AN EFFICIENT MATCHING APPROACH TO MOTION ANALYSIS OF IMAGES

Kostas Girtis*, Theodore Lilas** and Stefanos Kollias**

*Department of Informatics, University of Piraeus, Karaoli & Dimitriou 80, 18534 Piraeus, Greece.
e-mail: girtis@unipi.gr
** Computer Science Division, National Technical University of Athens, Zografou 15773, Greece.
e-mail:stefanos@cs.ntua.gr

ABSTRACT

Local information in not always enough for efficient motion analysis. Additive processing is required to get accurate results. This has been formulated as the aperture problem. Block matching algorithms are applied between successive images for motion estimation, assuming conservation of local intensity distribution. Matching approaches provide good results when the aperture problem does not exist. However, in regions when the aperture problem exists, additional constraints are required in order to recover the displacement of pixels between consecutive images.
In this paper we present a way to improve the performance of optical flow computation at the first early level. Morphological filters are introduced in the matching approach with which we overcome inherent problems of correlation based techniques.

1. INTRODUCTION

Recently morphological segmentation and motion analysis methods [1][2] have been applied to 3-D image sequence motion analysis. Such methods include four different steps: a simplification step, a marker or feature extraction step which detects presence of homogenous regions, a decision step which defines the contours of regions using a watershed algorithm, and a motion modeling one. Mathematical Morphology is used for segmentation purposes, because it deals with geometric features, such as size, shape, contrast or connectivity which can be considered segmentation oriented features. The original image sequences are considered as functions defined in 3-D space; segmentation is performed in this space and the produced regions are connected components which can be related with moving objects. This implies a complete knowledge about the position and shape of every segmented object of the scene in every time section. The performance of such methods is very good since they are based on precise estimation of the region contours; however they generally require an intensive computational load.
The problem of optic-flow estimation remains challenging, despite the progress achieved in computational vision. The goal is to derive the projection of 3-D velocities onto the imaging surface from image intensity. In most applications accurate and dense velocity maps are required, and this is usually hard to achieve.
The methods for computing optical flow can be grouped in differential methods[3], region-based matching[4][5], energy-based[6] and phase-based techniques[7]. Usually these techniques perform an early processing such as prefiltering, extraction of initial measurements based on locally available velocity information and integration of local measurements; where assumptions about the smoothness of the flow field or the underlying scene geometry are used. The last phase is required in order to overcome the inadequacy of local information, which has been formulated as the aperture problem. In the next we describe the proposed algorithm and in section 3 present some simulation results.

2. THE PROPOSED ALGORITHM

2.1 Existing Approach

In order to overcome the aperture problem in the above mentioned approaches, additional constraints are required.
Usually a smoothness constraint can be applied to adjacent pixels, since these pixels most probably belong to a rigid object and therefore have similar velocities. Applying a smoothness constrain uniformly causes the following problems: Motion boundaries are blurred and velocity is propagated to uniform areas where little information is available. An approach to overcome the problem is to filter the image edges and then not allow any smoothing over the edges. However this does not allow proper computation of optic flow at intensity edges that are not motion boundaries. Directional smoothness operators have been proposed to overcome the above problems.
An alternative approach which does not require any assumptions about the smoothness of the optical flow field, is based on the underlying geometry of the scene. If we consider a rigid object moving with a translational and a rotational velocity against an observer, then using a projective transformation we can compute the image velocity at each point. Assuming that locally points are planar, then if we know the optical flow at enough points, the equations computing the true velocity can be solved [8]. This approach has given very good results on synthetic images; however its performance on real images depends on the high level image segmentation which is required to avoid considering

neighbouring points, i.e., points that lie on different surfaces.

2.2 The Selective Matching Approach

The technique proposed in this paper is similar to matching based techniques, but provides an advantage over correlation methods by performing non-linear processing using morphological filters. The procedure that we use is the : For each pixel on one image a search is performed to find a matching pixel in the following image using an hierarchical multiresolution approach. Multiresolution representations are very effective in analysing the information content of images. In image processing the above technique is implemented using subband processing.

Image decomposition is performed with a filter bank of decimating QMF (quadrature mirror) filters. An appropriate bank of reconstruction QMF filters guarantee perfect reconstruction of the original image from its sub-band components. The decimating filter bank can successively be used in the first subband component resulting in a multiresolution image representation.

