Advances in Computational Sciences and Technology
ISSN 0973-6107 Volume 10, Number 1 (2017) pp. 147-158

http://www.ripublication.com

Encoder Optimization using Collaborative Filtering with Dictionary Learning to Improve PSNR

Shailendra Kumar[1], T. Usmani[2], Surya Prakash[3], S. H. Saeed[4], N. R Kidwai[5]

[1,2,4]Electronics & Communication Engineering Department, Integral University, Lucknow (UP), 226026 India.
[3]Computer Science & Engineering Department, Indian Institute of Technology, Indore (MP), 453441 India.
[5]Jahangirabad Institute of Technology, Barabanki (UP), 225203 India

Abstract

In this paper, we propose inculcation of collaborative and adaptive filtering for a H.264 /AVC deblocking filter. For this, an improved PSNR–B model including optimal adaptive filter with weight adjustment is designed for improved PSNR calculation. Block matching 3D algorithm is used for the image decomposition and form a dictionary of blocking and deblocking components for post processing. The dictionary is capable of removing the blocking artifacts effectively and avoids unwanted blurring and maintains edges found in the original image. Proposed method enhances the filtration and reduces deblocking time. The proposed technique is able to reduce about 46% deblocking time with improvement in PSNR.

Keywords: Collaborative Processing, Encoder Optimization, Data path Optimization.

1. INTRODUCTION

Day-by-day, the data on the network is increasing due the use of information technology by the people in the world. Due to this, incremented data traffic in the network has increased the significance of data compression specially in videos. Due to this, video compression has become almost a necessity in various consumer electronics goods like mobile phones, tablets and video cameras. Where data transmission and reception is required. There exist various coders for video compression in the literature. To make efficient video compression in these coders, various parameters are required to be taken care of like reduction in bit rate, amendment in the execution of applications and adjustment in compression according to latest real time video. H.264/AVC is the current coding standard which fulfils most of the requirements mentioned above. H.264/AVC standard has several features like integer transform, deblocking filter, multi-mode intra-prediction etc. It also exhibits multiframe variable block size quarter pixel motion vector accuracy as compared to previous coding standards. Amongst these features, deblocking filter plays an important role in improving the subjective quality of decoded video by decreasing blocking artifacts and discontinuities in frames which arise due to DCT, block base prediction and transform coefficient quantization. In H2.64/AVC, deblocking filter is used in both the coding and decoding path in order to take care of the loop in effects of the filter by taking the reference of the macroblocks. It has been seen that the use of deblocking filter, however increases the computational complexity of the process [1], [2].

In general, the proposed technique is the combination of two techniques; one is image enhancement [3], [4] and second is image restoration [5], [6]. With the consideration of human visual sensitivity, Image enhancement technique [3], [4] improves perceived image quality by smoothing visible artifacts in place of restoration of original pixel value. The advantage of this technique is to reduce computational complexities. Hsung et al. [3] proposed the extracted high frequency components from an image are applied to adaptive filter. The image deblocking technique proposed in [7] involves shape adaptive Discrete Cosine Transform in combination with intersection of confidence interval technique, which explain the shape of transform. Automatically detection of blocking artifacts at boundaries is proposed in [8], where four filtering modes are used to remove blocking artifacts. Dual non-local Kuan (DNLK) filter and over-complete dual non-local Kuan (OCDNLK) filter [4] is used to reduce blocking artifacts. In [6] [9], image restoration approach is generally used for image or video restoration and removing blocking artifacts. Image or video restoration is achieved at receiver side with the help of previous information and received information. This approach can be divided into three parts that is criterion based [5-9], constrained based [5] and constrained optimization based [6]. The constraint based approach creates several constraints for decoded image or video. This method is based on projection onto convex sets where prior information of original image is emphasized. Jeong et al. [5] proposed novel quantization constraint set, where reduction in deblocking artifacts is achieved by compliment the drawbacks of the smoothness constraint set. While, constrained optimization approach is based on previous information of an image or video. Jung, Jiao and Sun [6] proposed deblocking method using sparse representation. A general dictionary is created with the help of dictionary learning algorithm. In [10], Hai Wang and Naiyan Wang proposed a hierarchical Bayesian model called collaborative deep learning (CDL), which jointly performs deep representation learning for the content information and collaborative ltering for the ratings (feedback) matrix. Nikhil Rao, Hsiang-Fu [11], proposed collaborative filtering with graph information method. In this method, highly scalable alternating minimized method is used to speedup magnitude over competing approach and struggle to improve psnr. In this work, we propose a deblocking filter where encoder optimization technique is based on threshold detection. The proposed deblocking filter is faster as compared to normal existing deblocking filter.

