Groupwise Geometric and Photometric Direct Image Registration

Adrien Bartoli

BMVC'06 – IN PROCEEDINGS OF THE SEVENTEENTH BRITISH MACHINE VISION CONFERENCE, EDINBURGH, UK, SEPTEMBER 2006

Groupwise Geometric and Photometric Direct Image Registration

Adrien Bartoli ◇ Adrien.Bartoli@gmail.com
CNRS – LASMEA ◇ Clermont-Ferrand, France
http://comsee.univ-bpclermont.fr

Abstract

Image registration consists in estimating geometric and photometric transformations that align two images as best as possible. The direct approach consists in minimizing the discrepancy in the intensity or color of the pixels. The inverse compositional algorithm has been recently proposed for the direct estimation of groupwise geometric transformations. It is efficient in that it performs several computationally expensive calculations at a pre-computation phase.

We propose the dual inverse compositional algorithm which deals with groupwise geometric and photometric transformations, the latter acting on the value of the pixels. Our algorithm preserves the efficient pre-computation based design of the original inverse compositional algorithm. Previous attempts at incorporating photometric transformations to the inverse compositional algorithm spoil this property.

We demonstrate our algorithm on simulated and real data and show the improvement in computational efficiency compared to previous algorithms.

1 Introduction

Image registration is the task of applying some transformations to two images so that they match as best as possible. This can be seen as the computation of some geometric transformation, for example an homography, used to deform the images to model camera pose, and some photometric transformation, applied to the intensity or color of the pixels, to account *e.g.* for lighting change.

Image registration has been an important research topic for the past decades. It is central to many tasks in computer vision, medical imaging, augmented reality and robotics. Applications include image mosaicing [7, 10], object and feature tracking, *e.g.* [5, 8].

Broadly speaking, two approaches have been proposed: The feature-based and the direct approaches. The feature-based approach, see *e.g.* [11], relies on abstracting the input images by the geometric location of a set of carefully chosen, salient features. The direct approach, see *e.g.* [7], uses the value of all pixels of interest. The *inverse compositional algorithm* of Baker *et al.* [2] estimates groupwise geometric transformations such as homographies[1]. It has been shown to be one of the most reliable and computationally effi-

[1] To be precise, transformations parameterized such that there is a group structure on the parameter vector.

cient registration method. The efficiency stems from the so-called *inverse compositional trick*, making constant the Hessian matrix involved in the linear least squares problem to be solved at each iteration. This makes it possible to pre-compute its inverse.

This paper is about the registration of two images related by a geometric and a photometric transformation. An example of photometric transformation is 'gain and bias' which rescales and offsets the value of the pixels. We propose the *dual inverse compositional algorithm* which uses the inverse compositional trick for both the geometric and photometric counterparts of the registration, thereby preserving the possibility of pre-computing the inverse of the Hessian matrix.

Paper organization. We formally state the problem and review previous work in §2. We present as background material the inverse compositional algorithm of Baker *et al.* in §3. We propose the dual inverse compositional algorithm in §4. We report experimental results on simulated and real data in §5. A conclusion is provided in §6.

Notation. Vectors are denoted using bold fonts, *e.g.* $\mathbf{q}$, matrices using sans-serif fonts, *e.g.* E, and scalars in italics, *e.g.* a. The two-norm of a vector $\mathbf{r}$ is written $\|\mathbf{r}\|$. The gradient of a scalar-valued function f, in other words, its partial derivative vector, with respect to vector $\mathbf{x}$, is denoted $\nabla_{\mathbf{x}} f$. It is evaluated at $\mathbf{0}$, unless specified as in $\nabla f|_{\mathbf{x}_O}$. For vector-valued functions, $\blacktriangledown$ gives the Jacobian matrix, *i.e.* the matrix containing all the partial derivatives of the function. Columnwise matrix vectorization is written vect.

The source and target images to be registered are denoted $\mathscr{S}$ and $\mathscr{T}$ respectively. They are seen as functions from $\mathbb{R}^2$ to $\mathbb{R}^c$ where c is the number of channels, *i.e.* $c = 1$ in the grey-level case and $c = 3$ in the color case. For instance, $\mathscr{T}[\mathbf{q}]$ is the image value at pixel $\mathbf{q} \in \mathbb{R}^2$. Bilinear interpolation is used for sub-pixel coordinates. The unit column vector is denoted $\mathbf{1}$ with length obvious from the context. The geometric and photometric transformations are respectively denoted $\mathscr{G}$ and $\mathscr{P}$, with respective parameter vectors $\mathbf{g}$ and $\mathbf{p}$. The geometric transformation is also called the warp. By 'Hessian matrix' we mean the Gauss-Newton approximation to the true Hessian matrix.

