A NOVEL DE-INTERLACING TECHNIQUE BASED ON PHASE PLANE CORRELATION MOTION ESTIMATION

Mainak Biswas and Truong Nguyen

(mbiswas@ucsd.edu and nguyent@ece.ucsd.edu)

University of California, San Diego
9500, Gillman Drive
La Jolla, CA 92037, USA

ABSTRACT

In this paper, a de-interlacing technique using motion compensated interpolation is proposed. In the proposed scheme motion estimation is first performed between the same parity field i.e., reference and current field from either odd or even fields. Compared to other motion adaptive and motion compensated interpolation technique this method results in a better visual quality due to an accurate motion estimation technique. For further improvement a motion vector validation method based on motion parameters like pan, zoom or tilt is used to reduce the interpolation error caused due to the incorrect motion vector. Simulation results confirm that the proposed method has better visual performance compared to the conventional method.

1. INTRODUCTION

De-interlacing is a picture format conversion from interlaced to non-interlaced picture. De-interlacing converts each field from the interlaced picture into one frame. The process increases the vertical resolution per field by a factor of two. However the common TV signals do not satisfy the Nyquist criterion, one cannot rely on the linear sampling-rate conversion theory. Thus it is often fundamentally impossible to solve the de-interlacing problem under all circumstances.

The de-interlacing process is often complicated by the possibility that motion can change the scene contents from field to field. Relative motions between fields can be caused by movement of objects in the scene relative to the camera, or by camera changes such as pan, zoom or jitter. Existing de-interlacing methods can be generally classified into three basic categories: 1) Fixed interpolation techniques 2) motion adaptive techniques and 3) motion compensated techniques. Fixed interpolation techniques involve using data from a single field to produce a frame or merging two fields to form a frame. Motion adaptive methods use an adaptation to motion, that can be based on explicit motion detection or a non-linear median filtering operations. The motion compensated techniques use motion estimation and interpolate along the motion trajectory.

The fixed interpolation techniques include intra-field, inter-field or spatio-temporal interpolation. Intra-field interpolation is the simplest method where the vertical resolution is increased by line doubling. Disadvantage of this method is the vertical resolution gets limited to the line resolution of the field. The flicker or the alias artifacts can be improved by interpolating using the upper and lower lines. In inter-field interpolation the fields are merged without considering the time differences and thus artifacts are observed in motion areas. The spatio-temporal filtering doesn't degrade the image quality much in static or small motion area. However aliasing artifacts for the motion area exist because of incorrect temporal processing.

The motion adaptive methods were developed to solve the problems of fixed processing. The inter-field interpolation is used in static area that results in high resolution and flicker free images. The intra-field interpolation is used in motion area and the image quality remains the same as that of fixed intra-field interpolation. Often it is difficulty to detect a motion perfectly. Adaptive processing causes a change of resolution due to motion detection and switching of the de-interlacing methods.

In motion compensated de-interlacing methods, one first estimates motion trajectories and then performs filtering along them. Motion can either be modeled or estimated globally to compensate for the dominant motion only or can be estimated separately for each pixels by using different motion estimation algorithms available.

In this paper a robust motion estimation algorithm is proposed based on phase plane correlation technique. The motion compensated field interpolation technique interpolates the field composed of the missing lines in the current field. Then, the interpolated field is merged with the current field to generate the de-interlaced frames.

The paper is organized as follows: section 2 discusses the problems of motion estimation with the existing methods. The proposed de-interlacing method is described in section 3. The overview of the phase plane correlation method is described in section 4 followed by the algorithm in section 5. The simulation results are discussed in section 6 and the conclusion is drawn in section 7.

2. OVERVIEW OF THE EXISTING METHODS

In order to use motion compensated interpolation techniques, the motion vectors describing the true motion of the scene have to be available. The majority of the motion compensated de-interlacing systems use block matching motion estimator with added spatio-temporal smoothness constraint or initial prediction to reduce the operation count by recursion. This method works well with sequences where the motion is translational. The video sequences that we are interested are much more complex. The assumption of one motion vector describing all pixels in a 16x16 block does not hold true for most of the cases. These methods do not take into account multiple motions in a block resulting in annoying motion artifacts. For these types of motion estimators, the problem occurs in occlusion areas, i.e. at boundaries of moving objects. This is due to the fact that object boundaries do not coincide with the block boundaries used in the motion estimator. It is not possible to estimate motion in a block by matching it with the contents in a neighborhood for case where (un)covering occurs. To take into account of this problem some amount of post processing is needed with added field storage.

