ASYMMETRIC GRADIENT-BASED IMAGE ALIGNMENT

Jean-Baptiste AUTHESERRE, Rémi MEGRET, Yannick BERTHOUMIEU

IMS Laboratory, UMR 5218 CNRS, University of Bordeaux, 33405 Talence, France

ABSTRACT

A new method for template-based image alignment is presented in this paper. A gradient-based optimization of motion compensated error difference is addressed to solve the image alignment problem. Its novelty lies in the minimized error, which considers the bi-directional compensation of the reference template and the current image onto each other in a constrained asymmetrical fashion. The proposed approach is shown to be a generalization of previous state-of-the art gradient-based methods. An experimental evaluation is provided to show how the new method outperforms the former in presence of noisy images, and to give some insights into its properties.

Index Terms— Image registration, template matching, gradient-based methods, bi-directional motion compensation, asymmetric constraint.

1. INTRODUCTION

Motion estimation is a fundamental task of many vision applications such as object tracking, video compression, augmented reality or image mosaicking. Among the existing approaches to evaluate motion, template matching is a natural way to estimate the parameters that best warp one image onto the other. The optimum is conventionally found by minimizing the displaced frame difference between a reference template and a current image.

One of the reference algorithms was proposed by Lucas and Kanade for optical-flow computation [1]. Baker and Matthews summarized and compared many template-based approaches in [2]. Their taxonomy of the methods was extended in [3] in order to take into account recent approaches that proved to yield better convergence and robustness performances: the Efficient Second-Order Minimization (ESM) algorithm [4] and the Symmetric Gradient Method (SGM) [5]. These methods weight symmetrically the gradients of both images when estimating the update of motion parameters, which yields faster convergence and improved robustness. Both authors have given theoretical justification for these performances when the two images are equal up to motion compensation. In presence of noisy images, this assumption does not hold anymore.

When one or both images are corrupted by different amount of noise, the gradient of both images do not have to play a symmetric role. In this paper, a new method is developed that unifies and extends existing approaches. It allows us to improve the performances by weighting asymmetrically the gradients of the two images.

The paper is organized as follows. In section 2 we provide a short review of existing algorithms. In section 3 we introduce the new method and show how it subsumes the existing ones. In section 4, after recalling the principles of the performance evaluation, we show experimental results and discuss the properties of the proposed approach.

2. ANALYSIS OF RELATED WORK

2.1. Algorithmic scheme

Image alignment between a reference template T and a current image I can be formulated as finding motion parameter vectors (μ_I, μ_T) that minimizes a motion compensated error:

$$E(\mu_I, \mu_T) = \sum_i \left(e_i(\mu_I, \mu_T)\right)^2 \quad (1)$$

where μ_{ref} is fixed, and e_i represents the error at pixel x_i between the two motion compensated images I and T:

$$e_i(\mu_I, \mu_T) = I\left(f(\mu_I, x_i)\right) - T\left(f(\mu_T, x_i)\right) \quad (2)$$

In the sequel, the motion model f is assumed to exhibit group properties. Its group properties are extended to its parameter space using the following compositional notations:

$$f(\mu_1 \circ \mu_2, x) = f(\mu_1, f(\mu_2, x)) \quad (3)$$

$$f(\mu^{-1}, x) = y \Leftrightarrow x = f(\mu, y) \quad (4)$$

The gradient-based methods of interest [1-6] obey the general scheme of iterative Gauss-Newton approaches:

Step 1) Define initialization parameters μ^0 and μ_{ref}

Step 2) Express the motion parameters w.r.t. an incremental correction parameter vector $\delta\mu$: $\mu_I = \mu_I(\mu^k, \delta\mu)$ and $\mu_T = \mu_T(\mu_{ref}, \delta\mu)$. Compute $\delta\mu$ by using one step of gradient based optimization of E

$$\delta\mu = -\left(J^t J\right)^{-1} J^t e(\mu^k, \mu_{ref}) \quad (5)$$

where J is the Jacobian w.r.t. $\delta\mu$ of the pixelwise error vector e formed by stacking all e_i vertically.

Step 3) Use the generic update rule introduced in [3] to obtain the estimated parameters:

$$\mu^{k+1} \leftarrow \mu_I(\mu^k, \delta\mu) \circ \mu_T(\mu_{ref}, \delta\mu)^{-1} \circ \mu_{ref} \quad (6)$$

Step 4) Repeat steps 2) and 3) until convergence

2.2. Categories of approach

Gradient based approaches have been shown in [3] to be equivalent to the first order in the estimated $\delta\mu$ within each of the following categories: Forwards, Inverse, and Symmetric. We will therefore consider only one approach within each category.

