Motion Adaptive Deinterlacing with Accurate Motion Detection and Anti-Aliasing Interpolation Filter

Lejun Yu, Jintao Li, *Member*, IEEE, Yongdong Zhang and Yanfei Shen

Abstract — *Both the motion-detection and intra-field interpolation filter are important factors affect the efficiency of motion adaptive de-interlacing. New accurate motion detection (AMD) algorithm is proposed to improve the accuracy of motion detection, which reduces the possibility of error motion detection with a median filter. To improve the efficiency of intra-field interpolation deinterlacing in moving regions, an anti-aliasing interpolation filter (AAIF) is proposed, which is better than the typical windowed sinc function. The simulation results show that peak signal noise ratio(PSNR) of our proposed deinterlacing method is 0.5-7.5 dB higher than that of previous studies and attains the best quality of subjective view*[1].

Index Terms — deinterlacing, anti-aliasing interpolation filter, motion detection

I. INTRODUCTION

There are two kinds of video formats, i.e. progressive video and interlaced video. Most broadcasted television systems employ interlaced video to make a tradeoff between bandwidth and video quality. But a major drawback of the interlaced video on the current bright high-resolution displays leads to the line flicker and jagged effect of moving. Thus, various methods of deinterlacing, which convert field pictures in interlaced video into frame pictures in progressive video, have been presented to reduce those artifacts [1].

For the most part, existing methods can be categorized as intra-filed and inter-field methods. Various spatial interpolation filters are used for intra-field deinterlacing methods. The edge-based line average (ELA) [1] method is widely used since it involves simple calculation. However, the shortcoming of ELA is that it may use wrong edge information and is sensitive to small pixel values. The method of two lines linear interpolation blurs the edge in picture [1-3].

Inter-field deinterlacing methods include motion compensated (MC) algorithms [4-7] and motion adaptive methods. MC deinterlacing methods produce best reconstructed progressive in theory, with the cost of great computational complexity and complex buffer structure. So motion adaptive deinterlacing methods [3, 8-11] are proposed to make a tradeoff between the quality and complexity.

[1] This work was supported in part by the Beijing Municipal Science & Technology Project, under Grant No. Z0004024040231.

The authors are with Institute of Computing Technology, Chinese Academy of Sciences. Beijing, China

Lejun Yu is also with Graduate University of Chinese Academy of Sciences (e-mail: jlyu@ict.ac.cn).

Contributed Paper
Manuscript received April 15, 2006

In motion adaptive methods, both the motion detection and the intra-field interpolation are important. First, the moving and static regions are extracted with the motion detection algorithms, then static areas are deinterlaced by temporal interpolation and moving regions are deinterlaced by spatial interpolation. So the accuracy of motion detection is most important factor for motion adaptive deinterlacing. Error There are two kinds of motion detection error: mistaking the moving regions as static regions, which is called error A in this paper, mistaking the static regions as moving regions, which is called error B here. Some new motion detection methods [9, 10] improve the accuracy, but the computational complexity increases greatly. Further more, since the field pictures are from frame pictures subsampled line-by-line, the frequency in vertical is aliased at busy regions with many details [3]. To get the better reconstructed frame pictures, the interpolation filters should be anti-aliasing. The 6-tap filter driven from windowed sinc function in paper [8] is not the optimal for intra-field interpolation.

Considering the problems mentioned above, accurate motion detection (AMD) algorithm and anti-aliasing interpolation filter (AAIF) are proposed to improve the performance of motion adaptive deinterlacing. With a median filter, the new motion activity calculation method reduces the possibility of error A and error B, which means that the new motion detection algorithm is more accurate than traditional methods. To improve the performance of intra-field interpolation, base on the Wiener Filter theory, an anti-aliasing interpolation is proposed, which minimize the mean square error (MSE) between original frame pictures and reconstructed ones.

