A Fuzzy Motion Adaptive De-interlacing Algorithm Capable of Detecting Field Repetition Patterns

P. Brox, I. Baturone and S. Sánchez-Solano
Instituto de Microelectrónica de Sevilla
Consejo Superior de Investigaciones Científicas (CSIC)
Avda. Reina Mercedes S/N. Edificio CICA
41012 Sevilla (Spain)
email:brox@imse.cnm.es

Abstract **– A new motion adaptive algorithm for de-interlacing video is proposed in this paper. It employs two fuzzy systems to interpolate the missing lines of the transmission. One fuzzy system is used to evaluate the motion level at the current pixel, and a second one selects the most adequate temporal interpolation method. The combination of both systems provides an effective result with a low cost in term of hardware resources.**

Keywords **– De-interlacing, Detection of field repetition, Fuzzy logic, Motion adaptive**

I. INTRODUCTION

Analogue television broadcast systems in current use (PAL, NTSC and SECAM) employ interlacing to halve the bandwidth of video signal transmission. Interlacing consists of reducing the vertico-temporal sampling density by eliminating the odd/even lines of the frames in a sequence. Nowadays, the increasing demand of devices that require a progressive scanning format has encouraged the development of de-interlacing algorithms. These algorithms perform an interpolation to calculate the non-transmitted lines trying to minimize the artifacts caused by interlacing.

Two main categories are distinguished among de-interlacing algorithms: motion compensation (MC) and non-MC methods [1]. The first ones are the most advanced de-interlacing algorithms and, particularly, achieve very good results in moving areas of the image. The problem is that they require the calculation of a bi-dimensional motion vector and a displacement estimator at each pixel of the field, which means a high computational cost. Among non-MC methods, the motion adaptive algorithms offer a good trade-off between cost and quality [2]. This kind of algorithms combines an spatial or intra-field method with a temporal or inter-field method according to the presence of motion. They are based on the idea that temporal interpolation is very suitable for static areas, while spatial interpolation is more adequate when the level of motion is high.

Typical motion adaptive algorithms employ a threshold value to distinguish between small and high levels of motion. Selection of this threshold is difficult because usual problems, such as the presence of noise, produce wrong decisions. Inherently, transition between small and high motion is fuzzy rather than crisp and, hence, a fuzzy instead of a crisp motion detector is more robust. Several algorithms reported in the literature use fuzzy logic to detect motion in de-interlacing applications [3]-[4]. Van de Ville et al. describe in [3] how heuristic knowledge to detect motion is translated from 'if-then' rules expressed linguistically into an algorithm, where the presence or absence of motion is evaluated in a fuzzy way. This is performed by a complex fuzzy system, which uses five input variables since the level of motion is not only evaluated in the current pixel, but also pixels in the neighborhood are considered. The computational cost of this proposal is considerably high since it requires 26 maximum operators, 29 minimum operators, 60 defuzzification processes and one division. The output of the fuzzy motion detector ($\gamma(x,y,t)$) measures the motion in a pixel with the spatial coordinates (x,y) in the field (t) of the sequence. The luminance value of the interpolated pixel (X) is given by the following expression:

$$X = (1-\gamma(x,y,t))I_T + \gamma(x,y,t)I_S \tag{1}$$

where $\gamma(x,y,t)$ is a value in the interval $[0,1]$ instead of a crisp value of the conventional motion adaptive algorithms.