The main aim of the proposed deblocking filter is to improve the appearance of decoded pictures. The encoder optimization in the proposed filter results in faster calculation of boundary strength. Use of dictionary based approach in the filter development improves the visual quality of the video by smoothing the sharp edges in the macroblocks. Two main considerations which are to be taken care while proposing an efficient deblocking filter mechanism to improve the PSNR and to achieve a faster deblocking filter time are as follows.

(i) Visual perception is more vigorous in the plain region as compared to the flat regions.

(ii) Lower quality of the right image in stereoscopic pairs should not degrade the visual quality.

2. PROPOSED METHOD

To improve the PSNR and to reduce or maintain the deblocking time, encoder based threshold optimization deblocking mechanism is used in a stereoscopic image deblocking. In this mechanism, adaptive deblocking filter with slice level adaptivity is equipped with learning dictionary for deblocking. This technique provide improved PSNR-B model for efficient PSNR calculation, include data path optimization for weight adjustment of adaptive filter.

The overall framework of the proposed method is shown in Figure 1. The steps listed above are explained below in detail.

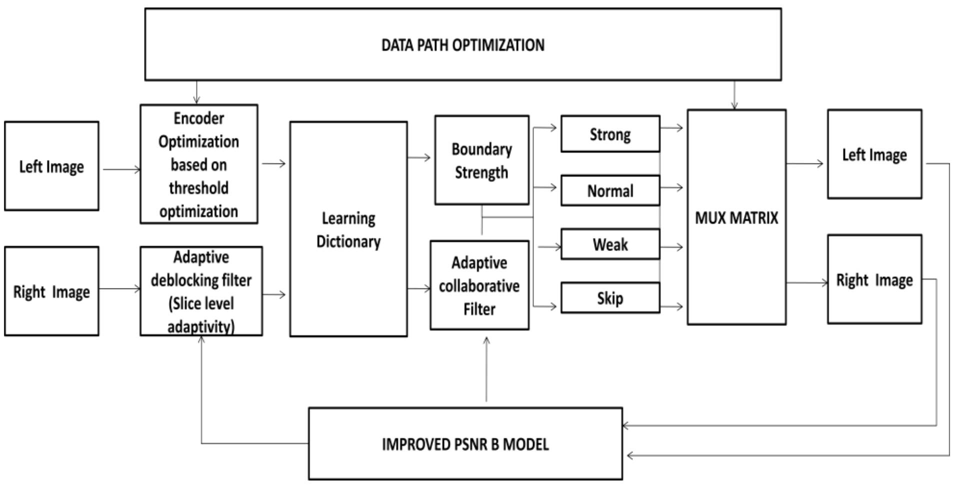

Figure 1: Framework. Proposed methodology for deblocking with added encoder optimization, learning dictionary modified PSNR B and Data path optimization

2.1 Adding an Adaptive Deblocking Filter Scheme

For improving the PSNR values it is important to preserve the sharpness and ensure that the true edges remain unaltered or unfiltered which impacts the visibility. Considering a 4x4 blocks by p_3, p_2, p_1, p_0, q_0, q_1, q_2, q_3 with actual boundary between the ranges p_0 and q_0. Considering quantization parameters thresholds as α and β which determines if the particular pixel is filtered or not. Filtering of line takes place if the following condition is satisfied [2].