The rest of this paper is organized as follows. Concepts and problems of motion adaptive deinterlacing algorithms are introduced in section II. In section III, the proposed motion adaptive deinterlacing with accurate motion detection and anti-aliasing interpolation filter are detailed. The effectiveness of the proposed algorithms is explained experimentally in section IV. Finally, our conclusions are given in Section V.

II. CONVENTIONAL MOTION DETECTION AND FREQUENCY ALIASING OF FIELD PICUTRES

The framework of deinterlacing used in this paper is illustrated in Fig. 1, where X fields are buffered for motion detection. Motion detection classes the current pixel as moving or static region, and chooses temporal or intra-filed interpolation deinterlacing with the motion detection result.

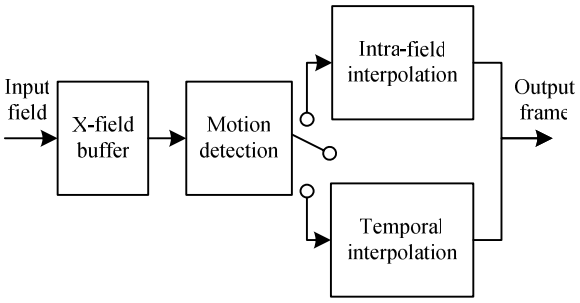

Fig. 1. Block diagram of motion adaptive deinterlacing

A. Conventional motion detection

In general, the motion information is obtained by comparing pixel values in two or more reference fields. If the differences exceed a threshold value, it is believed that the pixels are moving. Usually, the differences between two fields with the same parity are calculated, as shown in Fig. 2. Suppose the current field is n, and let X_n to be the pixel to be deinterlaced, whose neighboring pixels above and below are T_n and B_n respectively. The corresponding pixels in reference fields are denoted alike, as shown in Fig. 2.

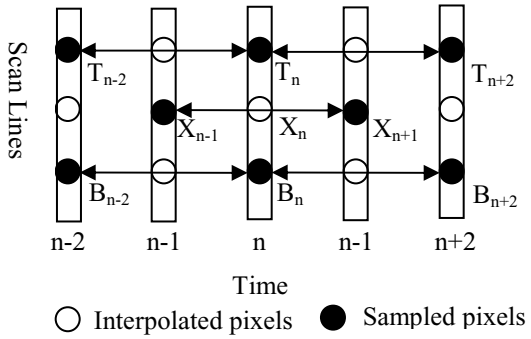

Fig. 2. Motion detection.

For simplicity, the motion activity P, Q and R are calculated respectively as

$$P = \left| \frac{T_n + B_n}{2} - \frac{T_{n-2} + B_{n-2}}{2} \right|, \tag{1}$$

$$Q = \left| \frac{T_n + B_n}{2} - \frac{T_{n+2} + B_{n+2}}{2} \right|, \tag{2}$$

$$R = |X_{n-1} - X_{n+1}|. \tag{3}$$

In the conventional 3-field motion detection method, the motion activity M of current pixel X_n is set to R, i.e.

$$M = R. \tag{4}$$

In (4), P and Q are not considered thought there are correlations between them, so the possibility of error A is higher than real case. For example, if P and Q are large, then T_n and B_n are in a moving region, X_n is in the same moving region with great possibility, but 3-field motion may mistake it is in static a region with the motion activity defined in (4).

As for conventional 5-filed motion detection, the current motion activity M is defines as [3]

$$M = \max(P, Q, R), \tag{5}$$

where max() is a function to get the max value. If the noise on T_n is large, not only motion detection result of current pixel in field n is error, but also those of the corresponding pixels in filed n-1, field n-2, field n+1 and field n+2 are error too . So noise will make the possibility of error B high in this case.

