

# Improved Image Registration using the Up-Sampled Domain

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

\*Communication Systems Group  
TU Berlin  
Berlin, Germany

{krutz, sikora}@nue.tu-berlin.de

†School of IT and EE  
Australian Defence Force Academy  
Canberra, Australia  
m.frater@adfa.edu.au

**Abstract**— We consider the warping problem which appears in well-known image registration algorithms that use higher-order motion models. An implementation inspired by recent work and a new image registration algorithm are used for the analysis. Both approaches rely on frame-to-frame estimation. The key technique is the well-established gradient descent approach for the estimation of higher-order motion parameters. We show that using up-sampled input images in the last step of the algorithms improves the accuracy of the estimated motion parameters. It can be seen in the experimental results that the performance of the image registration algorithms increases significantly only by applying the gradient descent on up-sampled images in comparison to recent algorithms developed.

## I. INTRODUCTION

Image registration algorithms using higher-order motion models are used widely in a number of areas, such as computer vision and video processing. One of the most common applications of these algorithms is video mosaicing or sprite generation. It is used in applications including video segmentation and video coding. In this work, the estimation of 2D motion parameters of consecutive images is related to video data. The goal is to estimate the global motion, i.e. the camera motion, during a number of frames of a video sequence. Recent works have been published in the field of development of global motion estimation algorithms and sprite generation [1], [2], [3]. All these methods rely on the gradient descent approach, which is also used in this work. Techniques have been proposed in [1], [2], [3] to prevent the influence of outliers due to, for example, foreground objects and to improve the performance of the algorithm. All approaches use original images in the last step, except in [3], which uses an up-sampled version of the original images. It has been shown here that applying the estimation algorithm for the global motion parameters produce better result in sprite generation. We take the idea of using up-sampling for improving the performance of the estimation of the global motion parameters for the considered algorithms and show why up-sampling is better. The significant part of the algorithm is the warping process. In almost all cases, the warped pixel does not match to the warping image pixel raster. This means that an interpolation has to be applied. Here, the widely used bilinear interpolation is used because it is fast and the results are reasonable. Applying another interpolation type at this point would increase the computational complexity too much. We show that the

warped image is resampled due to the warping step. Under-sampling can occur, and this leads to aliasing in the warped image. An aliased warped image causes an increased error and the performance and the accuracy of the algorithm to decrease. The suboptimal bilinear interpolation technique also produces errors. To avoid under-sampling during the warping process the images are up-sampled first. It is important that the up-sampling process is very accurate. It has been shown in [3] that the 7-tap wavelet synthesis filter is adequate for the interpolation of up-sampled images. Another approach to develop an optimal interpolation filter in the up-sampled domain has been proposed in [4]. Here the design of the filter coefficients relies on statistical methods. Both interpolation filters are used in this work. The paper is organized as follows. In the next section the global motion estimation algorithms used are introduced briefly. Afterwards, the up-sampling issue is outlined in detail. Section 3 shows the experimental results and the discussion. The conclusion and approaches for further work is given in Section 4.

## II. THE UP-SAMPLING ISSUE WITHIN IMAGE REGISTRATION ALGORITHMS

### A. Image Registration Algorithms

Two algorithms are considered for estimating the global motion parameters in video sequences. The first one is given in Fig. 1. This algorithm is inspired by [2]. Here, the input images are subdivided using an image pyramid. Then the estimation process is applied on the lowest resolution of the input images first. The achieved parameters initialize the motion parameters in the next upper stage of the pyramid and the estimation process begins with these initialized parameters. This procedure is repeated through the pyramid until at the highest resolution, (the original input frames) the final motion parameters are calculated. For the motion model the affine model is used in the first stage of the pyramid and the perspective motion model is used in the upper stages with highest resolution. The core technique is the Gauss-Newton gradient descent algorithm [5]. A simplified robust M-estimator [1] is applied on the gradient descent to prevent the influence of outliers. Moreover, the phase correlation method is used for the estimation of the translational motion parameters [6]. It has been shown that the Gauss-Newton gradient descent algorithm has a very good performance if the start point of the gradient descent is close