$$|p_0 - q_0| < \alpha(\text{Index}_A) \quad (1)$$

$$|p_0 - q_0| < \beta(\text{Index}_B) \quad (2)$$

During these conditions α and β are dependent on two factors namely the edge function and the offset values selected by the encoder for slice level. The table index can be defined as the

$$\text{Index}_A = \text{Min}(\max(0, QP + \text{Offset}_A), 51) \quad (3)$$

$$\text{Index}_B = \text{Min}(\max(0, QP + \text{Offset}_B), 51) \quad (4)$$

Where 0-51 represents the range of valid QP (Quantization Parameters) values. The values of the α and β are defined approximately according to the following relationships

$$\alpha(x) = 0.8\left(2^{\frac{x}{6}} - 1\right) \quad (5)$$

$$\beta(x) = 0.5x - 7 \quad (6)$$

2.2 Implementing the Slice Level Adaptivity

On the slice level adaptivity two offsets denoted as Offset_A and Offset_B are used for the adjustment of α and β to vary the amount of filtration.

2.3 Collaborative Filter Approach

A collaborative filter approach for a d dimensional signal groups are jointly filtered with following procedure to give a faster filtering time [2].

a) Linear transform is applied to a group and transform coefficients are minimized.

b) The collaborative filtering maps spatiotemporal volumes a separable linear transform Y_{AD} on each group of signals $G_z(x_0, t_0)$ and provides an estimate for each grouped volume.

$$G_z(x_0, t_0) = Y^{-1}(Y_{AD}\,(\Upsilon\left(Y_{3D}\left(G_z(x_0, t_0)\right)\right) \qquad (7)$$

Where Υ denotes a generic shrinkage operator. The filtered 3D group $G_y = (x_0, t_0)$ is composed of volumes $V_y(x, t)$

$$G_y(x_0, t_0) = \left\{V_y(x_i, t_i) : (x_i, t_i) \epsilon \operatorname{Ind}(x_0, t_0)\right\} \qquad (8)$$

With each V_y is an estimate of one of the corresponding unknown volume V_y in the original video y, where G_y represents stereoscopic video with multiple estimates at same coordinates (x, t). An aggregation has to be done on the signals for the computed estimate y for sequence which is given as

$$y = \frac{\sum_{(x_0,t_0)\epsilon X\,xT}\left(\sum_{(x_i,t_i)\epsilon \operatorname{Ind}(x_0,t_0)} w_{(x_0,t_0)} V_y(x_i, t_i)\right)}{\sum_{(x_0,t_0)\epsilon X\,xT}\left(\sum_{(x_i,t_i)\epsilon \operatorname{Ind}(x_0,t_0)} w_{(x_0,t_0)} X(x_i, t_i)\right)} \qquad (9)$$

2.4 Encoder Threshold Optimization Technique

Encoder optimization[3] can be done by identifying the visually important image regions in the currently decoded and blocky frame, including natural and artificial edge areas, and then optimizes the filters inherent thresholding decisions by optimization and analysis of possible threshold a threshold triples is assumed at the encoder ($\alpha_i, \beta_i, c_{oi}$) is tested. For this task, frame-wise objective function is defined as

$$F(\alpha_i \beta_i co_i) = Q_a(\alpha_i) + Q_\beta(\beta_i) + Q_{co}(c_{oi}) \qquad (10)$$

On which basis the optimization is performed.

$$Q_\theta(\theta_i) = S(X', O, M) \text{ with } X' : D(X, \theta_i) \qquad (11)$$

Edge–MSE is the mean square error relative to the total number of pixels that are positively marked, and

$$\text{Edge} - \text{SSE} = \sum (X'(m, n) - O(m, n))^2 \qquad (12)$$

Edge –SSE is the corresponding sum of squared errors between the X' and O with objective of Maximizing the edge –PSNR, but has the big advantage of never reaching infinite values. Therefore for analysis, employ the Edge –SSE to simplify and stabilize the optimization process.