B. Aliasing of field pictures

Given the original frame picture $F(x, y)$, the Fourier transform of which is $\Psi_F(\theta_x, \theta_y)$ [8, 11], i.e.

$$F(x, y) \Leftrightarrow \Psi_F(\theta_x, \theta_y). \tag{6}$$

The top field $f_e(x, y)$ or bottom filed $f_o(x, y)$ of $F(x, y)$ is composed of the even scanlines or odd scanlines respectively, and the Fourier transforms of them are

$$f_e(x, y) \Leftrightarrow \frac{1}{2}\left[\Psi(\theta_x, \frac{\theta_y}{2}) + \Psi(\theta_x, \frac{\theta_y}{2} - \pi) \right], \tag{7}$$

$$f_o(x, y) \Leftrightarrow \frac{1}{2} e^{j\frac{\theta_y}{2}} \left[\Psi(\theta_x, \frac{\theta_y}{2}) - \Psi(\theta_x, \frac{\theta_y}{2} - \pi) \right], \tag{8}$$

respectively. In the Fourier transforms of fields, there are a baseband signal $\Psi(\theta_x, \frac{\theta_y}{2})$ and an alias signal $\frac{1}{2}\Psi(\theta_x, \frac{\theta_y}{2} - \pi)$. The baseband signal is equal for both fields, except for a phase factor $e^{j\frac{\theta_y}{2}}$. The alias terms in both fields have the same magnitude but opposite signs. The aliasing is illustrated in Fig. 3, where only frequency in vertical direction is shown and the value of frequency is normalized by the folding frequency (half of the sampling frequency) for simplicity. In the following, all the frequency is normalized value. Considering the Kell factor K [3], the resolution can not reach the folding frequency, as shown in Fig. 3. When K s is set to 0.7, the baseband signal $\Psi(\theta_x, \frac{\theta_y}{2})$ and alias signal $\frac{1}{2}\Psi(\theta_x, \frac{\theta_y}{2} - \pi)$ overlap between [0.3 0.7]. To improve the quality of intra-filed interpolation, the interpolation filter should have the following qualities: first, the frequency-response between [0, 0.3] should equal to 1, without fast decaying; second, the frequency-response between [0.7, 1] should equals to 0, third, the frequency-response between [0.3, 0.7] should be anti-aliasing.

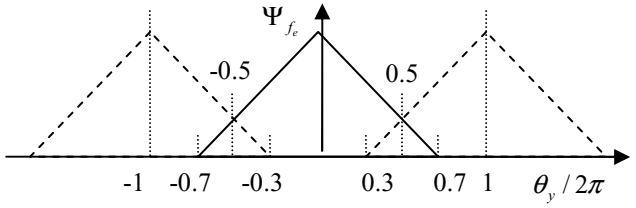

Fig. 3. Aliasing caused by interpolation in vertical direction

III. MOTION ADAPTIVE DEINTERLACING

A. Accurate motion detection (AMD)

To improve the accuracy of motion detection, an accurate motion detection (AMD) algorithm is proposed. The motion activity is defined as

$$M=\begin{cases} R & \text{if } P<T \text{ and } Q<T \\ \text{median}(P,Q,R) & \text{else} \end{cases}, \quad (9)$$

where T is a predefined motion detection threshold and median() is 3-tap median filter. The motion detection result is achieved by compare motion activity with the threshold T, i.e.

$$m=\begin{cases} 0 & \text{if } M<T \\ 1 & \text{else} \end{cases}. \quad (10)$$

If m=0, current pixel is in a static region, otherwise it is in a moving region. For the following analysis, the motion detection of P, Q and R are defined as

$$m_P=\begin{cases} 0 & \text{if } P<T \\ 1 & \text{else} \end{cases} \quad (11)$$

$$m_Q=\begin{cases} 0 & \text{if } Q<T \\ 1 & \text{else} \end{cases} \quad (12)$$

$$m_R=\begin{cases} 0 & \text{if } R<T \\ 1 & \text{else} \end{cases} \quad (13)$$

respectively. For deeply understand (9), all the cases of are analyzed as follows:

First, if the motion $m_P=0$ and $m_Q=0$, then $M=R$, that is to say $m=m_R$. Compared with 3-field motion detection, under the condition that P and Q are less than T, the possibility of error A is reduced.