Gutierrez et al. propose in [4] a simplification of the algorithm developed in [3]. This is performed by decreasing the number of input variables from five to one thanks to the use of a bi-dimensional convolution to measure the level of motion. The influence of the neighbors is considered by including a coefficients matrix (C), in which higher values correspond to closer neighbors. This proposal reduces the

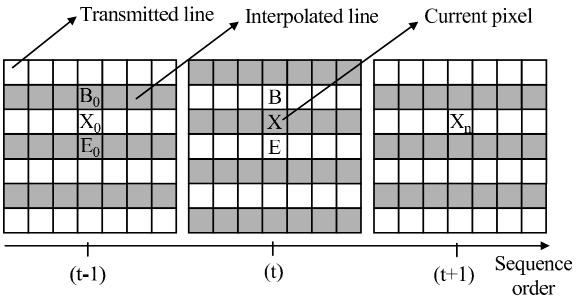


Fig. 1. PIXELS INVOLVED IN THE CALCULATION OF $motion(x,y,t)$

resources to 15 multipliers, 14 adders and 1 defuzzification process. Furthermore, the approach achieves a better quality in the reconstruction of the interpolated images than the proposal in [3].

Both fuzzy logic-based algorithms use a temporal aperture of four fields to measure motion and employ as temporal interpolation the value of the pixel in the previous field (field insertion method) or the mean value between the pixels in the previous and subsequent fields.

The algorithm proposed herein also uses fuzzy logic to interpolate between two processing modes, a spatial (I_S) and a temporal (I_T) interpolator. A relevant novelty of our proposal is the inclusion of a second fuzzy system which works with three instead of four fields and implements a smart temporal interpolation. Furthermore, this second fuzzy system is capable of detecting the presence of field repetition patterns. This ability is quite often demanded in video applications due to the arrival of hybrid material, that is, sequences which mixes different scanning formats. Section II includes a detailed description of the algorithm. The performance of the algorithm is compared in Section III with other well-known de-interlacing algorithms by extensive simulation of sequences. The obtained results with different scanning formats are measured quantitatively in terms of figures of merit such as Mean Square Error (MSE) and Peak Signal to Noise Ratio (PSNR), and qualitatively by visual inspection of the progressive frames. Finally some conclusions are expounded in Section IV.

TABLE I.
RULEBASE FOR INTERPOLATOR SELECTION

	if	*then*
1)	*motion* is S	I_T
2)	*motion* is L	I_S
3)	*motion* is M	$\lambda I_T + \delta I_S$

II. DESCRIPTION OF THE ALGORITHM

Our proposal uses as input value the bi-dimensional convolution of field difference signals that can be mathematically expressed as follows:

$$motion(x,y,t) = \frac{\Sigma C_{j,i} H_{i,j}}{\Sigma C_{j,i}} = \frac{(1\ 2\ 1)(H_{1,1}\ H_{1,2}\ H_{1,3})^T}{4} \quad (2)$$

where $C_{i,j}$ are the values of the weights and $H_{i,j}$ are described by the following differences of luminance values (see Fig. 1):

$$H_{1,1} = |B_0 - B| \quad (3)$$

$$H_{1,2} = |X_0 - X| \quad (4)$$

$$H_{1,3} = |E_0 - E| \quad (5)$$

Different weights and sizes matrix H have been studied to achieve a good trade-off between the computing resources and the quality of motion measurement [5]-[6]. As it can be seen in expression (2), the selected convolution only includes neighbors in vertical direction since a wide number of video sequence simulations shown a non-decisive influence of horizontal neighbors to measure the level of motion.

Unlike the proposals in [3]-[4], our algorithm uses a temporal aperture of three fields, as shown in Fig. 1. The interpolated values calculated in the previous field (B_0, E_0) are necessary to evaluate the $motion$ value in expression (2). A simple de-interlacing algorithm is employed to calculate the first progressive frame such as the vertico-temporal filter [7].

The influence of motion in selecting the kind of interpolation is evaluated by considering the following rules that are linguistically expressed as follows:

1. If $motion$ in the current pixel is small (S), the most adequate interpolated value is obtained by applying a temporal interpolation (I_T).
2. If $motion$ in the current pixel is large (L), the best result is obtained by performing a spatial interpolation (I_S).
3. If $motion$ in the current pixel is medium (M), then the value is better calculated by applying a linear combination of the temporal and spatial interpolators ($\lambda I_T + \delta I_S$).

