ISSN 0005–1144
ATKAAF 48(1–2), 21–25 (2007)

Mladen Tomić, Nino Stojković

Framework for Adaptive Postprocessing of Video Signals

UDK 621.397:004.932
IFAC 5.8.4

Preliminary communication

Inclusion of the latest research work to a real life applications and achieving adaptive video signal postprocessing is a complex and timely task. The proposed framework breaks the filtering process into several distinct phases and enables a researcher to move at a faster pace by narrowing the scope of the research and enabling seamless integration of this work into inherently adaptive filtering process. Some of the benefits of this modular and adaptive framework have been presented on the example of the MPEG-4 postprocessing improvements.

Key words: signal processing, video compression, MPEG-4, adaptive filtering, filtering framework, postprocessing

1 INTRODUCTION

As video signals became common in the most aspects of human activity, a lot of effort has been put in ensuring as high reproduced video quality as possible. This was vitally important in the case of lossy coded video signals, but it is also important for the other types of video signals. For that purpose, many different filtering techniques and filters were developed. Averbuch *et al.* in [1] proposed usage of weighted sums of symmetrically aligned pixels for deblocking of block-transform compressed images. They define the basic weights, to maximize efficiency at lower bit-rates but also adapt them at higher bit-rates to avoid edge ghosting effect. The masking effect of the Human Visual System (HVS) is taken into consideration by Chen *et al.* in [2]. They use large processing window in a low-activity areas to ensure proper image smoothing where is no background masking ability, but in a high-activity areas the window is reduced and the greatest weight is given to central pixel, therefore maintaining image details. Alter *et al.* in [3] proposed technique based on a weighted total variation minimization which allows them to efficiently deblock the image without too extensive oversmoothing. Different approach is taken by Liew *et al.* in [4] who proposed the method based on projection onto convex sets (POSC). Instead of using strict 8×8 blocks of pixels, their algorithm enforces smoothness in a more general region that might span several blocks. The thin-plate spline surface modeling is used for that purpose.

However, there are two basic problems that still show up with current research approach – it is very difficult to derive a method that will perform extremely well at the whole range of different images contained in a video signal, and also, the process of integration of research achievements in real life applications is too complex. The proposed framework is trying to address both of the issues by breaking the process of filtering into several distinct phases/processes. This way we facilitate adaptation to a specific video signal and enable the enhancements in some part of filtering process not to affect the other parts. In effect, research areas are narrowed to more specific topics where better results could be achieved, and integration with the existing applications is greatly simplified and made more transparent. This is all possible because of the clear and distinct rules for each process function and the way in which parameters are passed and used during different phases of the video signal processing. The example of MPEG-4 video signal postprocessing using the proposed framework will be presented.

2 FRAMEWORK FOR ADAPTIVE POSTPROCESSING

Postprocessing of the video signal is performed on every frame by any number of filters. Filter pool may contain more than one filter for the same purpose. For instance, there may be a single filter for the contrast enhancement and three filters for blocking artifacts reduction. To make postprocessing adaptive and independent of the video signal, the process of filtering is broken into six phases: Postprocessing Parameters Extraction (PPE), Signal Quality Estimation (SQE), Filtering Level Estimation (FLE), Filter Selection (FS), Filtering (F) and, finally, Filter Efficiency Verification (FEV). Each phase is implemented by means of respective Algorithms (PPEA, SQEA, FLEA, FSA, FA and FEVA, respectively) and in as many steps as needed, in different phases of the video signal processing. Informations between these steps are interchanged. The whole framework is shown in Figure 1.

During the PPE phase, video signal is analyzed and parameters for every filter are extracted. These parameters could mark the portions of the frame to be filtered, filtering mode to be performed or any other filter specific information. This process is influenced by filtering level, as determined by the FLE.

SQE should provide at least rough estimation of the current frame quality, in aspects that are meaningful for each of the filters.

FLE is the core phase for the adaptive postprocessing. By using informations from the SQE and other appropriate parameters, FLE derives optimal level of filtering for each filter. Apart from the SQE information, FLE takes as its input the measure of available system resources, output video buffer level, user preferences and previous level for each of the filters. Other inputs, such as the FEV output, could be added if necessary. The derived filtering level should provide the highest pos-

