Robust Image alignment using improved third-order global motion estimation

Y. Keller[2], A. Averbuch[1]

[1]School of Computer Science, Tel Aviv University, Tel Aviv 69978, Israel
[2]Department of Mathematics, Yale University, Connecticut, USA
Email: yosi.keller@yale.edu

Abstract

The estimation of parametric global motion using non-linear optimization is a fundamental technique in computer vision. Such schemes are able to recover various motion models (translation, rotation, affine, projective) with subpixel accuracy. The parametric motion is computed using a first order Taylor expansions of the registered images. But, it is limited to the estimation of small motions, and while large translations and rotations can be coarsely estimated by Fourier domain algorithms, no such techniques exist for affine and projective motions. This paper offers two contributions: First, we improve the convergence properties by an order of magnitude using a second order Taylor expansion. A third order convergence rate is achieved, compared to the second order convergence of prior schemes. Second, we extend the third order algorithm using a symmetrical formulation which further improves the convergence properties. The results are verified by rigorous analysis and experimental trials.

1   Introduction

Image registration plays a vital role in many image processing and computer vision applications such as optical flow computation [16, 6], tracking [4], video compression [8], layered motion estimation [3] and 3D reconstruction [22] to name a few. A comprehensive comparative survey by Barron et. al. [2] found the family of gradient-based motion estimation methods (GM) to perform especially well. The purpose of the GM algorithm is to estimate the parameters vector **p** associated with the *parametric image registration* problem: starting from pure global translation, rotation, affine, and projective motions. These models have been used extensively and are directly estimated from image spatio-temporal derivatives using coarse-to-fine estimation. They search for the best parametric geometric transform that minimizes the square of differences between image intensities over the whole image. Several formulations of the gradient methods were suggested. An updated comprehensive description of these methods was given in [11]. The registration is computed by relating a pair of images having some overlap using a first order Taylor series expansion. Each pixel in the common support contributes a linear constraint, denoted the constant brightness constraint. Thus, an over-constrained linear system is formulated yielding a robust estimate. Gathering and solving all the equations associated with pixels in the common support, estimates the *global motion* between the images [23].

The estimation of parametric global motion is the focal point of this work and has gained special attention in the computer vision community. As such schemes estimate a relatively small number of parameters (usually up to 8) based on a least squares formulation of the entire images, such schemes are robust and can be applied to a wide range of applications related to image stitching [23] and mosaicing [19]. This robustness was further used to estimate higher order parametric models that are able to register quadratic surfaces [21] (12 parameters), global nonlinear illumination changes [7] and barrel lense distortions [1]. A limited class of parametric motions was also applied to video compression in the MPEG4 compression standard [8]. The focal point in such works is to derive computationally efficient schemes, where due to the high frame-rate used (15-25 frames per second) the estimated motion is assumed to be small.

A critical implementation issue concerning the GM is their convergence when estimating large motions. As the estimated motion grows, the convergence rate decreases and the GM may converge to a local minima. Hence, GM algorithms are unable to estimate large motions and have to be bootstrapped for example by coarser and more robust Fourier schemes [15, 20]. But, while the estimation of translations and rotations can be bootstrapped by Fourier based methods there are no such reliable solutions for affine and projective motions. For example, [9] use affine invariant texture descriptors to bootstrap wide basis stereo.

We propose to improve the convergence properties of the GM algorithm using a second order Taylor expansion of the registered images. We show that the convergence properties of the proposed algorithm are superior to those of the regular GM for large and small motions. Convergence rates of 3 and $\frac{3}{2}$ are achieved for small and large motions, respectively, compared to 2 and 1 for the regular GM. Further improvement is achieved by combining the second order expansion with a symmetrical formulation [14]. While no gradient based scheme can achieve global convergence, we show rigorously and experimentally that the proposed scheme achieves improved convergence ranges compared to the standard GM schemes.

The paper is organized as follows: the optimization based GM approach to image registration is introduced in Section 2. The convergence properties of high order optimization schemes are analyzed in Section 3 and the third order (O3GM) and symmetric third order (O3GM) schemes are presented in Sections 4 and 5, respectively. Their convergence properties are derived in Section 6 and experimentally verified in Section 7. Concluding remarks and future work are discussed in Section 8.

2 Gradient methods based motion estimation

GM methodology [11] estimates the motion parameters $\mathbf{p}$ by minimizing the intensity discrepancies between I_1 and I_2

$$\mathbf{p}^* = \arg\min_{\mathbf{p}} \sum_{\left(x_i^{(1)}, y_i^{(1)}\right) \in S} \left(I_1\left(x_i^{(1)}, y_i^{(1)}\right) - I_2\left(x_i^{(2)}, y_i^{(2)}\right) \right)^2 \tag{1}$$

where $x_i^{(2)} = f(x_i^{(1)}, y_i^{(1)}, \mathbf{p})$, $y_i^{(2)} = g(x_i^{(1)}, y_i^{(1)}, \mathbf{p})$, S is the set of coordinates of pixels common to I_1 and I_2 in I_1's coordinates, $\mathbf{p}$ is the estimated parameters vector and f and g represent the motion model. In practice, solving Eq. 1 does not result in perfect intensity alignment due to relative intensity changes and pixels with non-corresponding pixels within the registered images. Next we follow the formulation of [23, 18] and solve Eq. 1 via a linearization scheme, based on a pixel-wise first order Taylor expansion of I_1 in terms of I_2 as a function of the parameter vector $\mathbf{p}$

$$I_1(x_i^{(1)}, y_i^{(1)}) = I_2(x_i^{(2)}, y_i^{(2)}, \mathbf{p}) + \sum_{p_k \in \mathbf{p}} \frac{\partial I_2(x_i^{(2)}, y_i^{(2)}, \mathbf{p})}{\partial p_k} (\varepsilon_{\mathbf{n}})_k + \frac{1}{2} \sum_{p_k \in \mathbf{p}} \sum_{p_m \in \mathbf{p}} \frac{\partial I_2(x_i^{(2)}, y_i^{(2)}, \widetilde{\mathbf{p}})}{\partial p_k \partial p_m} (\varepsilon_{\mathbf{n}})_k (\varepsilon_{\mathbf{n}})_m, \widetilde{\mathbf{p}} \in [0, \varepsilon_{\mathbf{n}}]. \tag{2}$$

