of images (Mpeg), captured at a particular time. We consider that the video is made of I-frames only. The structure of an I-Frame varies depending on the details of the image. It is expected that the compression algorithm has a maximum threshold (eg. 80 kbits), to not overload the network traffic per unit time. Commonly the time is measured in seconds, and the rate in kbps. The most recent image of the video is the frame number 16 and the oldest image is the frame number 1, the Fig. 1 shows how the size of each frame of the video varies. To understand the procedure, we must first examine how only one frame of the video is processed by the proposed storage. For example the image marked by the red arrow in Fig. 1.

The frame is saved and stored conventionally by a determined amount of time (eg. 32 hours), after which the storage go to zero because the frame is erased.

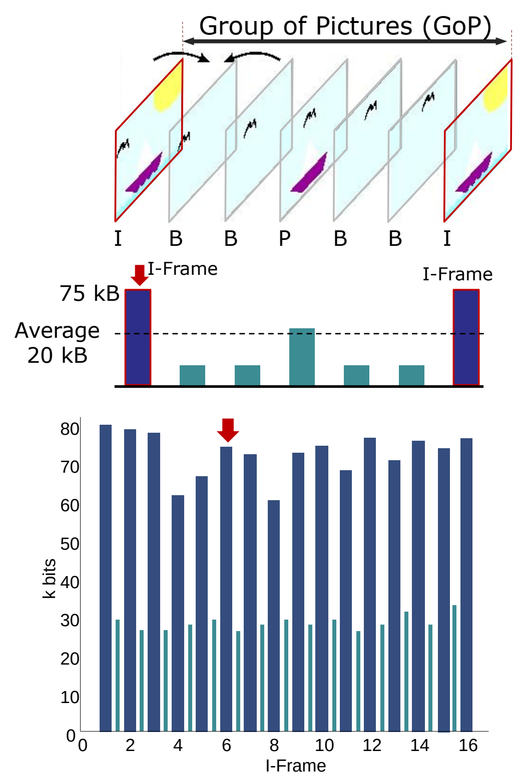


Fig. 1. Mpeg schema and Video stream with the length of each image (I-Frame in kilobits)

In Fig. 2 it is observed the lifetime of an image stored in a file, from now on the video frame will be analyzed as an image. It is important to understand that the bars in Fig. 2 are relative to a unique image or video frame, different from the

Manuscrito recebido em 4 de maio de 2011; revisado em 30 de setembro de 2011.

R. F. Larico (rlarico@decom.fee.unicamp.br), Y. Iano (yuzo@decom.fee.unicamp.br) e R. Higa (rhiga@decom.fee.unicamp.br) pertencem a Universidade Estadual de Campinas UNICAMP. Av. Albert Einstein, 400, Campinas, SP. O. Saotome (osaotome@ita.br) pertence a Instituto Tecnológico de Aeronáutica - ITA. São José dos Campos, SP.

Fig. 1 that shows multiple frames of various sizes. What is represented in Fig. 2 is the behavior of the size of the image during its lifetime in the storage memory.

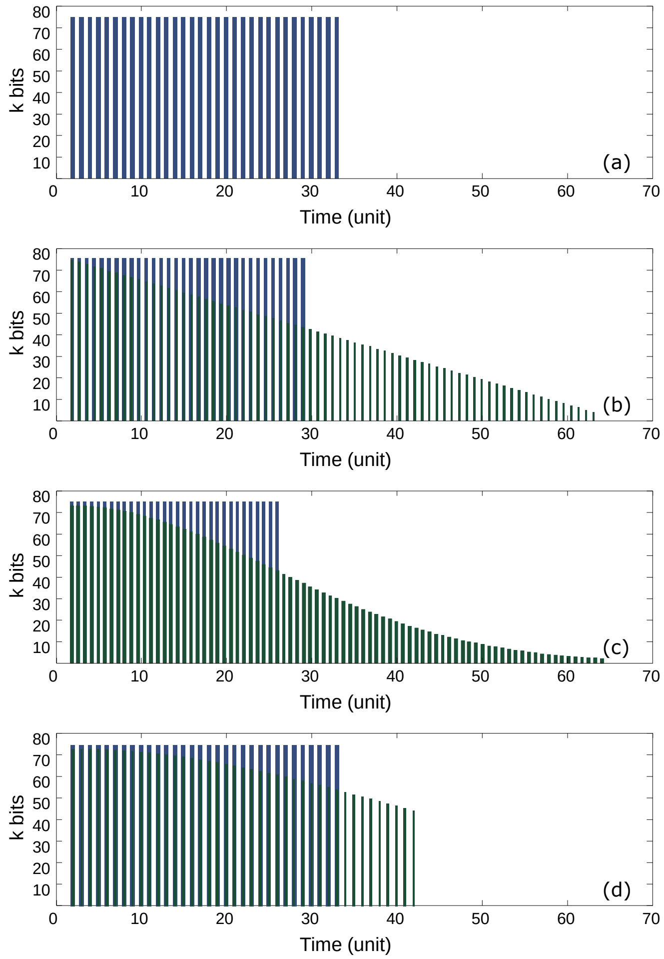


Fig. 2. Storage behavior of an image (i-frame) in the timeline, a) common app in blue, b), c), d) in green, the proposed behavior using differentiated and gradual degradation of the image

The proposed procedure allows that at each video frame in the timeline, a pruning or cutting of the encoded stream could be applied. This is possible due to hierarchical data trees process (SPIHT compression). The result of the procedure for only one frame is represented by the green bars in Fig. 2. This process must be controlled dynamically as shown in Fig. 2 (b), and may have a linear or Gaussian degradation process $g = e^{-x^2/2\times\sigma^2}$ (Fig.2 c, d). The sigma value and limits can be modified in the Gaussian curve $\int_0^{s1} g(x)dx$ and therefore guarantee the intended behavior, this will depend on the tags or user customization.