2 Problem Statement and Previous Work

The geometric registration problem is the minimization of a nonlinear least squares error function, given by the discrepancy in the value of the pixels, between the source image $\mathscr{S}$ and the target image $\mathscr{T}$ warped onto the first one by the unknown geometric transformation. The geometric transformation maps a pixel $\mathbf{q}$ in the region of interest $\mathscr{R}$ defined in the source image to the corresponding pixel $\mathscr{G}(\mathbf{q};\mathbf{g})$ in the target image. We expect that given an 'appropriate' parameter vector $\mathbf{g}$, $\mathscr{S}[\mathbf{q}]$ is 'close to' $\mathscr{T}[\mathscr{G}(\mathbf{q};\mathbf{g})]$, for all $\mathbf{q} \in \mathscr{R}$: This is the brightness constancy assumption, see *e.g.* [6]. The direct image registration problem is thus formally posed as:

$$\min_{\mathbf{g}} \sum_{\mathbf{q}\in\mathscr{R}} \|\mathscr{S}[\mathbf{q}] - \mathscr{T}[\mathscr{G}(\mathbf{q};\mathbf{g})]\|^2 . \tag{1}$$

Note that other error functions can be used, to deal for example with occlusions, see *e.g.* [5]. Most algorithms linearize each term in the transformation parameters $\mathbf{g}$, and iteratively update an initial guess by solving linear least squares problems. The popular

Lucas-Kanade algorithm [9] falls into this category. Sawhney *et al.* [10] show how lens distortion can be estimated along with homographies. In [2], Baker *et al.* propose the efficient *inverse compositional algorithm* for solving problem (1). More details are given in the next section.

Problem (1) does not take into account photometric changes, *i.e.* changes in the pixel values. These changes occur for example when lighting change between the acquisition of the two images or when two different cameras are used. Incorporating a photometric transformation gives:

$$\min_{\mathbf{g},\mathbf{p}} \sum_{\mathbf{q}\in\mathscr{R}} \|\mathscr{S}[\mathbf{q}] - \mathscr{P}(\mathscr{T}[\mathscr{G}(\mathbf{q};\mathbf{g})];\mathbf{p})\|^2. \tag{2}$$

A commonly employed photometric model $\mathscr{P}$ is an affine transformation modeling gain and bias, and accounting for global intensity change:

$$\mathscr{P}(\mathbf{v};\mathbf{p}) = a\mathbf{v} + b\mathbf{1} \quad \text{with} \quad \mathbf{p}^\mathsf{T} = (a\ b). \tag{3}$$

For instance, Jin *et al.* [8] uses this model for feature tracking in grey-level images.

Baker *et al.* extend the inverse compositional algorithm in [1] to deal with linear appearance variations of the source image. Their *simultaneous inverse compositional algorithm* can be straightforwardly specialized to estimate gain and bias. In their framework however, and since the photometric transformation is applied to the source image, the Hessian matrix varies accross the iterations, thereby spoiling the computational efficiency of the inverse compositional algorithm: The simultaneous inverse compositional algorithm re-estimates and inverts the Hessian matrix at each iteration. Baker *et al.* also propose several approximations to reduce the computational cost. They show that these approximations do not behave well for high gain values. Bartoli [3] shows that the Hessian matrix has a block structure with blocks constant up to some scale factors, depending on the gain. From this analysis, he derives an algorithm allowing one to pre-compute a blockwise inverse of the Hessian matrix. The normal equations are then solved by simply multiplying the right hand side by some constant, appropriately rescaled matrices, which is computationally efficient. Unfortunatelly, this algorithm does not extend to color images.

3 The Inverse Compositional Algorithm

This section is devoted to the description of the inverse compositional algorithm of Baker *et al.* [2] for the computation of groupwise geometric registrations. This algorithm forms the basis for our dual inverse compositional algorithm, presented in the next section.

The algorithm is illustrated in figure 1. Its advantages are two-fold. First, it converges rapidly compared to other optimization schemes. Second, as already mentioned, each iteration is performed efficiently.

