WINDOW-BASED IMAGE REGISTRATION USING VARIABLE WINDOW SIZES

Andreas Krutz, Michael Frater**, and Thomas Sikora**

*Commun. Systems Group
TU Berlin
Berlin, Germany

**School of IT and EE
University of New South Wales
Canberra, Australia

ABSTRACT

We present an unsupervised image registration algorithm to estimate the background object motion in a real video sequence. The algorithm is based on a Gaussian minimisation technique. It has been shown earlier that initialization of such an approach is very important to achieve the motion parameters of the background object precisely, and that the use of a windowing technique can give better background object motion estimation results, even with large background occlusions. In some cases, however, the fixed window size initializes the gradient descent algorithm in a sub-optimal way. Here, another window size would bring the desired estimation direction. In this paper, we present a technique where variable window sizes are used to prevent these outliers. Experimental results show that the technique works very well with the considered test sequences.

***Index Terms*—** Image Registration, Global Motion Estimation, Image Warping.

1. INTRODUCTION

The estimation of global motion parameters between consecutive images in video sequences is a fundamental technique that is used in applications such as sprite generation, camera calibration, and video analysis with their further applications. The high quality of a sprite generated by accurately computed motion parameters improves the performance of the applications that rely on sprites, for example, object segmentation and sprite coding. Several work has been done in this field. The fundamentals of global motion estimation and sprite generation and its applications have been proposed in [1], [2], [3]. The window-based approach given in [4] is based on the work proposed in [3] and some improvements from [5] and [2]. The key technique is a gradient-based energy minimization method. The well-known perspective motion model is used for an accurate estimation of the camera motion. Reliable performance of the gradient descent algorithm relies on the initialization of the motion parameters. It has been shown that operating on image pyramids brings better performance and lower computational complexity to the whole algorithm [3],[5]. Phase correlation is often used to initialize the translational motion parameters [1], [4].

The difference between the approach proposed in [4] and recent work is the use of windowing. On each pair of the input images windowing is performed at the coarsest level of the pyramid, and the phase correlation method and a gradient descent algorithm are then applied. The coordinates of the best window match are taken for further calculation during the upper stages of the pyramid. The goal of this technique is to initialize the gradient descent algorithm at the finest stage so that it calculates the motion of the background object exactly without distortion by any foreground objects. This approach often works with a static window size, and provides good results with the test sequences. However, in some cases outliers occur during the calculation over a whole test sequence, where the choice of an inappropriate window size leads to poor initialization of the algorithm. This means that the gradient descent does not converge to the desired minimum corresponding to the background motion. To prevent this, a method is presented in this paper where the window size can be adapted online during the frame-to-frame calculation of the motion parameters of a video sequence. The description of this idea embedded in the recently proposed algorithm is organized as follows. In the next section the windowed image registration algorithm is introduced in more detail. The automatic window size method and the whole updated image registration are presented in Section 3. Experimental results and the evaluation are shown in Section 4. Conclusions finalize the presentation of this method.

2. FRAMEWORK OF IMAGE REGISTRATION ALGORITHMS

The approach for the development of an image registration algorithm is described next in detail. For the core technique the Gaussian energy minimization method is taken because it is widely used in the field of motion estimation using higher-order motion models. It is so-called state-of-the-art [6]. Several additional techniques are needed for an improved perfor-

This work was developed within 3DTV (FP6-PLT-511568-3DTV), a European Network of Excellence funded under the European Commission IST FP6 programme.

M. R. Frater was supported in this work by the Australian Research Council under project DP0667074

mance of the whole algorithm. It has been shown in [6] that the Gauss-Newton approach performs very well if the start point is close to the desired minimum. This underlines the importance of the initialization. First, translational motion parameters have to be pre-estimated. Phase correlation is chosen for that because it is fast and robust. The achieved parameters are then included in the higher-order motion parameters (**m**) and the Gaussian minimization is applied. The second issue that brings better performance is the use of image pyramids. The input images are first separated in lower resolutions; three are most common. The algorithm then starts at the coarsest level and the achieved motion parameters initialize these in the upper lever until the original size is achieved. Using this the algorithm has lower computational cost and the error surface where the gradient descent algorithm operates is smaller and the minimum desired is easier to find. The whole algorithm is illustrated in Fig. 1.
There are some drawbacks, e.g. trapping in a not-desired minimum, which have to be improved. The main problem is the