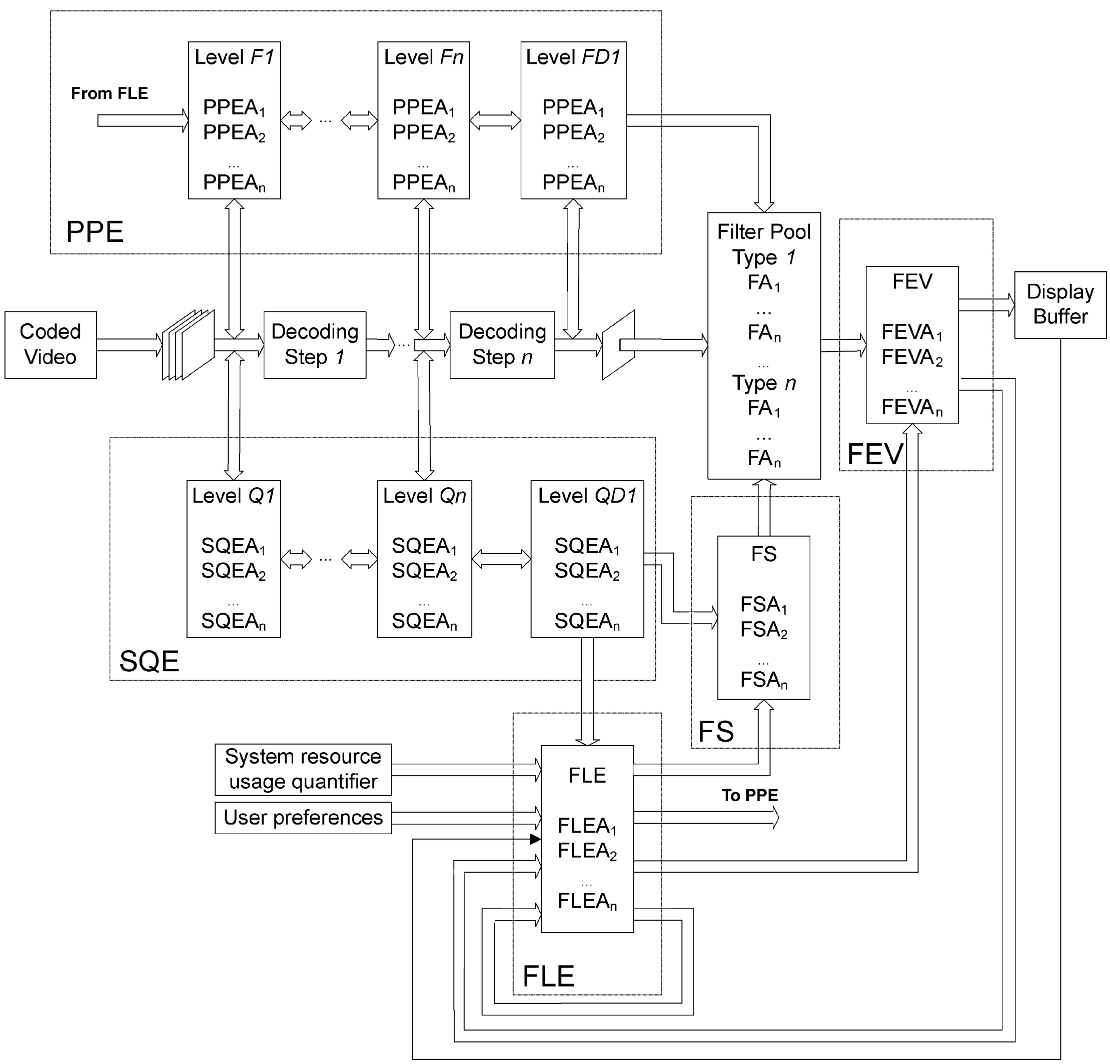

Fig. 1 Proposed framework for adaptive postprocessing

sible reproduced video quality that could be achieved with the available system resources and the current filter pool. In case of the lowered filtering levels, because of the insufficient system resources, FLE considers the priority of filters, as determined by the Human Visual System (HVS) characteristics or other factors included in the FLE algorithms.

With the FLE and SQE information available, the FS process is able to select the most appropriate filter from the filter pool, for each of the filter types. After the filtering, the FEV may be implemented for each filter type, which will verify that the filter efficiency was acceptable. Informations from the FEV may be used to request from the FLE increase or decrease of the filtering level for the specific filters, but also for the analysis of the entire process.

3 IMPLEMENTATION OF MPEG-4 POSTPROCESSING FILTERS

MPEG-4 uses deblocking and deringing filters to reduce the effect of blocking artifacts and mosquito noise, respectively. The filtering process is computationally very intensive and to achieve the speed gains with reasonable quality loss, MPEG-4 introduced blocking and ringing semaphores [5], which lead to the application of strong or weak deblocking or deringing filter in the case of the respective semaphore being 1 or 0. The efficiency of this approach is widely limited by the video signals bit-rate. By modifying this filtering process and implementing it through the proposed framework, the adaptive filtering will be achieved that will provide the highest possible reproduced video quality, with available system resources, at the broad range of bit-rates and with significant speed improvements at higher bit-rates.