3.1 Derivation

The inverse compositional algorithm iteratively updates an initial guess of the sought-after transformation. The key idea is to express the updated transformation as the composition of the current transformation $\mathscr{G}(\cdot;\mathbf{g})$ and an incremental transformation $\mathscr{G}^{-1}(\cdot;\boldsymbol{\delta}_g)$. The

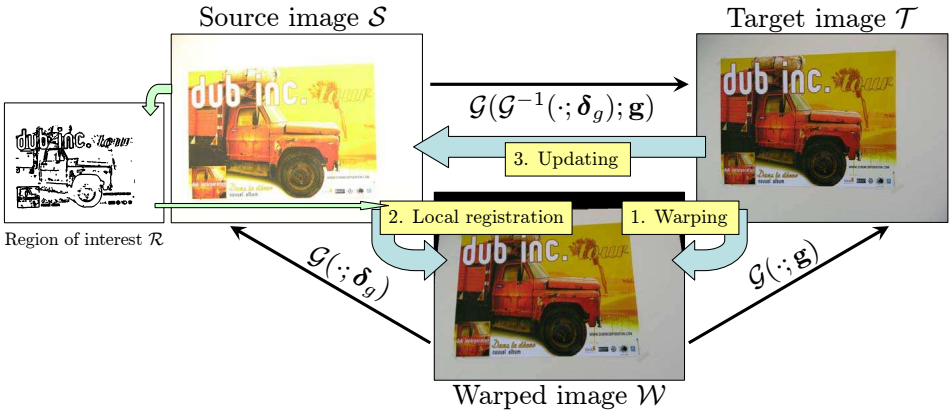

Figure 1: The *inverse compositional algorithm* efficiently computes a geometric transformation $\mathcal{G}(\cdot;\mathbf{g})$ by iterating the three main steps mentioned in the figure.

optimization is performed over $\boldsymbol{\delta}_g$, the parameter vector of the incremental warp, instead of $\mathbf{g}$. The geometric registration problem (1) is thus rewritten as:

$$\min_{\boldsymbol{\delta}_g} \sum_{\mathbf{q}\in\mathcal{R}} \|\mathcal{S}[\mathbf{q}] - \mathcal{T}[\mathcal{G}(\mathcal{G}^{-1}(\mathbf{q};\boldsymbol{\delta}_g);\mathbf{g})]\|^2.$$

Let $\mathcal{W}$ be the warped target image, *i.e.* $\mathcal{W}[\mathbf{q}] = \mathcal{T}[\mathcal{G}(\mathbf{q};\mathbf{g})]$. The incremental transformation is then applied to the source image, instead of the target one, leading to:

$$\min_{\boldsymbol{\delta}_g} \sum_{\mathbf{q}\in\mathcal{R}} \|\mathcal{S}[\mathcal{G}(\mathbf{q};\boldsymbol{\delta}_g] - \mathcal{W}[\mathbf{q}]\|^2. \tag{4}$$

Note that problem (4) is an approximation to the original problem (1) since the error function is expressed within the warped and not within the source image. The error function is linearized by first order Taylor expansion in $\boldsymbol{\delta}_g$, forming a Gauss-Newton approximation, $\min_{\boldsymbol{\delta}_g} \sum_{\mathbf{q}\in\mathcal{R}} \|\mathcal{S}[\mathbf{q}] + \mathsf{L}_g^{\mathsf{T}}[\mathbf{q}]\boldsymbol{\delta}_g - \mathcal{W}[\mathbf{q}]\|^2$. This is a linear least squares problem, that is solved via its normal equations. Defining $(\blacktriangledown\mathcal{S})|_{\mathbf{q}}$ as the $(2\times c)$ Jacobian matrix of the source image at $\mathbf{q}$ and $(\blacktriangledown_{\mathbf{g}}\mathcal{G})|_{\mathbf{q};\mathbf{0}}$ as the Jacobian matrix of the warp, evaluated at $\mathbf{q}$ and at warp parameters $\mathbf{0}$, representing the identity warp, see below, the Jacobian matrices $\mathsf{L}_g[\mathbf{q}]$ are obtained using the chain rule as $\mathsf{L}_g^{\mathsf{T}}[\mathbf{q}] = (\blacktriangledown\mathcal{S})|_{\mathbf{q}}{}^{\mathsf{T}}\,(\blacktriangledown_{\mathbf{g}}\mathcal{G})|_{\mathbf{q};\mathbf{0}}$. They only depend on the source image at the pixels of interest, and are thus constant over the iterations. Let $\mathcal{D}[\mathbf{q}] = \mathcal{W}[\mathbf{q}] - \mathcal{S}[\mathbf{q}]$ be the difference image, the normal equations are:

