Soft Computing Techniques for Video De-interlacing

Piedad Brox, Iluminada Baturone and Santiago Sánchez-Solano

Abstract— This paper presents the application of soft computing techniques to video processing. Specially, the research work has been focused on de-interlacing task. It is necessary whenever the transmission standard uses an interlaced format but the receiver requires a progressive scanning, as happens in consumer displays such as LCDs and plasma. A simple hierarchical solution that combines three simple fuzzy logic-based constituents (interpolators) is presented in this paper. Each interpolator specialized in one of three key image features for de-interlacing: motion, edges, and possible repetition of picture areas. The resulting algorithm offers better results than others with less or similar computational cost. A very interesting result is that our algorithm is competitive with motion-compensated algorithms.

Index Terms— Motion-adaptive De-interlacing, Edge-adaptive De-interlacing, Picture-repetition, Video Signal Processing, Soft Computing, Fuzzy Inference Systems.

I. INTRODUCTION

THE success of the interlaced video scan format is based on the technologies used to implement TV in the analog area, since CRTs display the interlaced video directly. Although present-day technologies are sufficiently powerful to achieve progressive video signal, interlacing is currently used by all the analog TV broadcast systems (NTSC, PAL and SECAM) and also by some of the modern digital transmissions as is described in [1].

The weakest point of the interlaced scanning is that scanning format coversion is considerable more complex if video is interlaced. Conversion between formats was required in the past by international programme exchange. Nowadays, this demand has recently increased due to the advent of new video scanning standards. Furthermore, interlaced video signal is unsuitable for current consumer displays such as plasma and LCD panels, which need a progressive scanning format. These facts have encouraged the development of video de-interlacing algorithms during the last years [2]. They perform the reverse operation of interlacing with a frame rate equal to the original field rate.

P. Brox, I. Baturone and S. Sánchez-Solano are with the Microelectronics Institute of Seville (CNM-CSIC), Américo Vespucio 41092, Seville (Spain). P. Brox and I. Baturone are also with the University of Seville, 41012 Seville (Spain), Phone:+34 954466666, Fax:+34 954466600 (e-mail:brox@imse-cnm.csic.es).

A rough classification of de-interlacing algorithms divides them into two categories: Motion Compensated (MC) and non-Motion Compensated (non-MC) algorithms. MC techniques offer the best solution, specially in moving areas but they require the calculation of motion vectors that usually involves a high computational cost. Among non-MC algorithms, feature-dependent de-interlacing algorithms are a good alternative since they reduce the computational complexity obtaining an overall picture improvement [3]- [8], [10]- [18], [27]- [28].

The application of soft computing techniques to solve video de-interlacing is described in this paper. Among them neuro-fuzzy logic-based systems are able to model uncertainty and subjective concepts in a better form than conventional methods. Image features such as motion, edges, or picture repetition often contain certain amount of uncertainty. For instance, sometimes it is difficult to define if a pixel belongs to an edge (i.e., object frontier or not). Especially this decision is not trivial when the pictures are noisy and/or contain a high number of details.

The 'divide-and-conquer' strategy is proposed to develop a local motion- and edge-adaptive de-interlacing algorithm capable of detecting repeated areas in the pictures. This algorithm consists of a hierarchical structure that combines three simple fuzzy logic-based systems, each of them tackling a relevant feature for de-interlacing: motion, edges, and possible picture repetition of areas.

This paper is organized as follows: Section II describes the three fuzzy systems that compose the algorithm. Its performance when de-interlacing a wide battery of sequences is presented in Section III. Finally, some conclusions are expounded in Section IV.

II. DESCRIPTION OF THE ALGORITHM

The block diagram of the complete algorithm is shown in Fig.1. It is composed of three fuzzy systems that are described in the following subsections.

A. *Fuzzy logic-based system for motion adaptation*

Motion-adaptive approaches are based on the fact that linear temporal interpolators are perfect in the absence of motion, whereas linear spatial methods offer a most adequate solution in case that motion is detected. Motion detection can be implicit, as in median-based techniques [3]- [4], or explicit, using a motion detector [5]- [8]. Explicit motion-adaptive algorithms calculate a new pixel value by interpolating between a spatial and a temporal de-interlacing method. For color transmission standards, the motion-adaptive de-interlacing process may be
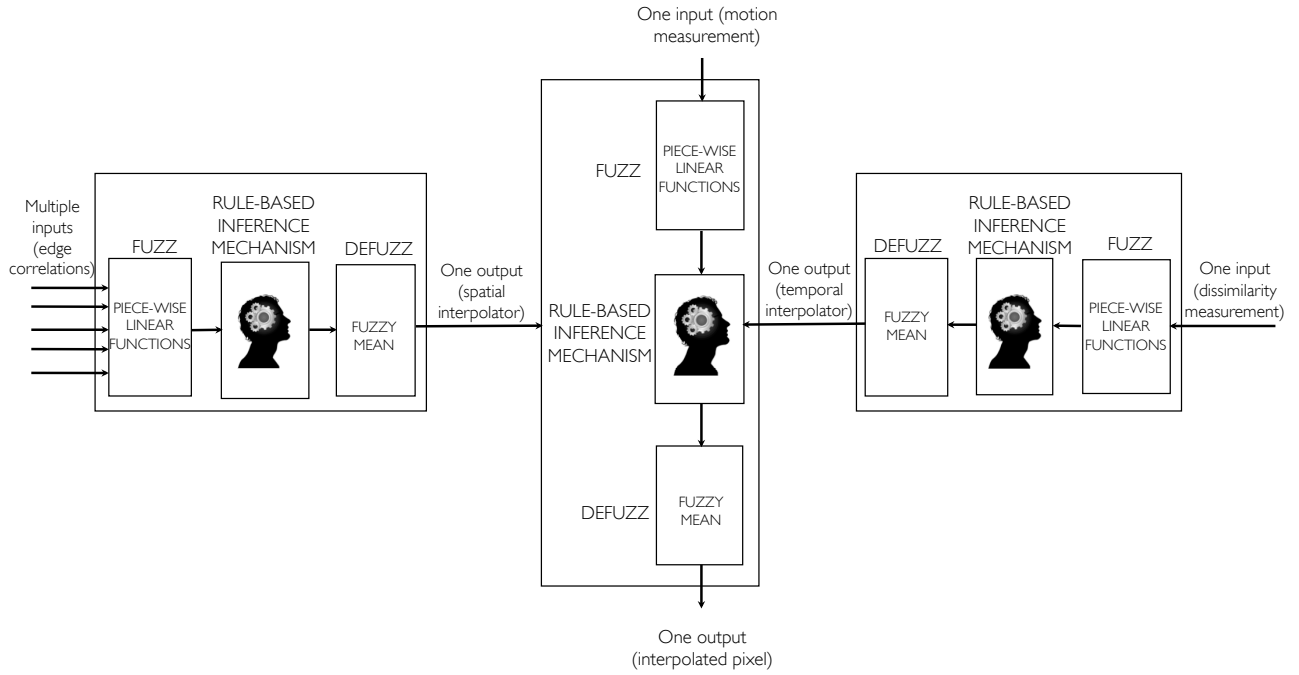
Fig. 1.   Block diagram of the complete algorithm that contains three fuzzy logic-based systems.