The metric used to compare the quality of the images are the PSNR and SSIM.

In addition, it is important that the coding gives priority to different visual information to ensure that our region of interest will be the last portion to be degraded.

III. Coding Scheme

The coding scheme can be divided in three main parts, as observed in Fig. 3.

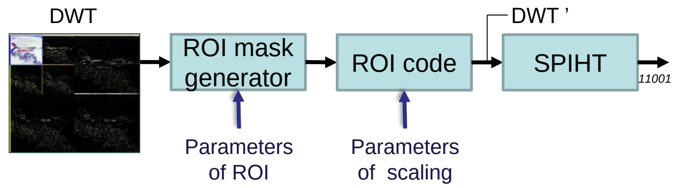


Fig. 3. Coding schema for stream generation of image compression

Wavelet transform has been used in a great variety of image coding problems [1]. As was described in [4], some of key features of wavelet transform which make it such an useful tool are as follows: spatial-frequency localization, energy compaction, decaying magnitude of wavelet coefficients across sub-bands [1]. The DWT of a signal x is calculated by passing it through a series of filters in two dimension, see Fig. 4.

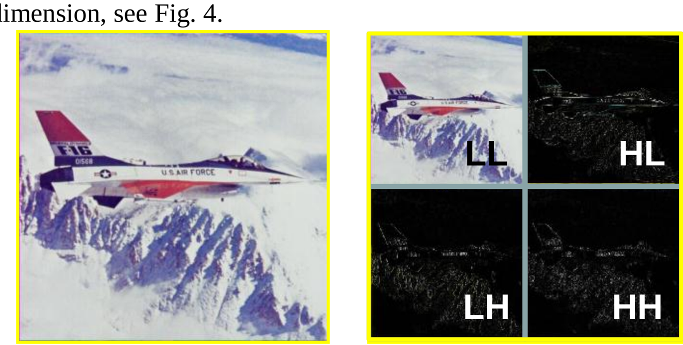


Fig. 4. Wavelet Transform 2D of one levels applied on the airplane image

SPIHT algorithm utilizes: 'searching for set in spatial orientation trees in a wavelet transform'; 'partitioning the wavelet transform coefficients in these trees into sets defined by the level of the highest significant bit in a bit plane representation of their magnitudes'(desirable for control rate); and 'coding and transmitting bits associated with the highest remaining bit planes first' (desirable for control quality). As was described in [2]. The algorithm shown in Fig. 5 is divided in four parts: initialization, sorting pass (set partition coding), refinement and threshold update.

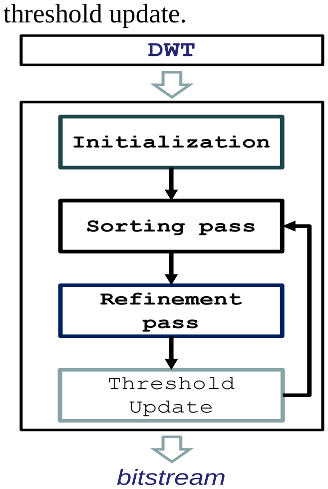


Fig. 5. Diagram for SPITH Algorithm

The main advantages are the simple implementation, fast coding, and low CPU usage. Its behavior is shown in more details on the site http://www.decom.fee.unicamp.br /~rlarico/ SPIHT/.

Region of interest (ROI). In many surveillance application ROIs are often captured with some information such motion, texture, etc. The references [1] and [3] are coded with 'tag' techniques (desirable for ROI detection). The wavelet coefficient in the box will be scaled up or the coefficients in the background scaled down, as depicted in [3]. There are several ways of coding the ROI as shown in Fig. 6. The Maxshift and Scaled methods are used according to the application, but other methods of hybrid ROI coding can lead to better subjective results (http://www.decom.fee.unicamp .br/~rlarico/ROI/).

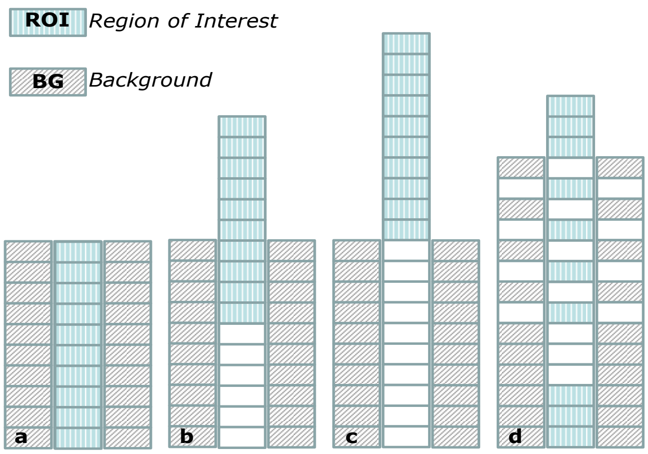


Fig. 6. ROI Scaling Methods a) without ROI; b) *ScaleShift* method of general ROI scaling; c) *Maxshift* method; d) hybrid method *HBShift.*

IV. Test and Simulation

The simulation tests used typical images such as the Airplane, Baboon, Barbara, Goldhill, Lena, Peppers, Satellite, and Tomography as observed in Fig. 7. The sizes of the images were 256x256 and 512x512. The storage of an image (I-Frame) is assumed to have a lifetime by a specified amount of time. Like the conventional storage the images are deleted after a certain period of time. This proposal will cut the stream to a minimum portion generated by the SPITH ROI according to the timeline. As a consequence the reduction in stream deteriorates the image until you reach a limit that is customized by the user.