The Forwards Compositional (FC) approach [6] uses $\mu_I = \mu^k \circ \delta\mu$ and $\mu_T = \mu_{ref}$. Plugging these expressions into (2) and computing its derivative yields the following Jacobian:

$$J^{FC}(x_i) = J_I^{FC}(x_i) = \left.\frac{\partial I(f(\mu^k \circ \delta\mu, x_i))}{\partial \delta\mu}\right|_{\delta\mu=0} \quad (7)$$

The Inverse Compositional (IC) approach [2] is related [3] to $\mu_I = \mu^k$ and $\mu_T = \mu_{ref} \circ \delta\mu^{-1}$, that yields

$$J^{IC}(x_i) = J_T^{IC}(x_i) = \left.\frac{\partial T(f(\mu_{ref} \circ \delta\mu, x_i))}{\partial \delta\mu}\right|_{\delta\mu=0} \quad (8)$$

The Symmetric Compositional (SC) approach [3] uses $\mu_I = \mu^k \circ (0.5\,\delta\mu)$ and $\mu_T = \mu_{ref} \circ (0.5\,\delta\mu)^{-1}$ which yields a symmetrical combination of the gradients on I and T:

$$J^{SC}(x_i) = \frac{1}{2}\left(J_I^{FC}(x_i) + J_T^{IC}(x_i)\right) \quad (9)$$

The Efficient Second-order Minimization (ESM) algorithm [4] is based on a Lie group parameterization: $\delta\mu$ is parameterized around $\delta\mu = 0$ by a Lie algebra vector δv_{lie}. The link between the two vectors is given by the exponential map "exp" relationship [7]:

$$\delta\mu = \exp(\delta v_{lie}) \quad (10)$$

This approach corresponds to a Symmetric Compositional Exponential-map approach (SCE) [3] with $\mu_I = \mu^k \circ \exp(0.5\,\delta v_{lie})$ and $\mu_T = \mu_{ref} \circ \exp(-0.5\delta v_{lie})$ which yields the following Jacobian:

$$J^{SCE}(x_i) = J^{SC}(x_i) \left.\frac{\partial \exp(\delta v_{lie})}{\partial \delta v_{lie}}\right|_{\delta v_{lie} = 0} \quad (11)$$

3. ASYMMETRIC COMPOSITIONAL APPROACH

We now introduce a new approach that unifies and extends the previous ones. The Asymmetric Compositional approach (AC) is based on the minimization of $E(\mu_I, \mu_T)$ with

$$\mu_I = \mu^k \circ ((1-\alpha)\,\delta\mu) \quad (12)$$

$$\mu_T = \mu_{ref} \circ (\alpha\,\delta\mu)^{-1} \quad (13)$$

where $\alpha \in [0, 1]$ is a tuning parameter corresponding to an asymmetric constraint imposed on the bi-directional parameter vector (μ_I, μ_T).

The Jacobian matrix is obtained by plugging equation (12) and (13) into (2) and computing its derivative:

$$J_\alpha^{AC}(x_i) = (1-\alpha) J_I^{FC}(x_i) + \alpha\, J_T^{IC}(x_i) \quad (14)$$

The special case α=0.5 corresponds to the Jacobian of the Symmetric Compositional approaches presented before. The two extreme α values correspond to the (FC) method [6] for α=0 and the (IC) method [2] for α=1.

Using instead a Lie group parameterization, we can also define the Asymmetric Compositional Exponential map (ACE) approach, based on:

$$\mu_I = \mu^k \circ (\exp((1-\alpha)\,\delta v_{lie})) \quad (15)$$

$$\mu_T = \mu_{ref} \circ (\exp(-\alpha\,\delta v_{lie})) \quad (16)$$

which leads to the following Jacobian:

$$J_\alpha^{ACE}(x_i) = J_\alpha^{AC}(x_i) \left.\frac{\partial \exp(\delta v_{lie})}{\partial \delta v_{lie}}\right|_{\delta v_{lie} = 0} \quad (17)$$

The ACE approach is therefore a generalization of the ESM algorithm [4], which can be obtained for α=0.5. The specials cases α=0 and α=1 can be added to the taxonomy, as Forward Compositional Exponential map (FCE) and Inverse Compositional Exponential map (ICE).

Table 1 shows a summary of how to integrate the proposed (AC) and (ACE) methods within the taxonomy scheme of [3] with its relationship to published works.

Table 1: Synoptic presentation of the proposed approach compared to existing algorithms.

Generic category	Jacobian	Direct param. eq. (6)	Lie group Param. eq. (10)
Forwards [2], [3]	$J^{FC} = J_0^{AC}$	(FC) [6]	(FCE)
Inverse [2], [3]	$J^{IC} = J_1^{AC}$	(IC) [2]	(ICE)
Symmetric [5], [3]	$J^{SC} = J_{0.5}^{AC}$	(SC) [3]	(SCE) [4]
α-Asymmetric	J_α^{AC}	(AC)	(ACE)

4. EXPERIMENTAL STUDY

To compare the performances of the presented estimation algorithms, we now use the benchmark proposed by Baker and Matthews [2].