2.5 Improved Method to Estimate PSNR

An improved method to estimate the PSNR can increase the PSNR values during the in loop deblocking mechanism of the filter. A new quality metric PSNR B type can be performed by considering a set of diagonal neighboring pixel pairs which are not falling on the horizontal and vertical neighboring pixel [4].

Defining the mean boundary pixel squared difference (D_B) and the mean non boundary pixel squared difference (D_{BC}) for image y to be

$$D_B(Y) = \frac{\sum(y_i, y_j) \in H_B(y_i - y_j)^2 + \sum(y_i, y_j) \in v_B(y_i - y_j)^2}{N_{H_B} + N_{V_B}} \quad (13)$$

$$D_B^C(Y) = \frac{\sum(y_i, y_j) \in R_B^c(y_i - y_j)^2 + \sum(y_i, y_j) \in L_B^C(y_i - y_j)^2}{N_{R_B^C} + N_{L_B^C}} \quad (14)$$

Where $N_{R_B^C}, N_{L_B^C}, R_B^c, L_B^C$ be the number of pixel pairs in the right and left blocking regions, blocking regions on right and blocking regions in the left diagonal. y_i, y_j are pixel values at the coordinates (i, j).

The above equation is applicable if only diagonal neighboring pixel pairs are considered

$$D_B^C(Y) = \frac{\begin{matrix}\{\sum(y_i, y_j) \in H_B(y_i - y_j)^2 + \sum(y_i, y_j) \in v_B(y_i - y_j)^2 + \\ \sum(y_i, y_j) \in R_B^c(y_i - y_j)^2 + \sum(y_i, y_j) \in L_B^C(y_i - y_j)^2\}\end{matrix}}{2 * (N_{H_B^C} + N_{V_B^C})} \quad (15)$$

$N_{H_B^C}$, $N_{V_B^C}$ are be the number of pixel pairs in the horizontal and vertical direction. The blocking effect factors which is dependent on the quantization steps and mean boundary pixels squared difference is given by

$$BEF(Y) = \eta[D_B(Y) - D_B^C(Y)] \quad (16)$$

Then the blocking effect factor (BEF) for macro block sizes is given for k^{th} block for an image is given by

$$BEF_k(Y) = \eta_k\left[D_{B_k}(Y) - D_{B_K}^{C}(Y)\right] \qquad (17)$$

For overall block sizes BEF is given by

$$BEF_{Tot}(Y) = \sum_{k=1}^{K} BEF_k(y) \qquad (18)$$

Finally the proposed PSNR-B is given as

$$PSNR_{B(x,y)} = 10\, log_{10} \frac{255^2}{MSE\_B(x,y)} \qquad (19)$$

The MSE measures the distortion between the reference and the test image, while the BEF specifically measures the amount of blocking artifacts using the test image.

2.6 Minimization of Time by Data Path Optimization

A local memory can be used for the storage of the temporary sampled data and then the data can be written back into the memory. To achieve a faster deblocking rate the number of attempts to access the bus can be reduced. During the sampled data filtration stage of data can be done in the raster scan model and rows of data can be stored in a row transposed form [5],[6]. A mechanism to increase the deblocking filter time is shown in the figure below. Four blocks namely R_{Read} (for storing a block of data read from main memory), R_{Trans} (performs transposing operation from main memory and acts as right block, R_{Write} (a block that stores all the data by writing into main memory). UP MEM acts as a block to store the data into a stack.