Second, if $m_P=1$ and $m_Q=1$, the result of median filter in (9) is definitely larger than T, then m=1, which means that the current pixel is a moving region. In this case, the condition $m_P=1$ and $m_Q=1$ indicates both T_n and B_n are moving region, so X_n is in the same moving object when the connectivity of moving object is considered. Thus the possibility of error A is also reduced, compared with 3-field motion detection.

Third, if only one of m_P and m_Q equals to 1, the result of median filter in (9) depends on the motion activity R. Though M is not equal to R, they are both greater or less than threshold T simultaneously, that is to say that, if $R>T$, then $M>T$; otherwise, $R\le T$ and $M\le T$, so $m=m_R$. Compared with 5-filed motion detection, possibility of error B is reduced when noise make m_P or m_Q large.

With the analysis above, we can translate (9) as: if P and Q are in moving regions, then the current pixel is also in moving regions, otherwise, the motion detection result of M is same as result of R, that is to say $m=m_R$. With this logical relation, equation (9) can be translated into

$$m=(m_p \ \& \ m_q) \mid m_R, \quad (20)$$

where "&" and "|" are logical "and" and logical "or" respectively. The pipeline of motion detection of proposed motion detection is illustrated in Fig. 4, where "1L" means one line delay, "1F" means one field delay, "AVE" is average of two inputs, "DIF" is absolute difference of two inputs, "CMP" is comparison of two inputs, "T" is predefined threshold, "AND" is logical "and" and "OR" is logical "or".

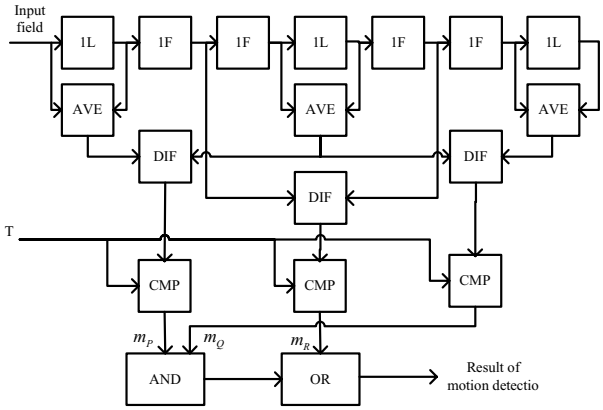

Fig 4. Pipeline structure for AMD

B. Anti-Aliasing Interpolation filter

Considering the requirements of intra-field interpolation in section II, an interpolation filter based the Wiener filter theory is designed to minimize the MSE between reconstructed frame and training sequences. The block diagram of the process of driving the filter is illustrated in Fig. 5.

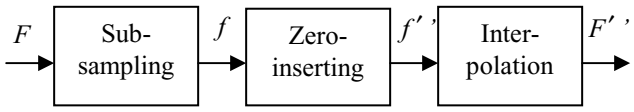

Fig 5. Block diagram for interpolation filter used in deinterlacing

The input original frame picture F is subsampled into field picture f, which is a top field f_e or bottom field f_o. Zero-inserting is performed for top and bottom filed respectively, that is to say,

$$f'_e(x,y)=\begin{cases} F(x,y) & y=0,2,4\cdots \\ 0 & y=1,3,5\cdots \end{cases} \quad (11)$$

$$f'_o(x,y)=\begin{cases} 0 & y=0,2,4\cdots \\ F(x,y) & y=1,3,5\cdots \end{cases} \quad (12)$$

Under the condition that the interpolation with 6-tap real coefficients, the interpolation filter can be defined as

$$\begin{cases} Filter=[a,0,b,0,c,1,c,0,b,0,a] \\ 2\times(a+b+c)=1 \end{cases} \quad (13)$$