$$C1 = \begin{pmatrix} 1 & 1 & 2 & 1 & 1 \\ 1 & 2 & 3 & 2 & 1 \\ 2 & 3 & 4 & 3 & 2 \\ 1 & 2 & 3 & 2 & 1 \\ 1 & 1 & 2 & 1 & 1 \end{pmatrix} \quad C2 = \begin{pmatrix} 1 & 2 & 3 & 2 & 1 \\ 1 & 3 & 5 & 3 & 1 \\ 1 & 2 & 3 & 2 & 1 \end{pmatrix} \quad C3 = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 3 & 2 \\ 3 & 4 & 3 \\ 2 & 3 & 2 \\ 1 & 2 & 1 \end{pmatrix}$$

$$C4 = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix} \quad C5 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad C6 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$$

Fig. 2.   Convolution masks options.

performed separately on each color space component. However, based on the fact that the details of an image are mainly determined by the luminance component of the video signal, and much less by the chrominance components, the motion detection is usually done with the luminance component. Furthermore, this is more efficient in terms of the number of computations required to implement it since only one of the three components is processed. When a de-interlacing method is selected for the luminance component of a pixel, the two chroma components of this pixel are also created by using the same de-interlacing method.

The primitive proposals that use the motion-adaptive idea de-interlace video by applying the following expression [5]:

$$I_p = (1 - \alpha)I_T + \alpha I_S \tag{1}$$

where I_S is the output of a spatial interpolator, and I_T is the output of a temporal interpolator. The variable α, which is the output of a motion detector, ranges from 0 to 1 and determines the level of motion.

The performance of explicit motion-adaptive algorithms relies on the quality of the motion detector, since it is strongly dependent on the combination of both de-interlacing algorithms. Our aim is to use a fuzzy system that performs the non-linear interpolation between a spatial and a temporal interpolator according to the presence of motion. The block diagram of this fuzzy system is shown in the middle part of Fig.1.

This fuzzy system mixes spatial and temporal sub-de-interlacers as a good trade-off between simplicity and efficiency. Anyway, the used methodology of soft computing would allow the more advanced mixing of three and more sub-deinterlacers.

The fuzzy system for motion adaptation has one input that is a motion measurement based on the use of bi-dimensional convolution between a matrix of picture differences, M, and a matrix of weights, C. For a current pixel of coordinates (x, y, t), the input of this fuzzy system is as follows:

$$motion(x, y, t) = \frac{\Sigma_{a=1}^3 (\Sigma_{b=1}^3 M_{a,b} C_{a,b})}{\Sigma_{a=1}^3 \Sigma_{b=1}^3 C_{a,b}} \tag{2}$$

where $M_{(i,j)}$ are the elements of the following matrix, M, of difference values among three consecutive fields or pictures:

$$M = \begin{pmatrix} M_{(-1,-1)} & M_{(0,-1)} & M_{(1,-1)} \\ M_{(-1,0)} & M_{(0,0)} & M_{(1,0)} \\ M_{(-1,1)} & M_{(0,1)} & M_{(1,1)} \end{pmatrix} \tag{3}$$

The elements of the matrix $M_{(i,j)}$ are calculated as follows:

$$M_{(i,0)} = \frac{|I(x+i,y,t+1)-I(x+i,y,t-1)|}{2} \tag{4}$$

$$M_{(i,-1)} = \frac{|I(x+i,y-1,t)-I_p(x+i,y-1,t-1)|}{2} \tag{5}$$

$$M_{(i,1)} = \frac{|I(x+i,y+1,t)-I_p(x+i,y+1,t-1)|}{2} \tag{6}$$

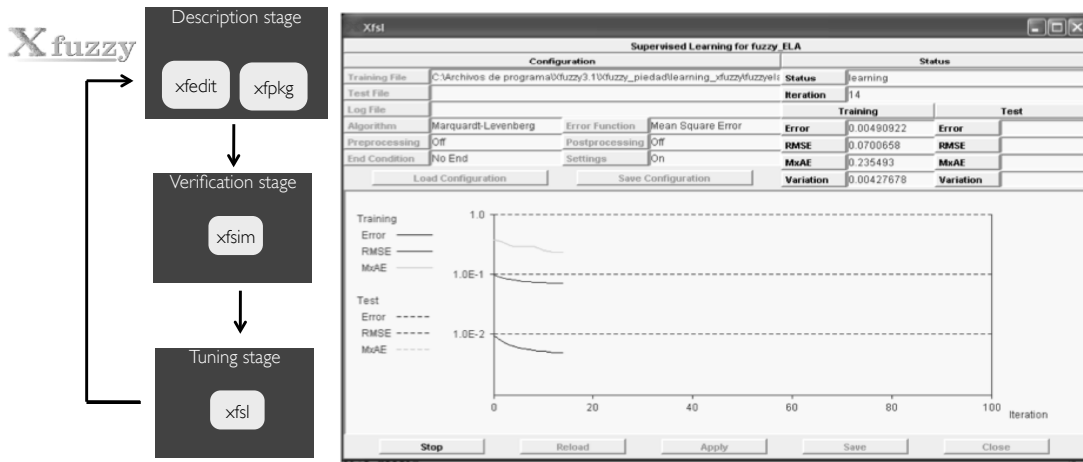where I_p are the interpolated luminances of the pixels in the

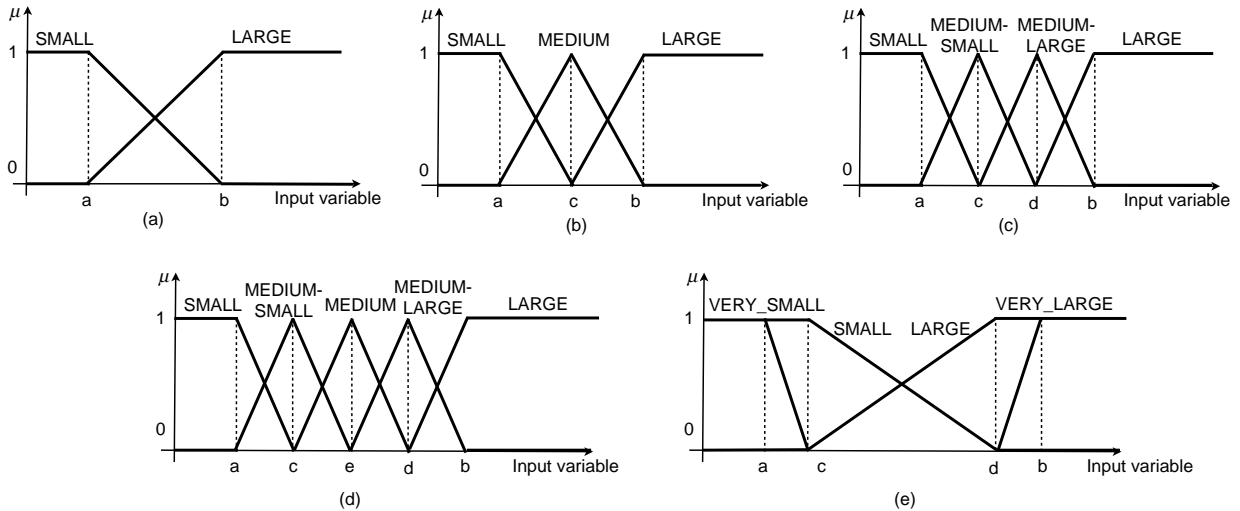Fig. 3.   A tuning stage is included in the design of the fuzzy systems.