This rule base is summarized in Table I. The fuzzy concepts small, large and medium used in the rules are modeled according to the membership functions shown in Fig. 2. Using the Fuzzy

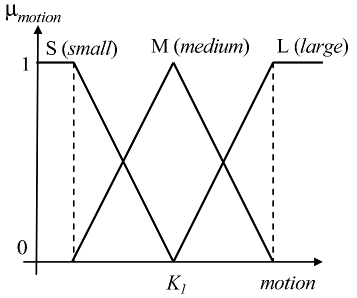


Fig. 2. MEMBERSHIP FUNCTIONS USED IN THE FUZZY SYSTEM TO EVALUATE MOTION

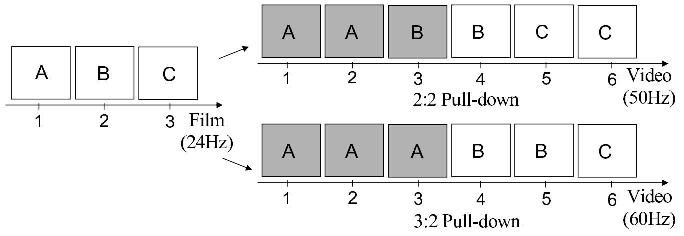


Fig. 3. STANDARD CONVERSION BETWEEN FILM AND VIDEO FORMATS

Mean as defuzzification method the new pixel value is calculated as follows:

$$X = \alpha_1 I_T + \alpha_2 I_S + \alpha_3(\lambda I_T + \delta I_S) \qquad (6)$$

where α_i is the corresponding activation degree of each rule in Table I.

Three degrees of motion are considered in this fuzzy system. After analyzing up to five degrees of motion [8], the rule base with three rules has been selected since it provides the most attractive solution in terms of hardware resources and quality of the interpolated image [8].

Let us consider the temporal aperture of three fields shown in Fig. 1. If the sequence was recorded by a video camera with a picture rate of 50 Hz (PAL) or 60 Hz (NTSC), the three fields of the aperture are different in moving areas of the image. However, if the material was registered with a cine-camera the picture rate is 24 Hz. A conversion of film material is necessary to display it on TV. The conversion to adapt both picture rates basically consists of repeating the fields twice (to achieve 50 Hz) or twice and three times alternatingly (to achieve 60 Hz) as it is shown in Fig. 3. This process is known as pull-down 2:2 and pull-down 3:2, respectively.

If the temporal aperture are composed by three fields from film material (for instance grey fields in Fig. 3), two or three of the fields in the aperture are similar. The detection of these cases is very interesting due to the risen presence of film and hybrid material on TV. Besides a perfect de-interlacing can be achieved by weaving at a expense of a minimal cost if the repeated field is detected.

To identify the presence of film material a simple fuzzy system is used. It selects the most adequate temporal interpolation depending on the *similarity* between two consecutive fields given by the following expression:

$$similarity(x,y,t) = \frac{|B-B_0|+|E-E_0|}{2} \qquad (7)$$

TABLE II.

RULEBASE FOR FIELD REPETITION DETECTION

	if	***then***
1)	*similarity* is S	X_0
2)	*similarity* is L	X_n

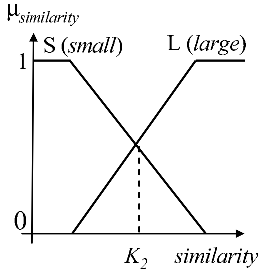


Fig. 4. MEMBERSHIP FUNCTIONS USED TO SELECT THE TEMPORAL INTERPOLATION

The heuristic knowledge of this fuzzy system is expressed by means of the following linguistic rules:

1. If $similarity$ between the fields *(t-1)* and *(t)* is small (S), the most adequate interpolated value is obtained by selecting the pixel value in the previous field at the same spatial position *(X_0)* in Fig. 1.
2. On the contrary, if $similarity$ is large (L), the pixel value in the previous field is not a good choice and is better to bet on the pixel in the next field *(X_n)* in Fig. 1.