Let the K frames used to train the filter to be $F=\{F_1, F_2, \ldots, F_K\}$, and the reconstructed ones to be $F'=\{F'_1, F'_2, \ldots, F'_K\}$, which can be driven by

$$F'_k(x,y)=\sum_{i=-5}^{5} Filter(i+5)\times f'_k(x,y+i) \quad (14)$$

So the MSE J between F and F' can be expressed as:

$$J = \frac{1}{KMN}\sum_{k=0}^{K-1}\sum_{y=0}^{M-1}\sum_{x=0}^{N-1}[F_k'(x,y) - F_k(x,y)]^2, \qquad (15)$$

where N and M is the width and height of image respectively. To make the filter easy to realized with 16-bit operators, it is assumed that

$$\begin{cases} a = \dfrac{\alpha}{128} \\ b = \dfrac{64 - \alpha - \beta}{128} \\ c = \dfrac{\beta}{128} \end{cases} \qquad (16)$$

where α and β are integers. With (16), J in (15) is minimized by the filter named Filter1, whose coefficients are shown in TABLE I (For simplicity the 0 and 1 coefficients are skipped.) The frequency-response curve of the filter is illustrated in Figure 4. For comparison, the filter of windowed sinc function in paper [8] and the simple linear filter, which are named Filter2 and Filter3 respectively, are shown in Table 1 and Figure 4 also. It can be seen that Filter3 decays in pass band [0, 0.3] rapidly, and frequency-response value is high in stop band [0.7, 1]. Compared with Filter2 and Filter3, the proposed interpolation filter Filter1 is the best one in the whole frequency band [0, 1].

TABLE I
INTRA-FIELD INTERPOLATIONS FILTERS.

Filter	Coefficients
Filter1	[3, -15, 76, 76, -15, 3]/128
Filter2	[3, -21, 146, 146, -21, 3]/256
Filter3	[1, 1]/2

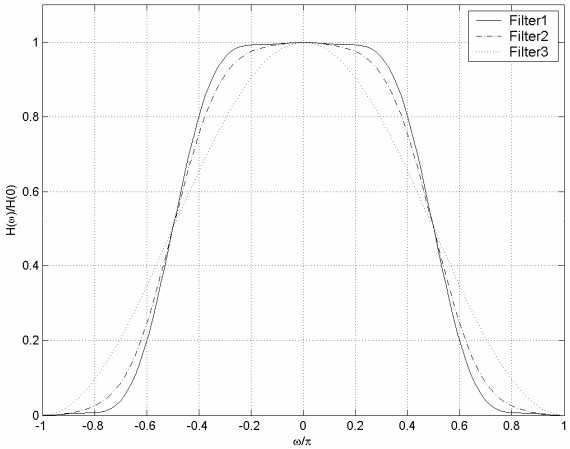

Fig 6. Frequency response curves of interpolation filters

The proposed intra-field interpolation can be realized by Fig. 7, where multiplications ("×3 ", "×15"and "×76") can be replaced with shift and sum operations for hardware realization. And right shift ">>" is used to replace division. "SUM" are "SUM3" are sum of 2 or 3 input data. Data in interpolation precess can be calculated with 16 bits operators, which make the algorithm is easy to implement.

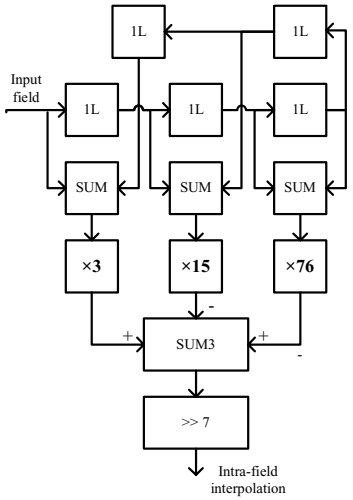

Fig. 7. Block diagram for AAIF

C. Motion adaptive deinterlacing

With the interpolation and motion detection algorithms proposed above, the block diagram of Fig. 1 can be realized as Fig. 8, where *A-F* are pixels referenced in intra-field interpolation. If current pixel is in moving region, intra-field interpolation result is outputted, otherwise temporal interpolation is outputted, that is,

$$X_n = \begin{cases} \dfrac{76(A+F) - 15(B+E) + 3(C+D)}{128} & if\ m = 1 \\ \dfrac{X_{n-1} + X_{n+1}}{2} & else \end{cases}. \qquad (17)$$