$$\mathsf{E}_g\boldsymbol{\delta}_g = \mathbf{b}_g \quad \text{with} \quad \mathsf{E}_g = \sum_{\mathbf{q}\in\mathcal{R}} \mathsf{L}_g[\mathbf{q}]\mathsf{L}_g^{\mathsf{T}}[\mathbf{q}] \quad \text{and} \quad \mathbf{b}_g = \sum_{\mathbf{q}\in\mathcal{R}} \mathsf{L}_g[\mathbf{q}]\mathcal{D}[\mathbf{q}].$$

The solution $\boldsymbol{\delta}_g = \mathsf{E}_g^{-1}\mathbf{b}_g$ for the local warp parameters is thus computed very efficiently since the Jacobian matrices $\mathsf{L}_g[\mathbf{q}]$ as well as the inverse of the Hessian matrix E_g can be pre-computed.

Once $\boldsymbol{\delta}_g$ has been computed, parameters $\mathbf{g}$ are updated by composing the current warp with the incremental warp. We write the warp update rule as $\mathbf{g} \leftarrow \mathcal{U}_g(\mathbf{g},\boldsymbol{\delta}_g)$, see below.

The process is iterated until convergence, determined, in our experiments, by thresholding $\|\delta_g\|$ by $\varepsilon = 10\text{e}-8$.

3.2 Parameterizing Homographies

Groupwise geometric transformations include translations, rotations, affinities and homographies. We describe the case of homographies. They have 8 degrees of freedom, and can be represented by (3×3) homogeneous matrices (*i.e.* defined up to scale). Let H be such an homography matrix, representing the current estimate of the geometric registration. Following [2], the local homography matrix is parameterized by an 8-vector δ_h as:

$$\Delta\mathsf{H} \sim \mathrm{I} + \begin{pmatrix} \delta_{h,1} & \delta_{h,2} & \delta_{h,3} \\ \delta_{h,4} & \delta_{h,5} & \delta_{h,6} \\ \delta_{h,7} & \delta_{h,8} & 0 \end{pmatrix}.$$

This parameterization is such that $\delta_h = \mathbf{0}$ gives $\Delta\mathsf{H} = \mathrm{I}$, as required. The inverse composition is performed by multiplying the current homography matrix to the right by the inverse of the incremental one: $\mathscr{U}_g(\mathsf{H}, \delta_h) = \mathsf{H} \cdot (\Delta\mathsf{H})^{-1}$. The (2×8) Jacobian matrix of the warp is straightforward to derive.

4 The Dual Inverse Compositional Algorithm

We extend the inverse compositional algorithm to estimate a groupwise photometric transformation along with a geometric one, as stated in problem (2). The algorithm is summarized in table 1 and illustrated in figure 2.

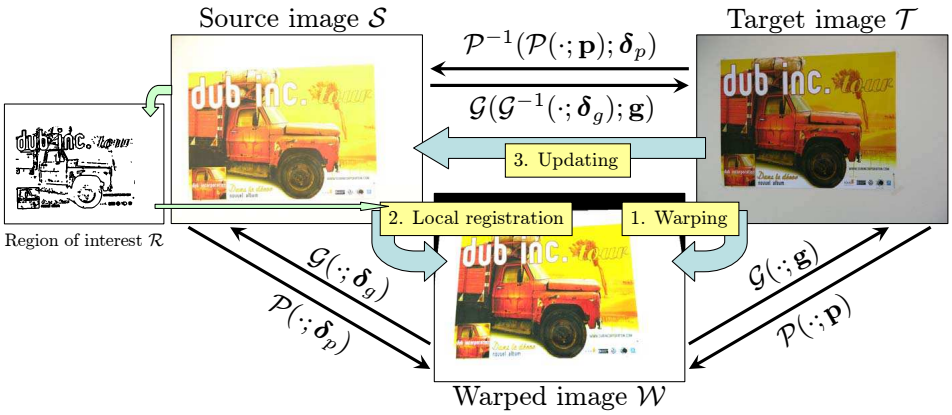

Figure 2: The *dual inverse compositional algorithm* extends the inverse compositional algorithm to jointly compute a geometric and a photometric registration , $\mathscr{G}(\cdot;\mathbf{g})$ and $\mathscr{P}(\cdot;\mathbf{p})$, by iterating the three main steps mentioned in the figure.

