

ENHANCEMENT OF VIDEO DATA USING MOTION-COMPENSATED POSTPROCESSING TECHNIQUES

Péter Csillag^{a)} and Lilla Böröczky^{b)}

^{a)} KFKI Research Institute for Measurement and Computing Techniques
P. O. Box 49, H-1525 Budapest, Hungary, Fax: +36-1-395-9027
<http://www.kfki.hu/~csillag>, E-mail: c-csillag@nov.mht.bme.hu

^{b)} IBM Corporation, Microelectronics Division
1701 North Street, Endicott, NY 13760, USA
E-mail: lboroczky@vnet.ibm.com

ABSTRACT

In many video coding schemes, especially at low bitrates, spatial and temporal subsampling of the image sequences is considered. This is realized by leaving out rows and columns from the images, and skipping whole frames at the transmitter. To get the best possible quality image sequence at the receiver side, the skipped portion of the video should be reconstructed using advanced motion-compensated (MC) postprocessing techniques.

Our paper mainly focuses on the restoration / generation of unknown frames of the sequence at time instances, where the original scene has not been sampled, or which were skipped from the original sequence in the transmitter. This enhancement of the temporal resolution is performed using our advanced MC interpolation algorithm, utilizing an accelerated motion model and motion-based segmentation with proper handling of covered and uncovered areas.

The algorithm can be used to avoid jerkiness and blurring of the restored image sequences.

2. Improvement of the spatial resolution of the sequence, interpolating gray (or color) values at fractional spatial positions both in the known (transmitted) and the unknown (interpolated) frames.
3. Removal of artifacts caused by subsampling, loss of data, distortion, etc.

Our paper mainly addresses the first problem.

Enhancement of the temporal resolution of an image sequence requires interpolation of image frames at time instances, where information was not transmitted, or has not been recorded. Good quality interpolation can only be done using advanced MC frame interpolation algorithms. In this paper an MC frame interpolation algorithm is presented, based on a pel-recursive motion estimation algorithm [1], and utilizing an accelerated motion model [4], motion-based segmentation [5] with proper handling of covered and uncovered areas [6].

The basic techniques used in our algorithm are described in the following sections.

1. INTRODUCTION

One area where motion-compensated video processing techniques can be successfully utilized is MC postprocessing of image sequences for the enhancement of video information.

Enhancement of the quality of video sequences includes:

1. Improvement of the temporal resolution of the sequence, by the insertion of MC interpolated frames, to avoid jerkiness caused by temporal subsampling of the original sequence, or inserting new frames into the sequence to decrease flickering, without introducing jerkiness.

2. MULTIREOLUTION PEL-RECURSIVE MOTION FIELD ESTIMATION

Our multiresolution motion estimation framework consists of three basic components: construction of multiresolution image representation, motion estimation and propagation strategy for refinement of the motion fields estimated in coarser resolution levels.

Multiresolution image decomposition is carried out using orthonormal wavelet transform. At each level of the pyramid a pel-recursive Wiener-based motion estimation [1] is performed. Refinement of the motion fields is done by a spatio-temporal prediction scheme [3].

The used multiresolution algorithm is much more robust, and results in significant improvement in comparison to the unilevel schemes, especially at large displacement vectors.

This work was supported in part by the E.C. Human Capital and Mobility grant "Motion analysis for advanced image communication systems" (MANADIX), Contract No: AC-ERBCIPDCT940033.

3. ACCELERATED MOTION MODEL

In our motion model constant acceleration was assumed for each moving pixel through 3 consecutive frames. The acceleration parameters can be determined based on two consecutive motion fields [4].

Frames at t_0-2 , t_0-1 and t_0 are fetched, and a motion field estimation is performed between frames at t_0-2 and t_0-1 , and between frames at t_0-1 and t_0 , resulting in two consecutive motion fields.