Fig. 4.   Piece-wise linear membership functions used by the fuzzy systems.

previous field[1].

And $C_{a,b}$ are the elements of the following weight matrix:

$$C = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 0 \end{pmatrix} \qquad (7)$$

Since a priori there is no restriction to select the convolution parameters (size and weights of the matrix), two different strategies have been studied. Firstly, a battery of six different convolution masks has been considered (see Fig. 2). The convolution masks were selected following the next two criteria: 1) the highest weight corresponds to the position of the current pixel; 2) the sum of the weights is a power of two in order to simplify the hardware implementation. Extensive simulations prove that the convolution mask does not have a significant influence on de-interlacing if two vertical neighbors are at least

---

[1]This expression requires that the previous field has to be de-interlaced. Hence, the first field in the sequence should be de-interlaced by a spatial interpolator method (in our case, the 'Improved Fuzzy-ELA 5+5' algorithm described in the next subsection is used).

considered. Therefore, we concluded that $C5$ and $C6$ are the best options. A second alternative is the use of supervised training algorithms to obtain the most suitable weights for the convolution mask. A set of data from standard progressive sequences have been employed to minimize an error function. A CAD tool called Xfuzzy 3 [9], which covers all the stages involved in the design process of fuzzy systems allows the application of this methodology. Xfuzzy 3 includes a specific tool called *xfsl* to ease the tuning process of fuzzy systems by applying learning algorithms. Fig. 3 shows a block diagram of the design methodology used in the design of the fuzzy system and the main window of *xfsl* tool. The well-known Marquardt-Levenberg approach is used as supervised learning algorithm. This second strategy corroborates the results obtained with the first one, and hence, one of the simplest convolution mask ($C5$) is employed to calculate the inputs of this fuzzy system.

The first step of the fuzzy inference process is called fuzzification (FUZZ). It consists of determining the degree to which the input belongs to each of the appropriate fuzzy sets via the chosen membership functions. The shape of the

TABLE I

RULE BASE OF THE FUZZY SYSTEM FOR MOTION ADAPTATION WITH 5 RULES

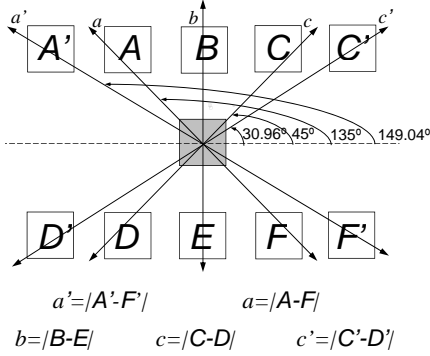| Rule	Antecedents	Consequent
1.	$motion(x, y, t)$ is $SMALL$	$I_T(x, y, t)$
2.	$motion(x, y, t)$ is $SMALL - MEDIUM$	$\gamma_1\, I_T(x,y,t) + \lambda_1\, I_S(x,y,t)$
3.	$motion(x, y, t)$ is $MEDIUM$	$\gamma\, I_T(x,y,t) + \lambda\, I_S(x,y,t)$
4.	$motion(x, y, t)$ is $MEDIUM - LARGE$	$\gamma_2\, I_T(x,y,t) + \lambda_2\, I_S(x,y,t)$
5.	$motion(x, y, t)$ is $LARGE$	$I_S(x, y, t)$

Fig. 5.　Aperture used in the ELA with 5+5 taps to interpolate the grey pixel, X.

membership functions is piece-wise linear since this selection eases the hardware implementation of the algorithm. Again there are several design options since there is no restriction to fix the number of membership functions. Four different options have been considered, from two up to five membership functions, as shown in Fig. 4(a), (b), (c) and (d). Because of linguistic coherence, the degree of overlapping between two consecutive sets is two.

The rule base of the inference system is strongly related to the selection of the number of membership functions. Table I shows the most complex situation that has been studied with five functions (see Fig. 4(d)). The first rule states that 'if $motion(x, y, t)$ is SMALL then the interpolated pixel is calculated by applying a temporal interpolator (I_T)'. On the contrary, the rule number five asserts that 'if $motion(x, y, t)$ is LARGE the interpolation is performed by applying a spatial interpolator (I_S)'. The other three rules consider intermediate situations that, when they are activated, perform different linear combinations of the spatial and the temporal interpolators [2].

Heuristic knowledge does not provide enough information neither to fix the constant values γ, γ_1, γ_2 and λ, λ_1, λ_2 of the rules' consequents, nor to determine the values a, b, c, d, and e, that describe the five possible linguistic labels (see Figure 4(d)). In order to choose these values, we have again employed a tuning process by using supervised learning algorithms as detailed previously.

Since one of our goals is to always consider a balanced

[2]Take note that the intermediate rules are not present in the case with two membership functions (see Fig. 4(a)). Only the rules number one, three, and five are present if three membership functions are selected (see Fig. 4(b)). Finally, the first, second, fourth, and fifth rules compose the rule base set of the system in case that four functions are selected (see Fig.4(c)).

solution between complexity and quality in the design of the fuzzy system, the introduction of more rules should be justified by a significant improvement of the de-interlacing performance. Extensive results show that using more than three rules, that is three membership functions, does not improve significantly the de-interlacing and even, sometimes, degrades the performance. Therefore, three rules have been selected as the best option.

The final step to calculate the value of the interpolated pixel is the defuzzification process (DEFUZZ) (see Fig. 1). Among the defuzzification methods, the Fuzzy Mean (FM) has been chosen since it is a simplified method that allows hardware simplicity. It consists of a weighted average of the rule consequents, c^r (see Table I), where the weights are the activation degrees, α^r, of the corresponding rules:

$$FM = \frac{\sum_r \alpha^r \cdot c^r}{\sum_r \alpha^r} \tag{8}$$

B. Fuzzy logic-based system for edge adaptation

This fuzzy system takes inspiration from well-known conventional edge-adaptive de-interlacing algorithms. This kind of techniques explores a neighborhood of the current pixel to extract information about the edge orientation [10]- [18]. Among them, Edge-based Line Average (ELA) algorithm interpolates the new pixel value by analyzing the luminance differences in the upper and lower lines. ELA looks for the most possible edge direction and then applies 'line average' along the selected direction. The pseudo-code of the ELA algorithm with 5+5 taps is as follows (see Fig. 5):

$$
\begin{aligned}
&if \quad min(a, b, c, a', c') = a \rightarrow X = (A + F)/2 \\
&elseif \quad min(a, b, c, a', c') = c \rightarrow X = (C + D)/2 \\
&elseif \quad min(a, b, c, a', c') = a' \rightarrow X = (A' + F')/2 \\
&elseif \quad min(a, b, c, a', c') = c' \rightarrow X = (C' + D')/2 \\
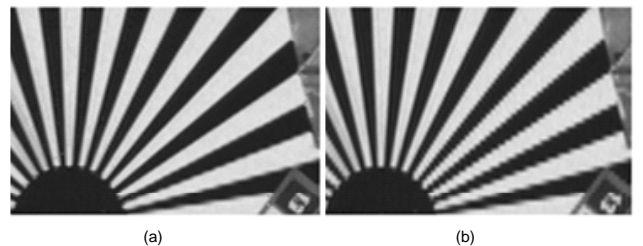&else \quad \rightarrow X = (B + E)/2
\end{aligned}
\tag{9}
$$

Fig. 6.　(a) ELA algorithm versus (b) line average.
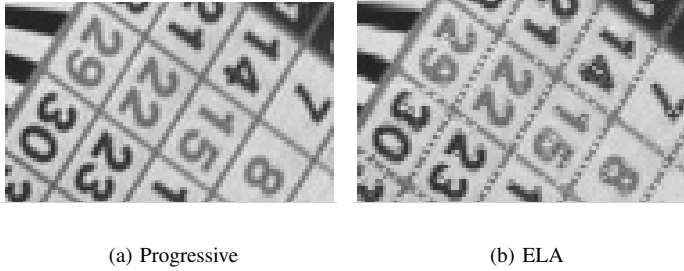
(a) Progressive   (b) ELA

Fig. 7. Progressive images versus ELA results in an area with details.