Another problem that the block based motion estimation methods other than [9] do not handle is the problem of false match from the regular structures. Simple DFD (displaced frame difference) calculation often results in mismatch in image type such as in Fig1. To avoid the false match some prior information is needed before starting the motion estimation. In this scope of paper we will present a novel motion estimation technique based on phase plane correlation to assuage all the above mentioned problems.

Fig1. A difficult video sequence that usually results in incorrect motion vectors due to the periodic structures.

3. PROPOSED MOTION COMPENSATED DE-INTERLACING

The block diagram of the proposed scheme is shown in Fig. 2. Let f_{n-1}, f_n and f_{n+1} denote the previous, current and next fields respectively.

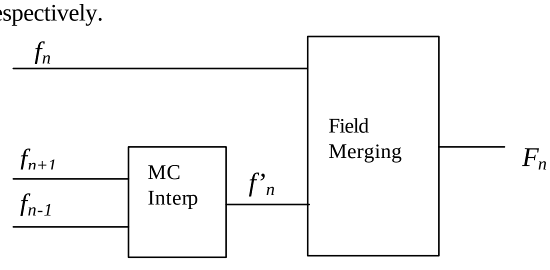

Fig. 2. The block diagram of the proposed algorithm.

The interpolated field f_n^i is obtained by the motion compensated field interpolation method using the fields f_{n-1} and f_{n+1}. Since f_{n-1} and f_{n+1} are the fields of the same parity the interpolation technique can be directly applied without any modification. The pixel values of the interpolated field f_n^i are determined using the straightforward motion compensated average given by

$$f_n^i(x) = \frac{1}{2}(f_{n-1}(x-v) + f_{n+1}(x+v)),$$

where v is the motion vector between f_{n-1} and f_n. The interpolated field f_n^i and the current field f_n are the fields of opposite parity with the same temporal position. Therefore the de-interlaced frame of f_n is easily obtained by merging these two fields. The scheme is composed of three processing units. First the phase plane correlation based motion estimation. Second, motion vector assignment based on camera parameters and finally field interpolation using the motion compensated averaging.

4. PHASE PLANE CORRELATION BASED MOTION ESTIMATION

Motion speeds in television images (those involve high action), often exceed 30 pixels per field. Thus a motion estimator for this application should ideally be able to compensate for motions well in excess of this speed. However the ability of eye, tracking the motion reduces with increased motion. Thus the accuracy is less critical in high speeds. The usual speeds that the standard video sequences have are close to 10-15 pixels per field. At this speed the motion vectors have to be accurate. Use of an incorrect motion vector in a temporal interpolation process can produce a major impairment that can easily outweigh the benefits gained by correct estimation of motion in other part of the image. Thus generation of "true" and "correct" motion vector is the key aim of this paper.
Use of phase correlation for accurate motion estimation was proposed in 70s and is used extensively by BBC R&D and "Snell and Wilcox" for their high quality de-interlacers and frame rate converters. The basic principle of the phase plane correlation approach can be found in [1, 5, 6, and 7]. The technique involves correlating two images by first performing a two dimensional Fourier transform on each image, multiplying together corresponding frequency components and performing inverse Fourier transform on the resulting phase differences. The result is a correlation surface that will have a peak at the coordinates corresponding to the shift between the two pictures.

The accuracy of the vectors can be as good as $1/10^{th}$ of a pixel using proper interpolator. ½ and ¼ pixel accuracy can be easily attained by interpolation using zero padding during FFT and IFFT. The advantage of the PPC technique over the block matching scheme is the ability to find local motion occurring within the block provided the moving objects are of significant size. The peak's location gives an estimate of the possible motions in the block. The fundamental strength of PPC is that it actually measures the direction and speed of the moving object rather than estimating, extrapolating or searching for them. Although PPC accurately measures the velocity, it cannot specify the position of the pixels having that motion vector. The important element of this system is the image correlation block that identifies the picture areas where the measured motion took place and establishes a level of confidence in the identification.

This involves shifting the image by the candidate motion vectors and evaluating their similarity with the reference frame. Similarities between pixel values indicate that an area with the measured motion has been found. The motion vectors that results in error below certain threshold are selected as motion vector for that part of the image. The vectors that do not result in a meaningful correlation with the reference image are identified as "spurious vector". The parts of the image that produces all spurious vectors are identified as "problem area". The problem can be caused due to difficult program material or due to occluding or appearing pixels. More processing is done on these blocks using information from the previous and the current images.