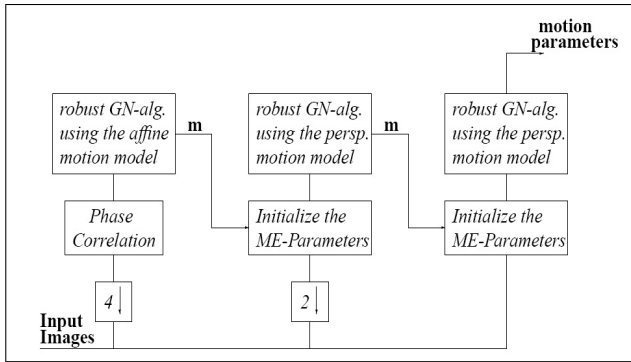


Fig. 1. Algorithm 1

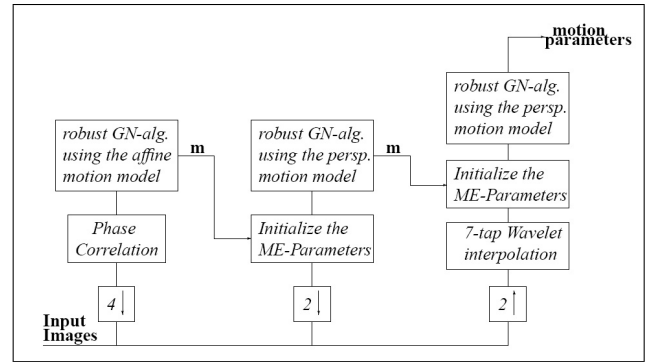


Fig. 2. Algorithm 2

to the desired minimum [5]. Therefore, the initialization of the translational parameters is very important and the estimation of the phase correlation method is robust and fast with good results [7]. The algorithm 2 shown in Fig. 2 has the same structure as the algorithm 1 except the third stage of the pyramid. Here, an up-sampling and interpolation is applied in the final stage of motion estimation. The 7-tap wavelet filter is used for the interpolation of the up-sampled images [3]. Instead of this filter other approaches are possible. The requirement is to produce accurate up-sampled versions of the original images. For this, a second filter is applied instead of the 7-tap wavelet (Wav). The optimal Inverse Wiener Filter (IWF) proposed in [4] is used for the interpolation. It will be shown that the use of the up-sampling first brings better results than using the original images in the final step. The third algorithm considered is a new approach for unsupervised global motion estimation [7], [8] and given in Fig. 4. The goal here is to achieve the background motion parameters even if large background occlusions occur due to large or multiple foreground objects. The algorithm relies on a windowing approach. The core of this algorithm is the same as the algorithms shown above. It will be shown that the up-sampling in the last step produce here better results in the estimation as well.

### B. The Up-sampling Issue

The perspective motion model is chosen for the higher-order motion model. The resulting warping matrix for the pixel warping is given in (1).

$$\begin{pmatrix} x' \cdot h' \\ y' \cdot h' \\ h' \end{pmatrix} = \begin{pmatrix} m_1 & m_2 & m_3 \\ m_4 & m_5 & m_6 \\ m_7 & m_8 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad (1)$$

The vector  $(x, y)$  is the pixel coordinate of the reference frame and  $(x', y')$  is the pixel coordinate of the warped frame. The matrix includes the motion parameters. Due to this warping process possible under-sampling can occur. Fig. 3 illustrates that case. If the warping process is applied at original input images for the final estimation and we assume that the original input images are optimally sampled, aliasing appears in the warped image due to under-sampling.