An accelerated motion trajectory can be described by a second order function, in the following way:

$$\mathbf{p}(t, \mathbf{s}) = \mathbf{s} + \mathbf{v}_0(\mathbf{s})(t - t_0) + \frac{\mathbf{a}(\mathbf{s})}{2}(t - t_0)^2 \quad (1)$$

where \mathbf{s} is the integer spatial position of the pixel at time t_0 , $\mathbf{v}_0(\mathbf{s})$ is the instantaneous velocity of that pixel at time t_0 and $\mathbf{a}(\mathbf{s})$ is the assumed constant acceleration of the same pixel over the time interval $[t_0-2, t_0]$.

To estimate the parameters of the motion (speed and acceleration), first the motion trajectory of each pixel has to be calculated. Each pixel of the frame at t_0 is tracked back to the frame at t_0-1 and frame at t_0-2 . The first step is easy, because if we consider a pixel in frame at t_0 at an integer spatial location \mathbf{s} , and the corresponding motion vector at this location is $\mathbf{d}(\mathbf{s})$, then the corresponding location in frame at t_0-1 is $\mathbf{s}-\mathbf{d}(\mathbf{s})$. However, while \mathbf{s} has integer coordinates, $\mathbf{d}(\mathbf{s})$ is a real vector, which means, that $\mathbf{s}-\mathbf{d}(\mathbf{s})$ points to a noninteger position, so we do not have a motion vector corresponding to this arbitrary location between frames at t_0-2 and t_0-1 . To overcome this problem a quadrolinear interpolation is performed on the previous motion field, which means, that the x and y components of the motion vector $\mathbf{d}_{pr}(\mathbf{s}-\mathbf{d}(\mathbf{s}))$ are calculated using bilinear interpolation from the motion field between frames at t_0-2 and t_0-1 , called \mathbf{d}_{pr} (previous). The positions of a moving pixel in frames at t_0-2 , t_0-1 and t_0 are $(\mathbf{s}-\mathbf{d}(\mathbf{s}))-\mathbf{d}_{pr}(\mathbf{s}-\mathbf{d}(\mathbf{s}))$, $\mathbf{s}-\mathbf{d}(\mathbf{s})$ and \mathbf{s} respectively.

From these three coordinates or two consecutive motion vectors of a moving point, final velocity and acceleration can easily be calculated using basic mechanics:

$$\mathbf{v}_0(\mathbf{s}) = \frac{3 * \mathbf{d}(\mathbf{s}) - \mathbf{d}_{pr}(\mathbf{s} - \mathbf{d}(\mathbf{s}))}{2} \quad (2)$$

$$\mathbf{a}(\mathbf{s}) = \mathbf{d}(\mathbf{s}) - \mathbf{d}_{pr}(\mathbf{s} - \mathbf{d}(\mathbf{s})) \quad (3)$$

Acceleration and final velocity are calculated at each pixel, forming an acceleration and a velocity vectorfield.

4. MOTION-BASED SEGMENTATION

To ensure good quality interpolation, especially at the object boundaries, the image sequence has to be segmented, in order to find and appropriately trace the moving objects.

An iterative algorithm has been developed for motion-based segmentation. It first determines a set of global motion vectors present in the image sequence from the estimated motion field by means of a vector quantisation procedure. Then segments the images of the sequence into moving rigid objects using these motion vectors.

The algorithm consists of the following steps.

1. First a motion estimation is carried out for the entire image. Two consecutive frames are taken, and a multiresolution pel-recursive Wiener-based motion estimation algorithm is performed out between them.
2. An LBG vector quantisation is carried out on the vectorfield, determining the number of global motion vectors as well as their coordinates the following way:
 - 2.1. The average motion vector is determined.
 - 2.2. This vector is split into two vectors, and an LBG iteration is carried out. The iteration consists of minimum two steps, and is stopped, when the decrease of the minimized error parameter is less than 0.5%.
 - 2.3. The vector, that is responsible for the largest part of the error parameter is split, and an LBG iteration is carried out as described in step 2.2
3. The whole algorithm is finished, when the number of vectors reaches a predefined maximum, or when the addition of a new motion vector decreases the error parameter by less than 15 %. If the second condition is true, this last vector (and the whole last iteration) is ignored, and the results of the previous iteration (giving a set of vectors less by one) is taken.
4. The images are segmented based on these motion vectors. There are different possibilities for this segmentation.
 - a) The estimated vectorfield can be quantised using the quantisation vectors determined by step 3. This method gives inaccurate boundaries, because the pel-recursive motion estimation algorithm requires a few pixels, to converge to a new motion vector after crossing object boundaries.
 - b) The first image of the image pair used for motion estimation can be displaced by the quantisation vectors, and the segmentation can be carried out comparing the displaced frame differences (dfd) corresponding to the vectors.