$I_1(x_i^{(1)}, y_i^{(1)})$ and $I_2(x_i^{(2)}, y_i^{(2)}, \mathbf{p})$ are the ith corresponding pixel in I_1 and $I_2(\mathbf{p})$, where $I_2(\mathbf{p})$ is the image I_2 transformed according to the parameters vector $\mathbf{p}$. $\frac{\partial I_2(x_i^{(2)}, y_i^{(2)}, \mathbf{p})}{\partial p_k}$ is the partial derivative with respect to the motion parameters given the current estimate $\mathbf{p}$ and $\varepsilon_{\mathbf{n}}$ is the estimation error of the parameters at iteration n. As $\widetilde{\mathbf{p}}$ is unknown, Eq. 2 can not be solved for $\varepsilon_{\mathbf{n}}$, instead we neglect the error term and solve for $\delta\mathbf{p}$

$$I_1(x_i^{(1)}, y_i^{(1)}) = I_2(x_i^{(2)}, y_i^{(2)}, \mathbf{p}) + \sum_{p_k \in \mathbf{p}} \frac{\partial I_2(x_i^{(2)}, y_i^{(2)}, \mathbf{p})}{\partial p_k} \delta\mathbf{p}_k. \tag{3}$$

By gathering the pixel-wise equations we get the system $\mathbf{H}\delta\mathbf{p} = \mathbf{I_t}$, where $\mathbf{H}_{i,j} = \frac{\partial I_2\left(x_i^{(2)}, y_i^{(2)}, \mathbf{p}\right)}{\partial p_j}$ and $\underline{I}_{t_i} = I_1\left(x_i^{(1)}, y_i^{(1)}\right) - I_2\left(x_i^{(2)}, y_i^{(2)}, \mathbf{p}\right)$, which is solved by least squares. Due to the omission of the error term in Eq. 2, in general we get $\delta\mathbf{p} \neq \varepsilon_{\mathbf{n}}$. Hence Eq. 1 is solved iteratively and in order to improve the convergence properties, the iterative process is embedded in a coarse-to-fine multiscale formulation [11, 14].

3 T-Order Convergence analysis

In this section we analyze the convergence properties of the non-linear least squares scheme used in GM based image registration. We formulate the optimization as a zero crossing problem solved by a Taylor series based approximations and derive the convergence rate when using a T-order Taylor approximation. Equation 1 is solved iteratively by approximating $f(\mathbf{p})$ at the zero crossing point $\mathbf{p}^*$ using a T-order Taylor series expansion around the current estimate of the solution $\mathbf{p}_k$

$$f(\mathbf{p}^*) = f(\mathbf{p}_n) + \sum_{k=1}^{N_P} \frac{\partial f(\mathbf{p}_n)}{\partial p_k} (\varepsilon_{\mathbf{n}})_k + \frac{1}{2} \sum_{k_1=1, k_2=1}^{N_P} \frac{\partial^2 f(\mathbf{p}_n)}{\partial p_{k_1} \partial p_{k_2}} (\varepsilon_{\mathbf{n}})_{k_1} (\varepsilon_{\mathbf{n}})_{k_2} + \ldots + R_T(\mathbf{p}_n, \widetilde{\mathbf{p}}), \widetilde{\mathbf{p}} \in [0, \varepsilon_{\mathbf{n}}]. \tag{4}$$

where $\varepsilon_{\mathbf{n}}$ is the estimation error in iteration n, N_P is the dimension of $\mathbf{p}$ and $R_T(\mathbf{p}_k, \widetilde{\mathbf{p}})$ is the T-order error term of the Taylor series approximation given by

$$R_T(\mathbf{p}_n, \widetilde{\mathbf{p}}, \varepsilon_{\mathbf{n}}) = \frac{1}{(T+1)!} \sum_{k_1=1,\ldots k_{T+1}=1}^{N_P} \frac{\partial^{T+1} f(\widetilde{\mathbf{p}})}{\partial p_{k_1} \ldots \partial p_{k_{T+1}}} (\varepsilon_{\mathbf{n}})_{k_1} \ldots (\varepsilon_{\mathbf{n}})_{k_{T+1}}, \widetilde{\mathbf{p}} \in [0, \varepsilon_{\mathbf{n}}]. \tag{5}$$

As both $f(\mathbf{p}^*)$ and R_T are unknown ($\mathbf{p}^*$ is a minimum point, but one can not assume that $f(\mathbf{p}^*) = 0$), Eq. 4 can not be solved directly. Hence it is approximated by neglecting R_T and $f(\mathbf{p}^*)$

$$H(\delta \mathbf{p}_n) = f(\mathbf{p}_n) + \sum_{k=1}^{N_P} \frac{\partial f(\mathbf{p}_n)}{\partial p_k} (\delta \mathbf{p}_n)_k + \frac{1}{2} \sum_{k_1=1,k_2=1}^{N_P} \frac{\partial^2 f(\mathbf{p}_n)}{\partial p_{k_1} \partial p_{k_2}} (\delta \mathbf{p}_n)_{k_1} (\delta \mathbf{p}_n)_{k_2}$$

$$+ \frac{1}{T!} \sum_{k_1=1,\ldots,k_T=1}^{N_P} \frac{\partial^T f(\mathbf{p}_n)}{\partial p_k \ldots \partial p_T} (\delta \mathbf{p}_n)_{k_1} \ldots (\delta \mathbf{p}_n)_{k_T} \tag{6}$$

and solve $H(\delta \mathbf{p}_n) = 0$ in the least squares sense. Using the above procedure with a first order Taylor expansion of $f(\mathbf{p})$ (T=1), is commonly known as the Gauss-Newton scheme [10, 5] and Eq. 6 reduces to a linear equation. This approach is utilized by the GM and SGM whose convergence properties are given by the following Lemma, proved in [14]

Lemma 1 *The convergence process of the GM and SGM schemes can be divided to two distinct phases, characterized by the deviation of the parameters $\mathbf{p}$ from their optimal value $\mathbf{p}^*$ and the convergence rate of the optimization scheme. **Near the minimum** $\mathbf{p_k} \to \mathbf{p}^*, \|\varepsilon_n\| \to 0$, a quadratic convergence rate is achieved. **Away from the minimum** $\|\varepsilon_n\| \gg 1$, a slow linear convergence rate is achieved.*

$$\|\varepsilon_{\mathbf{n}}\| \leq C_{GM}^L \cdot \|\varepsilon_n\| + C_{GM}^S \cdot \|\varepsilon_n\|^2$$

where C_{GM}^S and C_{GM}^L are the small and large deviation coefficients, respectively.