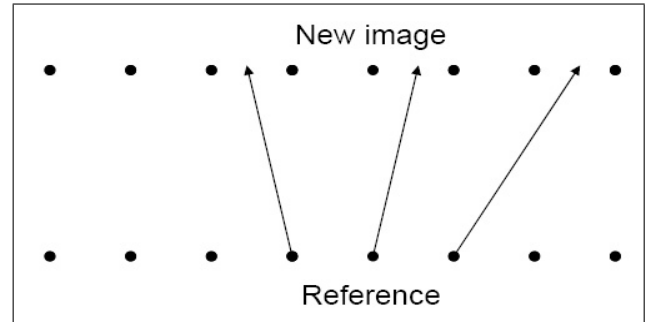


Fig. 3. Resampling due to warping

This aliasing produces an increased error and the performance of the gradient descent algorithm decreases. The robust M-estimator may remove some large errors, but cannot compensate for under-sampling. The design of the binary M-estimator is for prevention of errors due to outliers. The errors in consequence of aliasing in the warped image range through the whole image and affect the estimation process in general. To prevent under-sampling the input images can be up-sampled. Fig. 5 shows how up-sampling can avoid aliasing in the pixel domain. In Fig. 5 point spread functions are shown in each sampling case. It can be seen that in the over-sampled case (Fig. 5 (b)) the point spread functions are more overlapped which means that they are robust against possible under-sampling. Fig. 5 (a) shows the optimally sampled case (assumed the original images). If under-sampling arise here the point spreads of each pixels are spread like illustrated in Fig. 5 (c). As a result, aliasing appears in the warped image. The same issue can be explained in the frequency domain. The over-sampling can prevent the under-sampling resulting from image warping from causing aliasing, so long as the magnitude of the over-sampling is larger than that of the under-sampling.

### III. EXPERIMENTAL RESULTS AND DISCUSSION

The following experiments show the effect of applying the gradient descent algorithm on up-sampled versions of the original input images in the last stage of the global motion estimation algorithms mentioned above. Two test video sequences are chosen, the ‘‘Stefan’’-sequence and ‘‘Mobile&Calendar’’-

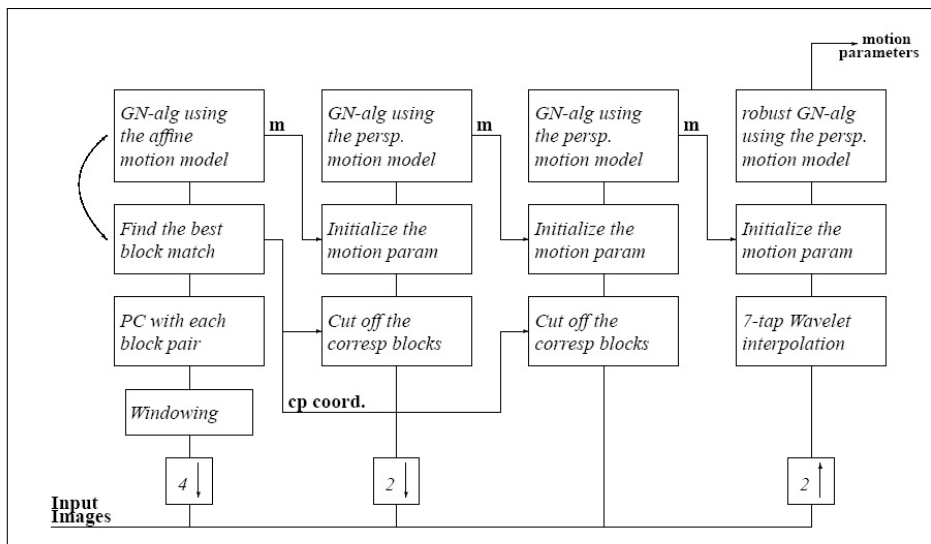


Fig. 4. Algorithm 3

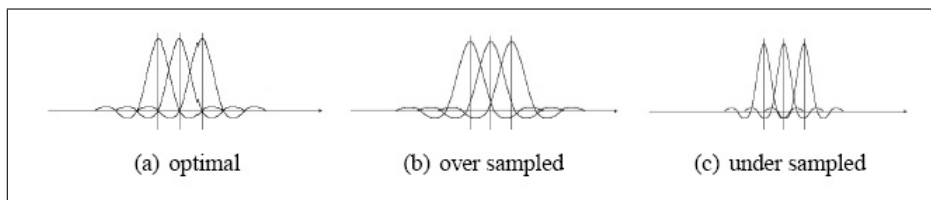


Fig. 5. Sampling issue (pixel domain)