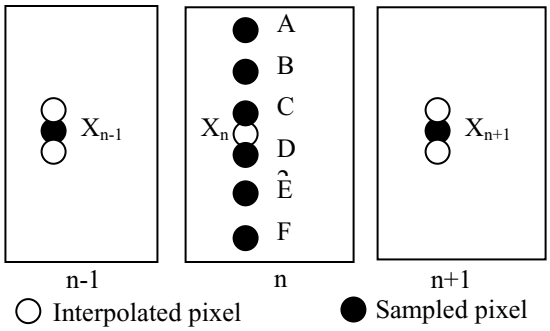

Fig. 8. Motion adaptive deinterlacing proposed

IV. SIMULATION RESULTS

A. Object performance

The intra-field deinterlacing algorithms, such as ELA, Bob and Merge [1] are tested for compared with the proposed method. And the motion adaptive deinterlacing with 3-, 4-, and 5-field motion detection methods are compared with the new motion detection algorithm, where Filter3 is used for intra-filed interpolation. Peak signal noise ratio (*PSNR*) is used as the measure of objective performance, which is defined as

$$PSNR = 10\log_{10} 255^2 / \left(\frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}(X_{m,n} - X'_{m,n})^2 \right) \text{(dB)} \qquad (18)$$

where M and N are the width and height of frame picture respectively, and $X_{m,n}$ and $X'_{m,n}$ are the original and deinterlaced pixel at position(m, n) respectively. The greater *PSNR* is, the better object performance is. The *PSNR* values of classic deinterlacing methods are shown in TABLE II, and the *PSNR* values of proposed deinterlacing are given in TABLE III.

To show the performance of different motion detection algorithms, (AMD+Filter3) is tested to compare with the traditional 3-, 4- and 5-filed motion detection, which are denoted as 3FMD, 4FMD and 5FMD in TALBE II respectively. The proposed motion detection algorithm improves *PSNR* about 0.20–04.0 dB 。 To show the performance of intra-field interpolation, the AMD motion detection is combined with Filter3, Filter2 and Filter1 respectively, the *PSNR* are given in TABLE III. On average, Filter1 improved *PSNR* by 0.48dB and 0.18dB respectively, compared with Filter3 and Filter2. In sum, proposed deinterlacing method (AMD+Filter1) improved PSNR by 0.5-7.5dB compared with methods in TABLE II.

TABLE II.
PSNR OF CLASSIC DEINTERLACING METHODS

Seqences	ELA	Bob	Merge	3FMD	4FMD	5FMD
Akiyo	35.85	39.92	44.24	46.12	45.79	45.48
CG	27.69	29.06	27.53	30.10	29.98	29.79
Foreman	32.66	32.31	28.77	34.04	33.94	33.64
Mobile	21.81	25.51	23.33	26.31	25.98	25.81
News	29.01	33.62	36.29	41.99	42.34	42.09
Paris	23.71	26.61	31.85	35.81	35.52	35.14
Tempete	25.93	29.39	28.42	30.79	31.02	30.68
Average	28.09	30.92	31.49	35.02	34.94	34.66

CG are abbreviation for sequence Coastguard and Mother-and-daughter respectively.