Let us illustrate the advantages of using an edge-adaptive algorithm by examining an image with a progressive picture with clear defined edges as shown in Fig. 6. The reconstruction of clear, sharp, and high-contrast edges are significantly improved by using a conventional edge-dependent algorithm such as ELA (Fig. 6(a)) versus a non-edge adaptive technique as line average that introduces an annoying stairs effect (Fig. 6(b)).

ELA performs well when the edge direction agrees with the maximum correlation, but otherwise introduces mistakes and degrades the image quality. These mistakes usually appear when the edges are not clear, the image is corrupted by noise, or there is a high number of details as shown in Fig. 7(b). In addition, ELA lacks the ability to detect horizontal edges.

Several proposals have been presented recently in the literature to avoid the above shortcomings of the ELA algorithm [11]- [18]. The approaches reported in [11] and [12] focus on enhancing the reconstruction of horizontal edges. Other algorithms use a larger neighborhood to get more information about the possible edge direction. For instance, the algorithm presented in [13] consists of a modified ELA module and a contrast enhancement module. The neighborhood is enlarged up to 6+6 taps in [14], 7+7 taps in [15], [16], 11+11 taps in [17], and 34+34 taps in [18]. A higher number of pixels in the aperture provides much information about edges but at expense of increasing the algorithm complexity.

The proposed fuzzy inference system is inspired by the ELA scheme but it uses a rule-based inference system to overcome ELA limitations. Thus, it has been named 'Improved Fuzzy-ELA 5+5' algorithm[3].

The inputs of this fuzzy system are the edge correlations in five directions as shown in Fig. 5. As the starting point to design the rules of the fuzzy system heuristic knowledge expressed linguistically has been applied, thus exploiting the ability of fuzzy logic to cope with symbolic knowledge. The following knowledge is employed to estimate correctly the edge direction (see Table II):

1) An edge is clear in direction a not only if a is small but also if b and c are large.
2) An edge is clear in direction c not only if c is small but also if a and b are large.

3) If a and c are very small and b is large, neither there is an edge nor vertical linear interpolation performs well; the best option is a linear interpolation between the neighbors with small differences: A, C, D, F.
4) If three antecedents are large, that is, no edge is found in the three directions (a, b, and c). An edge is clear in direction a', if a' is very small and c' is very large.
5) If three antecedents are large, that is, no edge is found in the three directions (a, b, and c). An edge is clear in direction c', if c' is very small and a' is very large.
6) Otherwise, a vertical linear interpolation would be the most adequate.

This heuristic knowledge is fuzzy since the concepts of 'small', 'large', and 'very small' are not understood as threshold values but as fuzzy ones. The concepts of *SMALL*, *LARGE*, *VERY_SMALL*, and *VERY_LARGE* are represented by fuzzy sets whose membership functions change continuously instead of abruptly between 0 and 1 membership values (μ), as shown in Fig. 4(e). The overlapping of the membership functions has been selected to ensure that no more than two rules are simultaneously activated. This strategy allows the use of the minimum operator as connective 'and' since a positive value for the activation degree of the sixth rule is always obtained. The particular breakpoints of the membership functions have been obtained after applying a tuning process following the design methodology that was explained in the previous subsection.

FM defuzzification method is used to calculate the output value. The output of this fuzzy system is the spatial interpolator, (I_S), which is used in the complete algorithm (see the left side of the complete block diagram in Fig. 1).

The first simulation results obtained with this algorithm offered few improvements in edges with inclination degrees in the directions a' and c'. The last two rules in Table II were frequently activated in cases of unclear edges. As result, the overall image quality was not as good as expected. To ensure the activation of these rules in situations where an edge is really placed in a' and c' directions, the following considerations should be fulfilled:

• An edge is not an isolated feature of the image, that is, an edge does not belong to an unique pixel in the image.

(a) Line Average   (b) Improved Fuzzy-ELA 5+5

Fig. 8. De-interlaced images for the Bicycle sequence in an area with clear edges.

---

[3]This algorithm is the result of a wide research work. Four different fuzzy systems for edge adaptation were analyzed. After extensive simulations the performance of this fuzzy system was better than the others.
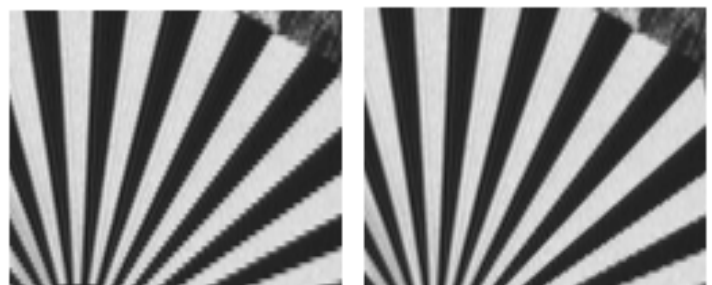
TABLE II
DESCRIPTION OF THE 'IMPROVED FUZZY-ELA 5+5' RULE SET

Rule	Antecedents	Consequent
1.	a is $SMALL$ and b is $LARGE$ and c is $LARGE$	$X = \frac{A+F}{2}$
2.	a is $LARGE$ and b is $LARGE$ and c is $SMALL$	$X = \frac{C+D}{2}$
3.	a is $VERY\_SMALL$ and b is $LARGE$ and c is $VERY\_SMALL$	$X = \frac{A+F+C+D}{4}$
4.	a' is $SMALL$ and a is $LARGE$ and b is $LARGE$ and c is $VERY\_LARGE$ and c' is $VERY\_LARGE$	$X = \frac{A'+F'}{2}$
5.	a' is $VERY\_LARGE$ and a is $VERY\_LARGE$ and b is $LARGE$ and c is $LARGE$ and c' is $SMALL$	$X = \frac{C'+D'}{2}$
6.	$Otherwise$	$X = \frac{B+E}{2}$

- Since a and c directions are close to a' and c', it seems logical that the rules to interpolate along these edge directions (the first and fourth rules for a and a', and the second and fifth rules for c and c') should frequently be activated simultaneously.

Fig. 8 and 9 show the images obtained for one snapshot of the illustrative sequence called 'Bicycle'. This video sequence contains a high number of edges that make evident the efficiency of the 'Improved Fuzzy-ELA 5+5' algorithm versus 'line average' (see Fig. 8). The superior performance of the edge-adaptive approach is noticed at naked eye. The comparison between both edge-adaptive proposals is established in Fig. 9. A zoom of a detailed area with unclear edges reveals ELA mistakes (white and black pixels). In general, conventional ELA algorithms are quite soft and spatially consistent however they make more mistakes. Since 'Improved Fuzzy-ELA 5+5' algorithm gives a much better visual impression, it has been selected as spatial interpolator.