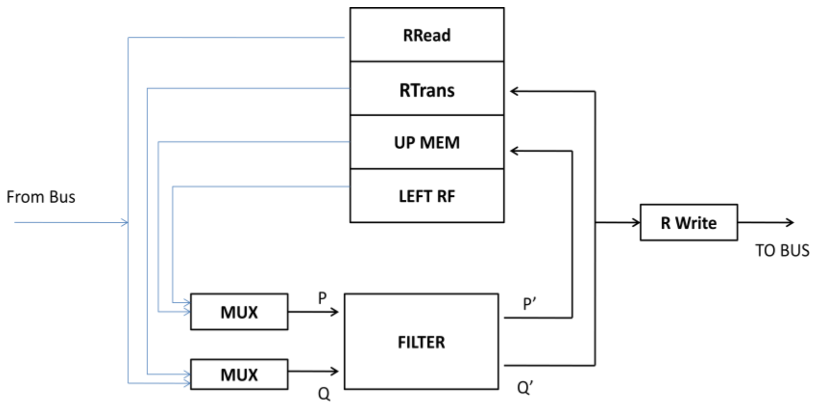

Fig 2: Minimizatin of data path schematic for fast a deblocking filter

2.7 Dictionary Based Deblocking for High Quality Fast Deblocking

For faster deblocking artifact reduction method in overcomplete 3D dictionary from natural depth images using the k singular value decomposition algorithm as shown in figure (3) [7]. A deblocking can then be performed using a 3D dictionary after estimating an error threshold of objective function by the third order polynomial fitting.

(i) Overall the steps involved are as follows applying the BM3D filtering and taking the LF and HF parts. Extracting a set of image patches from HF.

(ii) Applying the online dictionary learning algorithm to solve the operation minimization function.

(iii) Extract HOG features descriptors for each dictionary portions. Applying the K means algorithm to classify all the input samples into two clusters based on their modified HOG feature descriptor.

(iv) Identify one of the two clusters as non-blocking and blocking sub dictionary.

Apply image decomposition and reconstructing each patch to recover either non-blocking component or blocking component. Obtaining the blocking artifacts removed. Refining the artifacts removed version for each blocks. Return the blocking artifacts removed version.

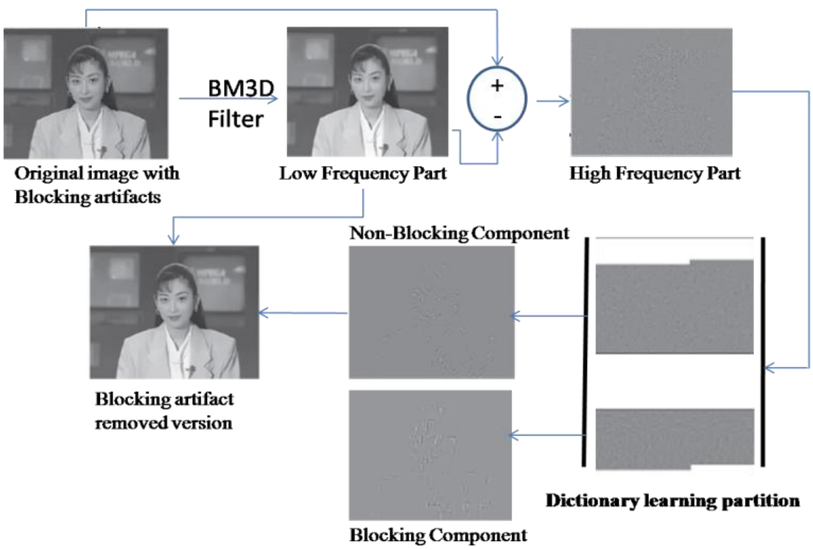

Fig 3: Modified overcomplete dictionary for deblocking filter.

4. RESULTS AND DISCUSSION

In this section, the performance of deblocking filter is analyzed using two parameters: first is the peak signal to noise ratio (PSNR) which is used to evaluate subjective video quality (as shown in (Table 1)). Second parameter is the deblocking time. The parameter speeds-up the filtering time (as shown in (Table 2)). Test sequence (shown in Figure 4) has been used with 200 frames in QVGA format. Although the improvement in PSNR is small however, significant difference in deblocking time is observed between the normal traditional deblocking filter and the proposed filter.

The proposed deblocking filter avoids distortion in frames as shown in Figure 4 (e). It changes the PSNR and speeds-up the filtering time significantly. The proposed technique comparatively performs better specially for lower bit rate hand held devices.