The way that the subband images are processed depends on the particular application. In our case we need a lower resolution image (e.g. the first subband component) which concentrates most of the signal energy of the higher resolution. In this case, we used the perfect reconstruction filter banks introduced in [8] in order to construct a pyramid for the image.

Matching is performed going from a coarse to a finer level. It is assumed that the intensity distribution around a pixel is kept at the different resolution levels and therefore, a correspondence can be established at all levels. Additionally a physical constraint is imposed on the maximum displacement that can take place between two successive images, constraining in this way the search area. The best match in the search window is considered the one that maximizes a similarity measure. The most commonly used measures, which we also consider, are correlation, sum of squared differences, mean and variance normalized correlation.

Following the above and depending on the morphology of the region, the matching algorithm either finds a best match or several candidates with similar matching measure. In the first case the displacement vector is accurately computed. This occurs in distinct regions like corners. The second case occurs in areas where the intensity gradient vectors are strongly oriented (near edges). There the true displacement vector cannot be accurately computed and only the component along the normal to the underlying intensity edge can be identified.

Rank order morphological filters have been used as non linear filters in order to improve the performance of region matching approaches. If we consider a discrete argument signal f(x), $x \in Z^v$ and a finite window $W \subseteq Z^v$ with $|W| =$ n points, then the k-th rank order of signal f w.r.t. W (k=1,2,..., n) $RO_\kappa(f;W)$ (x) is defined as the k-th largest value of f(x+y), $y \in W$. Some special cases seem interesting; if k=1, then the k-th rank order of a signal is the dilation $f \oplus \breve{W}$; if k=n then the operation equals the erosion f with W and if k=(n+1) /2 and n is odd then it becomes the median filter med(f;W).

In the proposed approach we have used as a similarity measure the sum of RO_κ (f;W) (x) where f(x) is the absolute difference of pixel intensities between the two images, W is the matching window and k lies in the range (1+l, n-m); n is the number of elements of W, while l and m depend on the size of W and the amount of noise in the image.

Rank order operators are non linear, translation invariant and increasing. Using rank order filters optical flow estimation is improved under noise, occlusion disclosure and motion boundaries. The underlying principle is that this technique can match any pattern that has moved provided that this pattern is partially visible. Therefore, we can significantly increase the size of the matching window without degrading performance, due to the fact that part of the window has moved and part of it has not.

Motion edges are usually intensity edges too. Although intensity edges provide good estimates of motion, in this case arises a problem that only part of the matching window has moved. In the case that the background is dull, the problem is eliminated. However, in general, this causes problems in estimating optical flow and as a result motion edges are blurred. Additionally, if we increase the size of the matching window in order to improve our confidence the error due to this problem becomes more significant. Therefore, the similarity measure used in other region matching techniques is not suitable.

3. EXPERIMENTS

Experiments have been conducted on synthetic and real images. Synthetic images have been generated by overlaying a pattern over a background image. A robotic system has been used for real life images. The robotic system provides accurate control over its motion, and therefore, enables us to derive accurate results.

The robotic system has nine degrees of freedom and consists of a six axes puma type robot and three external translational axes forming a Cartesian coordinate system along which the base of the puma type robot can be translated. The system is installed in a ship repair yard where the lighting conditions are not controlled and are far from ideal (figures1, 2). In the case presented here, the robot performs translational motion at a constant speed and the camera is placed ten meters away from the robotic arm. The comparison is done with the region matching approach, using the sum of square differences as a measure of similarity.

In the first example a part of the image of the robotic system is provided to the two algorithms. There is not any motion between the successive images. However we have added noise on both images which leads to errors in motion estimation which should have been zero at all points (figures 5,6). We have added two types of noise i) gaussian and ii) noise which causes some pixels to turn to white or black.

In the second example the algorithms are applied on two successive images of the moving robotic system. There we have examined different sizes of matching windows. However due to limited space only one comparison is given (figure 3,5).

In the third example we have used a combination of the two algorithms. At first we find the contours of the objects in the original frame using morphological methods, we process the picture 8x8 blockwise and we use the above described similarity measures depending on the existence of a contour in each block. We have examined different criteria for testing the existence of contours in the blocks. In the first case, using blocks of 8x8 pixels we examine if a contour of more than 4 pixels exists in the block. In this case we perform the selective matching approach , otherwise the similarity measure is the sum of absolute differences. So in 370 blocks of the picture the similarity measure is the sum of ROκ(f;W)(x) and in the rest 910 is the sum of absolute differences. In the second case where existence of contour requires more than 8/64 pixels, the number of blocks is 144 and 1136 respectively, while in the third case where the existence of contour is more than 16/64 pixels the blocks are 11 and 1269. Motion vectors derived from the second case are shown in figure 7.