We study the convergence properties of the T-order Taylor approximation. Recalling that $\mathbf{p}_{n+1} = \mathbf{p}_n + \delta \mathbf{p}_n$, the right-hand-side of Eq. 6 is a T-order Taylor order approximation of $f(\mathbf{p}_{n+1})$. Thus,

$$f(\mathbf{p}_{n+1}) = H(\delta \mathbf{p}_n) + R_T(\mathbf{p}_n, \widetilde{\mathbf{p}}, \delta \mathbf{p}_n), \widetilde{\mathbf{p}} \in [\mathbf{0}, \delta \mathbf{p}_n] \tag{7}$$

and

$$\|f(\mathbf{p}_{n+1})\| \leq \|R_T(\mathbf{p}_n, \widetilde{\mathbf{p}}, \delta \mathbf{p}_n)\| + \|H(\delta \mathbf{p}_n)\| = C_{T+1} \|\delta \mathbf{p}_n\|^{T+1} + \|H(\delta \mathbf{p}_n)\| \tag{8}$$

where

$$C_{T+1} = \frac{1}{(T+1)!} \left\| \sum_{k_1=1,\ldots,k_{T+1}=1}^{N_P} \frac{\partial^{T+1} f(\widetilde{\mathbf{p}})}{\partial p_{k_1} \ldots \partial p_{k_{T+1}}} \right\|, \widetilde{\mathbf{p}} \in [\mathbf{0}, \delta \mathbf{p}_n].$$

Next, we approximate $f(\mathbf{p}^*)$ based on the point $\mathbf{p}_{n+1}$ using a first order Taylor series expansion

$$f(\mathbf{p}^*) = f(\mathbf{p}_{n+1}) + \sum_{k=1}^{N_P} \frac{\partial f(\mathbf{p}_{n+1})}{\partial p_k} (\varepsilon_{\mathbf{n+1}})_k + \frac{1}{2} \sum_{k_1=1,k_2=1}^{N_P} \frac{\partial^2 f(\widetilde{\mathbf{p}})}{\partial p_{k_1} \partial p_{k_2}} (\varepsilon_{\mathbf{n+1}})_{k_1} (\varepsilon_{\mathbf{n+1}})_{k_2}, \widetilde{\mathbf{p}} \in (\mathbf{0}, \varepsilon_{n+1}) \tag{9}$$

Following Lemma 1, we study the convergence properties for large and small deviations separately.

Theorem 2 *The convergence rate of the T-order approximation based scheme for small deviation is given by $\|\varepsilon_{\mathbf{n+1}}\| \leq \frac{C_{T+1}}{C_1} \|\varepsilon_{\mathbf{n}}\|^{T+1} + \frac{\|f(\mathbf{p}^*)\|}{C_1}$, where C_1 and C_{T+1} are constants, and $\|f(\mathbf{p}^*)\|$ is the optimal least-squares error.*

Proof. For small deviations we have $\varepsilon_n \to 0$, hence

$$\left\| \frac{1}{2} \sum_{\substack{k_1=1 \\ k_2=1}}^{N_P} \frac{\partial^2 f(\widetilde{\mathbf{p}})}{\partial p_{k_1} \partial p_{k_2}} (\varepsilon_{\mathbf{n+1}})_{k_1} (\varepsilon_{\mathbf{n+1}})_{k_2} \right\| \ll \left\| \sum_{k=1}^{N_P} \frac{\partial f(\mathbf{p}_{n+1})}{\partial p_k} (\varepsilon_{\mathbf{n+1}})_k \right\|$$

Thus, Eq. 9 is reduced to

$$f(\mathbf{p}^*) - \sum_{k=1}^{N_P} \frac{\partial f(\mathbf{p}_{n+1})}{\partial p_k} (\varepsilon_{\mathbf{n+1}})_k = f(\mathbf{p}_{n+1}) \tag{10}$$

and

$$\|f(\mathbf{p}^*)\| + C_1 \|\varepsilon_{\mathbf{n+1}}\| \geq \|f(\mathbf{p}_{n+1})\| \tag{11}$$

where $C_1 = \left\| \sum_{k=1}^{N_P} \frac{\partial f(\mathbf{p}_{n+1})}{\partial p_k} \right\|$.

In the vicinity of the solution, we have that $\varepsilon_n \to 0$ and $\delta \mathbf{p}_n \to 0$. By substituting it in Eq. 6, we get that the quadratic zero-crossing equation is reduced to a linear equation, which can always be solved such that $\|H(\delta \mathbf{p}_n)\| = 0$. Equation 8 is reduced to $\|f(\mathbf{p}_{n+1})\| \approx C_{T+1} \|\delta \mathbf{p}_n\|^{T+1}$ and by substituting into Eq. 11 we get $\|f(\mathbf{p}^*)\| + C_1 \|\varepsilon_{\mathbf{n+1}}\| \geq C_{T+1} \|\delta \mathbf{p}_n\|^{T+1}$ and by Eq. 6 it follows that

$$\sqrt[T+1]{\frac{\|f(\mathbf{p}^*)\| + C_3 \|\varepsilon_{\mathbf{n+1}}\|}{C_{T+1}}} \geq \|\delta \mathbf{p}_n\|.$$

The convergence rate is then derived by

$$\|\varepsilon_{\mathbf{n}}\| + \sqrt[T+1]{\frac{\|f(\mathbf{p}^*)\| + C_1 \|\varepsilon_{\mathbf{n+1}}\|}{C_{T+1}}} \geq \|\varepsilon_{\mathbf{n}}\| + \|\delta \mathbf{p}_n\| \geq \|\varepsilon_{\mathbf{n+1}}\|. \tag{12}$$