(a)

(b)

(c)

(d)

(e)

Fig 4: (a) H.264 /AVC decoded frame with deblocking filter (b) H.264 /AVC decoded frame with encoder optimization (c) H.264 /AVC decoded frame with collaborative processing (d) H.264 /AVC decoded frame with overcomplete dictionary (e) H.264 /AVC decoded frame with adaptive filter, collaborative signal processing and overcomplete dictionary

Table I: PSNR -B values for the normal deblocking filter and proposed filter.

Quantization Parameter	PSNR -B (dB)		Difference in PSNR (dB)
	Deblocking filter	Proposed Filter	
20	49.25	51.22	1.97
25	48.29	50.14	1.85
30	47.35	49.34	1.99
35	46.34	48.04	1.70
40	45.29	46.90	1.61

Table II: Deblocking time for the normal deblocking filter and proposed filter.

Quantization Parameter	Deblocking time (ms)		% Difference in time
	Deblocking filter	Proposed Filter	
20	1145	616	-46.20
25	1042	619	-40.59
30	1039	624	-39.94
35	1034	617	-40.33
40	1028	603	-41.34

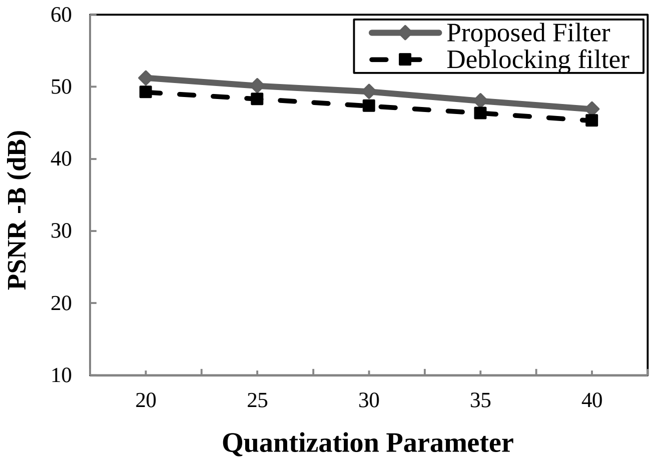

Fig 5(a): Performance results of table 1, Comparison of PSNR values with respect to quantization parameter.

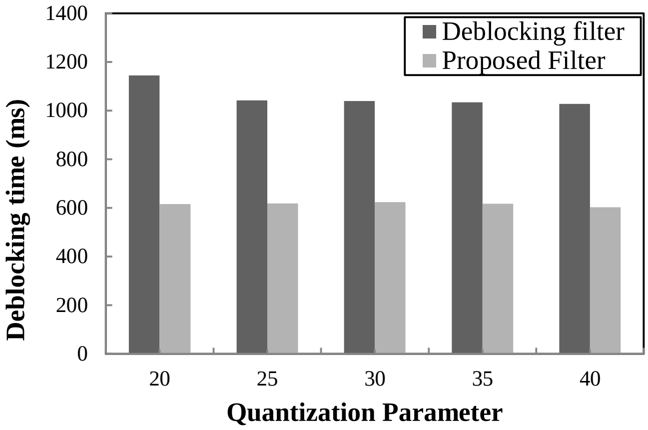

Fig 5 (b): Performance results of table 2 comparison of deblocking time with respect to quantization parameter.

5. CONCLUSIONS

In this paper, we have proposed inculcation of collaborative and adaptive filtering for a H.264 /AVC deblocking filter. The proposed technique has used a block matching 3D algorithm for the image decomposition and has formed a dictionary of blocking and deblocking components for post processing. The obtained dictionary is found to be capable of removing the blocking artifacts effectively and avoids unwanted blurring and maintains edges found in the original image. Experimental results show the improvement in PSNR values for different quantization parameters. Results also demonstrate that difference between traditional deblocking filter and proposed deblocking filter is reduced as quantization parameter increases. The proposed filter has achieved a noticeable reduction of about 46% in deblocking time. This reduction in deblocking time saves a way for hand-held devices like mobile phones and tablets to become faster.