Tests were done with "ROI scaling" with s = 0, s = 1 and s=2. It also could be applied to larger scales values. It is assumed that the conventional and the proposed system both use the same image coding for comparison purposes (In these tests the results were not compared with other conventional coding, however this analysis will be done by comparing the intra-frame images used by MPEG4. Part 10). The image is degraded automatically based on the custom bell curve. In Fig. 8 we observe that even with deterioration, if the system is able to generate labels and create an ROI, the most important details will be maintained. The details will not be lost abruptly and the image storage lifetime can be expanded.

Fig. 7. Original images airplane, Barbara, Lena, Goldhill, Peppers, Satellite and Tomography

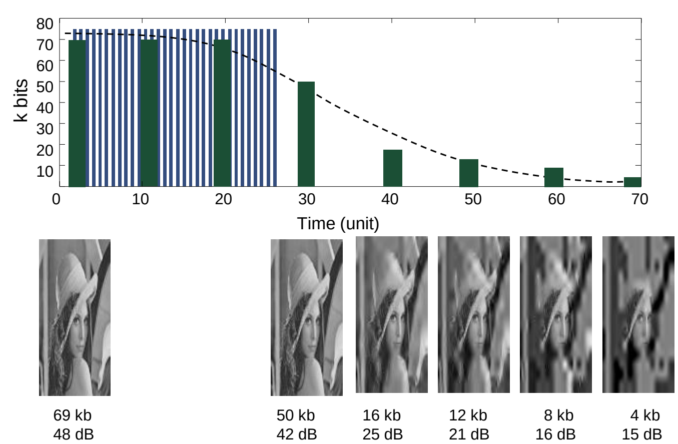


Fig. 8. Behavior of the degradation of the image Lena (128x128) reducing the storage space of the stream using the SPIHT with an ROI centered on the face

The Fig. 8 is relative to images of size 128x128 as a preliminary result.

The Fig. 9 shows an airplane image of size 512x512, with linear degradation without ROI. It is observed that the SPITH algorithm achieve low degradation in quality even without ROI.

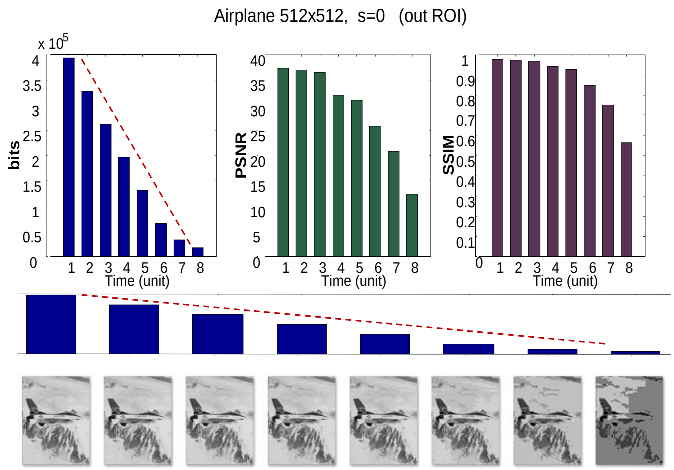


Fig. 9. Behavior of the degradation of the image airplane (512x512) reducing the storage space of the stream using the SPIHT without ROI s = 0

Because of security reasons there is a parameter to indicate when the image will start to degrade. Therefore until this threshold is reached the image is not degraded. After the time specified by the security parameter, the image stream will start to decrease. In real environments this threshold will be limited by the physical storage. We analyzed only the points where the degradation in dB is at an acceptable quality and has increased the storage time. The quality of image after degradation ranges from 25dB up to 15dB. The Fig. 10 shows the same image of Fig. 9 with ROI (s = 2). It is observed that the last image preserves the important features of the region of interest.

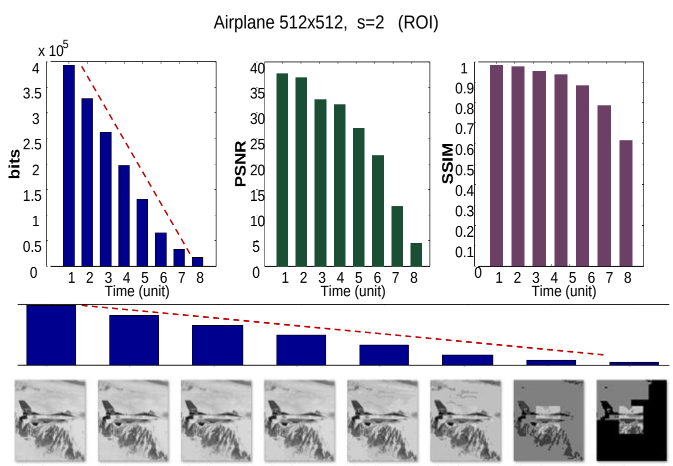


Fig. 10. Behavior of the degradation of the image airplane (512x512) reducing the storage space of the stream using the SPIHT with ROI s=2

For surveillance application it is normally used the 4CIF resolution which is better approximated by the 512x512 test images. In Fig. 11 and Fig. 12 is shown some results of acceptable degradation with high compression rate. The information in the filename is organized as [original image name]_[level]_[size in bits]_[used vshift 's'], if s=0 then no ROI is applied]. (eg. peppers512.bmp_rec_3_52429_bitneed0)

The expected behavior can be seen in all of the images tested. Additional plotted PSNR and SSIM results are available at http://www.lcv.fee.unicamp.br/rlarico/iwt2011/. Some curves are shown in Fig. 13-19. The images of interest here are the airplane, Barbara, Goldhill and Lena. These images are the closest a surveillance camera would acquire. That is the faces and body of the people, vehicles and buildings. It is also observed that because of the SPITH, the SSIM does not change like the PSNR, since at each pass of the algorithm the structure is preserved and enhanced.