Table II summarizes the rule base of this second fuzzy system. The shape of membership functions to model the fuzzy concepts small and large are shown in Fig. 4.

The output of this fuzzy system is given by the following expression:

$$I_T = \beta_1 X_0 + \beta_2 X_n \qquad (8)$$

where β_i is the activation of each rule in Table II.

Different methods can be selected to perform spatial interpolation. Specifically the algorithm reported in [9] has been selected due to its ability to interpolate pixels while preserving edges in the image.

A block diagram of the global algorithm is shown in Fig. 5. As it can be seen, two fuzzy systems are used (FS1 and FS2 boxes). FS_1 corresponds to the motion detection system whereas FS_2 selects the most adequate temporal interpolation.

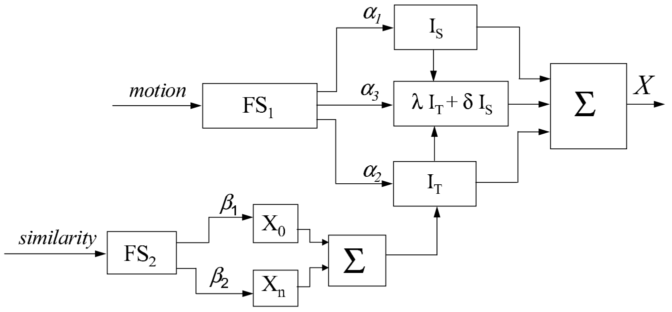


Fig. 5. BLOCK DIAGRAM OF THE GLOBAL ALGORITHM

TABLE III.

PSNR VALUES (IN DBS) FOR DIFFERENT DE-INTERLACING METHODS

Video Sequences	Missa	Paris	Trevor	Salesman	News	Mother	Carphone
Format	CIF	CIF	CIF	CIF	QCIF	QCIF	QCIF
Line Doubling	36.44	23.61	31.05	29.75	25.18	31.81	28.25
Line Average	40.47	26.67	35.04	33.53	29.25	35.94	32.61
ELA 3+3	39.49	25.53	34.11	32.11	26.63	35.39	32.65
ELA 5+5	38.56	24.64	33.31	30.17	25.92	34.2	31.51
Field Insertion	38.36	29.86	34.36	36.17	33.13	36.14	30.34
VT 2fields	40.25	30.73	36.61	36.54	35.46	39.61	34.08
VT 3fields [7]	40.52	31.37	37.16	36.95	35.67	40.89	34.54
Technique in [3]	40.01	33.12	35.38	37.62	34.73	39.49	32.27
Technique in [4]	40.18	35.28	36.69	38.29	37.51	41.87	34.78
Proposal	40.81	35.87	37.63	38.35	38.78	42.11	35.09

III. SIMULATION RESULTS

The performance of the proposed algorithm has been analyzed by de-interlacing several video sequences. They can be divided into two categories: a first group of standard video sequences and a second one of real film sequences from TV channels or movies.

The video sequences considered have widely been used as benchmarks in video processing applications. The interlaced video data have been obtained from these progressive sequences by eliminating lines. Many error measures have been proposed as figure of merit to evaluate the quality between the obtained interpolated frames and the original ones. However, as a consequence of the complexity of the human visual system, it is difficult to find an objective criterion to entirely describe the visibility of image distortions. Nevertheless, some measures seem to have a higher correlation than others with the perceived quality. A very popular quality criteria is the mean square error (MSE) given by the following expression:

$$MSE = \frac{1}{M\times N}\sum_{i=1}^{MxN}(X - X_{progressive})^2 \qquad (9)$$

where the image has a resolution of $M \times N$ pixels, X is the value of the interpolated pixel, and $X_{progressive}$ is the pixel value in the original progressive image.

