Parallel Predictive Motion Estimation using Object Recognition Methods

Holger Blume, Aishy Amer

Universität Dortmund, Lehrstuhl für Nachrichtentechnik, Tel.: +49 231 755 4535
AG Schaltungen der Informationsverarbeitung Fax: +49 231 755 3196
44221 Dortmund

1.) Introduction

Motion estimation is one of the key techniques in modern digital video signal processing. High quality vector fields are as well needed for vector based upconversion as for motion compensated coding or noise reduction [7]. Most of the estimation algorithms which are used today and which are implemented in dedicated hardware are block based algorithms (block matching algorithms). Although they are easy to implement and the quality of their resulting vector fields is better than that of the methods which were used formerly (gradient methods, phase plane correlation) they still suffer from some drawbacks. Most important is that using block based algorithms a block pattern will arise in the vector field (block pattern noise).

Real objects in real scenes do not coincide with block boundaries. Therefore object recognition methods shall be introduced into motion estimation. Within this work an object oriented motion estimation algorithm is presented which combines motion information and object information and yields nevertheless a good computational performance.

2.) The new object based algorithm

Within this algorithm object recognition is carried out by a morphological segmentation of the image followed by a contour analysis of the objects and a contour filling. This object information is then used to enhance the quality of the vector fields which are generated by a motion estimator. The basic idea of the new object oriented motion estimation algorithm is depicted in Fig. 1.

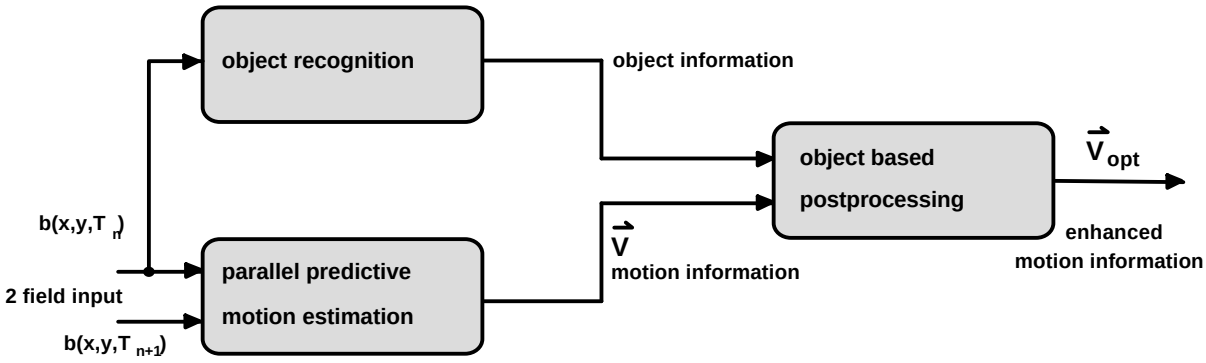


Figure 1: Block diagram of the new object oriented algorithm

Motion estimation based on regions coinciding with real objects leads to very good results, as motion is usually coherent inside each region in case of translatory motion [4]. Therefore many proposals have been made for object matching methods which carry out motion vector based clustering or segment matching starting from initial segments and then enhancing the vector field by iterative hierarchical steps. Within each of these steps the matching segments are refined [3]. Common to all of these methods is a huge amount of calculations because of the

several refinement steps. Therefore within this concept object recognition is taken through in parallel to the motion estimation. This object information is further used to enhance the quality of the vector fields by a homogenisation process within the object regions in a not hierarchical but regular way without raising the calculational costs too high.

Another prerequisition for this algorithm was ***robustness***. In case of non translatory motion (e.g. camera zoom, rotation, coverage and uncoverage) object oriented algorithms often yield worse results than standard motion estimators. Because of this non translatory motion is detected here by an analysis of the vector histogram in object regions. In such a case the homogenisation process is stopped and the original vector field of the motion estimator is kept.

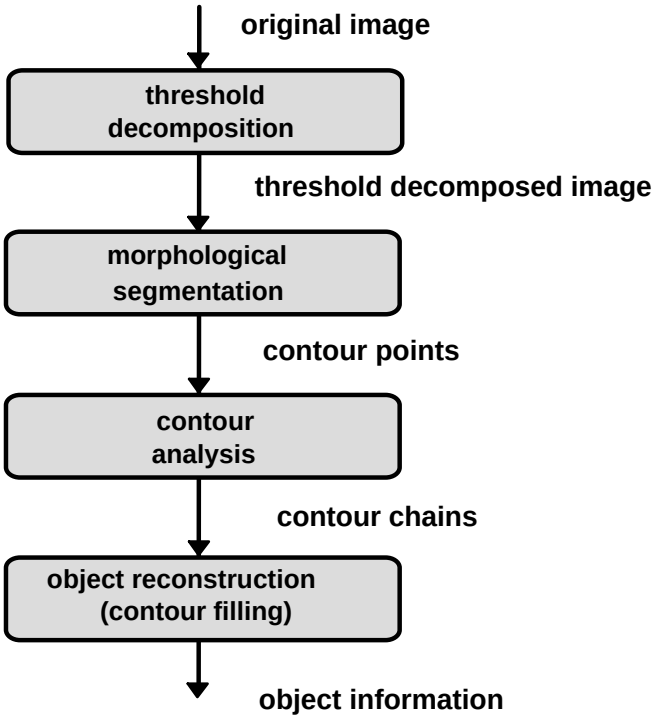
The incoming data stream is processed in two different ways:

Motion estimation

First of all an optimised ***block matching*** technique is applied (parallel predictive motion estimation) which is suitable for a VLSI relization [2]. The calculational cost of this method is about forty times lower than that of a standard Full-Search method. Further the number of input fields is reduced to two (only one field store required) as the implementation costs for digital video signal processing strongly depends on the number of field stores which are needed. Applying a raster adaptive reinterpolation to the two incoming fields this block matching method is able to calculate the motion between two successive fields.

