The Bi-directional Framework for Unifying Parametric Image Alignment Approaches

Rémi Mégret, Jean-Baptiste Authesserre, and Yannick Berthoumieu

IMS Laboratory, University of Bordeaux, France
{remi.megret,jean-baptiste.authesserre,
yannick.berthoumieu}@ims-bordeaux.fr

Abstract. In this paper, a generic bi-directional framework is proposed for parametric image alignment, that extends the classification of [1]. Four main categories (Forward, Inverse, Dependent and Bi-directional) form the basis of a consistent set of subclasses, onto which state-of-the-art methods have been mapped. New formulations for the ESM [2] and the Inverse Additive [3] algorithms are proposed, that show the ability of this framework to unify existing approaches. New explicit equivalence relationships are given for the case of first-order optimization that provide some insights into the choice of an update rule in iterative algorithms.

1 Introduction

Motion estimation is a fundamental task of many vision applications such as object tracking, image mosaicking, video compression or augmented reality. Image alignment based on template matching is a natural approach to image registration, by estimating the parameters that best warp one image onto the other. The optimum is conventionally provided by the minimization of the displaced frame difference between the template and an image. Since the Lucas and Kanade algorithm [4], many algorithms have been proposed to improve the performances. Baker and Matthews [1] summarized and compared experimentally template-based techniques divided into four classes (Forwards Additive, Forwards Compositional, Inverse Additive, and Inverse Compositional). Since then, methods such as Efficient Second-order Minimization (ESM) [2], Symmetrical Gradient Method (SGM) and Bi-directional Gradient Method (BDGM) [5], have been proposed that do not fit into the initial four classes. To our knowledge, no generic framework has been proposed yet in which all these methods can be classified.

In this paper, we develop a generic bi-directional framework which unifies the different template-based approaches. Thanks to this framework, the main contribution of this paper is to show how to rigorously define the alignment problem, and to propose a consistent set of well defined but generic classes. State of the art methods are expressed as instances of this generic formulation. For some of the methods, this formulation is new, or more general that initially.

In Sect. 2, the image alignment problem is formalized, and the bi-directional framework is explained. In Sect. 3, state of the art methods are classified and their interpretation within the framework is precised. In Sect. 4, a discussion addresses general issues concerning all approaches.

D. Forsyth, P. Torr, and A. Zisserman (Eds.): ECCV 2008, Part III, LNCS 5304, pp. 400–411, 2008.

2 Problem Formalization

2.1 Motion Model

The motion model is represented by a warp function $\mathbf{W}(\mu, \mathbf{x})$ of parameter vector $\mu \in \mathcal{P}$, applied at position $\mathbf{x} \in \mathbb{R}^2$. In order to facilitate the structured formulation of the framework, we require that the considered motion model form a group with respect to composition. This is the case of most models of interest [1], such as non degenerate affine motion and homographies.

This group property is extended to the parameter space $\mathcal{P}$:

$$\mathbf{W}(\mu_0 \circ \delta\mu, \mathbf{x}) = \mathbf{W}(\mu_0, \mathbf{W}(\delta\mu, \mathbf{x})) \qquad \text{Composition} \tag{1}$$

$$\mathbf{W}(\mu^{-1}, \mathbf{x}) = \mathbf{W}^{-1}(\mu, \mathbf{x}) \; = \{\mathbf{y} \mid \mathbf{W}(\mu, \mathbf{y}) = \mathbf{x}\} \qquad \text{Inverse} \tag{2}$$

$$\mathbf{W}(0, \mathbf{x}) = \mathbf{x} \qquad \text{Identity} \tag{3}$$

In order to apply gradient methods, the smoothness of the warp is additionnaly required. More precisely, it is assumed that $\mu \mapsto \mathbf{W}(\mu, \mathbf{x})$ is a $\mathcal{C}_1$-diffeomorphism, and that $\delta\mu \mapsto \mu \circ \delta\mu$ and $\delta\mu \mapsto \delta\mu^{-1}$ are $\mathcal{C}_1$-diffeomorphisms in a neighbourhood of $\delta\mu = 0$. Provided that the parameters are not close to a degenerate configuration, these constraints are again satisfied by the models of interest.