As $\varepsilon_n \to 0$ we have that $\|\varepsilon_n\| \ll 1$, hence $\|\varepsilon_{\mathbf{n+1}}\| \ll \sqrt[T+1]{\frac{\|f(\mathbf{p}^*)\| + C_1 \|\varepsilon_{\mathbf{n+1}}\|}{C_{T+1}}}$, and by substituting in Eq. 12 we have $\sqrt[T+1]{\frac{\|f(\mathbf{p}^*)\| + C_1 \|\varepsilon_{\mathbf{n+1}}\|}{C_{T+1}}} \leq \|\varepsilon_{\mathbf{n}}\|$ and

$$\|\varepsilon_{\mathbf{n+1}}\| \leq \frac{C_{T+1}}{C_1} \|\varepsilon_{\mathbf{n}}\|^{T+1} + \frac{\|f(\mathbf{p}^*)\|}{C_1}. \tag{13}$$

For example, the convergence analysis of the Gauss-Newton scheme in Theorem 1 uses $T = 1$ and $\|f(\mathbf{p}^*)\| \approx 0$

and by applying Eq. 13 we get $\|\varepsilon_{\mathbf{n+1}}\| \leq \frac{C_{T+1}}{C_1} \|\varepsilon_{\mathbf{n}}\|^2$ which corresponds to the expected quadratic convergence rate. ∎

Theorem 3 *The convergence rate of the T-order approximation scheme for large deviations is given by* $\|\varepsilon_{\mathbf{n+1}}\| \leq \frac{C_{T+1}}{C_2} \|\varepsilon_{\mathbf{n}}\|^{\frac{T+1}{2}}$*, where C_2 and C_{T+1} are constants and large deviations are characterized by* $\|\varepsilon_{n+1}\| \gg 1$*.*

Proof. For large deviations $\|\varepsilon_{n+1}\| \gg 1$ and $\|f(\mathbf{p}_{n+1})\| \gg \|H(\delta \mathbf{p}_n)\|$ in Eq. 8 and we get

$$\|f(\mathbf{p}_{n+1})\| \geq C_{T+1} \|\delta \mathbf{p}_n\|^{T+1}. \tag{14}$$

For Eq. 9 we get that for large deviations $\|f(\mathbf{p}_{n+1})\| \gg \|f(\mathbf{p}^*)\|$ and

$$\left\| \frac{1}{2} \sum_{k_1=1,k_2=1}^{N_P} \frac{\partial^2 f(\widetilde{\mathbf{p}})}{\partial p_{k_1} \partial p_{k_2}} (\varepsilon_{\mathbf{n+1}})_{k_1} (\varepsilon_{\mathbf{n+1}})_{k_2} \right\| \gg \left\| \sum_{k=1}^{N_P} \frac{\partial f(\mathbf{p}_{n+1})}{\partial p_k} (\varepsilon_{\mathbf{n+1}})_k \right\|$$

Thus, Eq. 9 becomes

$$\left\| \sum_{k_1=1,k_2=1}^{N_P} \frac{\partial^2 f(\widetilde{\mathbf{p}})}{\partial p_{k_1} \partial p_{k_2}} (\varepsilon_{\mathbf{n+1}})_{k_1} (\varepsilon_{\mathbf{n+1}})_{k_2} \right\| \approx \|f(\mathbf{p}_{n+1})\| \tag{15}$$

and $C_2 \|\varepsilon_{\mathbf{n+1}}\|^2 \approx \|f(\mathbf{p}_{n+1})\|$ where $C_2 = \left\| \sum_{k_1=1,k_2=1}^{N_P} \frac{\partial^2 f(\widetilde{\mathbf{p}})}{\partial p_{k_1} \partial p_{k_2}} \right\|$. Combining Eqs. 14 and 15 $C_2 \|\varepsilon_{\mathbf{n+1}}\|^2 \geq C_{T+1} \|\delta \mathbf{p}_n\|^{T+1}$

and

$$\frac{C_2}{C_{T+1}} \|\varepsilon_{\mathbf{n+1}}\|^{\frac{2}{T+1}} \geq \|\delta \mathbf{p}_n\|.$$

As $\varepsilon_{\mathbf{n+1}} = \varepsilon_{\mathbf{n}} - \delta\mathbf{p}_n$, we have that

$$\|\varepsilon_{\mathbf{n+1}}\| \leq \|\varepsilon_{\mathbf{n}}\| + \|\delta\mathbf{p}_n\| \leq \|\varepsilon_{\mathbf{n}}\| + \frac{C_2}{C_{T+1}}\|\varepsilon_{\mathbf{n+1}}\|^{\frac{2}{T+1}} \tag{16}$$

As $\|\varepsilon_{\mathbf{n+1}}\| \gg 1$ we get $\|\varepsilon_{\mathbf{n}}\| \approx \frac{C_2}{C_{T+1}}\|\varepsilon_{\mathbf{n+1}}\|^{\frac{2}{T+1}}$ and

$$\|\varepsilon_{\mathbf{n+1}}\| \leq \frac{C_{T+1}}{C_2}\|\varepsilon_{\mathbf{n}}\|^{\frac{T+1}{2}}. \tag{17}$$

For the Gauss-Newton scheme we have $T = 1$ and substituting in Eq. 16 $\|\varepsilon_{\mathbf{n+1}}\|\left(1 - \frac{C_2}{C_1}\right) \leq \|\varepsilon_{\mathbf{n}}\|$ which results

in the linear convergence rate discussed in Lemma 1. ∎

Theorem 4 *The range of convergence of the T-order approximation based scheme is related to the energy of the high-order derivatives of the objective function f.*

Proof. The convergence range ε_0 is given by computing the error $\varepsilon_{\mathbf{n}}$ such that $\|\varepsilon_n\| \leq \|\varepsilon_{n+1}\|$ for any $\|\varepsilon_n\| > \|\varepsilon_0\|$. Using Eq. 17 we get that for the convergence range ε_0 we have $\|\varepsilon_0\| \leq \|\varepsilon_{\mathbf{n+1}}\| \approx \frac{C_{T+1}}{C_2}\|\varepsilon_{\mathbf{n}}\|^{\frac{T+1}{2}}$ and we get

$$\|\varepsilon_0\| \approx \sqrt[\frac{T-1}{2}]{\frac{C_2}{C_{T+1}}}. \tag{18}$$