C. Fuzzy logic-based system for picture-repetition adaptation

The simplest temporal interpolator is named 'field insertion', and it consists of repeating the pixel with same spatial coordinates in the previous picture. This technique introduces annoying artifacts such as feathering when the previous pixel is not a good option. This is especially noticeable in film sequences when the previous field does not correspond to the same but to a different frame, which has motivated an active research field on film-mode detectors.

Different detectors have been proposed in the literature to identify the field-pairs originated by the pull-down process from the same film image. The primitive approaches perform a global identification for the entire field, that is, a control signal is activated to denote the presence of film material. One of these detectors, called zero-vector matching detectors, try to match the zero motion vector on a previous field [19]. This approach is used by the majority of current film detectors. However, there are some drawbacks in the performance of these film detectors since the sum of absolute differences and its comparison with an unique threshold value is not suitable. Firstly, an unique sum does not properly measure the level of motion. A second disadvantage is the determination of the threshold value. Its selection is extremely difficult since the properties of the image are very variable.

Other approaches try to analyze the frame characteristics to find out the repetition pattern [20]- [23]. Among them, several proposals have been reported in the literature to identify jagged edges in frames due to 'feathering' effect [20], [21]. Other detectors study the location of edges in the frames [22]. The idea is that, if two frames are similar, edges should be at the same spatial position. Therefore, the analysis of edges position can reveal the picture repetition pattern. Finally, a motion vector based approach has been proposed in [23]. The sum of the length of the motion vectors should indicate if two fields are identical or not, since fields from the same frame should provide a similar value of the sum.

In recent literature, the research works in the area of film-detection can be categorized into two groups. A first one is focused on the increase of the robustness in the detection of pattern repetition [24], [25].

The second group of research works is focused on the development of film-mode detectors that work locally [26], [27]. This kind of algorithms is very demanded due to the high increase of TV material that combines images from different origins in a single field, which is known as hybrid material. None of the techniques previously cited can locally deal with hybrid material, because all of them detect a single mode for the entire field.

Since our interest is focused on de-interlacing any kind of material: video, film and hybrid (currently very popular due to the explosion of multimedia market), the idea developed is to adapt locally temporal interpolator to the presence of repeated areas in the fields as spatial interpolator was adapted locally to the presence of edges. In this sense, area repetition, as edges and motion, is considered, in general, a local (as happens to video and hybrid sequences) and fuzzy feature of the image and, hence, two simple fuzzy rules are proposed to implement a pixel-by-pixel fuzzy selection between pixels in the previous (t-1) and posterior (t+1) pictures. The input of

(a) ELA   (b) Improved Fuzzy-ELA 5+5

Fig. 9. De-interlaced images for the Bicycle sequence in an area with non-clear edges.

TABLE III

RULE BASE FOR THE TEMPORAL INTERPOLATOR

Rule	Antecedents	Consequent
1.	$dissimilarity(x, y, t)$ is $SMALL$	$I(x, y, t-1)$
2.	$dissimilarity(x, y, t)$ is $LARGE$	$I(x, y, t+1)$

these rules is a measure of dissimilarity between consecutive fields. Taking advantage of previous experience, dissimilarity between two consecutive fields is calculated by using a bi-dimensional convolution:

$$dissimilarity(x, y, t) = \frac{\Sigma\Sigma M_{(i,j)} C'_{(i,j)}}{\Sigma\Sigma C'_{(i,j)}} \quad (10)$$

where $M_{(i,j)}$ are the elements of the matrix, M, defined in (3) and $C'_{(i,j)}$ are the coefficients of the convolution mask[4]:

$$C' = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad (11)$$

The influence of $dissimilarity$ in selecting the kind of temporal interpolation is evaluated by considering the following fuzzy rules (see Table III):

1) If $dissimilarity$ between the fields $(t$-1$)$ and (t) is *SMALL*, the most adequate interpolated value is obtained by selecting the pixel value in the previous field at the same spatial position ($I(x, y, t-1)$).
2) On the contrary, if $dissimilarity$ is *LARGE*, the pixel value in the previous field is not a good choice and is better to bet on the pixel in the next field ($I(x, y, t+1)$).

The shape of membership functions to model the fuzzy concepts *SMALL* and *LARGE* are piece-wise linear functions whose parameters are fixed by using a tuning process (see Fig. 4(a)). The output of this fuzzy system is given by applying the Fuzzy Mean defuzzification method.

This fuzzy logic-temporal interpolator system considers $dissimilarity$ between only previous and current fields as a good trade-off between simplicity and efficiency. Anyway, the used methodology of soft computing would allow considering more inter-field dissimilarity measures.

III. PERFORMANCE OF THE ALGORITHM

Since recursive information is employed to measure motion and dissimilarity between consecutive fields, morphological operators (MO) have been introduced to improve the reliability of both measures. Two basic operators are applied: erosion and dilation. The aim of the erosion process is to reduce the high-frequency noise. After performing the erosion, the dilation process is applied. Its purpose is to extend the motion and dissimilarity areas based on the fact that both features are property of an area of the picture rather than a property of an isolated pixel. The effect of including MO is shown in

[4]The convolution mask (C') was selected after analyzing extensively the performance of 6 different convolution masks following the same procedure explained for *motion*. The mask in eq. (10) obtains the highest values in PSNR.
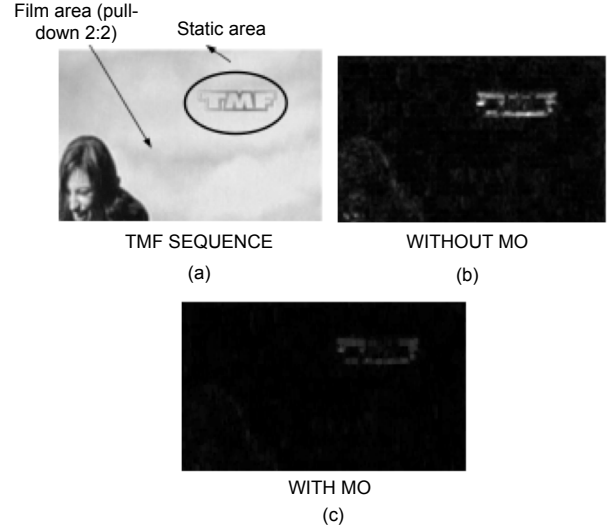

Fig. 10. Advantages of using morphological operators in the calculus of *dissimilarity* measure.

Fig. 10(a) that corresponds to a hybrid sequence. This part of the picture belongs to an interlaced field in which film area is similar to the previous field in the sequence, that is, *dissimilarity* measure should be small in film and static areas. As shown in Fig.10(c), the inclusion of MO reduces the false cases (in white) of *dissimilarity* measures in Fig. 10(b).