The experimental results on this benchmark did not exhibit any significant difference in performance between Lie group and direct parameterization methods. For the sake of clarity, we will now therefore provide experimental results for the Asymmetric Compositional (AC) approach, with comparisons with the following published approaches: Forwards Compositional (FC) [6], Inverse Compositional Reverse (IC) [2] and Symmetric Compositional Exponential map (SCE) [4].

4.1. Principle of the benchmark

The benchmark generates random disturbances by adding a spatial Gaussian noise of standard deviation σ (called Point Sigma) to three canonical point locations in the template (cf. figure 1): these three pairs of points (canonical and test points) define an affine warp parameter vector μ for the disturbance.

Using these parameters, a reference image (which is larger than the template) is warped onto image I (figure 1.b) and the image alignment algorithms are run in order to fit image I to T (figure 1.a). The convergence criteria is the root mean squared error of the distance between the test point locations and the destination locations of the canonical points projected with the estimated deformation (denoted RMS point error). Additionally, robustness to image noise can be evaluated by corrupting the template and the image with additive random Gaussian noise with respective variances σ_T and σ_I.

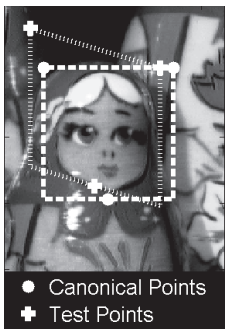
Two main performance criteria are considered:

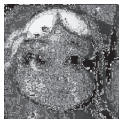
- Average frequency of convergence: percentage of tests where an algorithm converged to the correct estimate, (defined as a RMS Point Error less than 1 pixel).
- Average rates of convergence: for tests that converge for all methods, the average RMS point error is plotted against the algorithm iteration number.

4.2. Effect of image noise

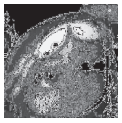
In the following experiments, the average frequency of convergence is plotted with varying spatial Point Sigma. The image I is corrupted with random noise of standard deviation σ_I=25. The parameter of the (AC) algorithm is fixed at α=0.70. The choice of the α value will be discussed in further details in next subsection.

From the point of view of the average frequency of convergence (Fig. 2), the (FC) approach has much worse results than the (IC) approach, which was already explained in [2] by the fact that the gradients of image I are corrupted by noise, yielding a suboptimal performance.





1.a Template T



1.b Disturbed Image at initialisation $I(f(\mu^0))$

Fig. 1: Principle of the benchmark [2]

The (SCE) approach has a larger region of convergence than both of them, by taking advantage of gradient information from both I and T, as was discussed in [4]. The new (AC) algorithm has even better results, by taking into account both images but weighting appropriately the gradients associated to the noisy image I. The results on the convergence rate (Fig. 3) lead to an analog analysis: the (AC) converges faster.

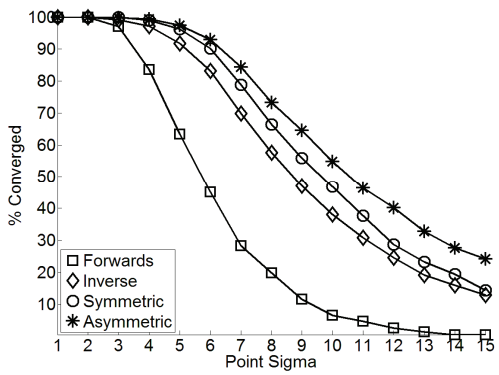


Fig. 2: Average frequency of convergence with corrupted image I, with σ_I=25 and α=0.7.

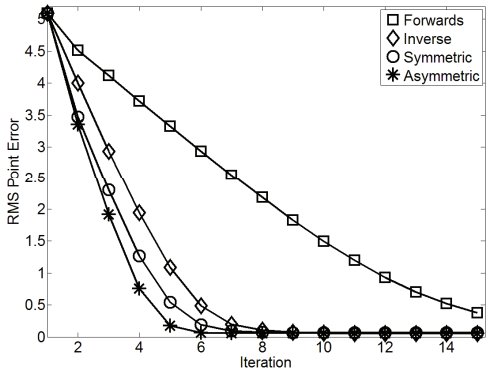


Fig. 3: Convergence rate, with corrupted image I, with σ_I=25 and α=0.7 and σ=8.

4.3. Influence of the α value

Since the Forwards, Inverse and Symmetric approaches can be considered as special cases of the Asymmetric approach, we now study the influence of the choice of the α value on the performances.

The results of figure 4 are obtained in the same configuration as in figures 2 and 3, but restricting to a fixed Point Sigma $\sigma=8$, and varying the α parameter, from 0 to 1. For the rate of convergence $\sigma=4$ (Fig 4, right), all points on the same vertical represent the RMS point error for a fixed α, and for the successive iterations (only the first five iterations are shown).