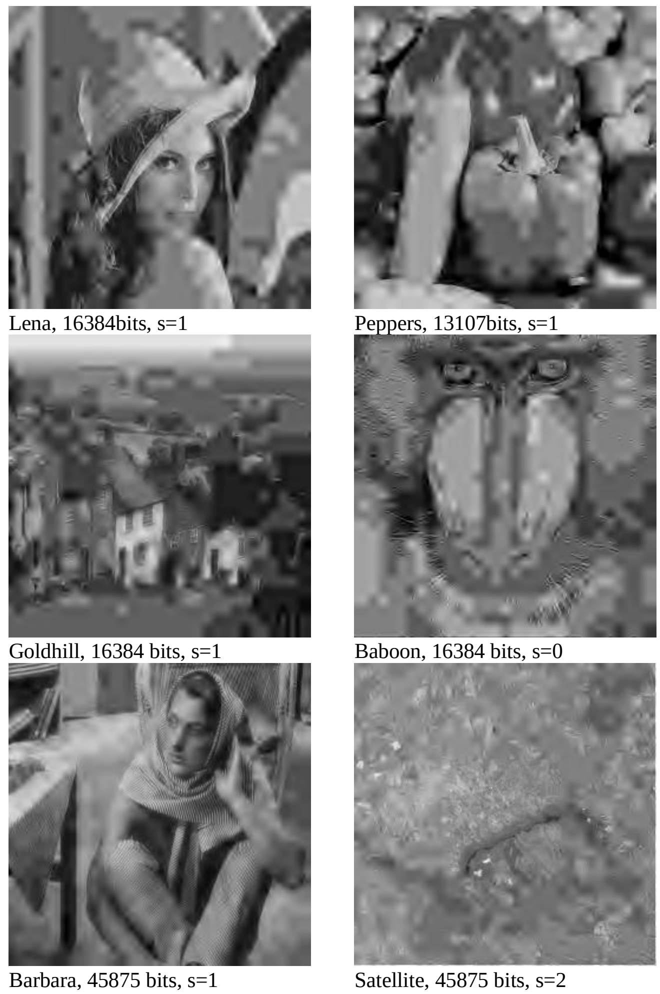


Fig. 11. Behavior of the degradation of some images 256x256 pixel encoded SPIHT with ROI

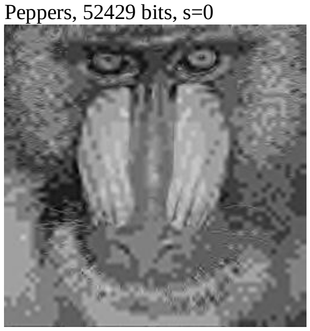
V. CONCLUSIONS

The images are stored using SPITH coding. This removes the need of re-encode the frame to reduce the size. The image can be stored for a longer time while degrading the quality. This enhances the number of images stored without changing the storage devices. The proposed procedure allows lower rates for video applications. Therefore improving the storage and increasing the number of images in the stored video while keeping a reasonable quality of the image.


Lena, 52429 bits, s=1


Peppers, 52429 bits, s=0


Goldhill, 52429 bits, s=1


Baboon, 65536 bits, s=1


Barbara, 65536 bits, s=0


Satellite, 131072 bits, s=2

Fig. 12. Behavior of the degradation of some images 512x512 pixel encoded SPIHT with ROI

In Fig. 8 we can see that the storage time is doubled. Then the initial quality is deteriorated at the time 32. However, the ROI can save important information with small loss, 23dB at most. The same behavior could be seen on Fig 12 plus the PSNR and SSIM plotted results.

In the Fig. 9 and Fig. 10 is shown thumbnails of the images resulted by the pruning algorithm. We can see that the subjective quality of these images in relation to the image size is very good.

The proposed system aims to improve the management of stored frames from a hard erase to a soft erase procedure. So this will allow better quality images in the region of interest, increase of the number of images stored while keeping the storage capacity. This scheme could work after a conventional storage of video acquisition (based on hierarchical coding), ensuring the expansion of the Timeline.

This study shows that a schedule for decrease the quality by a known fixed rate is possible. This enhances the efficiency of the video storage.

It is necessary to have a minimum quality parameter control for the general case to make sure that the images are still usable. This will avoid that storage space is wasted saving unusable images.

The proposed next steps would to find a subjective analyzer algorithm that is suitable for security applications to complement the objective metrics PSNR and SSIM. And to make the pruning control based on the quality instead of the size of the image.

VI. ACKNOWLEDGEMENTS

The authors would like to thanks the following institutions for the support on this research: ITA-ITASAT, CAPES-RHTVD, CNPq, CAPES, FAPESP, FAEPEX/UNICAMP.