4.1 Derivation

Considering problem (2), and applying the inverse compositional trick as in the previous section, but for both the geometric and photometric transformations gives:

$$\min_{\delta_g,\delta_p} \sum_{\mathbf{q}\in\mathscr{R}} \|\mathscr{S}[\mathbf{q}] - \mathscr{P}^{-1}(\mathscr{P}(\mathscr{T}[\mathscr{G}(\mathscr{G}^{-1}(\mathbf{q};\delta_g);\mathbf{g})];\mathbf{p});\delta_p)\|^2. \tag{5}$$

The optimization is now to be performed on the incremental parameters δ_g and δ_p, the latter representing the updating transformation for the photometric registration. Note that the update rule employed for the photometric registration is different from the one used for the geometric registration, *i.e.* the composition is in the reverse order. Let $\mathscr{W}$ be the warped image, in the geometric and photometric sense, *i.e.* $\mathscr{W}[\mathbf{q}] = \mathscr{P}(\mathscr{T}[\mathscr{G}(\mathbf{q};\mathbf{g})];\mathbf{p})$. Applying the incremental transformations to the source image instead of the target image gives:

$$\min_{\delta_g,\delta_p} \sum_{\mathbf{q}\in\mathscr{R}} \|\mathscr{P}(\mathscr{S}[\mathscr{G}(\mathbf{q};\delta_g)];\delta_p) - \mathscr{W}[\mathbf{q}]\|^2. \tag{6}$$

We show below that the normal equations induced by the Gauss-Newton approximation have a constant Hessian matrix. Linearizing using first order Taylor expansion yields:

$$\min_{\delta_g,\delta_p} \sum_{\mathbf{q}\in\mathscr{R}} \|\mathscr{S}[\mathbf{q}] + (\mathsf{L}_g^\mathsf{T}[\mathbf{q}]\ \mathsf{L}_p^\mathsf{T}[\mathbf{q}])\delta_{gp} - \mathscr{W}[\mathbf{q}]\|^2.$$

We denote the joint incremental parameter vector $\delta_{gp}^\mathsf{T} = (\delta_g^\mathsf{T}\ \delta_p^\mathsf{T})$ and the joint Jacobian matrices by $\mathsf{L}_{gp}^\mathsf{T}[\mathbf{q}] = (\mathsf{L}_g^\mathsf{T}[\mathbf{q}]\ \mathsf{L}_p^\mathsf{T}[\mathbf{q}])$. The Jacobian matrices $\mathsf{L}_p[\mathbf{q}]$ for the photometric parameters are $\mathsf{L}_p[\mathbf{q}] = (\blacktriangledown_\mathbf{p}\mathscr{P})|_{\mathscr{S}[\mathbf{q}];\mathbf{0}}$. As in the original inverse compositional algorithm, the Jacobian matrices only depend on the source image at the pixels of interest. They are thus constant as well as the Hessian matrix E_{gp} of the normal equations:

$$\mathsf{E}_{gp}\delta_{gp} = \mathbf{d}_{gp} \quad \text{with} \quad \mathsf{E}_{gp} = \sum_{\mathbf{q}\in\mathscr{R}} \mathsf{L}_{gp}[\mathbf{q}]\mathsf{L}_{gp}^\mathsf{T}[\mathbf{q}] \quad \text{and} \quad \mathbf{d}_{gp} = \sum_{\mathbf{q}\in\mathscr{R}} \mathsf{L}_{gp}[\mathbf{q}]\mathscr{D}[\mathbf{q}].$$

The warp is updated as in the inverse compositional algorithm: $\mathbf{g} \leftarrow \mathscr{U}_g(\mathbf{g},\delta_g)$. The photometric transformation update rule is written $\mathbf{p} \leftarrow \mathscr{U}_p(\mathbf{p},\delta_p)$. Note that the incremental transformation is composed to the left and not to the right of the current transformation, contrarily to the case of the warp. As an example, the update rule and the Jacobian matrices for the gain and bias photometric registration (3) are:

$$\begin{pmatrix} a \\ b \end{pmatrix} \leftarrow \frac{1}{1+\delta_a}\begin{pmatrix} a \\ b-\delta_b \end{pmatrix} \quad \text{and} \quad \mathsf{L}_p^\mathsf{T}[\mathbf{q}] = (\mathscr{S}[\mathbf{q}]\ \mathbf{1}).$$