REFERENCES

[1] References 1. List, P., Joch, A., Lainema, J., Bjontegaard, G., Karczewicz, M. Adaptive deblocking filter. IEEE Transactions on Circuits and Systems for Video Technology 13(7), 614–619 (2003).

[2] Mastani, S.A., Shilpa, K.: New approach of estimating PSNR-B for de-blocked images. CoRR abs/1306.5293 (2013).

[3] T.C. Hsung, W.S. D.P. Lun: A deblocking technique for block-transform compressed images using wavelet transforms modulus maxima. Journal of Visual Communication and Image Representation 25(5), 1488–1498 (1998).

[4] R. Zhang, W.-K.C. W. Ouyang: Image post processing by non-local kuan's lter. J. Visual Commun. Image Represent 22(3), 251–262 (2011).

[5] Y. Jeong, H.K. I. Kim: A practical projection-based post processing of block-coded images with fast convergence rate. IEEE Trans. Circuits Syst. Video Technol. 10(04), 617–623 (2000).

[6] C. Jung, H.Q.T.S. L. Jiao: Image deblocking via sparse representation, signal process. Image commun. Signal Process. Image Commun. 27(06), 663–677 (2012).

[7] A. Foi, K.E. V. Katkovnik: Pointwise shape-adaptive dct for high- quality denoising and deblocking of grayscale and color images. IEEE Trans. Image Process 6(5), 1395–1411 (2007).

[8] C.-H. Yeh, S.-J.F.J.M.-J.C.J.-A.J. T.-F. Ku: Post-processing deblocking lter algorithm for various video decoders. IET Image Process 6(5), 534–547 (2012).

[9] S.W. Wu, A.G.: Improved decoder for transform coding with application to the jpeg baseline system. IEEE Trans. Commun. 40(02), 251–254 (1992).

[10] Hao Wang, N.W., Yeung, D.-Y.: Collaborative deep learning for recommender systems. In: Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 235–244 (2015).

[11] Nikhil Rao, Y.P. Hsiang-Fu, Inderjit, R., Dhillon, S.: Collaborative filtering with graph information: Consistency and scalable methods. In: Advances in Neural Information Processing Systems NIPS-2015 (2015).

[12] Seung Ho Shin, Y.C., Kim, T.: Fast deblocking filter for stereoscopic video coding in mobile broadcasting. IEEE Transactions on Consumer Electronics 57(2), 811–816 (2011).

[13] Maggioni, M., Boracchi, G., Foi, A., Egiazarian, K.: Video denoising, deblocking, and enhancement through separable 4-d nonlocal spatiotemporal transforms. IEEE Transactions on Image Processing, 21(9), 3952–3966 (2012).

[14] Hanke, K., Hosten, P., Jager, F.: Content-adaptive encoder optimization of the h. 264/avc deblocking filter for visual quality improvement. In: Proceedings of IEEE Conference on Visual Communications and Image Processing (VCIP 2011), pp. 1–4 (2011).

[15] Kun, Y., Chun, Z., Zhihua, W.: Design of adaptive deblocking filter for h. 264/avc decoder soc. In: Proceedings of IEEE Asia Pacific Conference on Circuits and Systems, (APCCAS 2006), pp. 109–112 (2006).

[16] Li, Y., Han, N., Chen, C.: A novel deblocking filter algorithm in h. 264 for real time implementation. In: Proceedings of IEEE Third International Conference on Multimedia and Ubiquitous Engineering,(MUE 2009), pp. 26–30 (2009).

[17] R. Roy, and T. Kailath, "ESPRIT _ estimation of signal parameters via rotational invariance techniques," IEEE Trans. Acoust. Speech Signal Process, Vol. 37, no. 7, pp. 984_95, Jul. 1989.