The performance of the complete algorithm has been proved by de-interlacing a battery of progressive standard sequences. They can be divided into two categories: a group of video sequences and a second one of real film sequences from TV channels or movies. These sequences have artificially been interlaced in order to measure the difference between the interpolated and original frames. The average Mean Squared Error is a very popular measure:

$$\overline{MSE} = \Sigma_t \frac{MSE(t)}{L} = \Sigma_t \frac{\frac{\Sigma_{x,y}(I_P(x,y,t)-I(x,y,t))^2}{MN}}{L} \quad (12)$$

where the sequences contains L frames with a resolution of MxN pixels, $I_P(x, y, t)$ is the value of the interpolated pixel, and $I(x, y, t)$ is the pixel of the original progressive image. Strongly related to the $\overline{MSE}$ is the $\overline{PSNR}$[5]:

$$\overline{PSNR} = \Sigma_t \frac{PSNR(t)}{L} = \Sigma_t \frac{20log\frac{255}{\sqrt{MSE(t)}}}{L} \quad (13)$$

The proposed algorithm has been also compared with other de-interlacing algorithms with less or similar computational

[5]A unique value of PSNR is not meaningful, but an improvement of 0.5 dBs in PSNR is quite perceptible by human visual system. Pixels that compose the border are excluded to calculate PSNR.

TABLE IV

COMPARISON WITH OTHER DE-INTERLACING ALGORITHMS

Sequence	News	Mother	Missa	Paris	Salesman	Phantom Menace	Die Another Day	Fire Rose	Fargo1	Fargo2	Tokyo	Body guard
Type of material	Video	Video	Video	Video	Video	Video	Video	Film	Film	Film	Film	Film
Line doubling	25.18	31.81	36.44	23.61	29.75	29.57	31.76	34.51	30.48	28.79	27.22	31.45
Line average	29.25	35.94	40.47	26.67	35.04	34.14	35.18	38.76	35.92	34.31	31.46	35.61
ELA 3+3 [2]	26.63	35.39	39.49	25.53	32.11	33.91	34.53	35.55	35.28	33.66	30.02	35.08
ELA 5+5 [2]	25.92	34.21	38.56	24.64	30.17	33.11	33.56	33.61	34.33	32.16	28.53	33.93
Field insertion	33.13	36.14	38.36	29.86	36.17	25.78	39.63	36.41	31.23	33.07	36.49	26.31
VT-2 fields [2]	35.46	39.61	40.25	30.73	36.54	32.11	39.51	40.32	35.87	40.99	36.84	37.88
VT-3 fields [2]	35.67	40.89	40.52	31.37	37.16	34.92	37.82	41.16	38.43	38.91	35.13	37.77
Median based-method [2]	33.51	38.49	39.44	30.27	36.61	32.46	41.44	39.45	35.32	39.89	34.92	38.5
Proposal in [28]	34.73	39.49	40.01	33.12	37.62	35.01	39.52	39.36	36.64	40.11	34.88	39.05
Proposal	39.91	42.37	40.82	36.91	38.94	34.92	42.76	41.45	38.96	43.65	38.17	41.11

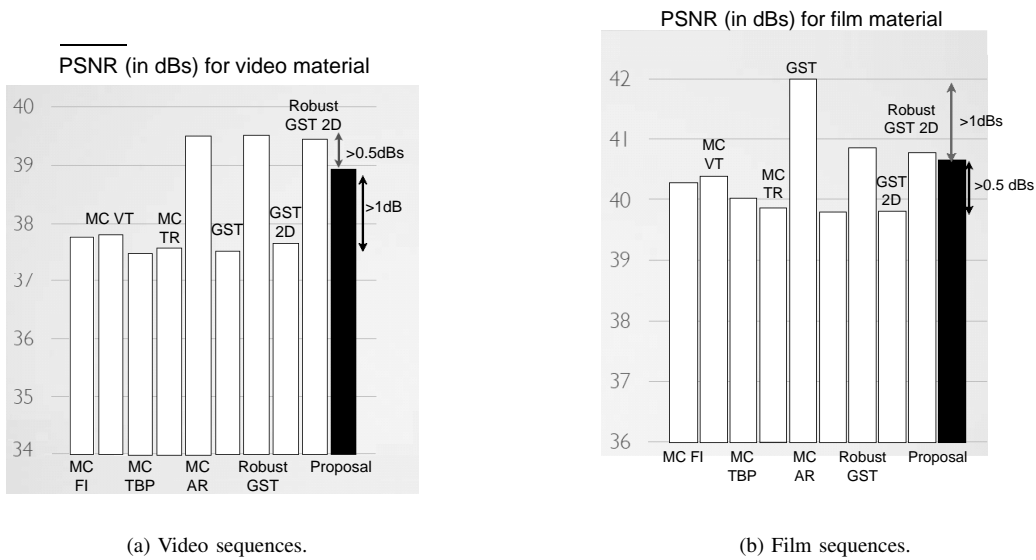(a) Video sequences.

(b) Film sequences.

Fig. 11. Average PSNR after de-interlacing all the sequences.

cost: four spatial method such as line doubling, line average, and conventional ELA (edge-adaptive interpolation algorithms [2]) 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 [2], a median-based method [2]; and, finally the fuzzy motion adaptive algorithms reported in [28]. As shown Table IV, the proposed algorithm performs better than the other algorithms since it achieves the highest values.

To show up the efficiency of using fuzzy sets to represent the antecedents of the rules, the three fuzzy systems have compared with similar ones that employ crisp definitions. The procedure to select the crisp values is illustrated in Fig. 12 for two concepts: SMALL and MEDIUM. The results of this crisp approach have been included in Table V. After performing this study, our conclusion is that the approach with fuzzy concepts always performs better than the crisp version of the algorithm.

TABLE V

AVERAGE PSNR IN dBs. FUZZY VERSUS CRISP ALGORITHM

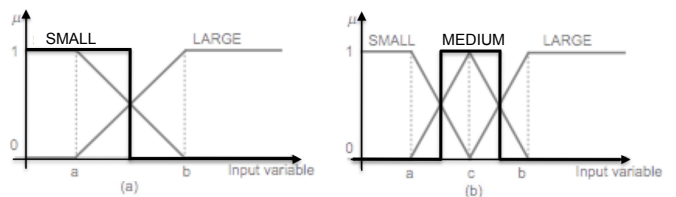| Sequence	Paris	Salesman	Phantom Menace	Fire Rose	Tokyo
Crisp	36.46	38.33	34.47	41.21	37.86
Fuzzy	36.91	38.94	34.92	41.45	38.17

Fig. 12. Crisp membership functions versus fuzzy ones.

Fig. 13.   Deinterlaced images obtained with MC de-interlacing algorithms and our proposal in the Paper 3D hybrid sequence.

A. Comparison with MC de-interlacing algorithms

The final proposal has demonstrated that it is clearly superior to other evaluated de-interlacing algorithms with less or similar computational cost. Consequently, its comparison with other kind of algorithms, the MC de-interlacing algorithms, can be interesting.

Fig. 11 shows the average PSNR results provided by nine MC de-interlacing algorithms [2][6]. All the MC methods use the 3-D RS matcher to calculate motion vectors without the inclusion of an explicit film detector. The parameters used by 3-D RS algorithm are the same in all the MC de-interlacing algorithms. Furthermore, a meandering processing of the pixels and a reverse scan of the fields is performed to improve the quality of this technique [2].