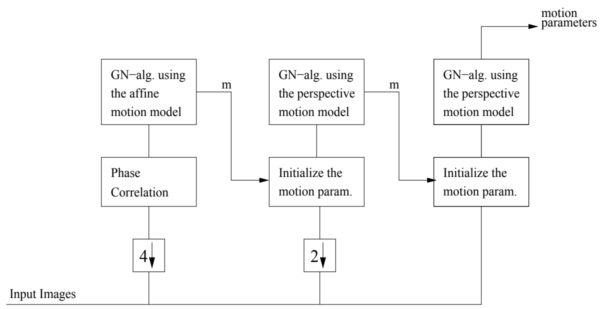


Fig. 1. Core image registration algorithm (inspired by [3])

influence of the motion of foreground objects. Due to the operation on the whole image size it is possible that the algorithm in Fig. 1 can converge with any of the minima that occur from several moving objects. An approach to address this problem is the use of a windowing technique. It has been shown in [4] that, using windowing, the gradient descent can be forced in one direction exactly. In practice, windowing is applied at the coarsest level of the pyramid. The phase correlation method and the gradient-based error minimization method using the affine motion model is applied on each window pair. The best window match is taken for further calculation. The window size grows through the pyramid with respect to the current image size, and the gradient descent algorithm using the perspective motion model is used on the considered window pair until the original size of the input images is achieved. The input images are then up-sampled using the 7-tap wavelet filter [5]. The last step of the algorithm is the calculation of the motion parameters on the whole up-sampled input images using motion parameters of the window pair as initialization. Up-sampling is used to prevent aliasing in the warping process [7]. The window-based image registration algorithm is pictured in Fig. 2. Experimental results have shown the improvement of this approach in comparison with recent algorithms with the considered test sequences [4]. A drawback is the use of a fixed window size because outliers can be occur due to an inappropriate choice of the window size.

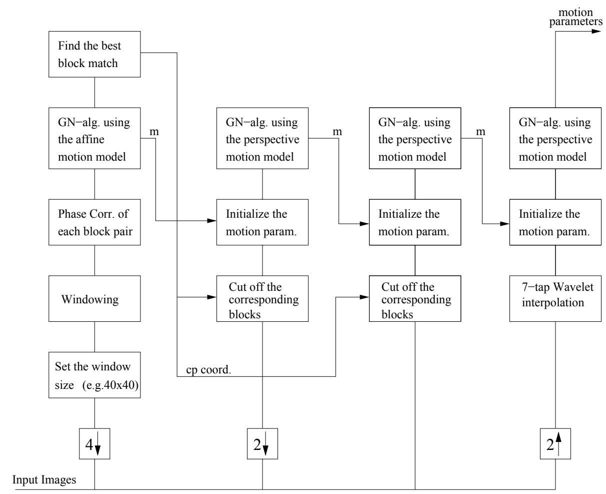


Fig. 2. Recently proposed image registration algorithm [4]

3. UPDATED IMAGE REGISTRATION ALGORITHM USING AUTOMATIC WINDOW SIZES

The failure of the initialization can be prevented if variable window sizes are used. Two sizes (32x32 and $40x40$ pixel)are used for the first approach. The algorithm starts with one window size and the task is to detect the appearance of an outlier online during the calculation over a sequence. The $RMSE$-value (Root Mean Square Error) of the current image pair is calculated at the third stage of the algorithm, at the original size of the input images. If an outlier occurs this $RMSE$-value increases significantly in comparison to the previous value. This can be emphasized if the first numerical derivative of the $RMSE$-values is considered. The calculation is not that difficult because only the $RMSE$-value of the previous image pair has to be stored. The differential $RMSE$ is then calculated as :

$$RMSE_{diff} = RMSE_{curr} - RMSE_{prev},$$

A threshold defines if a peak in the differential $RMSE$ appears and the algorithm goes back and start with the second window size for the current image pair. After the calculation of the motion parameters the window size is changed to the previous size for the next input image pair. The threshold can be calculated using the variance of the differential $RMSE$:

$$T = \frac{1}{N} \sum_{k=1}^{N} (e_k - \mu)^2,$$

where e_k holds the diff. $RMSE$-value at the point k, μ is the average, and N is the current number of the differential