This method gives correct boundaries, but results in a very noisy segmentation mask, even using a 3x3 window for matching the two images. The use of a 5x5 window for comparing the images gives satisfying result.

- c) The two above approaches can be combined using linear combination of the errors in the quantisation process. The coefficients for this purpose are determined experimentally.
- d) An advanced nonlinear combination of the above methods can be constructed, that gives larger priority to the vectorfield (method a)) inside the objects, and to the dfd (method b)) near the object boundaries.

In our experiments method c) was used.

5. The results of step 4. are filtered by a 5x5 mode filter, and the small regions are eliminated in a similar way to the method used by Thoma and Bierling [2], but using other thresholds and changing the color of each small region to the color of its neighbour it has the largest common border with.

5. THE INTERPOLATION FILTER

Using the result of the segmentation algorithm and the acceleration and velocity fields, the unknown frames are reconstructed by an MC interpolation filter algorithm. It is based on the assumption that the pixels in the frames to be interpolated are moving with the calculated constant acceleration along a motion trajectory (which is actually a parabola in the image plane) from their corresponding pixels in the previous transmitted frame to their corresponding pixels in the consecutive transmitted frame.

Considering the pixels belonging to the moving area in the unknown frames, at first the spatial positions of these pixels are determined by using the corresponding pixel in frame at t_0 , (the frame, where the final velocity was calculated) and the calculated acceleration and final speed by equation (1).

The spatial position is calculated as the nearest integer value of the appropriate displaced position. This position is assumed to have the same motion parameters as the exact fractional one, and from this sampling grid position using the same estimated displacement vectors, the spatial positions of the two corresponding pixels in the two transmitted frames are calculated.

For the pixels of covered and uncovered parts of the background and other objects in the interpolated frames the above presented algorithm is applied with some modifications described in section 6.

6. HANDLING OF COVERED AND UNCOVERED AREAS

In order to properly treat covered and uncovered areas of the background and the objects, the spatial order of the objects (which one is in front of the other) has to be determined [6]. Based on motion information, this can only be done by examining three consecutive frames, and two consecutive motion vector fields. Thus a second motion estimation has to be carried out between the 2nd and a 3rd frames.

If an uncovered area appears between two moving objects, then it is examined on the third frame. If this area moves together with Object 1 between the 2nd and 3rd frames, then Object 2 is in front of Object 1, if it moves with the same vector as Object 2, then Object 1 is in front of Object 2.

The construction of a frame based on the results of the segmentation, and the spatial order of the objects is shown in Fig. 1.

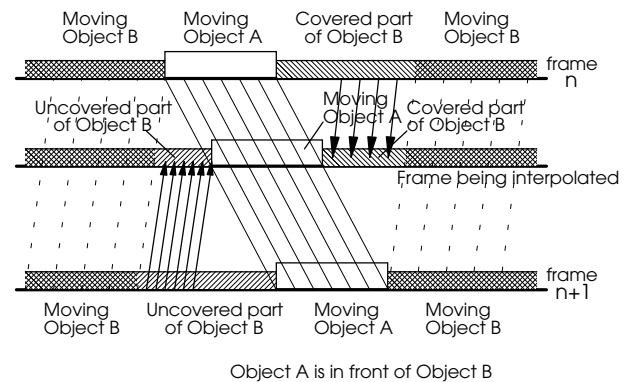


Fig. 1. Interpolation with proper handling of covered and uncovered areas of objects

7. EXPERIMENTAL RESULTS

In order to test the performance of our algorithms, experiments were carried out for artificially generated and real life image sequences. The used video sequences ranged from artificially moved images through simple videophone and video conferencing images like "Miss America" and "Trevor White" to complicated sequences including the well-known "Calendar and Train" (caltrain) sequence.