REFERENCES

[1] LiuXuedong, "A Coding Scheme for Surveillance Video," IEEE, 2007.
[2] William A. Pearlman and Amir Said, "Set Partition Coding: Part I of Set Partition Coding and Image Wavelet Coding Systems," Foundations and Trends in Signal Processing, 2008.
[3] Ye Shujiang; Gu Yanfeng; Zhang Ye; "An automatic ROI coding algorithm based on multi-resolution target detection," Systems and Control in Aerospace and Astronautics, 2008.
[4] Stéphane G. Mallat, "A wavelet tour of signal processing: the Sparse way," Academic Press , Elsevier, 2009.


Roger Fredy Larico Chávez, Systems Engineer for San Agustin University of Arequipa. Peru (UNSA) in 2002. Masters and Doctor degree in Electrical Engineering from State University of Campinas (UNICAMP) in 2012. Currently, this researcher in electrical engineering at Visual Communication Laboratory (LCV) Department of Communications, School of Electrical and Computer Engineering, UNICAMP. Acting on the following topics: Image processing (multispectral analysis, enhancement, segmentation, detection, compression and storage).


Yuzo Iano, received his PhD. in electrical engineering in 1986. Currently he is an associate Professor in electrical engineering at Unicamp (State University of Campinas, Brazil). He also works at Visual Communication Laboratory (LCV) on the same University. He is responsible for some digital signal processing (sound and image) projects. His research interests include video and audio coding, digital video and audio compression and nission.


Osamu Saotome, bachelor's at Engenharia Eletrônica from Instituto Tecnológico de Aeronáutica (1974), master's at Processamento Digital de Sinais from Tokyo Institute Of Technology (1984) and doctorate at Processamento Digital de Sinais from Tokyo Institute Of Technology (1987). Has experience in Electric Engineering, focusing on Electronic Circuits, acting on the following subjects: Real time systems, electronic devices and systems, algorithms, digital signal processing and radar reciver.


Rogério Seiji Higa, is currently a PhD. candidate at UNICAMP (State University of Campinas, Brazil). He received his Masters Degree in Electrical Engineering at UNICAMP in 2008. He also works at Visual Communication Laboratory (LCV) on the same University. His research interests include image processing and computer graphics.

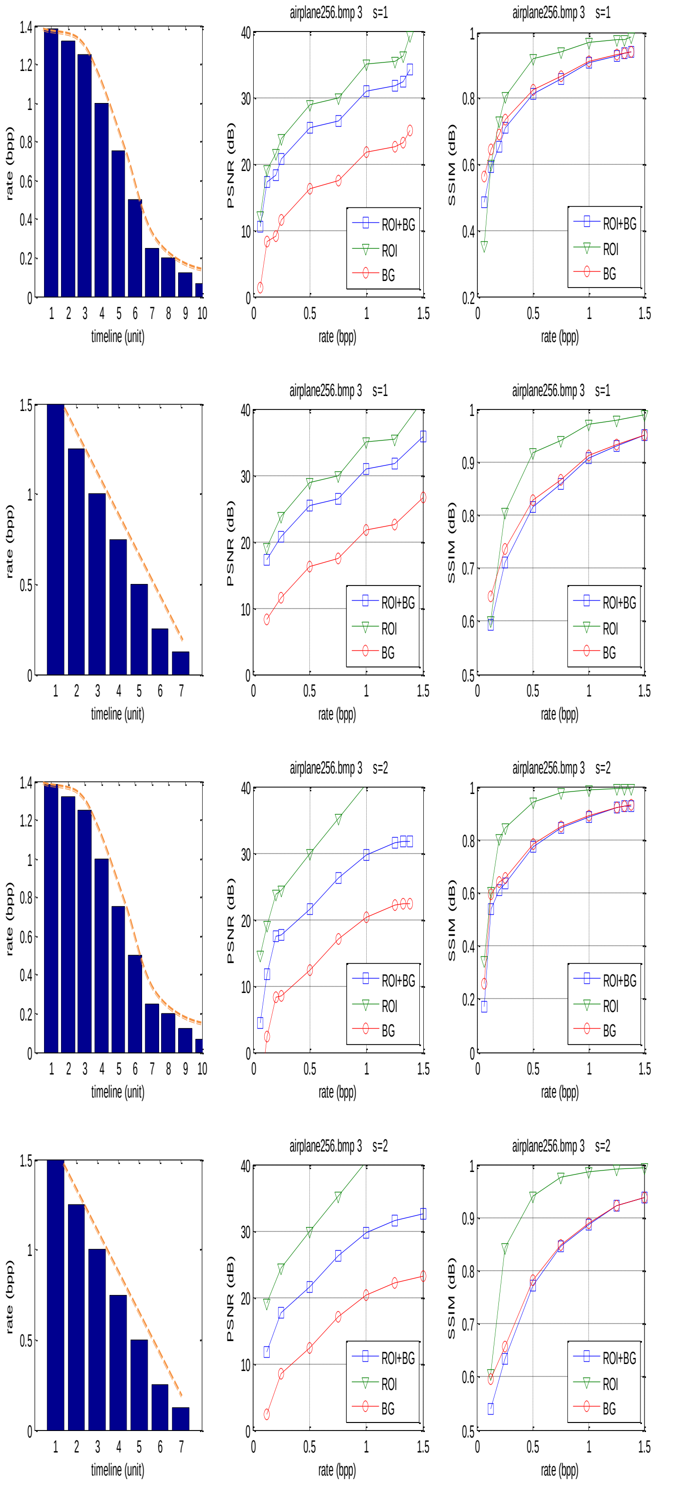


Fig. 13. Behavior of the degradation PSNR SSIM of airplane 256x256 pixel encoded SPIHT with ROI

