Artifacts suppression in images and video. Non-Local Means as algorithm for reducing image and video distortions.

Fedak V.I, Nakonechny A.Y.

***Abstract*— Images and video are often coded using block-based discrete cosine transform (DCT) or discrete wavelet transform (DWT) that cause a great deal of visual distortions. This paper reviews a range of image and video artifacts and highlights general ways for their reduction. Methods for reducing image and video artifacts are classified, their pros and cons are described. The authors introduce and compare several existing algorithms. Non-Local Means (NLM) algorithm is chosen by means of comparing complexity and quality of different algorithms and is considered to be the better algorithm for artifacts reduction. Besides, implementation details of this algorithm are given and possible ways for optimization are offered.**

I. Introduction

Due to purpose of communication, we need to obtain, process and deliver information. This information is not limited to text files or sample messages; nevertheless various visual pieces of information could be transmitted including image and video files.

However, transmission channels have limited bandwidth and storage devices hold limited capacity. Digital video is broadcast and stored in an encoded form, so it requires less information (bits) than the original. Acceptable quality for standard definition video can be obtained using a bit rate of 6 megabits per second. Artifacts are the result of a lossy data compression applied to image and video. Section II presents an artifact classification, meanwhile Section III describes an overview of existing methods for reducing image and video distortions. Section IV represents NLM algorithm and details of implementation. Section V compares some existing algorithms with NLM and Section VI introduces possible NLM improvements.

II. Artifacts classification in Images and video

Compression artifact is a particular class of data errors that are usually the consequence of quantization in lossy data compression. In this paper, types of artifacts that could be frequently observed due to image or video processing are presented. These distortions can be classified into the following types:

Blocking artifacts:

They are the most visible image and video degradation of all artifacts. This effect is caused by all block-based coding techniques. It is a well-known fact that all compression techniques divide image into small blocks and then compress them separately. Due to the coarse quantization, the correlation among blocks is lost, and horizontal and vertical borders appear.

Ringing artifacts:

Ringing artifacts are visible for all compression techniques especially when image or video is transformed into frequency domain. Ringing effect is caused by the quantization or truncation of the high frequency coefficients and can also come from improper image restoration operations. Moreover, it appears as distortion along sharp edges in the video sequence. This artifact occurs very often when DWT encoder is used. Furthermore, it could be observed after image or video has been de-coded using a frequency coder.

Blur effect:

Blurring is another artifact resulting from the absence of high frequencies in the low bit rate video. It appears around the sharp edges, and all image details become blurred. This effect is very similar to ringing artifact, and sometimes it is hard to distinguish between them. The difference between these two effects is that they appear on different sides: horizontal or vertical (Fig. 1).

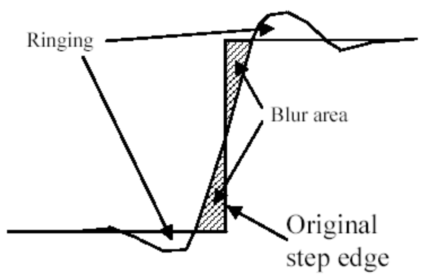

Fig. 1 Ringing and edge blurs in one dimensional signal.

Manuscript received June 8, 2009.

V.I. Fedak is PhD student of the National University: “Lviv Polytechnic”, Department of computer systems and automatic (author phone: 380973637284; e-mail: fedak77@rambler.ru, vfedak@softservecom.com).

A.Y. Nakonechny, prof. of the National University: “Lviv Polytechnic”, Department of computer systems and automatic (author; e-mail: adrnakon@polynet.lviv.ua).

Flickering:

Flickering is one of the most annoying temporal artifacts in predictive video coding. As it is widely known, modern algorithms encode video as a sequence of images. The first frame from this sequence is a key frame (I), others are additional (previous [P] and subsequent [B]) frames. All sequences are encoded by motion-compensated algorithms. When an observer watches the de-coded video, the flickering effect is noticeable due to the difference between key frames (I) and other frames (P, B).

III. Overview of existing methods for image and video artifacts reduction

All postprocessing methods and algorithms for reducing image and video artifacts could be divided into the following types:

- spatial-temporal algorithms;
- algorithms that transform signal to frequency domain;
- motion-compensated algorithms;
- iterative approaches based on the theory of projections onto convex set (POCS).

Figure 2 depicts the flow chart for postprocessing algorithms.