Strongly related to the MSE is the PSNR, which is defined as follows:

$$PSNR = 20\log\frac{255}{\sqrt{MSE}} \qquad (10)$$

The proposed algorithm has been also compared with other de-interlacing algorithms with less or similar computational cost: four spatial method such as line doubling, line average, and conventional ELA (edge-adaptive interpolation algorithms [1]) using 3+3 and 5+5 taps; the simplest temporal de-interlacing algorithm called field insertion, and two vertico-temporal filtering with two and three fields [7]; and, finally the fuzzy motion adaptive algorithms reported in [3] and [4].

Table III shows the average PSNR values obtained when de-interlacing fifty fields of seven video sequences. The PSNR results show that the proposed algorithm performs better than the other algorithms since it achieves the highest values. Furthermore its computational complexity is very low since it only requires 3 multipliers and 2 adders to implement the first fuzzy system (FS_1).

The algorithm has also been tested to de-interlace the real film sequences shown in Table IV. In this case, the proposal is compared with one of the best well-known algorithm (VT 3fields [7]), and two proposals that use the same fuzzy system to detect motion (FS_1) but they employ two different temporal interpolators (I_T). The $Proposal_{previous}$ employs the value of the previous field (X_0), whereas the $Proposal_{average}$ uses the average value of pixels in the previous and next field ($0.5(X_n + X_0)$). The results are shown in Table IV and prove the advantages of the inclusion of the second fuzzy system FS_2. It decreases the average MSE value by a high factor if the results are compared with VT 3fields.

The MSE values at each field of 'Fire-Rose' film sequence are shown in Fig. 6. As it can be seen, the $Proposal_{previous}$ shows an asymmetrical behavior. The sequence corresponds to a pull-down 2:2 pattern and the previous field coincides with the current field alternatively. This effect is reduced when the $Proposal_{average}$ is applied but our proposal achieves the lowest MSE error as shown in Fig. 6.

Finally, the superior performance of the proposed algorithm can be corroborated by the visual inspection of the progressive frames shown in Fig. 7.

IV. CONCLUSIONS

The proposed algorithm performs better that other motion adaptive algorithms reported in the literature. It uses two simple fuzzy systems: one considers the level of motion to employ spatial and temporal interpolation and a second one that selects the best solution for a temporal interpolation. A relevant