2.2 Bi-directional Image Alignment

The fundamental assumption used for image alignment correspond to the gray-level constancy equation (see Fig. 1 for an illustration of the concepts presented in this section):

$$\forall \mathbf{x} \in R_{\text{ref}} \qquad \mathbf{I}(\mathbf{W}(\bar{\mu}, \mathbf{x})) = \mathbf{T}(\mathbf{W}(\mu_{\text{ref}}, \mathbf{x})) \quad , \tag{4}$$

where $\bar{\mu}$ represents the parameters of the true displacement between images $\mathbf{I}$ and the reference coordinate frame, and μ_{ref} represents a fixed transformation between the template image $\mathbf{T}$ and the reference coordinate frame. R_{ref} corresponds to the region of interest, expressed in the reference coordinate frame.

If both images are considered to be continuous functions with respect to the position in the image, then the change in variables $\mathbf{x} = \mathbf{W}({\mu_{\text{ref}}}^{-1} \circ \mu_T, \mathbf{z})$ using an arbitrary μ_T leads to a more generic equation:

$$\forall \mathbf{z} \in R \qquad \mathbf{I}(\mathbf{W}(\mu_I, \mathbf{z})) = \mathbf{T}(\mathbf{W}(\mu_T, \mathbf{z})) \quad , \tag{5}$$

where $R = \mathbf{W}({\mu_T}^{-1} \circ \mu_{\text{ref}}, R_{\text{ref}})$ is the transformed region of interest and

$$\mu_I = \bar{\mu} \circ {\mu_{\text{ref}}}^{-1} \circ \mu_T \quad . \tag{6}$$

The bi-directional image alignment framework can therefore be formalized as finding a pair (μ_I, μ_T) that minimizes the discrepancy between the right and the left hand side of (5).

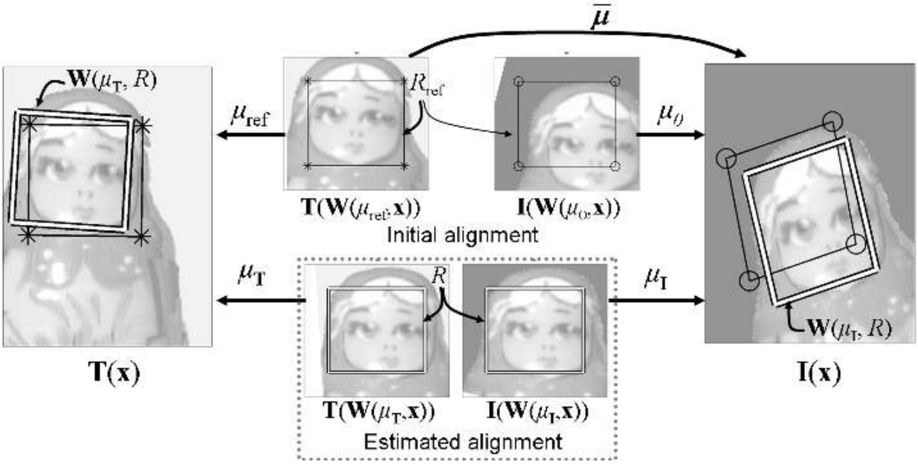

Fig. 1. General principle of the bi-directional framework, when aligning two images $\mathbf{I}$ and $\mathbf{T}$. The initial parameters μ_0 and μ_{ref} are shown in parallel with the parameters μ_I and μ_T leading to a correct alignment. The support region R is emphasized on the common compensated frames $\mathbf{T}(\mathbf{W}(\mu_T, \mathbf{x}))$ and $\mathbf{I}(\mathbf{W}(\mu_I, \mathbf{x}))$, as well as its corresponding regions on $\mathbf{T}(\mathbf{x})$ and $\mathbf{I}(\mathbf{x})$. For the Forwards approach, $\mu_T = \mu_{\mathrm{ref}}$. For the Inverse approach, $\mu_I = \mu_0$. In the general case shown here (Dependant and Bi-directional approaches) both μ_I and μ_T are varying during the optimization.