Equation 18 implies that $\|\varepsilon_0\|$ is maximized as $C_{T+1} \rightarrow 0$. This corresponds to using an approximation order which is higher than the order of the function f. For example, solving a quadratic equation set (in the least squares sense) using $T = 2$, allows an infinite convergence range. In practice, given a function f, C_{T+1} can be reduced by smoothing f. ∎

4 Third order gradient methods

This section introduces the third order GM formulation (O3GM) which is integrated into the regular GM algorithm presented in Section 2. The O3GM replaces *only* the basic GM step of the GM. The iterative refinement and multiscale steps are left intact. We rigorously showed in Section 3 that for nonlinear optimization problems, such as the GM, the higher the order of approximation, the better the convergence rate and range. In particular, the registration of images related by large motions, results in large parameter deviations. Thus, the approximation error grows, lowering the convergence rate down to the point of divergence. The focal point of this work is to improve the estimation of large deviations, which might otherwise be impossible, where using a higher order scheme lowers the approximation error, making the estimation of larger motions possible and small deviations need fewer iterations to converge. The second order approximation results in a set of parabolic equations solved iteratively. Thus, the proposed algorithm uses two iterative cycles - the first is identical to the iterative refinement step of the GM, while the second is used to solve the parabolic equations. We minimize Eq. 1 by expanding it in a second order Taylor series expansion and solving for δp, an update to the motion parameters vector $\mathbf{p}$, by solving

$$I_1(x_i^{(1)}, y_i^{(1)}) = I_2(x_i^{(2)}, y_i^{(2)}, \mathbf{p}) + \sum_{j=1}^{N_p}\delta p_j\frac{\partial I_2(x_i^{(2)}, y_i^{(2)}, \mathbf{p})}{\partial p_j} + \frac{1}{2}\sum_{j,s=1}^{N_p}\delta p_j\delta p_s\frac{\partial^2 I_2(x_i^{(2)}, y_i^{(2)}, \mathbf{p})}{\partial p_j\partial p_s} \tag{19}$$

where N_p is the number of motion parameters. While one can use a higher order approximation than the second order one in our approach, such schemes will result in an increased computational complexity that might prove exhaustive. The first and second partial derivatives with respect to the motion parameters are computed using the chain rule. Equation 19 is evaluated for each pixel common to the images I_1 and I_2 forming the quadratic equation set $\{r_i\}_{i=1,...,N}$ where N is the size of the common area. Equation 19 is solved by Newtonian methods where the iterative solution is given Gauss-Newton optimization

$$\mathbf{p_k} = \mathbf{p_k} - \left(\left(\mathbf{J^{2D}}\right)^T\mathbf{J^{2D}}\right)^{-1}\left(\mathbf{J^{2D}}\right)^T\mathbf{r}, \, k = 0,.. \tag{20}$$

where $r = (r_1, \ldots, r_N)^T$, $\mathbf{J^{2D}}$ is the Jacobian and N is the number of corresponding pixels. As the polynomial equations set is given explicitly, the computation of $\mathbf{J^{2D}}$ is fast and accurate up to machine precision.

5 Symmetric third order GM formulations

The symmetric GM (SGM) formulation was introduced in [14] to improve the convergence properties of GM algorithms. It utilizes the symmetry of the image registration problem to reduce its approximation error. The error reductions achieved by the O3GM and SGM are complementary, thus, we integrate both in a unified framework we call the Symmetric third order GM (SO3GM). The image registration problem is formulated symmetrically using a parametric motion model defined by $\mathbf{p}$

$$I_2\left(x_i^{(2)}, y_i^{(2)}, \mathbf{p}/2\right) = I_1\left(x_i^{(1)}, y_i^{(1)}, -\mathbf{p}/2\right). \tag{21}$$

and

$$I_1(x_i^{(1)}, y_i^{(1)}) = I_2\left(x_i^{(2)}, y_i^{(2)}, \mathbf{p}\right) + \frac{1}{2}\sum_{j=1}^{N_p} \delta p_j \left(\frac{\partial I_2(x_i^{(2)}, y_i^{(2)}, \mathbf{p})}{\partial p_j} + \frac{\partial I_1(x_i^{(1)}, y_i^{(1)})}{\partial p_j}\right)$$
$$+ \frac{1}{8}\sum_{j,s=1}^{N_p} \delta p_j \delta p_s \left(\frac{\partial^2 I_2(x_i^{(2)}, y_i^{(2)}, \mathbf{p})}{\partial p_j \partial p_s} - \frac{\partial^2 I_1(x_i^{(1)}, y_i^{(1)})}{\partial p_j \partial p_s}\right). \tag{22}$$

where N_p is the number of motion parameters. Given that the images I_1 and I_2 are related by an unknown motion $\mathbf{p}$, and as both sides of Eq. 21 are expanded using a second order Taylor expansion, they approximate a middle point (in the parameters space) which is $\delta_1 = \pm\mathbf{p}/2$ apart from both images. In contrast the O3GM scheme used a single approximation over the interval $\delta_2 = \mathbf{p}$. Recalling that the approximation error is related to $\|\delta_i\|^3$, we get that both sides of Eq. 21 are associated with an error of $\frac{\|\mathbf{p}\|^3}{8}$ and the overall error is bounded by $\frac{\|\mathbf{p}\|^3}{4}$ compared to $\|\mathbf{p}\|^3$ for the O3GM. Similar results were derived for the GM scheme in [14]. By constructing the above equation for all of the pixels in the common support between I_1 and I_2 we derive a quadratic equation set which is solved by either of the two schemes described in Section 4. In practice, the only implementation difference between the O3GM and the SO3GM is the construction of the quadratic equation set.