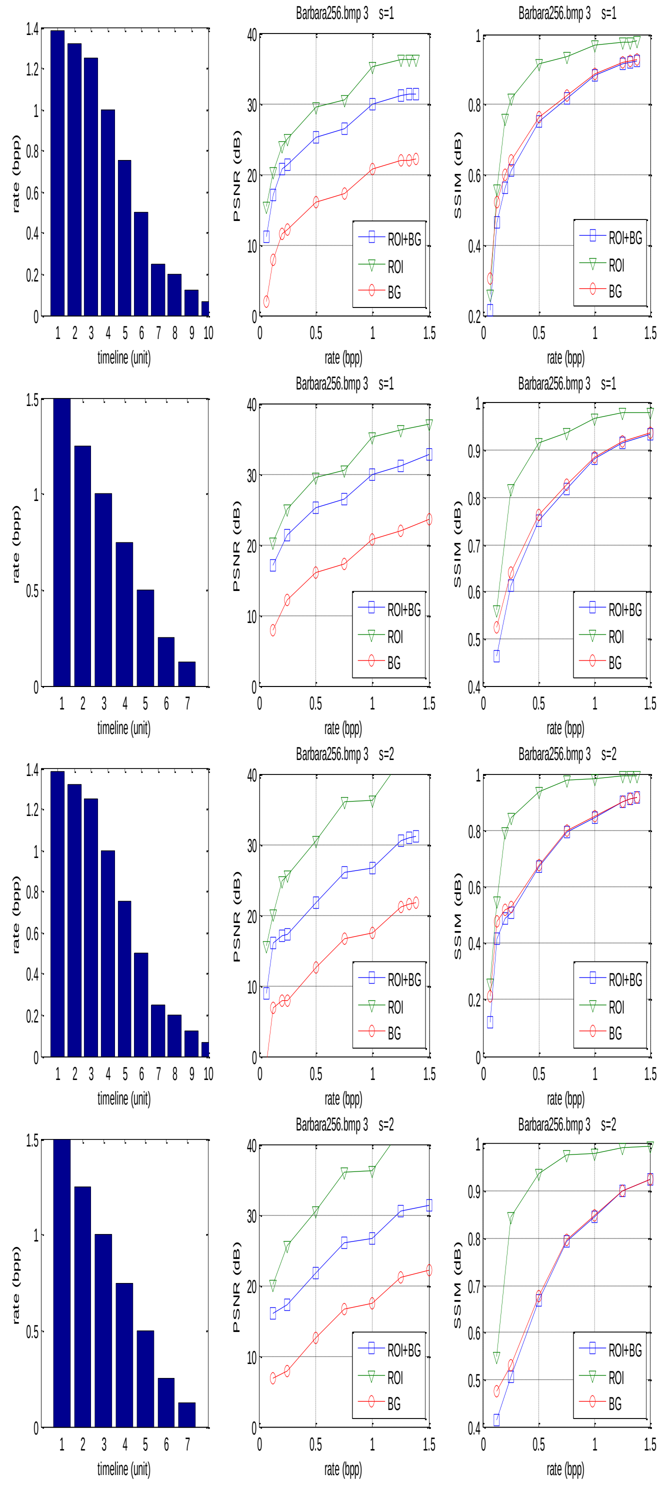


Fig. 12. (b) Behavior of the degradation PSNR SSIM of Barbara 256x256 pixel encoded SPIHT with ROI