We denote $\mathbf{e}(\mu_I, \mu_T)$ the N-dimensional vector obtained by concatenating the pixelwise differences $\mathbf{e}_i$ between the compensated images over a spatial sampling $(\mathbf{x}_i)_{i=1..N}$ of R,

$$\mathbf{e}(\mu_I, \mu_T) = \mathbf{I}(\mathbf{W}(\mu_I, R)) - \mathbf{T}(\mathbf{W}(\mu_T, R)). \tag{7}$$

Using the L_2 norm of $\mathbf{e}$, the bi-directional error function corresponds to:

$$E(\mu_I, \mu_T) = \sum_{\mathbf{x} \in R} \Big(\mathbf{I}(\mathbf{W}(\mu_I, \mathbf{x})) - \mathbf{T}(\mathbf{W}(\mu_T, \mathbf{x})) \Big)^2 \tag{8}$$

Once the optimal bi-directional parameters μ_I and μ_T have been estimated, the equivalent estimate $\hat{\mu}$ to the true forwards displacement $\bar{\mu}$ is then computed by applying the update rule derived from (6):

$$\hat{\mu} \leftarrow \mu_I \circ {\mu_T}^{-1} \circ \mu_{\mathrm{ref}} \tag{9}$$

The region of interest R that appears in (8) is considered to be constant. We delay the discussion on the implications of this choice to Sect. 4.1.

Many different error metrics may be used as a replacement of the L_2 norm, based on pixelwise difference [6], but that could for instance also use color distributions [7]. In this paper, the focus is on the motion model aspects of image alignment. We will therefore limit ourselves to errors similar to (8).

2.3 Criteria for the Classification of Methods

Composing μ_I and μ_T to the right by the same parameter μ yields an infinity of pairs $(\mu_I \circ \mu, \mu_T \circ \mu)$ that satisfy (5). Four main image alignement categories can be proposed, depending on the constraints enforced on μ_I and μ_T.

In the following, μ_0 represents an initial estimate of $\bar{\mu}$, and μ_{ref} the motion parameter vector between the template image and the reference coordinate frame. Aligning the image and the template consists in finding a corrective parameter vector $\delta\mu$, whose nature depends on the approach.

Forwards (F). The image is warped onto the template with respect to $\delta\mu \in \mathcal{P}$: $\mu_T = \mu_{\text{ref}}$ and μ_I depends on the increment $\delta\mu$.

Inverse (I). The template is warped onto the image with respect to $\delta\mu \in \mathcal{P}$: $\mu_I = \mu_0$ and μ_T depends on the increment $\delta\mu$.

Dependent (D). The image and the template are warped using parameter vectors μ_I and μ_T that are both dependent from a common source $\delta\mu \in \mathcal{P}$. The **Symmetric (S)** approaches are a special case of dependent approaches, which correspond to applying symmetric corrections to μ_T and μ_I. These will be the only dependent approaches detailed in this paper.

Bi-directional (B). The image and the template are respectively warped using independant corrective parameters $\delta\mu_I$ and $\delta\mu_T$ that can be concatenated into a bi-directional parameter vector $\delta\mu = (\delta\mu_I, \delta\mu_T)^t \in \mathcal{P}^2$.

The first three approaches consider a corrective vector in a single parameter space $\delta\mu \in \mathcal{P}$, even though the (D) approach warps both images. On the opposite, the (B) approach instead considers that the optimization of (8) takes places inside the complete bi-directional space $\delta\mu \in \mathcal{P}^2$.

Inside each category, the methods can be further characterized by:

Meta-Parametrization of the motion parameters, which expresses the functional relationship $\mu_I = \mu_I(\delta\mu)$ and $\mu_T = \mu_T(\delta\mu)$ of both parameter vectors with respect to the corrective vector $\delta\mu$. A corrective parameter vector equal to the identity parameter $\delta\mu = 0$ should correspond to the initial alignment parameters: $\mu_I(\delta\mu) = \mu_0$ and $\mu_T(\delta\mu) = \mu_{\text{ref}}$. It is called meta-parametrization in order to differentiate it from the parametrization, which consists in choosing the parametric model $\mathbf{W}$.

Optimization method used to minimize $E(\mu_I(\delta\mu), \mu_T(\delta\mu))$, the error function of (8). This may involve gradient-based optimization, such as the Gauss-Newton (GN) and the Newton (N) methods [1] or even higher-order methods [8], but also learning approaches, such as the learning of a Linear Estimator (LE) [9]. We emphasize that a particular meta-parametrization is not restricted to any type of optimization.