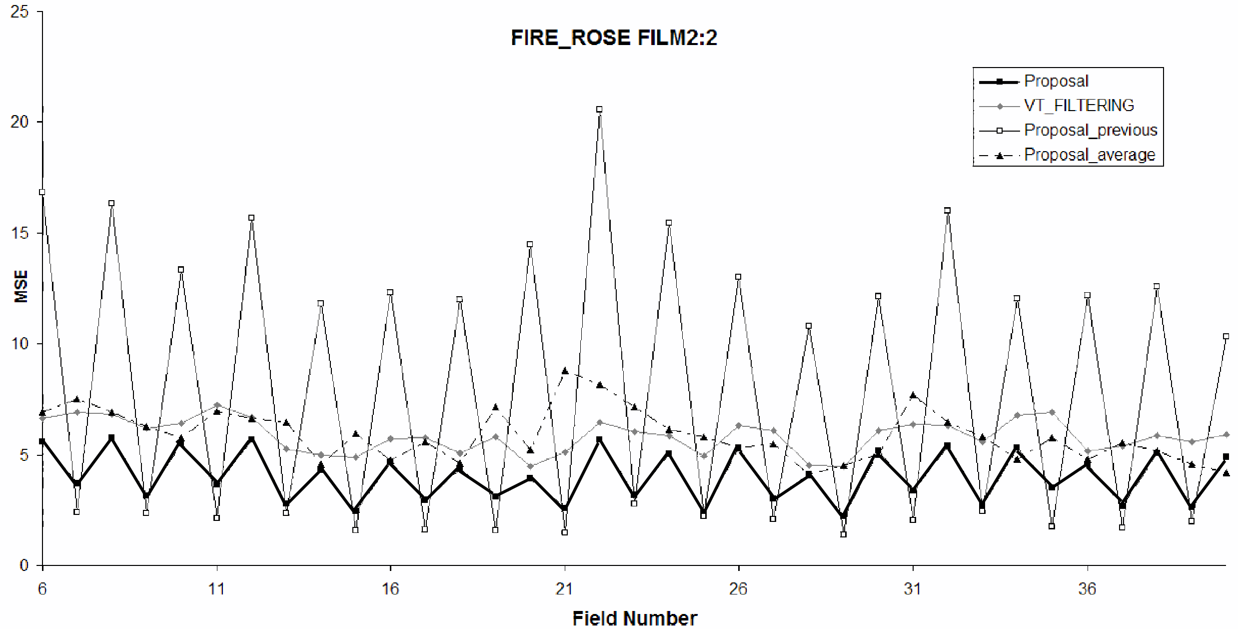


Fig. 6. MSE RESULT AT EACH FIELD OF THE 'FIRE-ROSE' FILM SEQUENCE

TABLE IV.
MSE RESULTS OF FILM SEQUENCES

Film Sequence	Fire-rose	Chop-hunt	Fargo-repair	Fargo	Tokyo
VT 3fields [7] $Proposal_{previous}$ $Proposal_{average}$ Proposal	5.84 8.047 5.901 4.006	1.48 13.89 11.84 0.85	9.35 5.44 7.03 4.45	10.403 17.36 14.92 7.96	517.26 516.11 482.24 383.59

advantage of including the second fuzzy system is that film mode and hybrid material are detected.

ACKNOWLEDGMENT

This work has been supported in part by the Spanish CICYT Projects TEC2005-04359 and DPI2005-02293, and by the Projects TIC2006-635 and TEP2006-375 from the Andalusian regional Government.

REFERENCES

[1] G. de Haan and E. B. Bellers, "De-interlacing: an overview," in *Proc. of the IEEE*, Sep. 1998, pp. 1839–57.

[2] A. M. Bock, "Motion adaptive standards conversion between formats of similiar field rates," *Signal Processing: Image Communication*, vol. 6, no. 3, pp. 275–80, Jan. 1994.

[3] D. Van de Ville W. Philips and I. Lemahieu, "Fuzzy-based motion detection and its application to de-interlacing," in *Fuzzy techniques in image processing. Book Series of Studies in Fuzziness and Soft Computing*, 2000.

[4] J. Gutiérrez-Ríos F. Fernández-Hernández J. C. Crespo and G. Trevino, "Motion adaptive fuzzy video de-interlacing method based on convolution techniques," in *Proc. of Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU)*, Jul, 2004.