PPE is implemented in two steps for both the deblocking and the deringing filter. One step is used to extract the semaphores of the intra frame, while the other is used to extract the semaphores of the inter frame. MPEG-4 defines that the horizontal blocking semaphore of intra-frame should be set to 0 if there is at least one non-zero coefficient in the top row of 8×8 matrix of the inverse quantized DCT coefficients. The same way, vertical blocking semaphore of the intra-frame should be set to 0 if there is at least one non-zero coefficient in the far left column of the coefficients matrix. Blocking semaphores for inter-frames are extracted in a similar way by using not only the residual signal but also the blocking semaphores of the reference block (as indicated by the motion vector). These rules are altered in a way that the number of coefficients needed to set the blocking semaphore to 0 becomes variable, and that variable will be called DeblockingLevel, with the range of valid values [0, 8]. For the value of 0, deblocking filter will be completely disabled, while for the value of 1 the blocking semaphores will be set in the same way as in the MPEG-4 standard. The remaining values are used to fine-tune the appropriate level of filtering, with the value of 8 meaning that all block boundaries in the frame will be processed.

The ringing semaphore should be set to 0 if there is at least one non-zero coefficient in the 8×8 DCT coefficients matrix in the positions other then the DC component or the first vertical, and the first horizontal AC component. The same way as in the case of the blocking semaphore this number is set to be variable and this variable will be called DeringingLevel, with a range of valid values [0, 62]. The value of 0 means that all blocks will be processed, while the value of 1 results in the ringing semaphores being set the same way as defined in the MPEG-4 standard. Again, the remaining values are used to fine-tune the level of processing, with the value of 62 being used to completely disable deringing filter. Both the DeblockingLevel and the DeringingLevel are derived by the FLE and input to the PPE.

SQE for this purpose is very simple and it purely extracts the quantization parameter (QP) of the current frame, but it still provides a solid estimation of the level of blocking artifacts and mosquito noise that will arise. FLE takes three inputs: SQE quality estimation (QP of the current frame), available system resources measure and the output video buffer level. Considering latter two inputs, FLE will select appropriate values for DeblockingLevel and DeringingLevel to achieve the highest possible quality.

A number of tests, using subjective quality assessments, were carried out to determine the appropriate levels for these variables and the results obtained were subject to objective assessment through the FEV to further confirm its validity. Since there was only one filter for each filter type, the FS needed not to be implemented. FEV was used to provide objective assessment of the filtering process. It was implemented using the recently proposed Structural Similarity (SSIM) index [6, 7] which proved to provide a much better correlation between the calculated and the perceived quality than the usual PSNR measure.

Figure 2 shows SSIM values for the CIF format *Foreman* sequence for the standard implementation and the framework implementation, relative to the full filtering of every block and block boundary

$$\frac{SSIM - SSIM_{full}}{SSIM_{full}},$$

while Figure 3 shows processing times for different CIF format sequences by using the framework implementation, relative to the standard implementation

$$\frac{t_{fram} - t_{stand}}{t_{stand}}.$$

Because of the adaptive nature of the framework implementation, the highest quality is retained at a wide range of bit-rates, and still in the worst case

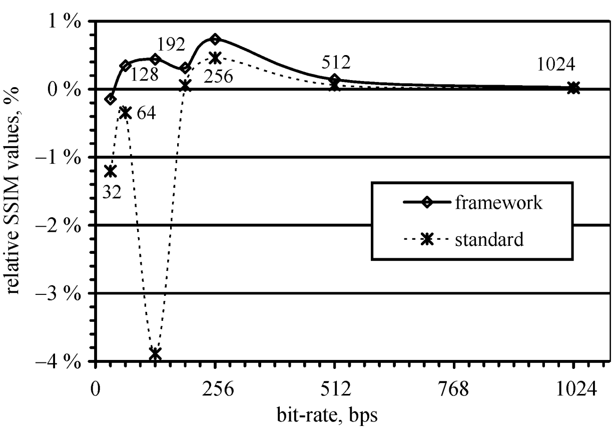

Fig. 2 Comparison of SSIM values, relative to the full filtering, for the CIF format Foreman sequence