For video sequences (see Fig. 11(a)), the longest bars correspond to '*MC AR*' and '*Robust GST*' algorithms, which achieve the highest PSNR results in the majority of the sequences, whereas the bar of '*Robust GST-2D*' method is slightly smaller than these two methods. A very interesting result is that our final proposal is clearly superior (more than 1 dB) than a group of MC methods composed by the following algorithms: *MC field insertion*, *MC VT-filtering*, *MC TBP*, *MC TR*, *GST*, and *GST-2D*.

Fig. 11(b) shows the average PSNR results obtained after de-interlacing film sequences. The ranking of the methods is the same in video sequences and, again, our final approach achieves the fourth best result. A difference is that in film

---

[6]MC FI means *MC field insertion* algorithm, MC VT means *MC Vertico-temporal* algorithm, MC TR means *MC Time Recursive* algorithm, MC AR means *MC Adaptive-Recursive*, GST means *Generalized Sampling Theorem* [2].
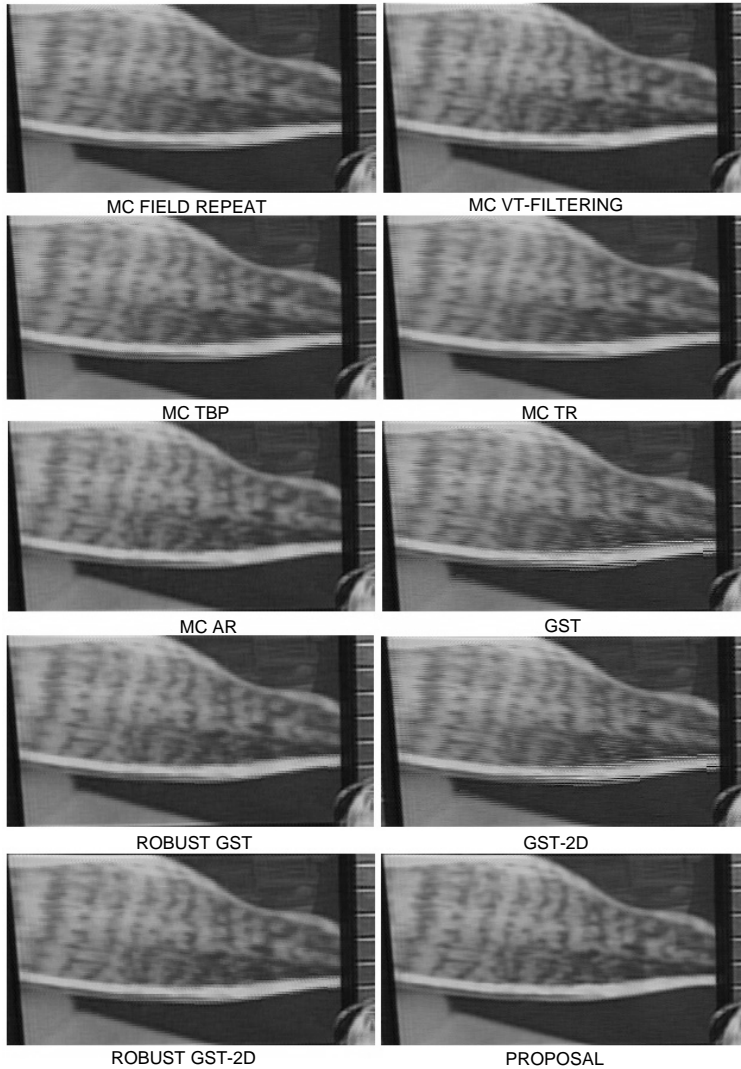
Fig. 14.   Deinterlaced images obtained with MC de-interlacing algorithms and our proposal in the Donna hybrid sequence (video area)

sequences the group of *MC field insertion*, *MC VT-filtering*, *MC TBP*, *MC TR*, *GST* and *GST-2D* are closer to our approach.

The performance of MC de-interlacing algorithms has also been analyzed by de-interlacing four hybrid sequences [7].

Fig. 13 shows a de-interlaced area of the Paper3D hybrid sequence. This sequence contains video material (the letters written on the paper) with a tricky background of film material. Among MC de-interlacing techniques, *MC VT-filtering* and *MC AR* achieve the best results. The rest of MC techniques can not appropriately reconstruct the video text and the ghosting effect is clearly visible. Our proposal even slightly improves the results of *MC VT-filtering* and *MC AR* as shown in Figure 13.

The hardware implementation of the complete algorithm has been carried out on a modern FPGA (Virtex-4 Xilinx). A demonstrator that has validated the hardware implementation of the algorithm has been performed by using a development board from Xilinx. This demonstrator is able to accept a

progressive video signal coming from the camera or the VGA input, interlace it, and send the de-interlaced signal to the VGA output in real-time.

IV. Conclusions

The following conclusions have been obtained after performing the research work:

- The algorithm is the result of combining three interpolators, each of them tackling a relevant feature for de-interlacing: motion, edges, and possible repetition of areas in fields.
- The feature of motion has been considered by a fuzzy motion-adaptive interpolator that uses a simple convolution to measure the motion at each pixel and a fuzzy inference system to evaluate how this measurement should influence on the interpolation decisions.
- The spatial interpolator is an edge-adaptive algorithm that uses five potential edge directions and six fuzzy rules to adapt the interpolation to the presence of edges. Extensive analysis let us conclude that its performance

---

[7]These hybrid sequences are not included in Table IV since their corresponding progressive sequences are not available.

Fig. 15.   Deinterlaced images obtained with MC de-interlacing algorithms and our proposal in the Donna hybrid sequence (film area).

(especially in terms of visual inspection) is better than other evaluated edge-dependent interpolators in both clear and unclear edges. Its performance for very low angles can be improved by increasing the processing window. The size of 5+5 taps was selected since it offers a good trade-off between complexity and performance.

- The third interpolator is a fuzzy area-repetition-dependent temporal interpolator that uses a simple convolution to measure the dissimilarity between consecutive fields and employs two simple fuzzy rules to adapt interpolation to repetition. Despite its simplicity, it reduces considerably annoying artifacts such as feathering.

- A tuning process of the values of the rule parameters has been successfully performed by using a supervised learning algorithm that minimizes the mean square error between a set of data corresponding to progressive and de-interlaced results of different sequences.

- The resulting algorithm, which uses the three interpolators and tuned parameters, not only performs better than other algorithms of similar complexity (such as vertical-temporal and median-based approaches) but also improves the results obtained by algorithms of much greater complexity (such as several motion- compensated ones). It offers better results (in terms of PSNR and visual inspection) in areas of the images with small and large

motion, with clear and unclear edges, and with film and video material mixed.

- A computer-aided design methodology has been employed to cross the way easily and rapidly from high-level algorithm development to synthesis. At algorithmic level, Matlab and its Image Processing Toolbox have been employed to develop the algorithms. Xfuzzy 3 and its $xfsl$ tool have facilitated tuning the parameters of the fuzzy rule bases proposed.

REFERENCES

[1] J. Whitaker, *Television transmissions systems*, chapter book of 'Standard handbook of video and television engineering'. McGraw-Hill Editorial, Blacklick OH (USA), 2002.

[2] G. de Haan and E. B. Bellers, *De-interlacing - an overview*. Proc. of the IEEE, vol. 86, no. 9, pp. 1839-1857, Sep. 1998.