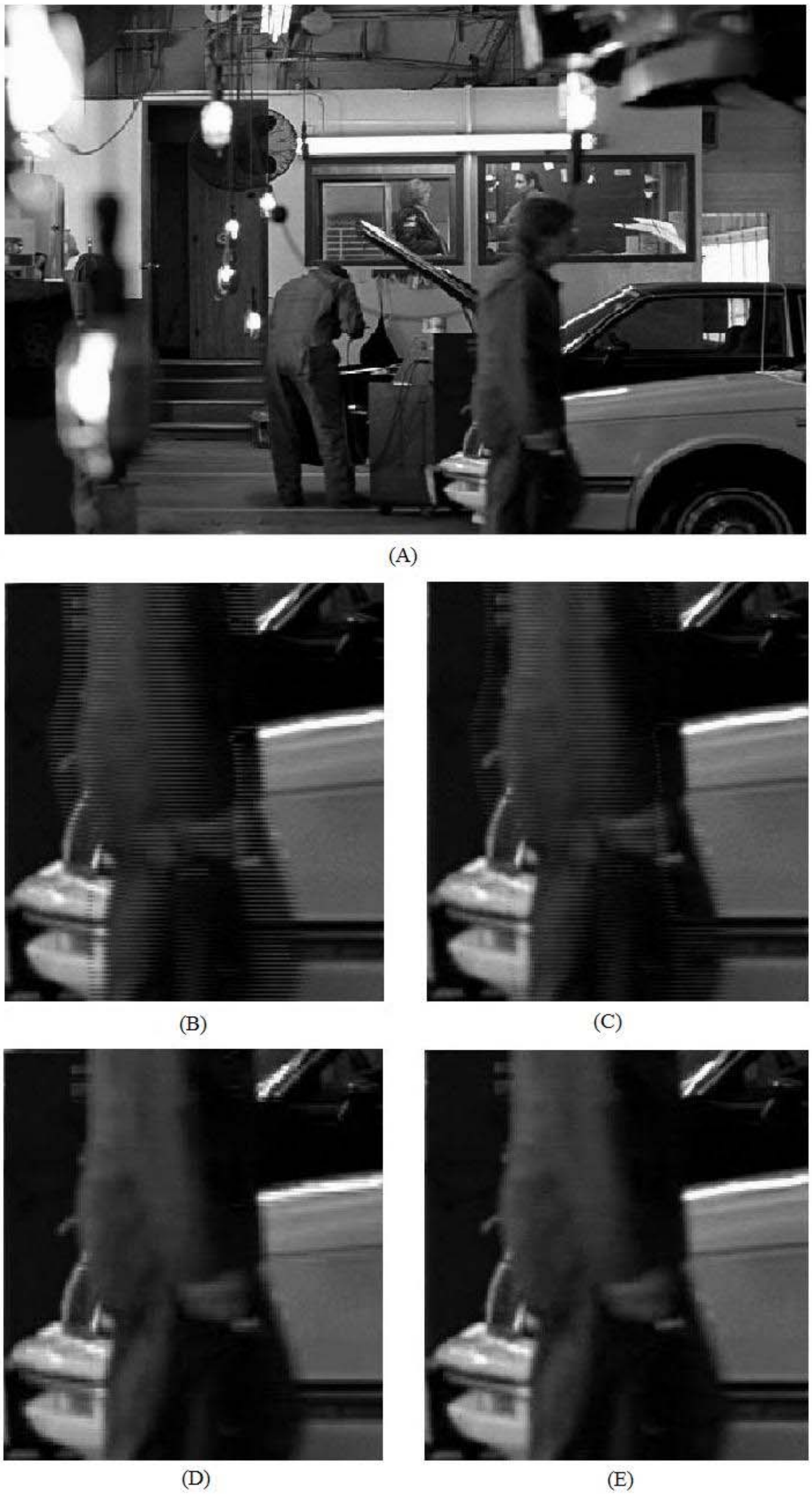
[5] P. Brox I. Baturone S. Sánchez-Solano J. Gutiérrez-Ríos and F. Fernández-Hernández, "A fuzzy edge-dependent motion adaptive algorithm for de-interlacing," *Fuzzy Sets and Systems. Special Issue: Image Processing*, vol. 158, no. 3, pp. 337–347, Feb. 2007.

[6] P. Brox I. Baturone and S. Sánchez-Solano, "A fuzzy motion adaptive algorithm for interlaced-to-progressive conversion," in *Proc. of Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU)*, Jul, 2006.

[7] M. Weston, "Interpolating lines of video signals," in *Asignee: British Broadcasting Corporation*, Dec, United States Patent Office Us 4,789,893 1998.

[8] P. Brox I. Baturone and S. Sánchez-Solano, "Fuzzy motion adaptive algorithm for video de-interlacing," in *Proc. of International Conference on Knowledge-Based and Intelligent Information and Engineering Systems (KES)*, Oct, 2006.

[9] P. Brox I. Baturone and S. Sánchez-Solano, "A fuzzy edge-dependent interpolation algorithm," in *Soft Computing in Image Processing: Recent Advances. Heidelberg, Germany. Springer*, 2007.



(A)

(B)

(C)

(D)

(E)

Fig. 7. DETAILED AREA AFTER DE-INTERLACING THE FIELD IN (A) BY APPLYING: (B) $Proposal_{previous}$ (C) $Proposal_{average}$ (D) VT 3FIELDS (E) PROPOSED ALGORITHM