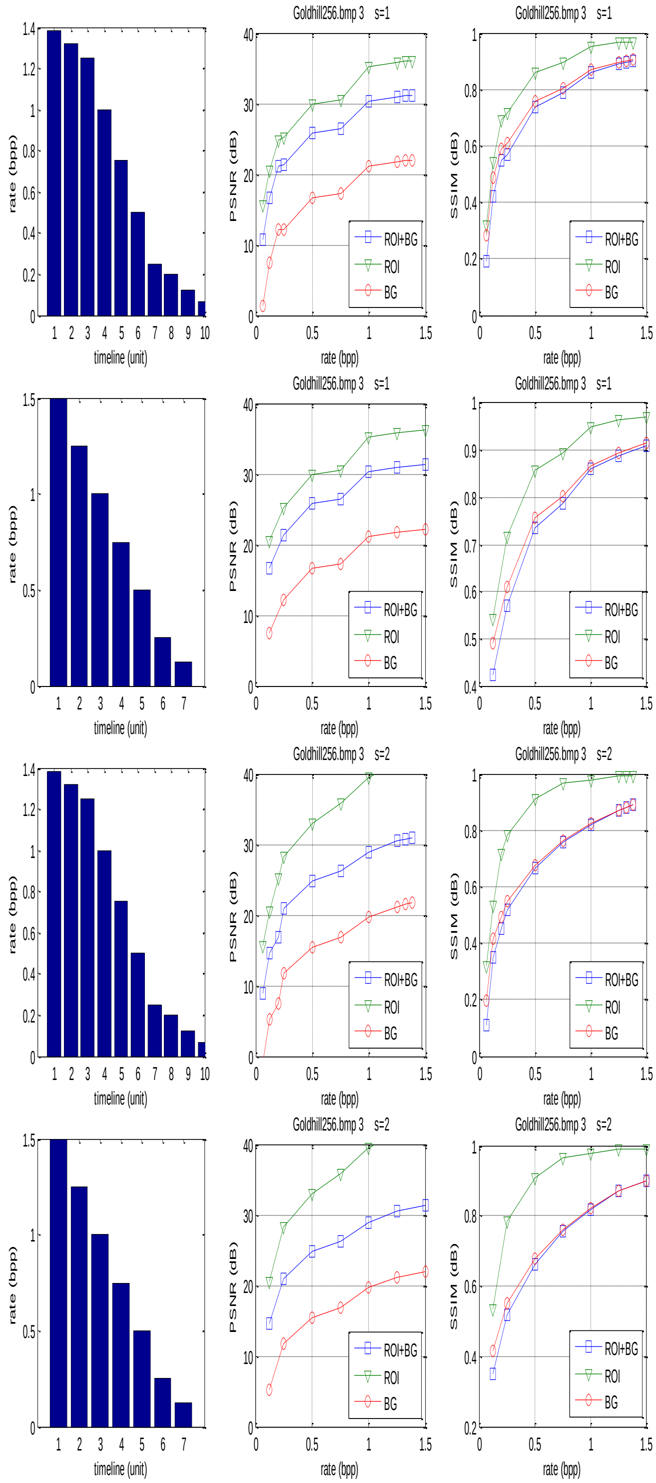


Fig. 14. Behavior of the degradation PSNR SSIM of Goldhill 256x256 pixel encoded SPIHT with ROI

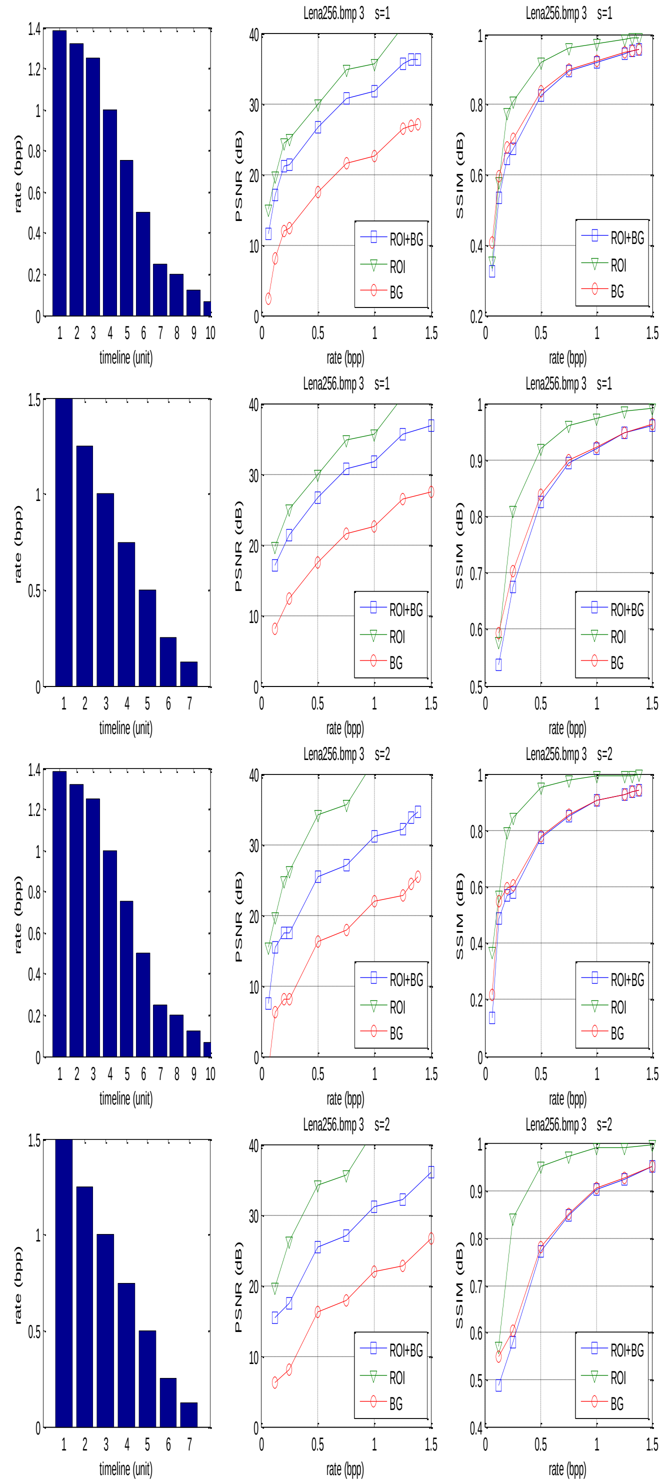


Fig. 15. Behavior of the degradation PSNR SSIM of Lena 256x256 pixel encoded SPIHT with ROI