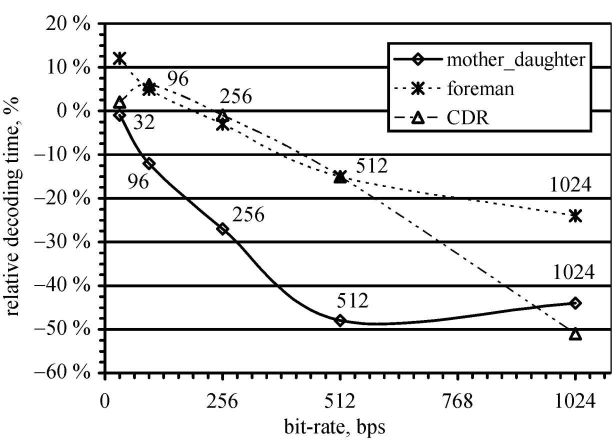

Fig. 3 Comparison of framework implementation's processing times, relative to the standard approach

processing is just over 10 % slower than the standard implementation, but as the bit-rates rise, in some situations, it becomes more than 50 % faster. More benefits occur in the case of very limited system resources. In the most current applications the lack of available resources is solved by completely disabling certain filters, which inevitably causes the loss of quality. Framework implementation introduced adaptive fine-grained tuning of filtering level and made it possible to use virtually all of the available resources to achieve as higher quality as possible in the case of either temporary or permanent lack of system resources.

4 CONCLUSION

The proposed powerful and modular framework opens many research possibilities for all of the basic building blocks – PPE, SQE, FLE, FS and FEV. By strictly separating these research areas it becomes possible to form a pool of postprocessing filters, each of them being superior in a certain area of application and being selected by the FS decision and its application level being determined by the FLE. This way, the framework may be applied to a video signal of any given quality, to be enhanced in a number of ways. Example presented showed how seamless integration of research work to a real life application can be and what improvements this framework can introduce. Being inherently adaptive, it crosses the boundary of being applicable only to a specific video signals and the innovative approach that was taken makes it possible to use it for a wide range of different video signals and applications, which promotes it to a valuable tool for the researchers and the developers in the extensive area of video signal processing and reproduction.

ACKNOWLEDGEMENT

This work is supported by Croatian Ministry of Science and Technology under grant 0069015.

REFERENCES

[1] A. Z. Averbuch, A. Schlar, D. L. Donoho, **Deblocking of Block-Transform Compressed Images Using Weighted Sums of Symmetrically Aligned Pixels**. IEEE Transactions on Image Processing, vol. 14, no. 2, pp. 200–212, Feb. 2005.

[2] Tao Chen, Hong R. Wu, Bin Qiu, **Adaptive Postfiltering of Transform Coefficients for the Reduction of Blocking Artifacts**. IEEE Transactions on Circuits System Video Technique, vol. 11, no. 5, pp. 594–602, 2001.

[3] F. Alter, S. Durand, J. Froment, **Deblocking DCT-Based Compressed Images with Weighted Total Variation**. In Proceedings of International Conference on Acoustics, Speach and Signal Processing ICASSP 2004, Montreal, Canada, vol. 3, pp. 221–224, 2004.

[4] A. W. C. Liew, H. Yan, N. F. Law, **POCS-Based Blocking Artifacts Suppression Using a Smoothness Constraint Set with Explicit Region Modeling**. IEEE Transactions on Circuits and Systems for Video Technology, vol. 15, no.6, June 2005, pp. 795–800.

[5] ..., **Information Technology – Coding of Audio-Visual objects – Part 2: Visual, International Organisation for Standardisation**, ISO/IEC 14496-2: 1999/Amd. 1:2000(E), July 2000.

[6] Z. Wang, L. Lu, A. C. Bovik, **Video Quality Assessment Based on Structural Distortion Measurement, Signal Processing: Image Communication**. special issue on Objective video quality metrics, vol. 19, no. 2, pp. 121–132, February 2004.

[7] Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, **Image Quality Assessment: From Error Visibility to Structural Similarity**. IEEE Transactions on Image Processing, vol. 13, no. 4, pp. 600–612, Apr. 2004.

Sustav za adaptivno postprocesiranje videosignala. Uključivanje najnovijih istraživačkih dostignuća u svakodnevne primjene i postizanje adaptivnog postprocesiranja videosignala zahtjevan je zadatak za koji je potrebno mnogo vremena. Predloženi sustav dijeli postupak filtriranja u nekoliko jasno odvojenih faza te na taj način sužava opseg svakog pojedinog istraživanja. Time je omogućena jednostavna integracija novih dostignuća u inherentno adaptivan proces filtriranja, što omogućava osjetno brži napredak u istraživanju. Neke prednosti ovog modularnog i adaptivnog sustava prikazane su na primjeru poboljšanja procesa postprocesiranja videosignala kodiranog prema MPEG-4 normi.

Ključne riječi: obrada signala, kompresija videa, MPEG-4, adaptivno filtriranje, filtarski sustav, postprocesiranje

AUTHORS' ADDRESSES

mr. sc. Mladen Tomić
University of Rijeka, Faculty of Engineering
Vukovarska 58, 51000 Rijeka
HR – Croatia

Nino Stojković, associate professor
University of Rijeka, Faculty of Engineering
Vukovarska 58
51000 Rijeka
HR – Croatia

Received: 2007-02-15