Object recognition

The second processing block of Fig. 1 the ***object recognition*** is carried out by a luminance based segmentation of the image. This concept is further refined in Figure 2.



After a ***threshold decomposition*** of the original image the decomposed grey scale image is segmented by morphological operators. Then contour chains are generated which further will serve as homogenisation boundaries for the post processing step. The single elements shall be described now in more detail:

Today ***morphological segmentation*** is widely used for image segmentation and these image segments are applied as well for image coding as for motion estimation [4, 5, 6].

Figure 2: Object recognition algorithm

Within this concept the morphological segmentation is carried out in order to generate the contour points following the relation:

$$C(S) = (S\ XOR\ (S \oplus E_1)) \ominus E_2 \quad (1)$$

with

$C(S)$ the calculated contour point image
S the threshold decomposed image
E_1 disc element for dilation
E_2 square element for erosion
$\oplus$ dilation symbol
$\ominus$ erosion symbol

Morphological segmentation has been applied as it could be shown that this yields segmentation contours which are more appropriate for the demands of TV-systems. This means:

- stability concerning noise influence
- by choice of the structuring element the size of the resulting segments can be adjusted
- most regular implementations for morphological operators are available

The segmented image is further processed by a ***contour analysis*** which generates contour chains of the contour points which were produced by the morphological segmentation. Within this processing step contours which are smaller than a given size are eliminated. This is due to the fact that no vector field homogenisation is necessary within very small moving objects which are not important for the perception of an image.

Within the ***object reconstruction*** step the interior of each region which is left after the contour analysis is marked by a constant grey level value. Therefore each of these contours is filled by an unique grey level which performs as a registration sign for each object in the final object based post processing step.

Object based Post processing

The ***object based post processing*** (third processing step in Fig. 1) can be subdivided in three main elements

- homogenisation of background motion
- completion of object points
- homogenisation of object motion

which are described in more detail in [1].

As a result of this improved motion estimation higher interpolation quality is achieved for a vector based upconversion. Applying these vector fields for upconversion moving textured picture content isn't disrupted any more as it was the case using standard non homogeneous vector fields. Especially the convergence behaviour of the predictive motion estimation algorithm could be enhanced by the object information as this information is calculated field by field and so even after a scene cut when the convergence process of the motion estimator just begins a homogeneous vector field is available.

3.) Calculational costs

The following table depicts the calculational costs for the object recognition and the object based post processing in comparison to the calculation time required for the block-matching.

The data given in the table is measured in seconds by the clock command of "C" on a SUN4 machine with a SPARC2 processor. The data is given for four different sequences:

seq 1: "BBC-car" seq 2: "TV-speaker"
seq 3: "prlcar" seq 4: artificial test sequence
by calculation of one single standard TV-field (720 * 288).

Algorithm	seq 1	seq 2	seq 3	seq 4
morphological segmentation (incl. thresh. decomp.)	7.599	7.066	5.283	6.616
contour analysis	1.166	1.433	0.899	0.849
object reconstruction	2.449	4.366	1.449	2.549
object based post processing	3.683	3.499	3.149	3.799
total amount for object based processing (OP)	15.082	16.516	10.966	13.813
parallel predictive block-matching (PPBM)	208.841	210.141	208.641	212.058
relation OP / PPBM	1 : 13	1 : 12	1 : 19	1 : 15

Table 1: Calculational Costs of the algorithm elements for several sequences

As can be seen in the table the calculational cost for the object based processing is about 1/14 of the calculational cost of the block-matching. The PPBM itself has a complexity which is about forty times lower than that of a Full-Search algorithm.

Because of these reasons this new algorithm seems to be propitious for a realisation as it improves the quality of the vector fields and raises the calculational costs of the whole concept only by a small amount. Further very regular operations (morphological operators) are applied which can easily be implemented.

Acknowledgment

The authors would like to thank the DFG (Deutsche Forschungsgemeinschaft) for their support funding the work carried out under this subject.

References

[1] Amer, A. :"Bewegungsschätzung mit Methoden der Objekterkennung", Diplomarbeit D3-94, Lehrstuhl für Nachrichtentechnik, Universität Dortmund, 1994

[2] Blume, H. :"Bewegungsschätzung in Videosignalen mit parallelen örtlich zeitlichen Prädiktoren", Beitrag zum 5. Dortmunder Fernsehseminar 29.9.-1.10.1993, Dortmund

[3] Guse, W. :"Objektorientierte Bewegungsanalyse in Bildfolgen", Fortschritt-Berichte VDI, Nr. 223, VDI-Verlag, 1992

[4] Pardas, M.; Salembier, P. :" 3D morphological segmentation and motion estimation for image sequences", Signal Processing, Vol. 38, 1994

[5] Pei, S. C.; Fei-Chin, C. :"3D spatiotemporal subband decomposition for hierarchical compatible video coding by mathematical morphology", Image Com. Vol. 6, 1994

[6] Salembier, P. :"Morphological multiscale segmentation for image coding", Signal Processing Vol. 38, 1994

[7] Schröder, H. :"Image Processing for TV-Receiver Applications" IEE International Conference on Image Processing and its applications, Maastricht, April 1992