4.2 Some Groupwise Photometric Models

We mention some global photometric transformations that can be employed within our framework. The most common photometric transformation for grey-level images is the aforementioned gain and bias. In the color image case, we use affine transformations, *i.e.* transformations that can be witten as $\mathscr{P}(\mathbf{v};\mathbf{p}) = \mathsf{A}\mathbf{v} + \mathbf{b}$, where A is a (3×3) matrix combining the three color channels, and $\mathbf{b}$ is a 3-vector, modeling a per-channel bias. Finlayson *et al.* [4] show that linear transformations are well adapted for color constancy in practice. We have tried several variants, summarized below:

Model	$\mathscr{P}(\mathbf{v};\mathbf{p})$	$\mathbf{p}^\mathsf{T}$	Number of parameters
Single gain and bias	$a\mathbf{v}+b\mathbf{1}$	$(a\ b)$	2
Per-channel gain and bias	$\mathrm{diag}(\mathbf{a})\mathbf{v}+\mathbf{b}$	$(\mathbf{a}^\mathsf{T}\ \mathbf{b}^\mathsf{T})$	6
Full affine channel mixing	$\mathsf{A}\mathbf{v}+\mathbf{b}$	$\left((\mathrm{vect}(\mathsf{A}))^\mathsf{T}\ \mathbf{b}^\mathsf{T}\right)$	12

The update rules are formulated by multiplying the incremental transformation to the left of the current one. Deriving the Jacobian matrices is then straightforward.

OBJECTIVE

Register a target image $\mathscr{T}$ to a source image $\mathscr{S}$ by computing the parameters $\mathbf{g}$ of a geometric registration $\mathscr{G}(\cdot;\mathbf{g})$ and the parameters $\mathbf{p}$ of a photometric registration $\mathscr{P}(\cdot;\mathbf{p})$ by minimizing the error in intensity or color. Other inputs are the region of interest $\mathscr{R}$ in the source image and an initial value for $\mathbf{g}$ and $\mathbf{p}$. Upon convergence, $\sum_{\mathbf{q}\in\mathscr{R}}\|\mathscr{S}[\mathbf{q}]-\mathscr{P}(\mathscr{T}[\mathscr{G}(\mathbf{q};\mathbf{g})];\mathbf{p})\|^2$ is minimized.

ALGORITHM

Pre-computations

1. The Jacobian matrices $\mathsf{L}_{gp}[\mathbf{q}]=\left((\blacktriangledown\mathscr{S})|_{\mathbf{q}}^\mathsf{T}\ (\blacktriangledown_{\mathbf{g}}\mathscr{G})\big|_{\mathbf{q};\mathbf{0}}\quad(\blacktriangledown_{\mathbf{p}}\mathscr{P})\big|_{\mathscr{S}[\mathbf{q}];\mathbf{0}}\right)$ for $\mathbf{q}\in\mathscr{R}$
2. The Hessian matrix $\mathsf{E}_{gp}=\sum_{\mathbf{q}\in\mathscr{R}}\mathsf{L}_{gp}[\mathbf{q}]\mathsf{L}_{gp}^\mathsf{T}[\mathbf{q}]$ and its inverse

Iterations

1. Warp the target image $\mathscr{T}$ to $\mathscr{W}$ using the estimates $\mathbf{g}$ and $\mathbf{p}$, *i.e.* $\mathscr{W}[\mathbf{q}]=\mathscr{P}(\mathscr{T}[\mathscr{G}(\mathbf{q};\mathbf{g})];\mathbf{p})$
2. Compute the incremental transformations:
   - Compute the error image $\mathscr{D}[\mathbf{q}]=\mathscr{W}[\mathbf{q}]-\mathscr{S}[\mathbf{q}]$
   - Compute the right hand side of the normal equations $\mathbf{b}_{gp}=\sum_{\mathbf{q}\in\mathscr{R}}\mathsf{L}_{gp}[\mathbf{q}]\mathscr{D}[\mathbf{q}]$
   - Solve for the incremental transformations $\delta_{gp}=\mathsf{E}_{gp}^{-1}\mathbf{b}_{gp}$
3. Update the estimates: $\mathbf{g}\leftarrow\mathscr{U}_g(\mathbf{g},\delta_g)$ and $\mathbf{p}\leftarrow\mathscr{U}_p(\mathbf{p},\delta_p)$

Table 1: The proposed dual inverse compositional algorithm for groupwise geometric and photometric registration of grey-level or color images.