TABLE III.
PSNR OF PROPOSED DEINTERLACING METHODS

Seqences	AMD+Filter3	AMD+Filter2	AMD+Filter1
Akiyo	46.12	46.79	46.89
CG	30.20	30.37	30.52
Foreman	34.23	34.02	34.35
Mobile	26.15	26.32	26.55
News	42.51	43.36	43.61
Paris	35.88	36.20	36.19
Tempete	31.29	31.45	31.68
Average	35.20	35.50	35.68

B. Subject performance

PSNR is not the only measure for evaluation the performance of deinterlacing quality, subjection performance is important too. For example, there a two dancers moving fast in the background of test sequence News, but the reporters in the foreground is almost static. With the 3FMD deinterlacing, there are error motion detections in the fast moving regions around the dancers, and mistake to use temporal interpolation to deinterlacing the pixels. These error motion detections make the jagged effects obvious, as shown in Fig. 9b. Though 5-field motion detection result is right, but interpolation by Filter3 blurs the border, as show in Fig. 9c. Compared with the original frame image Fig. 9a, the reconstructed picture with proposed algorithm has the best quality, as shown in Fig. 9d. These pictures confirm proposed deinterlacing algorithm improve the subjective quality of reconstructed frames.

Fig. 9a. Enlarged part of original frame picture

Fig. 9b. Deinterlaced result by 3-field motion detection

Fig. 9c. Deinterlaced result by 5-field motion detection

Fig. 9d. Deinterlaced result by proposed algorithms

V. CONLUSIONS

The simulation above shows that the proposed motion adaptive deinterlacing improves both objective and subjective performance. The new motion detection algorithms reduce the possibility of error A in 3-field motion detection method and that of error B in 5-filed motion detection method. And the anti-aliasing intra-field interpolation is better than windowed sinc function.

REFERENCES

[1] G.D.Haan, E.B. Bellers, "Deinterlace-an overview," *Proceeding of the IEEE*, vol.86, no. 9, Sep. 1998, pp. 1839-1857.

[2] A.M. tekalp, *Digital Video Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1995.

[3] B. Bhatt, F. Templin and A. Cavallerano, "Grand Alliance HDTV multi-format scan converter," *IEEE Trans. Consumer Electronics*, vol. 4, no. 4, Nov. 1995, pp. 1020-1031

[4] G. Thomas, "A comparison of motion-compensated interlace-to progressive conversion methods," Signal Process: Image Commun., vol. 12, 1998, pp. 209-229

[5] X. Gao, J. Gu and J. Li, "De-interlacing algorithms based on motion compensation," *IEEE Tran. Consumer Electronics*, vol. 51, no. 2, May 2003, pp. 589-599.

[6] D. Wang, A. Vincent, and P. Blancchfield, "Hybrid de-interlacing algorithm based on motion vector reliability, " *IEEE Trans. Trans. Circuits Syst. Video Technol*, vol. 15, no. 8, Aug. 2005, pp.1019-1025

[7] Y.L Chang, S.F. Lin, C.Y. Chen, L.G. Chen, "Video de-interlacing by adaptive 4-field global/local motion compensated approach," *IEEE Trans. Trans. Circuits Syst. Video Technol*, vol.15, no.12, 2005, pp.1569-1579

[8] R. Li, B..Zeng, and M. Liou, "Reliable motion detection/compensation for interlaced sequences and its applications to deinterlacing," *IEEE Trans. Trans. Circuits Syst. Video Technol*, vol. 10, no. 1, Feb, 2000, pp.23-29

[9] Y.Y. J, S.Yang and P.Yu, "An effective de-interlacing technique using two types of motion information," *IEEE Trans. Consumer Electronics*, vol. 49, no.3, Aug, 2003, pp.493-39.

[10] S.G. Lee, D.H. Lee, "A motion-adaptive de-interlacing method using an efficient spatial and temporal interpolation", *IEEE Trans. Consumer Electronics*, vol. 49, no.4, Nov. 2003, pp. 1266-1271.

[11] R. A. Beuker and I.A. Shah, "Analysis of interlacevideo signals and it applications," *IEEE Trans. Image Processing*, vol.3, no.5, Sep. 1994, pp. 501-512

Lejun Yu was born in Jiangxi, China in 1977. He received B.S and M.S degrees computer engineering and electrical engineering from Beijing Normal University in 1999 and 2002 respectively. Now, he is pursuing the PH.D degree at Institute of Computing Technology, Chinese Academy of Sciences. His research interests are in video coding and image processing.