sequence. The background of the “Stefan”-sequence is very significant. Thus algorithm 1 and 2 can be applied without any pre-segmentation. Fig. 6 depicts the PSNR-curves over nearly 300 frames. It can be seen that algorithm 2 provides a better estimation at least in every step only by using up-sampled images for the final estimation. In Fig. 7 algorithm 2 is compared with the PSNR-curve taken from [1]. The estimation of the motion parameters is more accurate even in difficult regions of the sequence. The results achieved with the second test sequence are given in Fig. 8. Here, large background occlusion occurs due to multiple foreground objects. An unsupervised global motion estimation algorithm (algorithm 3) [7], [8] is applied to compute the motion parameters from the background object exactly. Fig. 9 shows the difference between using original and up-sampled input images. It shows that using up-sampled images the algorithm produces better results as well. Tab. I and II summarize the mean PSNR-values of each algorithm and video sequence. The up-sampling brings a better result up to 1.3 dB (compare to alg. 1) and 0.75 dB (compare to Smolic), respectively. These PSNR values are computed frame-to-frame consecutively.

If the short-term parameters are accumulated to calculate long-term motion parameters for sprite generation, the accuracy of the long-term parameters achieved with e.g. algorithm 2 or 3 is much better as shown in [3]. Moreover, the performance of every other application using sprites, e.g. video segmentation, increases with more accurate higher-order motion parameters.

TABLE I  
MEAN PSNR VALUES “STEFAN”

Avg. PSNR	Alg.1	Alg.2 (Wav)	Alg.2 (IWF)	Alg.3	Smolic
“Stefan”	28.08	29.33	29.33	29.34	28.59

TABLE II  
MEAN PSNR VALUES “MOBILE&CALENDAR”

Avg. PSNR	Algorithm 3 with original images	Algorithm 3 up-sampled images
“Mobile&Calendar”	25.37	25.92

A disadvantage is the higher computational complexity due to the up-sampled versions of the input images.

#### IV. CONCLUSION AND FURTHER WORK

We propose improved image registration algorithms to estimate more accurate higher-order motion parameters. The algorithms are applied to real video sequences for the accurate estimation of the background object. The performance of the algorithm is increased using up-sampled versions of input images. This issue was described in detail theoretically and experiments have outlined the better estimation results. This method increases the performance of applications based on

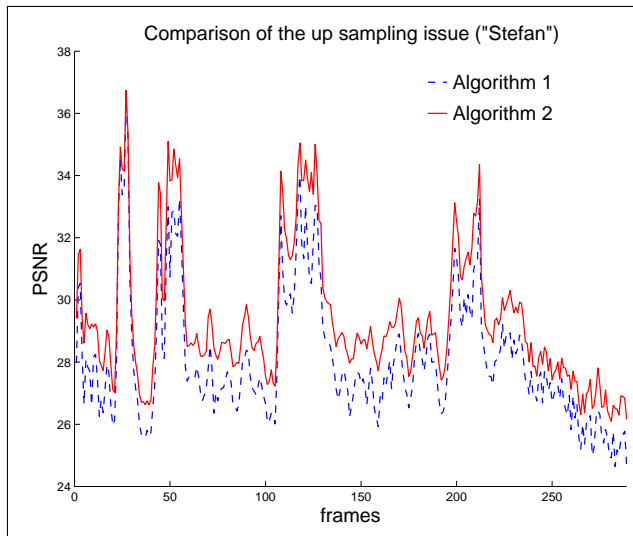


Fig. 6. PSNR-curves over "Stefan" (1)

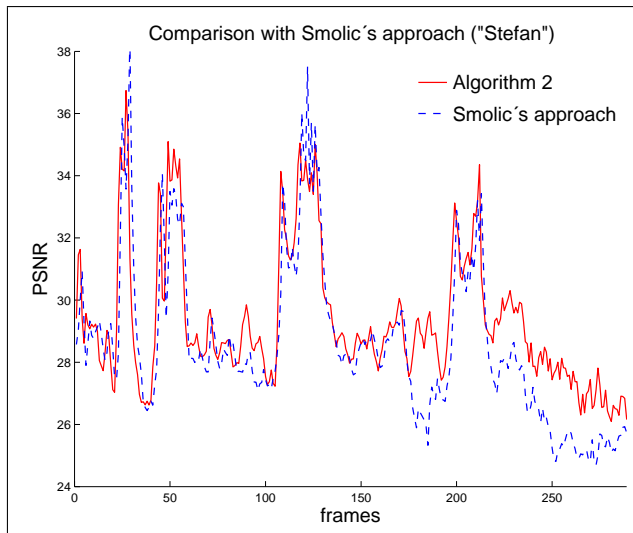


Fig. 7. PSNR-curves over "Stefan" (2)