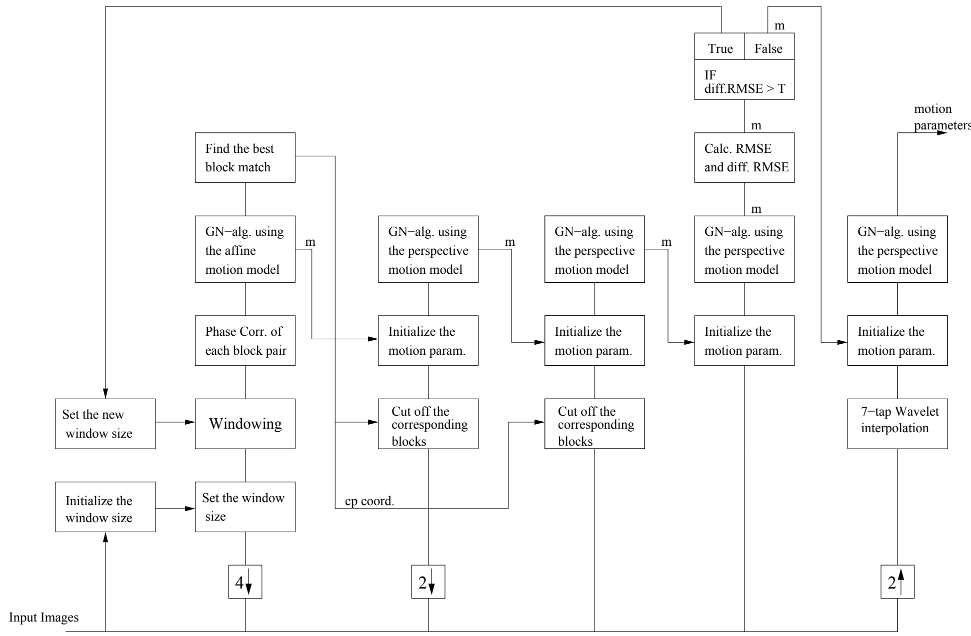
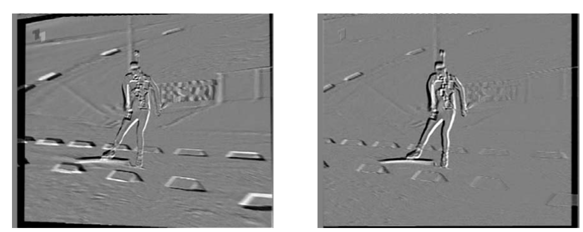


Fig. 3. Updated image registration algorithm

$RMSE$-values. The definition of the start window size is calculated very easy. At the beginning the first two frames of the sequence are taken and both sizes are applied. The size that produces the lower $RMSE$ is taken for further calculation. This can be extended when more than on example of the test sequence are considered. The technique described is now integrated in the present windowed image registration algorithm. The modification at the beginning and at the third level of the pyramid brings the new updated algorithm which can be seen in Fig. 3. It has been found experimentally that the aligned windowing technique works at its best at the third stage of the algorithm. The computational complexity does not increase so much because calculating the $RMSE$-value using initialized motion parameters takes only a few more steps. It can be seen in the next section that only a few outliers occur and the re-computation during the algorithm at the outlier-case using a new window size has to be accomplished rarely. This technique using only two different window sizes works very well with the considered test sequences as shown in the next section.

4. EXPERIMENTAL RESULTS

The updated windowed image registration algorithm is compared with the recently proposed algorithm using fixed window sizes by three test sequences chosen. The sequences are the well-known “Stefan” and “Mobile&Calendar” and a new sport sequence “Biathlon” with several moving foreground objects and difficult textures. (The “Biathlon”-sequence is taken from the German first TV-station.) For a comprehensive evaluation the recently proposed algorithm and the updated version are compared using a frame-to-frame motion parameter calculation. Figure 4 shows an example for registration improvement at an outlier case. The $PSNR$-, $RMSE$-, and differential $RMSE$-value are computed for each input image pair. The curves are illustrated in Fig.5 - 7 for each of the three test sequences. It can be seen that at each sequence several outliers occur when a fixed window size is used. The new algorithm using adaptive window sizes intercepts the outliers. Table 1 shows mean $PSNR$-values for the outlier cases. The updated algorithm significantly improves the performance of the algorithm at this stage. Table 2 shows average $PSNR$-values comparing the proposed and recent algorithms. The average $PSNR$-value of the proposed algorithm increases up to about 2.7 dB in comparison to recently proposed algorithms.



(a) Error frame (outlier with fixed window size)

(b) Error frame (window size changed)

Fig. 4. Comparison of a fixed window size and variable window size at the outlier case (frame 94, “Biathlon” sequence)

5. CONCLUSION

A significant improvement of a window-based image registration algorithm has been shown. Instead of fixed window sizes a technique has been presented where the window size

Table 1. Mean PSNR values at the oulier cases

Avg. PSNR in dB	fix window size (40x40)	changed window size
“Stefan”	25.32	29.36
“Mobile”	19.80	31.57
“Biathlon”	24.56	29.94

Table 2. Mean PSNR values of the compared algorithms

Avg. PSNR in dB	Core Alg.	Smolic Smolic	Proposed (40x40)	Proposed (arbitrary)
“Stefan”	28.08	28.49	29.24	29.52
“Mobile”	28.60	-	31.18	31.37
“Biathlon”	28.44	-	28.89	29.08

can be change online during the motion parameter calculation over a test sequence. Two window sizes are considered. Further work can be done in extension to more than two window sizes for an accurate analysis of the background motion even with difficult test material. The algorithm sets up the accurate motion parameters by applications such as mosaicing and camera calibration.