6 Convergence properties of the third order GM

In this section we compute the convergence properties of the proposed scheme based on the analysis given in Section 3. For the proposed scheme we use $T = 2$ and compute Eqs. 13 and 17 to evaluate the convergence rates for small and large deviations, respectively. For large deviations we evaluate Eq. 17 and get

$$\|\varepsilon_{\mathbf{n+1}}\| \leq \frac{C_{T+1}}{C_2}\|\varepsilon_{\mathbf{n}}\|^{\frac{T+1}{2}} = \frac{C_3}{C_2}\|\varepsilon_{\mathbf{n}}\|^{\frac{3}{2}} \tag{23}$$

which improves the linear convergence rate achieved by the GM and SGM. For small deviations we evaluate Eq. 13 and have

$$\|\varepsilon_{\mathbf{n+1}}\| \leq \frac{C_{T+1}}{C_1}\|\varepsilon_{\mathbf{n}}\|^2 + \frac{\|f(\mathbf{p}^*)\|}{C_1} = \frac{C_3}{C_1}\|\varepsilon_{\mathbf{n}}\|^3 + \frac{\|f(\mathbf{p}^*)\|}{C_1} \tag{24}$$

Hence, for images with low optimal alignment error ($\|f(\mathbf{p}^*)\| \to 0$), the proposed scheme achieves third-order convergence compared to the second order rate of the GM and SGM schemes. While for $\|f(\mathbf{p}^*)\| \gg 0$ the scheme might diverge. In practice, the input images I_1 and I_2 can be normalized to lower $\|f(\mathbf{p}^*)\|$ and the cases where the changes in intensities are extreme ($\|f(\mathbf{p}^*)\| \gg 0$) are handled by multi-sensor [13] and multi modality [17] registration schemes. Equations 23 implies that $\|f(\mathbf{p}^*)\|$ influences only the small deviation phase, as for large deviations the convergence rate is dominated by the approximation error.

To analyze the convergence range $\|\varepsilon_0\|$, we study Eq. 18 and compare the range for the GM ($T = 1$) and O3GM schemes ($T = 2$). For the GM we get

$$\|\varepsilon_0^{GM}\| \approx \sqrt[\frac{1}{2}]{\frac{C_2}{C_2}} = 1. \tag{25}$$

while for the O3GM we get

$$\left\| \varepsilon_0^{O3GM} \right\| \approx \frac{C_2}{C_3}. \tag{26}$$

An improved convergence rate is achieved for $C_3 < C_2$, which corresponds to registering smooth input images. As the registered images are smoothed to avoid spatio-temporal aliasing [11, 12] and to allow accurate derivative computation, this condition is fulfilled and no extra smoothing is needed.

The convergence properties of the SO3GM can be derived by considering Eq. 22, which defines the derivatives of $\widehat{f}$, the objective function minimized by the SO3GM. Since the SO3GM used the same order of approximation as the O3GM, we get the same orders of convergence for both small and large deviations. The difference lies in the ratio of C_3 and C_2, the overall energy of the derivatives. Denote by $\widetilde{C}_3$, $\widetilde{C}_2$ and $\widetilde{C}_1$ the corresponding derivatives norms for the SO3GM. We compare their value to C_3, C_2 and C_1 used by the O3GM and show that the convergence rate is improved. Using Eq. 22 we get that the norm of $\widetilde{C}_2$ is bounded by

$$\widetilde{C}_2 = \frac{1}{4} \left\| \frac{\partial^2 I_2(x_i^{(2)}, y_i^{(2)}, \mathbf{p})}{\partial p_j \partial p_s} - \frac{\partial^2 I_1(x_i^{(1)}, y_i^{(1)})}{\partial p_j \partial p_s} \right\| \leq \frac{1}{4} \left\| \frac{\partial^2 I_2(x_i^{(2)}, y_i^{(2)}, \mathbf{p})}{\partial p_j \partial p_s} \right\| + \frac{1}{4} \left\| \frac{\partial^2 I_1(x_i^{(1)}, y_i^{(1)})}{\partial p_j \partial p_s} \right\| \approx \frac{C_2}{2}.$$

For $\widetilde{C}_3$ and $\widetilde{C}_1$ we have

$$\widetilde{C}_3 = \frac{1}{8} \left\| \frac{\partial^3 I_2\left(x_i^{(2)}, y_i^{(2)}, \widetilde{\mathbf{p}}\right)}{\partial p_j \partial p_s \partial p_t} + \frac{\partial^3 I_1\left(x_i^{(1)}, y_i^{(1)}, \widehat{\mathbf{p}}\right)}{\partial p_j \partial p_s \partial p_t} \right\| \leq \frac{1}{8} \left\| \frac{\partial^3 I_2\left(x_i^{(2)}, y_i^{(2)}, \widetilde{\mathbf{p}}\right)}{\partial p_j \partial p_s \partial p_t} \right\| + \frac{1}{8} \left\| \frac{\partial^3 I_1\left(x_i^{(1)}, y_i^{(1)}, \widehat{\mathbf{p}}\right)}{\partial p_j \partial p_s \partial p_t} \right\| = \frac{C_3}{4},$$

and

$$\widetilde{C}_1 = \frac{1}{2} \left\| \frac{\partial I_2(x_i^{(2)}, y_i^{(2)}, \mathbf{p})}{\partial p_j} + \frac{\partial I_1(x_i^{(1)}, y_i^{(1)})}{\partial p_j} \right\| \leq \frac{1}{2} \left\| \frac{\partial I_2(x_i^{(2)}, y_i^{(2)}, \mathbf{p})}{\partial p_j} \right\| + \frac{1}{2} \left\| \frac{\partial I_2(x_i^{(2)}, y_i^{(2)}, \mathbf{p})}{\partial p_j} \right\| = C_1.$$

Hence, we compute the convergence rates and ranges for the SO3GM. For large deviations we get

$$\left\| \varepsilon_{\mathbf{n+1}} \right\| \leq \frac{\widetilde{C}_3}{\widetilde{C}_2} \left\| \varepsilon_{\mathbf{n}} \right\|^{\frac{3}{2}} = \frac{C_3/4}{C_2/2} \left\| \varepsilon_{\mathbf{n}} \right\|^{\frac{3}{2}} = \frac{1}{2} \frac{C_3}{C_2} \left\| \varepsilon_{\mathbf{n}} \right\|^{\frac{3}{2}} \tag{27}$$

and for small deviations

$$\left\| \varepsilon_{\mathbf{n+1}} \right\| \leq \frac{\widetilde{C}_3}{\widetilde{C}_1} \left\| \varepsilon_{\mathbf{n}} \right\|^3 + \frac{\|f(\mathbf{p}^*)\|}{\widetilde{C}_1} = \frac{C_3/4}{C_1} \left\| \varepsilon_{\mathbf{n}} \right\|^3 + \frac{\|f(\mathbf{p}^*)\|}{C_1} = \frac{1}{4} \frac{C_3}{C_1} \left\| \varepsilon_{\mathbf{n}} \right\|^3 + \frac{\|f(\mathbf{p}^*)\|}{C_1} \tag{28}$$