The regular structures discussed in section 2 can still result in partial correlations at incorrect distances that differ from the correct distance by the pitch of the structure. The effect of this is observed on the correlation surface as a large peak and will be flanked by smaller peaks at uniform spacing in a straight line. This problem can be solved as the motion information in the macro scale is known. The information from the initial resolution (24x24) is processed to estimate the motion parameters. The similar vectors in different window will show that there is a pan in the image. Geometrically varying mirrored motion vector will imply a zoom in the image sequence. In reality most of the sequences contain both pan and zoom in addition to rotation and tilting.

5. ALGORITHM FOR MOTION ESTIMATION

The frame is divided in macro blocks of size 24x24, this can be varied depending on the amount of motion in the sequence. The correlation surface is generated by doing the windowed two dimensional FFT and followed by IFFT. For simplicity the pixel accurate motion estimation technique is explained. The correlation surface is initially filtered to remove any noise. The filtered surface is then threshold and normalized to sharpen the peaks. The "peak hunting" algorithm then searches for "N" top peaks in the surface. The motion parameter estimation is done on all the major peaks that are the measure of the global motion in the blocks. The information from the parameter estimator block is then used in the image correlation block.

$\vec{X}_1$	$\vec{X}_2$	$\vec{X}_3$
$\vec{X}_4$	$\vec{X}_0$	$\vec{X}_5$
$\vec{X}_6$	$\vec{X}_7$	$\vec{X}_8$

Fig.3. The neighbourhood of the block of interest. The shaded ones are non-causal neighbours and are taken from previous frame

The image correlation block gets all the peak information from the block of current interest and the neighboring blocks as described in Fig.3. The non-causal block information is taken from the previous frame. It is assumed that the motion is linear for this short interval. For the 9 blocks there are 9N motion vectors that have to be evaluated. Due to the spatio-temporal smoothness, the unique motion vectors are much fewer than 9N. The pre-processing block finds the unique motion vectors and arranges them in order of significance. Motion vectors from the current frame have larger weights than from the previous frame. Moreover the "spurious" motion vectors value that were not used in the previous blocks that falls into neighborhood are not used. The search process thus refines itself as it progresses and thus reduces the total number of candidate vectors. The main block is then divided into smaller sub-blocks. The size can be chosen to be 2x2 or 4x4 depending on the difficulty of the video. Each of the blocks is then symmetrically motion compensated.

The significance of a candidate motion vector is calculated using the MSE and the smoothness factor as described in [4]. A threshold is used to identify the spurious vectors at this stage. The output of this block is the validated vector field that is fed to the motion compensated field interpolator. A vector interpolation stage is used that computes the position of the interpolated field between the input fields. The "non-problematic" parts of the image are interpolated using a spatio-temporal interpolator. The problem areas are then further processed. It is assumed that vectors converge in the area where the background is being occluded and diverge where it is revealed. With this information, the solution becomes trivial. For the occluding objects the motion vector is taken from the previous block lying in the direction of motion and then interpolated in the forward direction using the previous frame. For the revealing objects the motion vectors is taken from the next block in the direction of motion and then interpolated in the reverse direction. This result in moving objects being placed in the correct position.

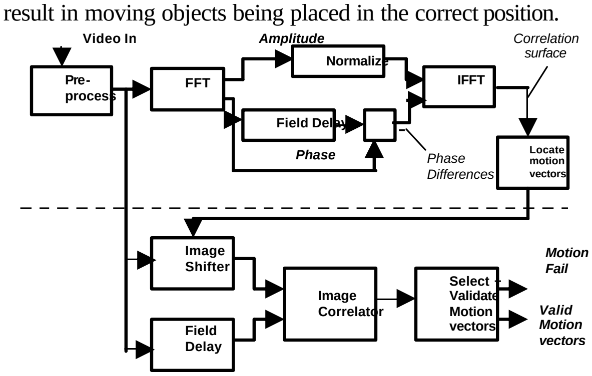

Fig.4. Block diagram of the phase plane motion estimation and motion validation method

6. SIMULATION RESULTS

In this section the performance of the proposed de-interlacing algorithm is compared with other existing algorithms on various complex motions.

6.1 Measurement System

For performance comparison, the progressive sequence is converted into interlaced sequence by eliminating even or odd lines alternatively and then the algorithm is applied to the synthetically made interlaced sequence as shown in Fig. 5. Vertical LPF (0-170 cph) were applied before sub sampling. The PSNR is measured for the interpolated error of the pixels in the missing line.

6.2 Test Sequences

Four sequences are used for performance comparison. There are sequences with fast panning motion over a flat background (Heli), slow zoom (Zoom), irregular motion with illumination changes (Lights) and small motion with low illumination (Apollo).