4. CONCLUSIONS

Experiments have shown that the proposed approach provides more accurate results compared to other region matching approaches. Rank-order filters provide non linear processing resulting in better response of the algorithm under noise. Additionally, the optimum level is achieved when larger windows are used; this effect increases the matching confidence, as we overcome the aperture problem. However, the proposed algorithm is computationally more intensive than correlation matching approach and requires additional tuning. So in the third experiment we used a combination of the two algorithms and we have taken better results. In regions where there are not long contours we use as similarity measure the sum of absolute differences which is computationally less intensive. This method gives accurate results in less time compared to approach where the similarity measure is the sum of ROκ(f;W) (x) for the total number of blocks.

REFERENCES

[1] P. Salembier and J. Serra, 'Morphological Multiscale Image Segmentation' , Proc. of Visual Communication and Image Processing, Boston 1992.

[2] P. Salembier, ' Morphological Multiscale Segmentation for Image Coding' Signal Processing (Special Issue on Nonlinear Signal Processing), vol.38, no.3, pp. 359-386, 1994.

[3] H. Nagel, 'Displacement vectors derived from second order intensity variations in image sequences', Computer Vision, Graphics and Image Processing, 21:85-117, 1983.

[4] P.J. Burt, C. Yen, and X. Xu, 'Multi-resolution flow-through motion analysis.' In Proc. IEEE CVPR Conf., pp. 246-252, 1983.

[5] P. Anandan. 'Measuring Visual Motion from Image Sequences.' PhD thesis, COINS Dept, Univ. of Massachusetts, Amherst, 1987.

[6] S.P. Liou and R. Jain, 'Motion detection in spatiotemporal space', Computer Vision, Graphics and Image Processing, 45:227-250, 1989.

[7] M. Ziegler. 'Hierarchical motion estimation using the phase correlation method in 140mb/s HDTV coding.' In Proc. Third Int' l Workshop on HDTV, Torino, Italy, 1989.

[8] A. Tirakis, A. Delopoulos and S. Kollias, '2-D Filter Bank Design For Optimal Reconstruction Using Limited Subband Information', IEEE Image Processing, vol.4, no.8, pp 1160-1165, Aug 1995.

figure 1

figure 2


figure 3

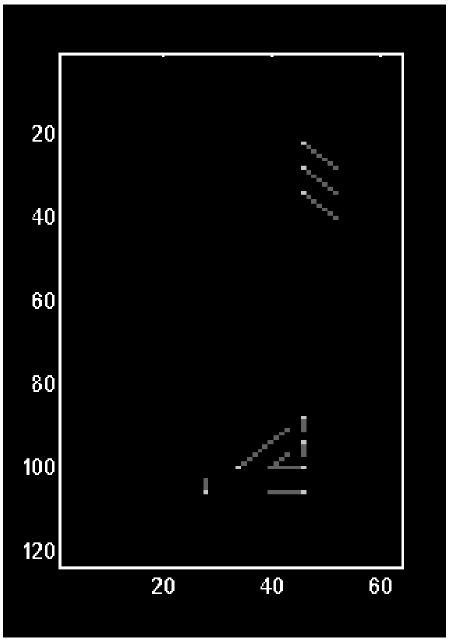

figure 4


Figure 5

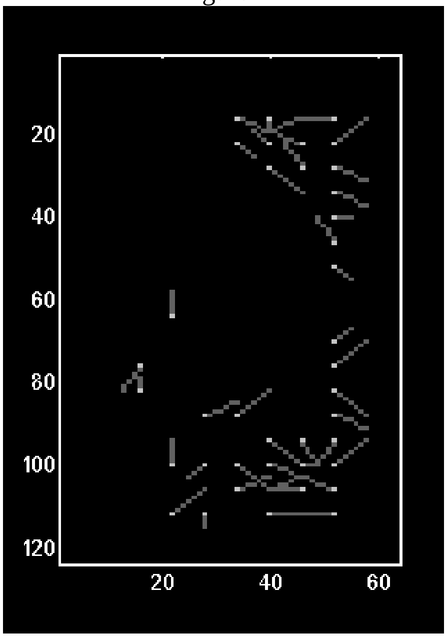

Figure 6


figure 7