Thus, we see that the SO3GM allows better convergence rates, especially for low deviations. Next we show that the convergence range $\left\| \varepsilon_0^{SO3GM} \right\|$ is also improved. By substituting $\widetilde{C}_2$ and $\widetilde{C}_3$ in Eq. 26 we have

$$\left\| \varepsilon_0^{SO3GM} \right\| \approx \frac{\widetilde{C}_2}{\widetilde{C}_3} = \frac{C_2/2}{C_3/4} = 2\frac{C_2}{C_3} = 2 \left\| \varepsilon_0^{O3GM} \right\|. \tag{29}$$

7 Experimental Results

This section describes the performance of the proposed algorithms and verifies the convergence analysis given in Sections 3 and 6. The same implementations of the *iterative refinement* and *multiscale embedding* were used for the O3GM, SO3GM, SGM and GM algorithms. Thus, the only difference between the schemes is the *single iteration module*. The affine and projective motion were tested using real images. The images in Fig. 1 were taken by a photogrammetric aerial camera, while the images in Fig. 2 were taken by a hand held 35mm camera. The GM algorithm was implemented according to [11] which is considered a state-of-the-art implementation. The first and second order derivatives were computed using central differences. In order to avoid spatiotemporal aliasing and allow accurate computation of the spatial derivatives, the images were initially smoothed by a 5 tap Gaussian filter with a bandwidth of $\sigma = 2$. The same filter was used to construct the multiscale pyramid whose scales were 1 (original scale) and 1/3.

The common support of I_1 and I_2 was computed in each iteration by applying the current motion estimate to the parametric equations defining the bounding rectangular of I_2. Next we identify the intersection points between

the bounding rectangles of I_1 and (the transformed) I_2. Finally we compute the convex hull of their mutual support and substitute the pixels in it in Eqs. 19 and 22, for the O3GM and SO3GM, respectively. The focal point of the experiments is to illustrate the improved convergence properties of the proposed scheme when estimating large motions. Although we also show that the scheme is accurate, in practice, due to its high computational complexity, it should only be applied when the estimated motion is large. In the vicinity of the true solution (when the estimated motion becomes small) one should switch to the regular GM which provides a better trade-off of computational complexity Vs. convergence. Finally, as higher order derivatives are known to be sensitive to noise, we address the convergence in the presence of noise in Section 7.2.

7.1 Affine and projective motion estimation

The registration of real images using the affine and projective motion models is given in Figs. 1 and 2, respectively. The initial estimate of the motion was given as translations, computed by aligning the X marks in both images. We intentionally chose an inaccurate initial estimate, making the residual motion (estimated by the various GM algorithms) large. The same initial motion was used by all the different schemes. For these real images, the final alignment error results from the lack of perfect matching, and the existence of non corresponding objects in both images. In the affine case in Fig. 1, the O3GM outperformed the GM by converging twice as fast. Better convergence was achieved by the symmetric motion models (SGM and SO3GM) and the best convergence was achieved by the SO3GM. The initial estimate (based on the translation between the X signs in Fig. 1) was $(\delta x, \delta y) = (231, -17)$. The results of registering the projective images are given in Fig. 2. These images have a significantly different brightness due to the auto-exposure of the camera. We intentionally left the brightness as is, to make registration more difficult. The initial estimate was $(\delta x, \delta y) = (583, -43)$. The timing results for the affine case are given in Table 1, where the measurement were taken on a 2.8 GHz PC computer and the algorithm was implemented using non optimized C++. For these high-order models, the complexity of the proposed algorithms is higher than the computational complexity of the GM and SGM. Yet, it is applicable in cases where the GM and SGM diverge, and such examples are shown in Fig. 2.
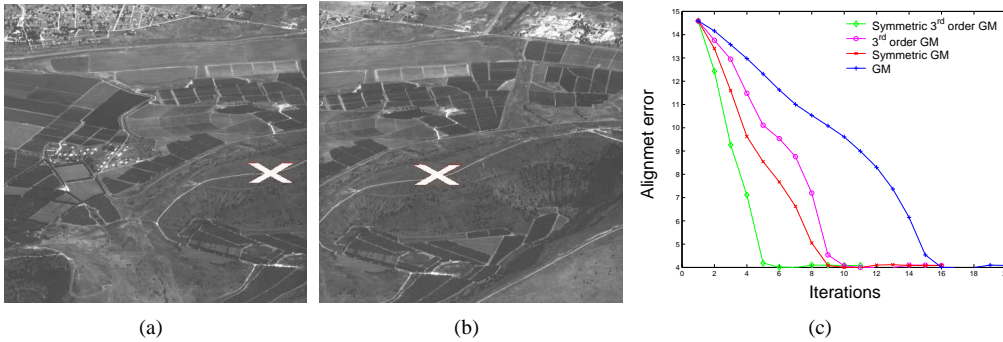

| (a) | (b) | (c) |

Figure 1: Registration results for affine motion. The initial estimate of the motion was given by marked by the red X. The SO3GM converged 4 times faster than the GM and twice as fast as the O3GM and SGM.

	Iteration#	Total timing [s]	Average iteration timing [s]
GM	22	0.8	0.036
SGM	17	1.1	0.064
O3GM	17	1.2	0.070
SO3GM	12	1.4	0.117

Table 1: Timing results for the affine registration given in Fig. 1

7.2 Sensitivity to noise

Since higher order schemes are known to be sensitive to noise, we tested the convergence in the presence of noise. White Gaussian noise (WGN) was added to both the image in Fig. 1 with $\sigma = 0, 30, 60$. The GM and SGM

(a)   (b)   (c)

Figure 2: Registration results for projective motion. The initial estimate of the motion was given by the red X. The SGM converged better than the O3GM while the SO3GM showed the best convergence properties.