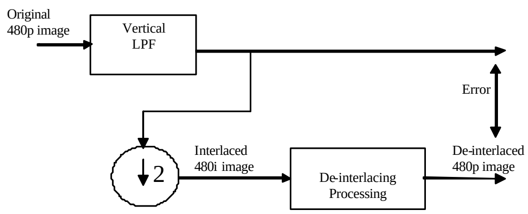

Fig.5. Performance comparison using synthetically made interlaced video sequence.

6.3 Comparison with conventional de-interlacing methods

The performance is compared with that of field repetition, field average, spatio-temporal filtering, edge based line average (ELA) , motion adaptive edge based line average and motion compensated interpolation using full search block match.

Sequence	Lights	Heli	Apollo	Zoom
Field Rpt	22.3641	20.3530	34.2984	23.9232
Field avg	25.7323	23.2776	37.6670	27.5768
V-T filt	25.6891	23.0077	36.8225	27.4367
ELA	25.4152	23.0545	35.8934	27.7225
MA-ELA	24.2721	23.6106	37.6740	26.2287
MCI	26.2295	23.8041	36.4724	28.6582
Proposed Method	**28.693**	**25.4355**	**39.1930**	**31.7422**

Table 1. Average PSNR performance of the discussed methods

It can be seen that the proposed de-interlacing scheme performs around 3-4 dB better than the existing de-interlacing methods. The PSNR value is not the true measurement of the image quality. Subjective picture quality is also important. The resolution, alias and blocking artifact on many images were checked. There was minimal blocking artifact in the processed video. Alias is much reduced but still remains in small amount at vertical motion area. The reader is urged to visit the following link to download and watch the de-interlaced video. http://videoprocessing.ucsd.edu/demo/deinterlacing. It is difficult to compare the subjective quality of the video from the printed images in this paper. Moreover rescaling the original 720x480 video images for thumb nailing introduces artifacts that were not present originally.

7. CONCLUSION

A motion compensated de-interlacing method using phase plane correlation technique was proposed. The problem of motion vector inaccuracy was solved by estimating the motion parameters from the motion vector data and using this to validate the computed motion vectors. A smooth motion vector switching scheme was proposed by comparing the current vector with its immediate neighbor by using statistical relevance weights to reduce the blocking artifacts.

The de-interlacing performance was measured by re-conversion of the synthetic interlaced video to progressive. From the simulation results and the demo videos it can be asserted that the proposed method can be used to generate high resolution progressive images without the blocking artifacts.

8. REFERENCES

[1] G.A Thomas, Television motion measurement for DATV and other applications. BBC RD 1987/11

[2] Phil Tudor, Progressive transmission of interlaced pictures. BBC Research & Development Technical Memorandum No. T-1231, August 1993.

[3] G.A Thomas, Distorting the time axis: Motion Compensated image processing in the studio. Broadcasting Convention, 1988. IBC 1988 pg 256-259.

[4] G. de Haan and E.B Bellers, De-interlacing: A Key technology for scan rate conversion. Advances in Image communication, Vol 9, Elsevier.

[5] V.G Devereux, Standards conversion between 1250/50 and 625/50 TV systems. Broadcasting Convention, 1992. IBC., International pg 51-55.

[6] Thomas, G.A, Application of motion compensation to video processing in the TV studio, IEE Colloquium on video processing in TV studio, 1991 pg 5/1-5/4.

[7] G.A Thomas, A comparison of motion compensated interlace to progressive conversion methods, BBC R&D 1996/9.

[8] S.Erturk and T.J Dennis, Image Sequence stabilisation based on DFT filtering, IEE Proc-Vis , Image Signal Process, Vol 147,No 2 , April 2000 pg 95 -102.

[9] G. de Haan , P.W.A.C Biezen, H. Hujigen and O.A Ojo, True motion estimation with 3D recursive search block matching motion estimation, IEEE Tr. On circuits and systems for video technology, vol 3, Oct 1993 pg 368-379.

[10] B.T Choi, S.H Lee and SJ Ko - New frame rate up-conversion using bi-directional motion estimation, IEEE Tr On Circuits and Systems for video technology, Vol 46,Issue 3 , Aug 2000 pg 603-609.

[11] Sugiyama and Nakamura, A method of de-interlacing with motion compensated interpolation, IEEE Tr on Consumer Electronics, Vol 45,No.3 , Aug 1999 pg 611-616.

[12] A.J Patti, M.I Sezan and A.M Tekalp, Robust methods for High-quality stills from interlaced video in the presence of dominant motion, IEEE Tr on Circuits and Systems for video technology, Vol 7 , No 2 , April 1997 pg 328-342.