In order to ensure the objective evaluation of the proposed algorithms, spatial and temporal subsampling of the sequences was performed before the experiments, and the interpolated pixels and frames were compared to the skipped ones.

7.1. Motion estimation

An example of a motion field is shown in Fig. 2.



Fig. 2. Motion field obtained for frames 0 and 3 of sequence "Calendar and Train" (caltrain)

7.2. Accelerated motion

The incorporation of acceleration into the motion model resulted typically in a gain of 0.5 to 1 dB in the signal to noise ratio of the interpolated images, over the traditional constant speed model.

The SNR of the accelerated algorithm is compared to the constant speed one in Fig. 3.

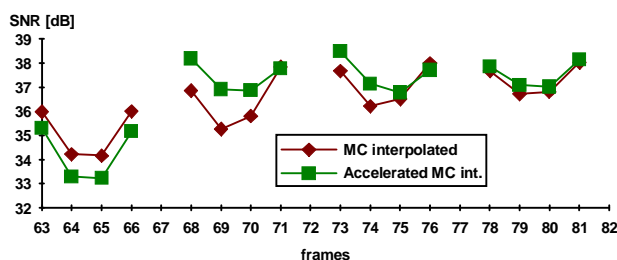


Fig. 3. Comparison of the SNR of the two algorithms for image sequence "Miss America", skipping 4 frames

7.3. Motion-based segmentation



Fig. 4. Segmentation of frames 0 and 3 of caltrain, combined method (4.b) after eliminating the small regions

An example of the results of segmentation according to the algorithm described in section 4 is shown in Fig. 4.

The proper handling of the covered and uncovered regions especially improved the visual quality of the reconstructed sequence.

8. CONCLUSION

Applications of the described techniques include:

- enhancement of low bitrate video sequences in the receivers of video telephony/conferencing equipment
- in consumer electronics: 100Hz and/or progressive scan display of traditional 25/50 Hz interlaced television broadcast in high quality receivers
- in special applications, like transmission of sign language over low bitrate video channels, etc.

Results of our simulations proved, that all the above described postprocessing algorithms are applicable and advantageous in enhancement of the quality of the video material. The algorithms should be combined into one interpolation framework, to achieve the best results.

9. REFERENCES

- [1] J. Biemond, L. Looijenga, D. E. Boeke, R. H. J. M. Plompen, "A Pel-Recursive Wiener-Based Displacement Estimation Algorithm", *Signal Proc.*, Vol. 13, No. 4, 1987, pp. 399-412.
- [2] R. Thoma, M. Bierling, "Motion Compensated Interpolation Considering Covered and Uncovered Background", *Signal Proc.: Image Comm.*, Vol. 1, No. 2, Oct, 1989, pp. 191-212.
- [3] P. Csillag, L. Böröczky, K. Fazekas, "Motion-Compensated Frame Interpolation for Low Bitrate Video Coding Based on Wavelet Decomposition", *Proceedings of the International Picture Coding Symposium PCS '94*, Sacramento, California, Sept. 21-23, 1994. pp. 6-8.
- [4] P. Csillag, L. Böröczky, K. Fazekas, "Motion-compensated Frame Interpolation for Low Bitrate Video Coding Considering Accelerated Motion", *The European Symposium on Advanced Networks and Services, Conference on Compression Technologies and Standards for Image and Video Communication*, 20-23 March, 1995, Amsterdam, The Netherlands, pp. 130-138.
- [5] P. Csillag, L. Böröczky: "MC Frame Interpolation Applying Motion-Based Segmentation and an Accelerated Motion Model", *Proceedings of the International Picture Coding Symposium PCS'96*, Melbourne, Australia, March. 13-15. 1996. pp. 547-552.
- [6] P. Csillag, L. Böröczky: "Frame Rate Conversion Based on Acceleration and Motion-Based Segmentation" *Proceedings of the European Symposium on Advanced Imaging and Network Technologies, Conference on Digital Compression Technologies and Systems for Video Communications*, 7-9 October, 1996, Berlin, pp. 438-448.