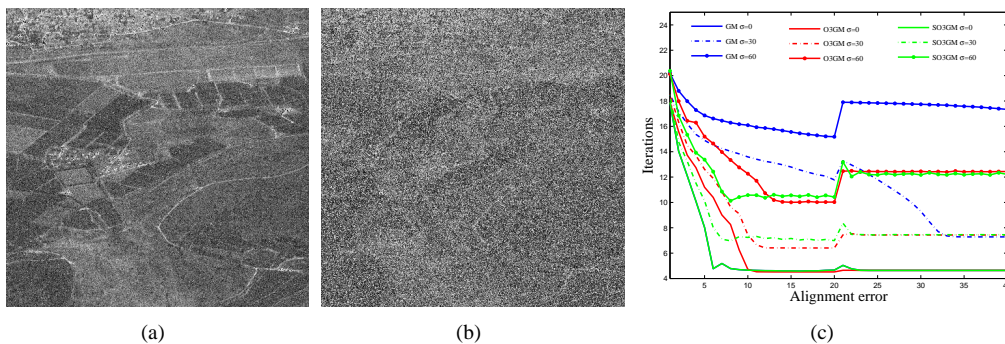
(a)   (b)   (c)

Figure 3: Estimating affine motion of noisy images. The O3GM and SO3GM allow better registration of noisy images, compared to the regular GM. The initial estimate is the same as in Fig. 1.

schemes were applied and the results are depicted in Fig. 3. The same initial motion was assumes as in the prior section. Figs. 3a and 3b show Fig. 1a after adding WGN with $\sigma = 30$ and $\sigma = 60$, respectively. The initial estimate of the motion, was chosen such that the residual motion was large. The convergence shown in Fig. 3c, shows that the O3GM was able to converge better for all noise levels and thus verify the convergence analysis - in all cases the O3GM and SO3GM converged faster (iteration-wise) than the regular GM and in most cases, it was able to converge where the regular GM diverged.

8   Conclusions and future work

In this work we presented the O3GM and SO3GM image registration algorithms which enhance the performance of gradient based registration methods. These algorithms extend the current state-of-the-art schemes and were shown to have superior convergence range and rate. They are especially suitable for the estimation of large affine and projective motions which can not be bootstrapped by Fourier domain methods. Future work includes the application of the O3GM and SO3GM to numerically ill-posed computer vision problems which are based on the gradient methods, such as wide baseline stereo and 3D reconstruction.

References

[1] Y. Altunbasak, R.M. Mersereau, and A.J. Patti. A fast parametric motion estimation algorithm with illumination and lens distortion correction. *IEEE Transactions on Image Processing*, 12(4):395–408, April 2003.

[2] J.L. Barron, D.J. Fleet, and S.S. Beauchemin. Performance of optical flow techniques. *International Journal of Computer Vision*, 12(1):43–77, February 1994.

[3] J.R. Bergen, P. Anandan, K.J. Hanna, and R. Hingorani. Hierarchical model-based motion estimation. In *European Conference on Computer Vision*, pages 237–252, 1992.

9

[4] M.J. Black and A.D. Jepson. Eigentracking: Robust matching and tracking of articulated objects using a view-based representation. In *European Conference on Computer Vision*, pages 329–342, 1996.

[5] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[6] T. Brox, A. Bruhn, N. Papenberg, and J. Weickert:. High accuracy optical flow estimation based on a theory for warping. In *Proceedings of the ECCV'04*, volume 3024, pages 25–36, 2004.

[7] F.M. Candocia. Simultaneous homographic and comparametric alignment of multiple exposure-adjusted pictures of the same scene. *IEEE Transactions on Image Processing*, 12(12):1485–1494, December 2003.

[8] F. Dufaux and J. Konrad. Efficient, robust, and fast global motion estimation for video coding. *IEEE Transactions on Image Processing*, 9(3):497–501, March 2000.

[9] Vittorio Ferrari, Tinne Tuytelaars, and Luc Van Gool. Wide-baseline multiple-view correspondences. In *Proceedings, IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pages I–718 – I–725, June 2003.

[10] P. Gill. *Practical Optimization*. Academic Press, 1982.

[11] M. Irani and P. Anandan. *Vision Algorithms: Theory and Practice: International Workshop on Vision Algorithms, Corfu, Greece, September 1999. Proceedings*, chapter All about direct methods, pages 267–278. Springer-Verlag, 1999.

[12] M. Irani and S. Peleg. Motion analysis for image enhancement: Resolution, occlusion and transparency. *Journal of Visual Communication and Image Representation*, 4(4):324–335, December 1993.

[13] Michal Irani and P. Anandan. Robust multi-sensor image alignment. In *Proc. IEEE Int. Conf. Computer Vision*, pages 959–966, January 1998.

[14] Y. Keller and A. Averbuch. Fast motion estimation using bi-directional gradient methods. *IEEE Transactions on Image Processing*, 13(8):1042–1054, August 2004.

[15] C. D. Kuglin and D. C. Hines. The phase correlation image alignment method. *IEEE Conference on Cybernetics and Society*, pages 163–165, September 1975.

[16] B. D. Lucas and Takeo Kanade. An iterative image registration technique with an application to stereo vision. In *International Joint Conference on Artificial Intelligence*, pages 674–679, 1981.

[17] Frederik Maes, André Collignon, Dirk Vandermeulen, Guy Marchal, and Paul Suetens. Multimodality image registration by maximization of mutual information. *IEEE Trans. Med. Imaging*, 16(2):187–198, 1997.

[18] Steve Mann and Rosalind Picard. Virtual bellows: constructing high quality stills from video. pages 363–367, Austin, TX, November 13-16 1994. IEEE International Conference Image Processing.

[19] Shmuel Peleg, Benny Rousso, Alex Rav-Acha, and Assaf Zomet. Mosaicing on adaptive manifolds. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(10):1144–1154, 2000.

[20] S. Reddy and B. N. Chatterji. An FFT-based technique for translation, rotation, and scale-invariant image registration. *IEEE Transactions on Image Processing*, 3(8):1266–1270, August 1996.

[21] Amnon Shashua and Yonatan Wexler. Q-warping: Direct computation of quadratic reference surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(8):920–925, 2001.

[22] G.P. Stein and A. Shashua. Model-based brightness constraints: on direct estimation of structure and motion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(9):992–1015, September 2000.

[23] R. Szeliski. Image mosaicking for tele-reality applications. In *Proc. of IEEE Workshop on Applications of Computer Vision*, pages 44–53, May 1994.