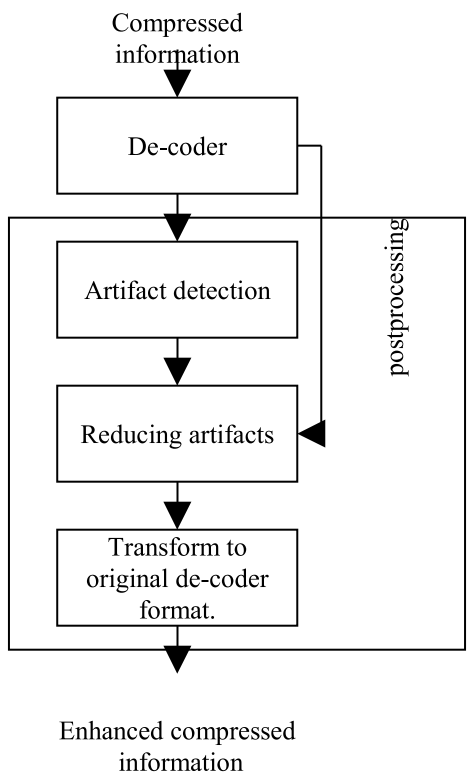

Fig. 2 General scheme for postprocessing algorithms

Many approaches have been proposed in the literature aiming at alleviation of the blocking artifacts in the images and video.

Spatial-temporal algorithms modify image pixel values. These approaches are usually used together with the edge detection algorithms to prevent the blurring effect. As a great number of algorithms have been developed nowadays, it would be rational to overview these approaches due to which completely versatile solutions can be reached.

With the purpose of improving image and video quality, Steven and Choy [6] propose the algorithm that uses local statistics of transform coefficients. The authors investigated that pixel brightness diversity among blocks is greater than within one block, and border pixels are filtered by spatial algorithm. This approach reduces the blocking effect from the image and simultaneously introduces the additional blur to the image's edges.

Ramamurthi et al. [7] who used local statistics as means of differentiation between monotone and edge blocks introduced a generic filter for the removal of blocking artifacts and staircase effect. Monotone blocks contain less spatial details than edge blocks. They propose to use two-dimensional filtering that is applied for monotone blocks and one-dimensional directional filtering for edge blocks.

Vinh and Kim [8] present a new pixel classification-based approach for block artifact reduction. Instead of classifying each block of fixed size to smooth region or edge region, they distinguish each pixel using the binary edge map from edge detection process. (Fig. 3)

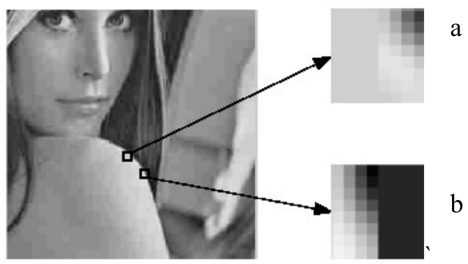

Fig.3 The block (a) is classified as the smooth region, but it contains some edge pixels. The block (b) is classified as the edge region, but it contains some smooth pixels

They reduce grid noise in the smooth region using an adaptive filter.

Frequency algorithms transform image or video (sequence of images) to frequency domain and modify DCT or DWT coefficients. These approaches are very efficient, but of high complexity, because image and video signal has to be transformed from spatial to frequency domain and vice versa. Wang [9] proposes adaptive algorithm of blocking artifacts reduction in DCT domain. He proposes efficient algorithm for blind measuring the blocking artifact in DCT domain. Wang's algorithm works as follows:

- Divide image to edge and monotone areas. Sobel edge detector [17] is used for these purposes.

- Reduce blocking artifacts in non-edge areas. Horizontal and vertical smoothing filter in spatial domain is used.
- Apply Filter Tao [10] in the edge areas.
- Transform image to the original format. Quantization Constraints.

In [11], Alan and Liew introduce wavelet-based deblocking and de-ringing algorithm for artifacts suppression. Based on a theoretical analysis of the blocking artifacts, the proposed algorithm is able to take into account the statistical characteristic of block discontinuities, as well as the behaviour of wavelet coefficients across scales for different image features to suppress both the blocking and ringing artifacts.

Motion-compensated algorithms are used to diminish different types of artifacts. Furthermore, these techniques are very often used with spatial-temporal and frequency algorithms. In this article, motion-compensated technique is combined with spatial-temporal Non-Local Means algorithm.

Another class of postprocessors using iterative image recovery methods based on the theory of projections onto convex sets (POCS) is proposed in [3-4]. POCS are effective in eliminating blocking artifacts but less practical for real-time applications, since the iterative procedure adopted increases the computation complexity.

IV. Non-Local Means details and implementation

NLM algorithm removes the noise while retaining all this meaningful image information. For this purpose, the NLM algorithm tries to take advantage of the redundancy and self similarity of the image. Most image details occur several times; every small window has many similar windows within the same image.
See Figure 4 [5].