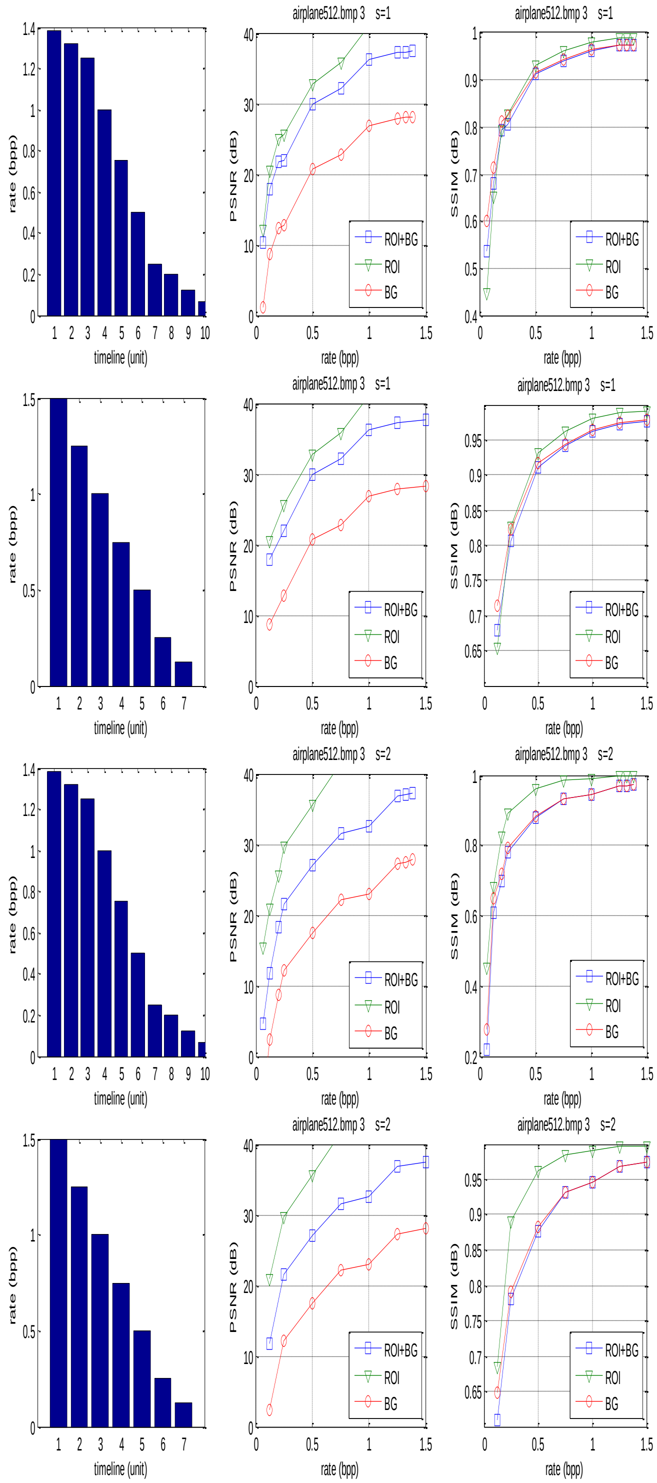


Fig. 16. Behavior of the degradation PSNR SSIM of airplane 512x512 pixel encoded SPIHT with ROI

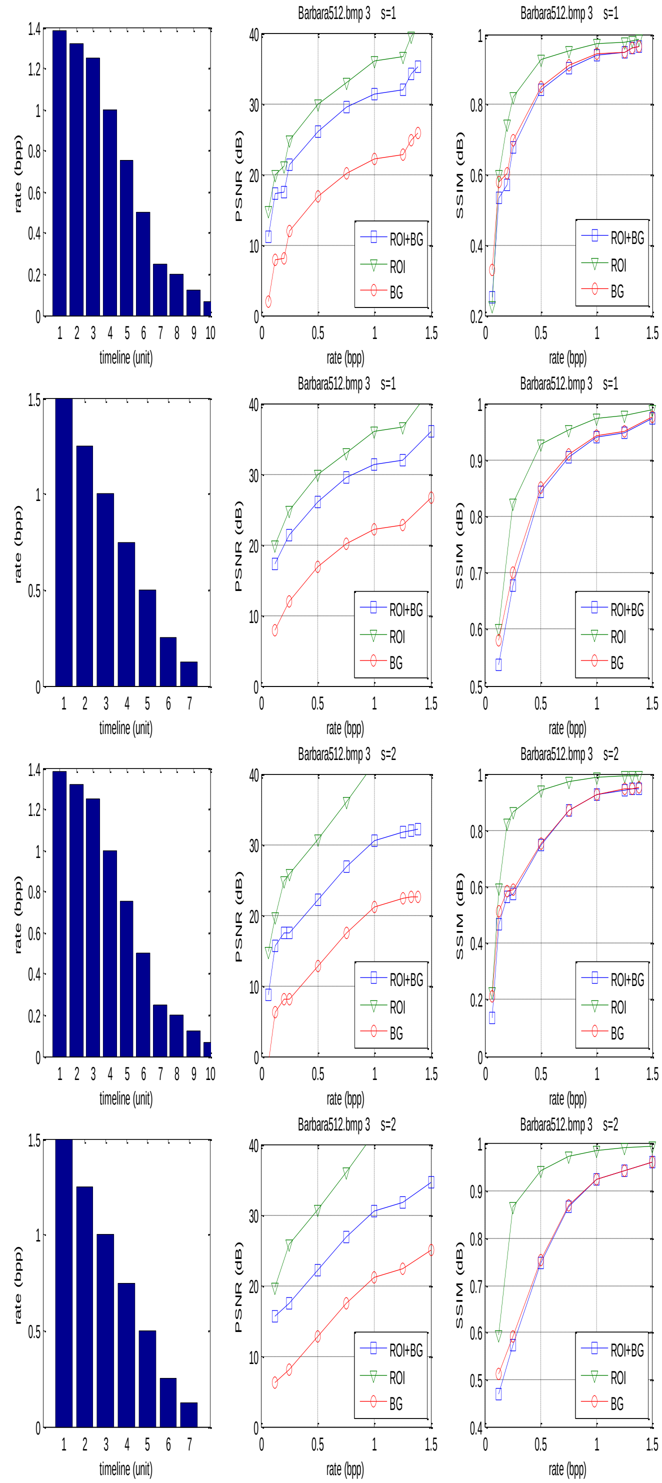


Fig. 17. Behavior of the degradation PSNR SSIM of Barbara 512x512 pixel encoded SPIHT with ROI