5 Experimental Results

Our experiments are designed to compare the converge properties and the computational cost of the proposed algorithm compared to other algorithms in various conditions. All comparisons are done by estimating homographies. The region of interest is found by extracting edges and dilating them using a circular element with 2 pixel radius.

5.1 Simulated Data

Algorithms compared. This set of experiments aims at comparing the behaviour in terms of rate of convergence and computational cost of 'SIC-LAV', the simultaneous inverse compositional algorithm for linear appearance variation of Baker *et al.* [1], 'SIC-GB', the simultaneous inverse compositional algorithm specialized to gain and bias by

Bartoli [3] and 'DIC', the proposed dual inverse compositional algorithm. Note that SIC-LAV and SIC-GB produce *exactly* the same results but with different computation times.

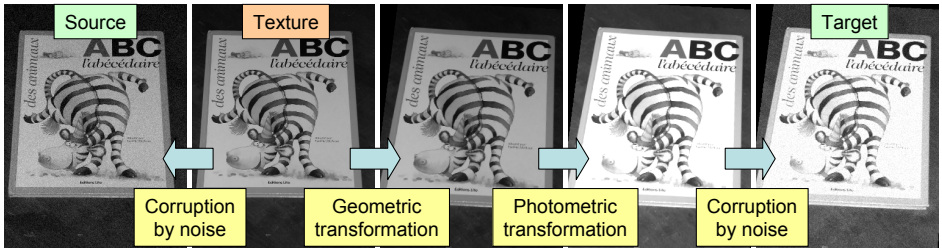

Figure 3: Synthetic data generation scheme.

Simulation setup. The image generation process is illustrated on figure 3. Given a texture image, we simulate a 2D homography by displacing four points in random directions by some magnitude γ with default value $\gamma = 5$ pixels. This is used to generate the target image in conjunction with a gain α and a bias β with default values $\alpha = 1.2$ and $\beta = 15$. Finally, centred Gaussian noise with variance σ is added to the pixel value in the source and target images, with default value $\sigma = 10\%$, *i.e.* 25.5 over 255. Finally, the pixel values are clamped between 0 and 255 in order to simulate sensor saturation. We measure the *geometric error*, defined as the RMS of the distance between the four displaced points and the points transferred by the computed homography. Convergence to the right solution is declared if the geometric error is below one pixel within 20 iterations. We also measure the *photometric error* defined as the mean of the difference image $\mathscr{D}$. The source image is 600×800, and the region of interest is made of 25,392 pixels.

Results. We performed comparisons of the algorithms for numerous combinations of the parameters governing the simulated images. We show part of these results concerning computational time in figure 4.

We make the following observations:

- DIC converges slightly more often than SIC-LAV and SIC-GB, in particular for large geometric transformation magnitudes ($\gamma \geq 8$ pixels).
- When they converge, the three methods give the same estimate.
- DIC requires slightly less iterations than SIC-LAV and SIC-GB.
- DIC has computation time slightly lower than SIC-GB, and much lower than SIC-LAV.

5.2 Real Data

We tried the algorithms on several datasets. We report detailed results for the 'poster' image pair shown in figures 1 and 2. The source and target image size is 640×480. The region of interest is made of 42,295 pixels.

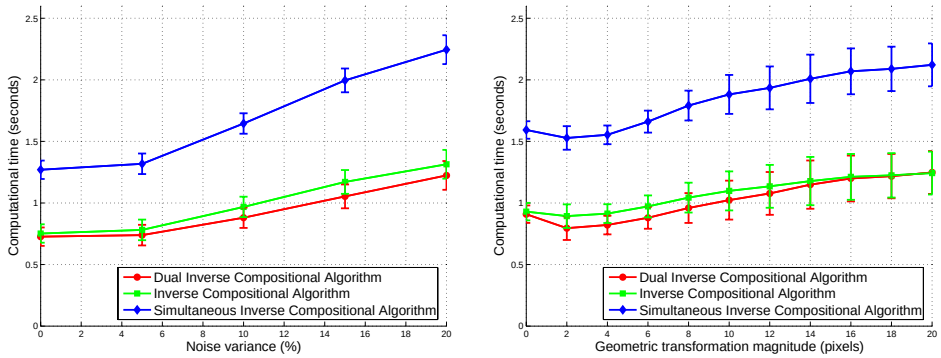

Figure 4: Computational time versus (left) the variance of added noise on the pixel values and (right) the magnitude of the geometric transformation.