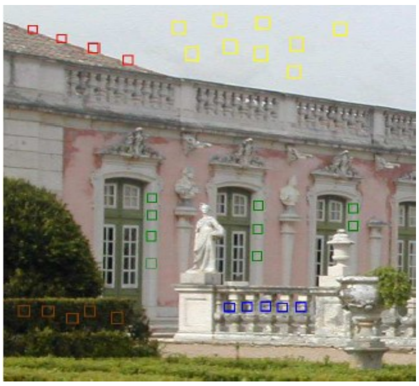

Fig. 4 The similar pieces of image

The NLM algorithm is an improvement of bilateral filtering. The bilateral and the NLM filters are two very successful image de-noising filters. Both bilateral and NLM filters are based on the assumption that image contents are likely to repeat themselves within some neighbourhood. Therefore, de-noising each pixel is achieved by averaging all pixels in its neighbourhood.

The NLM algorithm estimates the value of pixel x as an average of all pixels values. The probability that pixel y is similar to pixel x is defined by calculating the difference in luminance and position values between pixels x and y in the neighborhood filters.

The estimated value $NL(v)(i)$ from the discrete noisy image $v = \{v(i) \mid i \in I\}$ is computed as a weighted average of all the pixels in the image:

$$NL(v)(i) = \sum_{j \in I} w(i,j)v(j) \quad (1)$$

The neighborhood of pixel x is defined as the set of pixels in a sequence in which each pixel has a surrounding window similar to the window around pixel x. All pixels in this neighborhood can be used for predicting pixel x. The NLM filter is determined by the formula:

$$NL_h\{\underline{x}\} = \frac{1}{C(\underline{x})} \cdot \sum_{\underline{y} \in Q(\underline{x})} z(\underline{x}) \cdot e^{-\frac{\|\underline{N}(\underline{x}) - \underline{N}(\underline{y})\|_2^2}{h^2}} \quad (2)$$

(2.6)

where

$$C(\underline{x}) = \sum_{\underline{y} \in Q(\underline{x})} e^{-\frac{\|\underline{N}(\underline{x}) - \underline{N}(\underline{y})\|_2^2}{h^2}} \quad (3)$$

is a normalizing constant, N(x) is a vector which contains the pixels in the window surrounding pixel x. Q(x) is a search window around pixel x, in which the neighborhood of pixel x is searched. The window N(x) contains Sx·Sy pixels and the search window Q(x) contains Ax·Ay pixels.

The NLM filter is exactly implemented as described in Equation 2. All parameters used in implementation are presented in Table 1.

Table1

Parameter	**Description**
h	Parameter which determines the amount of averaging in Non-Local Means.
Sx	The match window/patch size of a pixel in the horizontal direction.
Sy	The match window/patch size of a pixel in the vertical direction.
Ax	The search window size in the horizontal direction.
Ay	The search window size in the vertical direction.

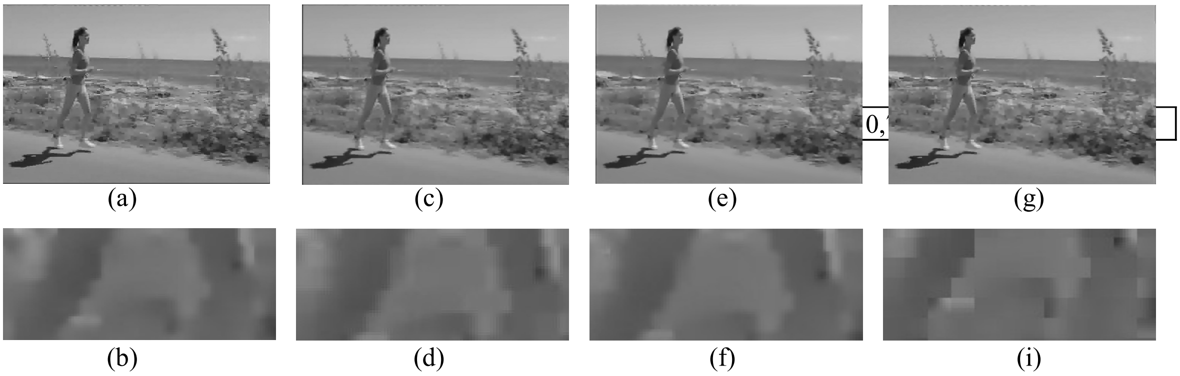

Fig. 7 Girlsea frame 25. Comparasion between different postprocessing techniques. (a) – NLM implementation; (c) – algorithm proposed by Steven and Choy [6]; (e) - wavelet-based algorithm proposed by Alan and Liew in [11]; (g) – image without applying any postprocessing tecniques. (b), (d), (f), (i) – their corresponding zoomed fragments

Every pixel is restored by the weighted average of all pixels in its (temporal) search window. The level of averaging is determined by the filtering parameter h.