The performances of the (FC), (IC) and (SCE) approaches displayed in figures 2 and 3 can be retrieved in figure 4 for $\alpha=0$, 1 and 0.5 respectively. The optimal α value in such a configuration corresponds to $\alpha\approx0.7$, which could not be obtained using previous approaches.

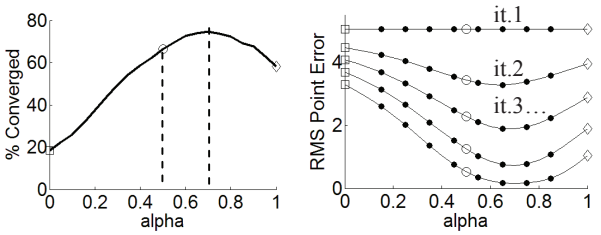


Fig. 4: Left: Average frequency of convergence. Right: Rate of convergence (first 5 iterations), with noise in image I.

When there is no noise in the images $\sigma_I=\sigma_T=0$ (see fig 5. left), the optimal value is $\alpha=0.5$, which puts in perspective the results shown in [4] and [5] that when the two images are identical up to motion compensation, using a Symmetrical approach yields much improved performance.

When noise is increased in one of the images, the optimal α value is shifted away (see Fig.6), naturally weighting down the corrupted gradients. The optimal α value indeed comes from a tradeoff between the use of both image gradients and the quality of those gradients. From a practical point of view, the (IC) approach ($\alpha=1$) uses only one Jacobian which can be computed off-line. The use of both gradients requires additional computations per iteration which are compensated by a fewer number of iterations (Fig 3, 4, 5), and a better robustness (Fig 4 left, Fig 6).

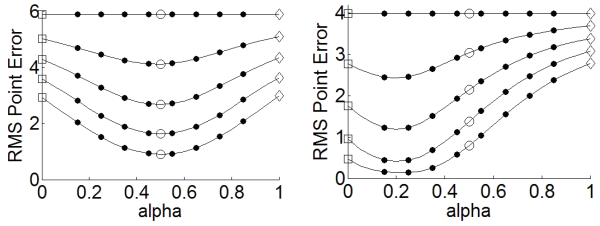


Fig. 5: Convergence rate for the first 5 iterations. Left: without image noise. Right: with noise in image T, $\sigma_T=25$.

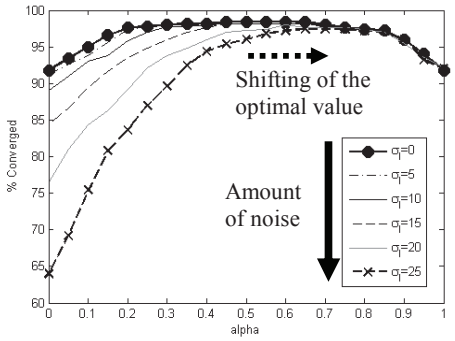


Fig. 6: Average frequency of convergence w.r.t. parameter α, for increasing amount of noise σ_I in image I ($\sigma_T=0$).

5. CONCLUSION

In this paper, we have presented a new approach for parametric image alignment, the Asymmetric Compositional approach, which has been shown to generalize existing state-of-the art gradient-based image alignment approaches. Benchmark experiments have highlighted situations where the new method outperforms the existing ones, especially when the template or the image is corrupted by noise. At this stage, the method requires the parameter α to be defined. Its adaptive estimation is the subject of future work.

6. REFERENCES

[1] B. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision", In Proceedings of the International Joint Conference on Artificial Intelligence, pp. 674-679, 1981.

[2] S. Baker and I. Matthews, "Lucas-Kanade 20 Years On: A Unifying Framework", International Journal of Computer Vision, 56, pp. 221-255, 2004.

[3] R. Mégret, J.B. Authesserre and Y. Berthoumieu, "The Bi-directional framework for unifying parametric image alignment approaches", European Conference on Computer Vision, 2008.

[4] S. Benhimane and E. Malis, "Homography-based 2D Visual Tracking and Servoing", The International Journal of Robotics Research, Vol. 26, pp. 661-676, 2007.

[5] Y. Keller and A. Averbuch, "Fast motion estimation using bidirectional gradient methods", The IEEE Transactions on Image Processing, 13, pp. 1042-1054, 2004.

[6] H.Y. Shum and R. Szeliski, "Construction of panoramic image mosaics with global and local alignment", International Journal of Computer Vision, 36(2), pp. 101-130, 2000.

[7] E. Bayro-Corrochano, J. Ortegón-Aguilar, "Lie algebra approach for tracking and 3D motion estimation using monocular vision", Image and Vision Computing, 25, pp. 907-921, June 2007.