6. REFERENCES

[1] R. Szeliski, “Video mosaics for virtual environment,” *IEEE Computer Graphics Appl.*, pp. 22–30, 1996.

[2] A. Smolic and J.-R. Ohm, “Robust global motion estimation using a simplified m-estimator approach,” in *Int. Conf. on Image Processing (ICIP00)*, Vancover, Canada, 2000.

[3] F. Dufaux and J. Konrad, “Efficient, robust, and fast global motion estimation for video coding,” *IEEE Trans. Image Process.*, vol. 9, pp. 497–501, Mar. 2000.

[4] A. Krutz, M. Frater, M. Kunter, and T. Sikora, “Windowed image registration for robust mosaicing of scenes with large background occlusions,” in *Int. Conf. on Image Processing (ICIP06)*, Atlanta, USA, Oct. 2006.

[5] G. Ye, M. Pickering, M. Frater, and J. Arnold, “A robust approach to super-resolution sprite generation,” in *Int. Conf. on Image Processing (ICIP05)*, Genova, Italy, Sept. 2005.

[6] S. Baker and I. Matthews, “Lucas-kanade 20 years on: A unifying framework,” *International Journal of Computer Vision*, vol. 56, pp. 497–501, Feb. 2004.

[7] A. Krutz, M. Frater, and T. Sikora, “Improved image registration using the up-sampled domain,” in *Int. Conf. on Multimedia Signal Processing (MMSP06)*, Victoria, Canada, Oct. 2006.

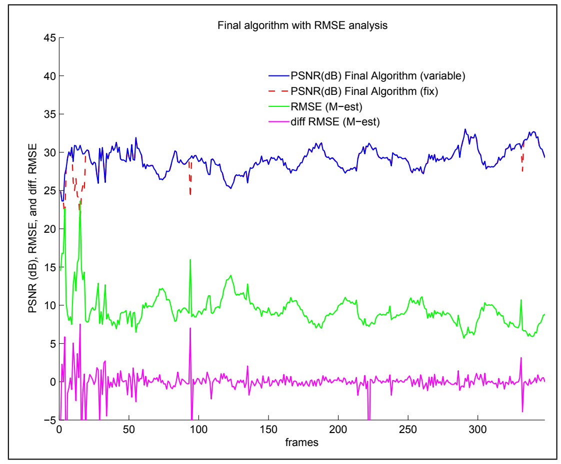


Fig. 5. PSNR-curves over “Biathlon” with and without $RMSE$-analysis

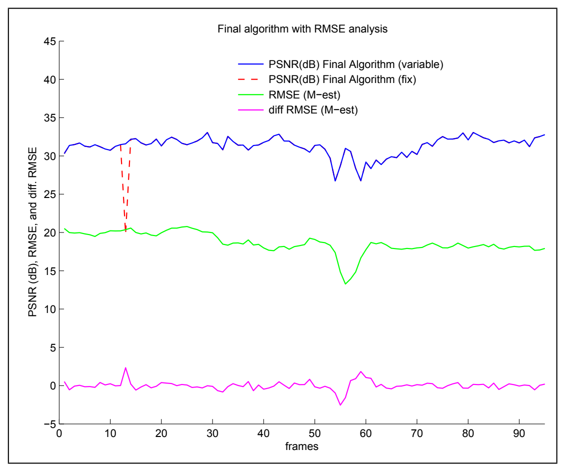


Fig. 6. PSNR-curves over “Mobile& Calendar” with and without $RMSE$-analysis

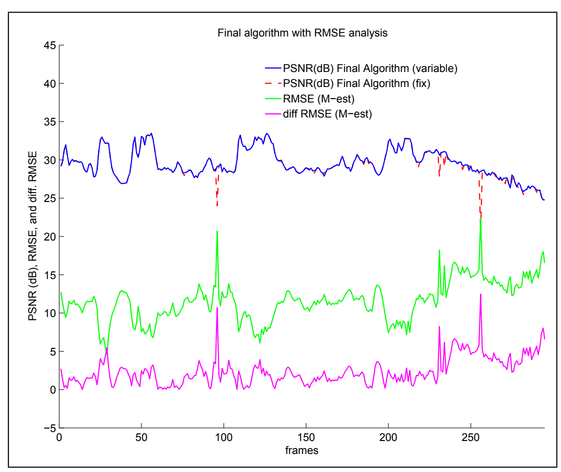


Fig. 7. PSNR-curves over “Stefan” with and without $RMSE$-analysis