Complexity

Complexity is an indication of the algorithm efficiency. The complexity of the NLM algorithm is defined as the amount of multiplication required to process a single frame in a sequence. If the size of similarity window is Sx·Sy, the amount of multiplication is equal to Sx·Sy. The search of similar windows is performed in the larger "search window" Ax·Ay. So, the overall complexity of the algorithm for one frame is $N^2 \cdot Sx \cdot Sy \cdot Ax \cdot Ay$, where N^2 is the number of the image pixels (the search is performed in the whole image, the size of image is N·N).

In case of video, this complexity must be multiplied by the number of frames At . The complexity can be used as the speed comparison of different NLM implementations.

V. Experiments

In this article, the ten video sequences are used to estimate and determine better NLM parameters for perceiving better video quality (test sequences are provided by Royal Philips Electronics). The sequences are chosen to have varying content and motion. All sequences are encoded and decoded by MPEG-2 codec. The video quality metrics (PSNR, SSIM [18], BIM [19]) for Soccer (Fig. 5) and Girlsea (Fig. 6) sequences with different compression level are presented in Table 2.

Table 2

Sequence	Metric	1Mbps	3Mbps	Original
Soccer	PSNR	32,33	36,34	-
	BIM	2,24	1,46	1,06
	SSIM	0,790	0,895	1
Girlsea	PSNR	28,86	33,95	-
	BIM	4,52	2,1	1,01

Test sequences are processed with different filtering parameter h to obtain its optimal value, and a better image/video quality and complexity is achieved by using the following parameters:

Ax=Ay=5, Sx=Sy=3 and h = 10.

Fig. 5 Soccer, Frame 2

Fig. 6 Girlsea, Frame 2

When NLM is used as a static filter, a significant reduction in BIM is obtained if the filtering parameter h is increased. When h is increased, PSNR is lowered, image details are blurred, sharpness and contrast are decreased in video sequences. Additional de-blocking can be obtained by increasing the match window length. The advantage of increased match window length is that it lowers the BIM while preserving PSNR and details. Increasing the search window size blurs out details and leads to a lower contrast and a higher complexity.

NLM is compared with spatial-temporal and frequency algorithms that are presented in Section III. In Fig.7 Girlsea video sequence is shown after applying different postprocessing techniques (compression level corresponds to 3Mbps bit rate). NLM is applied with optimal parameters that are determined in Section IV. The results received after subjective evaluation and comparison of objective metrics indicate that NLM has better performance than wavelet-based algorithm proposed by Alan and Liew in [11], and provides better quality results than algorithm proposed by Steven and Choy [6]. Having analyzed quality metrics, it can be concluded that NLM is suitable for reduction of compression artifacts.

VI. Optimization of NLM approach and further ways of research

The several improvements are proposed for NLM algorithm [12-16]. Heated disputes arise concern whether it is possible or not to use NLM with motion-estimation techniques. Nevertheless, it is obvious that NLM can be extended for de-noising purposes to the temporal domain to obtain a spatial-temporal filter. It is very hard to use the motion-estimation approach together with NLM, because we face such a well-known motion-estimation problem as the aperture problem. On the other hand, we can benefit from it. If more than one image is available for the processing, more pixels can be used in de-noising process. This form of spatial-temporal filtering might also reduce temporal flickering which is in compressed sequences. So, for receiving better video quality, we propose to use frame temporal buffer T which contains input (unprocessed in_{t-2}, in_{t-1}) and output (after NLM processing out_{t-2}, out_{t-1}) of previous frames $\{in_{t-2}, in_{t-1}, in_t, out_{t-2}, out_{t-1}\}$, where t indicates the frame number.

Hierarchical block matching algorithm is proposed to find similar windows for speeding-up NLM (Fig. 8.). First a larger block is chosen to obtain a rough estimate to find similar windows. Afterwards NLM, in which similar windows from previous iteration are used, is applied to the original image sequence.