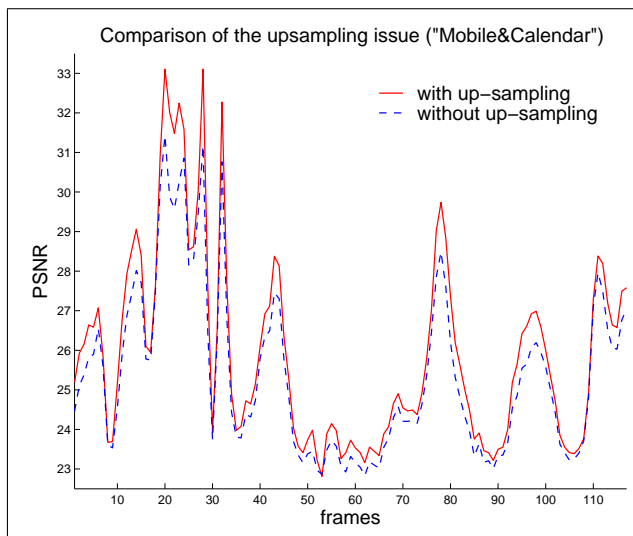
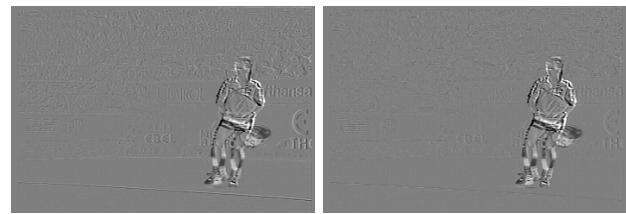
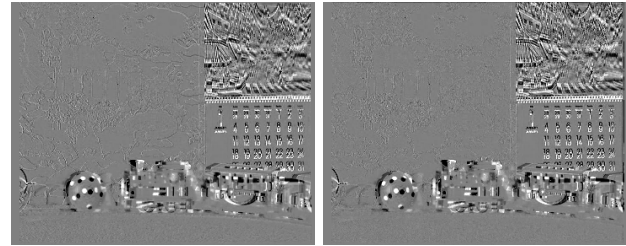


Fig. 8. PSNR-curves over "Mobile&Calendar"



(a) without up-sampling

(b) with up-sampling



(c) without up-sampling

(d) with up-sampling

Fig. 9. Example error frames with and without up-sampling input images

higher-order motion parameters. In this work, the suboptimal bilinear interpolation technique is used for computing the warped pixel values. A further step can be the examination of the proposed up-sampling issue using a more accurate interpolation technique, e.g. the spline interpolation.

#### ACKNOWLEDGMENT

We would like to thank Dr. A. Smolic for providing his PSNR-data resulted in [1] for a direct comparison. This work was developed within 3DTV (FP6-PLT-511568-3DTV) a European Network of Excellence funded under the European Commission IST FP6 programme.

#### REFERENCES

- [1] A. Smolic and J.-R. Ohm, "Robust global motion estimation using a simplified m-estimator approach," in *Int. Conf. on Image Processing (ICIP00)*, Vancouver, Canada, 2000.
- [2] 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.
- [3] 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. 2000.
- [4] T. Sikora, "Optimal wiener interpolation filters for multiresolution coding of images," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 7, no. 2, pp. 443–447, Apr. 1997.
- [5] 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.
- [6] C. Kuglin and D. Hines, "The phase correlation image alignment method," in *Proc. IEEE 1975 Int. Conf. Cybernetics and Society*, Sept. 1975, pp. 163–165.
- [7] A. Krutz, "Unsupervised global motion estimation with application to video mosaicing," Diploma thesis, TU Berlin, Berlin, Germany, Jan. 2006.
- [8] 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.