Figure 5 shows the evolution of the photometric error through the iterations for several methods and photometric models. We observe that the error increases, usually over the first few iterations, and then decreases. The magnitude of error variation, both at the increasing and decreasing phases, is strongly related to the number of parameters in the photometric model. In other words, the more flexible the photometric model, the steepest the error variation.

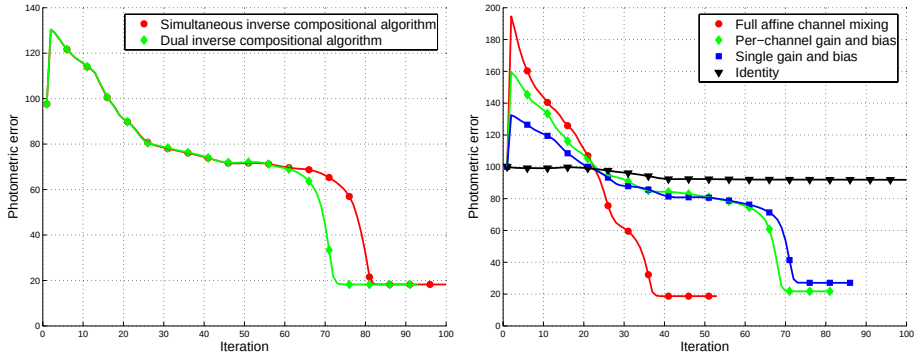

Figure 5: Photometric error through the iterations for the 'poster' image pair, (left) using intensity only and (right) using color, with the dual inverse compositional algorithm.

We see that for grey-level images, DIC performs slightly better than SIC-GB and SIC-LAV. For color images, DIC successfully estimates various photometric models. In particular, it estimates a full affine channel-mixing transformation. We observe from figure 6 that this model fits the images clearly better than the simpler ones, namely the single and per-channel gain and bias, since, in addition to the fact that the photometric error is lower, convergence requires far less iterations.

Figure 6: Difference image, number of iterations and photometric error for the proposed dual inverse compositional with various photometric models.

6 Conclusion

The proposed dual inverse compositional algorithm extends the inverse compositional algorithm to deal with photometric transformations. Efficiency stems from the fact that, as the geometric transformation, the photometric one is dealt with using inverse composition.

Experimental results shows that the proposed algorithm slightly outperforms the simultaneous inverse compositional algorithm optimized for, but limited to, gain and bias computation in convergence rates and terms of computational cost. Its main advantage is thus its ability to deal with color and any groupwise geometric and photometric models.

The MATLAB code used to produce the experimental results in this paper is available on the web homepage of the author.

References

[1] S. Baker, R. Gross, and I. Matthews. Lucas-Kanade 20 years on: A unifying framework: Part 3. Technical Report CMU-RI-TR-03-35, Carnegie Mellon University, November 2003.

[2] S. Baker and I. Matthews. Lucas-Kanade 20 years on: A unifying framework. *International Journal of Computer Vision*, 56(3):221–255, February 2004.

[3] A. Bartoli. Direct image registration with gain and bias. In *Topics in Automatic 3D Modeling and Processing Workshop*, 2006.

[4] G. D. Finlayson, M. S. Drew, and B. Funt. Color constancy: Generalized diagonal transforms suffice. *Journal of the Optical Society of America A*, 11(11):3011–3019, November 1994.

[5] G. D. Hager and P. N. Belhumeur. Efficient region tracking with parametric models of geometry and illumination. IEEE *Pattern Analysis and Machine Intelligence*, 20(10), 1998.

[6] K.P. Horn and G. Schunck. Determining optical flow. *Artificial Intelligence*, 17, 1981.

[7] M. Irani and P. Anandan. About direct methods. In *Workshop on Vision Algorithms: Theory and Practice*, 1999.

[8] H. Jin, P. Favaro, and S. Soatto. Real-time feature tracking and outlier rejection with changes in illumination. In *Proceedings of the International Conference on Computer Vision*, 2001.

[9] B. D. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. *International Joint Conference on Artificial Intelligence*, 1981.

[10] H. Sawhney and R. Kumar. True multi-image alignment and its application to mosaicing and lens distortion correction. IEEE *Pattern Analysis and Machine Intelligence*, 21(3), 1999.

[11] P. H. S. Torr and A. Zisserman. Feature based methods for structure and motion estimation. In *Workshop on Vision Algorithms: Theory and Practice*, 1999.