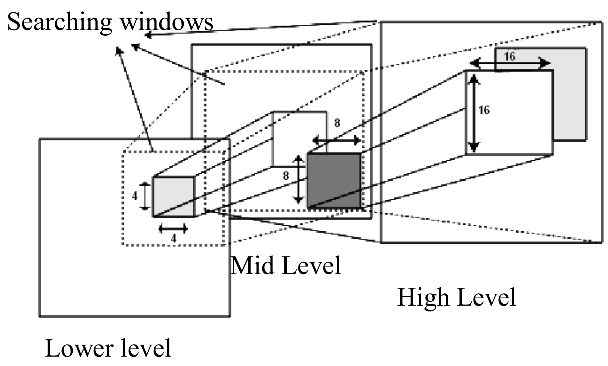

Fig. 8 Hierarchical block matching algorithm

So, this approach can greatly improve performance of NLM.

References

[1] Artifacts in Image and Video Systems: Classification and Mitigation Amal Punchihewa and Donald G. Bailey Institute of Information Sciences & Technology, Massey University g.a.punchihewa@massey.ac.nz d.g.bailey@massey.ac.nz

[2] Objective Evaluation of Edge Blur and Ringing Artifacts: Application to JPEG and JPEG 2000 Image Codecs G. A. D. Punchihewa, D. G. Bailey, and R. M. Hodgson Institute of Information Sciences and Technology, Massey University, New Zealand.

[3]Y. Yang, N. P. Galatsanos, and A. K. Katsaggelos, "Projection based spatially adaptive reconstruction of block-transform compressed images," *IEEE Trans. Image Processing*, vol. 4, no. 7, pp. 896–908, 1995.

[4] H. Paek, R.-C. Kim, and S.-U. Lee, "A DCT-based spatially adaptive postprocessing technique to reduce the blocking artifacts in transform coded images," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 10, no. 1, pp. 36–41, 2000.

[5] Antoni Buades "Image and film denoising bynon-local means". Ph. D. Thesis.

[6] S. S. O.Steven S. O. Choy, Y.-H.Yuk-Hee Chan, andW.-C.Wan-Chi Siu, "Reduction of block-transform image coding artifacts by using local statistics of transform coefficients," IEEE Signal Processing Letters, vol. 4, no. 1, pp. 5–7, Jan. 1997.

[7] B. Ramamurthi and A. Gersho, \Nonlinear space-variant postprocessing of block coded im- ages", in IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. 34, No. 5, pp. 1258{1268, Oct. 1986.

[8] Truong Quang Vinh, Young-Chul Kim Block Artifact Reduction Based on Pixel Classification Using Binary Edge Map 2008

[9] C. Wang et al.: Adaptive Reduction of Blocking Artifacts in DCT Domain for Highly Compressed Images. Adaptive Reduction of Blocking Artifacts in DCT Domain for Highly Compressed Images. 2003

[10] Postprocessing technology yields best results and achieves C. Tao, H. R.Wu, and Q. Bin, "Adaptive postfiltering of transform coefficients for the reduction of blocking artifacts," IEEE Trans. Circuits Syst. Video Technol., vol. 11, no. 5, pp. 594-602, May. 2001.

[11] Alan W.-C. Liew, Member, IEEE, and Hong Yan, Senior Member, IEEE Blocking Artifacts Suppression in Block-Coded Images Using Overcomplete Wavelet Representation IEEE Trans. Circuits Syst. Video Technol., vol. 11, no. 14, no. 4, APRIL 2004

[12] Pierrick Coupy, Pierre Yger, Sylvain Prima, Pierre Hellier, Charles Kervrann and Christian Barillot "An Optimized Blockwise Non Local Means Denoising Filter for 3D Magnetic Resonance Images "

[13] M. Protter & M. Elad, H. Takeda & P. Milanfar "Generalizing the Non-Local-Means to Super-resolution Reconstruction " April 13, 2008.
[14] N.A. Thackera_, P.A. Bromileya and J.V. Manjonb "A Quantitative Theory of the Non-Local Means Algorithm "
[15] N. A. Thacker and J.V. Manjon. "A Statistical Interpretation of Non-Local Means. " Tina Memo No. 2008-004, China, 2008.
[16] Mona Mahmoudi and Guillermo Sapiro "Fast image anf video denosing via Non-Local Means of similar neighborhoods " Minneapolis, Minnesota 55455–0436" June 2005
[17] O. R. Vincent, O. Folorunso "A Descriptive Algorithm for Sobel Image Edge Detection" Proceedings of Informing Science & IT Education Conference (InSITE) 2009.
[18] Yubing Wang "Survey of Objective Video Quality Measurements." EMC Corporation Hopkinton, MA 01748, USA.
[19] H. R. Wu and M. Yuen, "A generalized block edge impairment metric for video coding", IEEE Signal Processing Letters, vol. 4, pp. 317-320, November 1997.