[3] P. Haavisto, J. Juhola, Y. Neuvo, *Fractional frame rate up-conversion using weighted median filters*. IEEE Trans. on Consumer Electronics, vol.35, no.3, pp.272-278, 1989.

[4] P. Haavisto, Y. Neuvo, *Motion adaptive scan rate up-conversion*. Multi-dimensional Systems Signal Processing, no.3, pp.113-130, 1992.

[5] A. M. Bock, *Motion-adaptive standards conversion between formats of similar field rates*. Signal Processing: Image Communication, vol.6, no.3, pp.275-280, 1994.

[6] H. Jiang, D. Huu, E. Tinyork, *Motion adaptive deinterlacing*. United States Patent (US 6,459,455), Oct. 2002.

[7] W-K. Lin, C-Y. Lu, *Method for motion pixel detection*. United States Patent (US 7,034,888), Apr. 2006.

[8] W-K. Lin, C-Y. Lu, *Method for motion pixel detection with adaptive thresholds.* United States Patent (US 7,242,435), Jul. 2007.

[9] I. Baturone, F. J. Moreno-Velo, S. Sánchez-Solano, A. Barriga, P. Brox, A. Gersnoviez, M. Brox, *Using Xfuzzy environment for the whole design of fuzzy systems.* Proc. IEEE Int. Conference on Fuzzy Systems (FUZZ-IEEE), pp.517-522, London (United Kingdom), Jul. 2007.

[10] T. Doyle, M. Looymans, *Progressive scan conversion using edge information.* Signal Processing of HDTV, II. L. Chiariglione, ED., Elsevier Science Publishers, pp.711-721, 1990.

[11] M. H. Lee, J. H. Kim, J. S. Lee, K. K. Ryu, D. Song, *A new algorithm for interlaced to progressive scan conversion based on directional correlations and its IC design.* IEEE Trans. on Consumer Electronics, vol.40, no.2, pp.119-129, May 1994.

[12] C. J. Kuo, C. Liao, C. C. Lin, *Adaptive interpolation technique for scanning rate conversion.* IEEE Trans. on Circuits and Systems for Video Technology, vol.6, no.3, pp.317-321, Jun. 1996.

[13] H. Y. Lee, J. W. Park, T. M. Bae, S. U. Choi, Y. H. Ha, *Adaptive scan rate up-conversion system based on human visual characteristics.* IEEE Trans. on Consumer Electronics, vol.46, no.4, pp.999-1006, Nov. 2000.

[14] J. Salonen, S. Kalli, *Edge adaptive interpolation for scanning rate conversion.* Signal Processing of HDTV IV, Elsevier, pp.757-764, Nov. 1993.

[15] R. Simonetti, A. P. Filisan, S. Carrato, G. Ramponi, G. Sicuranza, *A deinterlacer for IQTV receivers and multimedia applications.* IEEE Trans. on Consumer Electronics, vol.39, no.3, pp.234-240, Aug. 1993.

[16] G. De Haan, R. Lodder, *De-interlacing of video data using motion vector and edge information.* Proc. IEEE Int. Conf. on Consumer Electronics (ICCE), Los Angeles (USA), Jun. 2002.

[17] Y. L. Chang, S. F. Lin, L. G. Chen, *Extended intelligent edge-based line average with its implementation and test method.* Proc. IEEE Int. Symposium on Circuits and Systems (ISCAS), vol.2, pp.341-344, Vancouver (Canada), May 2004.

[18] H. Yoo, J. Jeong, *Direction-oriented interpolation and its application to de-interlacing.* IEEE Trans. on Consumer Electronics, vol.48, no.4, pp.954-932, Nov. 2002.

[19] T.C. Lyon, J.J. Champbell, *Motion sequence pattern detector for video.* Asignee: Faroudja, Yves C., Los Altos Hills, CA. US, Jan. 1, 1991. United States Patent Office US 4,982,290.

[20] C. Correa, R. Schweer, *Film mode detection procedure and device.* Asignee: Deutsche Thomson-Brandt GMBH, Villingen-Schwennigen (DE), Jul. 1, 1998. European Patent Office, European Patent EP 0567072B1.

[21] H. Y. W. Lucas, *Progressive/interlace and redundant field detection for encoder.* Applicant: STMICRO Electronics Asia Pacific PTE LTD, Singapore, Jun. 8, 2000. World Intellectual Property Organization, International Publication Number: WO 00/33579.

[22] P. Swan, *System and method for reconstructing noninterlaced captured content for display on a progressive screen.* Assignee: ATE Technologies, Inc. Thornhill, Canada, Apr. 25, 2000. United States Patent Ofiice US 6,055,018.

[23] G. de Haan, H. Huijgen, P. Biezen, O. Ojo, *Method and apparatus for discriminating between movie film and non-movie film and generating a picture signal processing mode control signal.* Asignee: U.S. Philips Corporation, New York, USA, Nov. 15, 1994. United States Patent Office US 5,365,280.

[24] A. Dommisse, *Film detection for advanced scan rate converters.* M. Sc. Thesis, Technische Universiteit Eindhoven (TUE), Eindhoven, The Netherlands, Aug. 2002.

[25] C.-C. Ku, R.-K. Liang, *Robust layared film-mode 3:2 pulldown detection/correction.* IEEE Trans. on Consumer Electronics, vol.15, no.4, pp.1190-1193, Nov. 2004.

[26] G. de Haan, R. B. Wittebrood, *Recognizing film and video object occuring in parallel in sigle television signal fields.* Asignee: Koninklijke Philips Electronics N. V., Eindhoven, NL, Aug. 2005. United States Patent Office US 6,937,655.

[27] L. He, H. Zhang, *Motion object video on film detection and adaptive de-interlace method based on fuzzy logic.* Assignee: nDSP Corporation, Campbell,CA, Sep. 258, 2004. United States Patent Office US 6,799,168.

[28] D. Van de Ville W. Philips and I. Lemahieu, *Fuzzy-based motion detection and its application to de-interlacing.* Fuzzy techniques in image processing. Book Series of Studies in Fuzziness and Soft Computing, pp. 337-368, 2000.

Piedad Brox received the degree in Physics from the University of Córdoba in 2002, and the PhD. degree in Physics (with honors) from the University of Seville in 2009. Since 2002, she has been with the Instituto de Microelectrónica de Sevilla, CNM-CSIC. Currently, she is a Postgraduate Research Fellow. Her research areas are fuzzy image processing algorithms and their hardware implementation (starting from high-level descriptions to prototyping in FPGAs).

Iluminada Baturone received the Physics degree (with honors) and the Ph.D. degree in Physics (with honors) in 1991 and 1996, respectively, both from the University of Seville, Spain. Since 1990 she has been with the Instituto de Microelectrónica de Sevilla, CNM-CSIC. Currently she is an Associate Professor at the University of Seville. Her current research interests include hardware/software codesign, neuro and fuzzy systems, and applications in image processing, autonomous robots, and cripto-biometric systems.

Santiago Sánchez-Solano received the Physics degree (with honors) in 1980 and the Ph.D. degree in Physics in 1990, both from the University of Seville, Spain. After six years as a System Analyst at the Computer Center of the University of Seville, he joined the Instituto de Microelectrónica de Sevilla, CNM-CSIC, where he is currently a Senior Researcher. His main research interests include VLSI design, CAD tools for microelectronic design, and hardware implementation of neuro-fuzzy systems.