The forwards and inverse categories overlap the categories of the same name proposed by Baker and Matthews [1] with one special case that will be discussed in more details in Sect. 3.5. The symmetric and the bi-directional approaches were first used by Keller and Averbuch [5]. To our knowledge, our proposal is the first systematic and generic classification covering all of these approaches as special cases.

3 Classification of Existing Methods

In this section, state of the art image alignment methods are going to be classified, and expressed with respect to the criteria we presented. Table 3 sums up the proposed categories and the associated meta-parametrizations, which are used in Table 3 to provide a synoptic view on the classification of existing algorithms. The justification of this classification is done on a case by case basis in the indicated subsections. The methods marked with the symbol * are of particular interest, and are discussed more specifically.

Although we give more details related to the Gauss-Newton optimization (GN), because it allows us to give new insights for the Inverse Additive algorithm [3] [1], the Efficient Second-order Minimization algorithm [2] and the Symmetric Gradient Method approach [5], we recall that the choice of a meta-parametrization is distinct from the choice of an optimization method.

3.1 Gauss-Newton Optimization (GN) within the Framework

The Gauss-Newton optimization of the generic error function (8) yields :

$$\delta\mu = -(J^t J)^{-1} J^t \mathbf{e}(\mu_0, \mu_{\text{ref}}) \tag{10}$$

where $J = \left.\frac{\partial \mathbf{e}(\mu_I(\delta\mu), \mu_T(\delta\mu))}{\partial \delta\mu}\right|_0$ corresponds to the Jacobian matrix of the error vector defined in (7). The considered functions $\mathbf{I}(\mathbf{x})$, $\mathbf{T}(\mathbf{x})$, $\mathbf{W}(\mu, \mathbf{x})$, $\mu_I(\delta\mu)$ and $\mu_T(\delta\mu)$ are assumed to be differentiable w.r.t. to $\mathbf{x}$ and $\delta\mu$.

Table 1. Categorized bi-directional meta-parametrizations and corresponding update rules. Categories are defined in Sect. 2.3. The following naming conventions are used: 1st letter: F=Forwards, I=Inverse, S=Symmetric, B=Bi-directional; 2nd letter: C=Compositional, A=Additive; 3rd letter: R=Reverse, D=Direct, M=Midway, E=Exponential map, O=Opposite. Approaches marked with a symbol * correspond to a new or more generic formulation of the problem.

C.	App.	μ_I	μ_T	$\bar{\mu} = \mu_I \circ {\mu_T}^{-1} \circ \mu_{\text{ref}}$	Sec
F	FC	$\mu_0 \circ \delta\mu$	μ_{ref}	$\mu_0 \circ \delta\mu$	3.2
	FA	$\mu_0 + \delta\mu$	μ_{ref}	$\mu_0 + \delta\mu$	3.2
I	ICR	μ_0	$\mu_{\text{ref}} \circ \delta\mu$	$\mu_0 \circ \delta\mu^{-1}$	3.3
	ICD*	μ_0	$\mu_{\text{ref}} \circ \delta\mu^{-1}$	$\mu_0 \circ \delta\mu$	3.3
	IAR	μ_0	$\mu_{\text{ref}} + \delta\mu$	$\mu_0 \circ (\mu_{\text{ref}} + \delta\mu)^{-1} \circ \mu_{\text{ref}}$	3.4
	IAD*	μ_0	$\mu_{\text{ref}} \circ (\mu_0 + \delta\mu)^{-1} \circ \mu_0$	$\mu_0 + \delta\mu$	3.5
D/S	SCM*	$\mu_0 \circ (\frac{1}{2}\delta\mu)$	$\mu_{\text{ref}} \circ (\frac{1}{2}\delta\mu)^{-1}$	$\mu_0 \circ (\frac{1}{2}\delta\mu) \circ (\frac{1}{2}\delta\mu)$	3.6
	SCE*	$\mu_0 \circ \mu(\frac{1}{2}\delta v)$	$\mu_{\text{ref}} \circ \mu(-\frac{1}{2}\delta v)$	$\mu_0 \circ \mu(\delta v)$	3.6
	SCO*	$\mu_0 \circ (\frac{1}{2}\delta\mu)$	$\mu_{\text{ref}} \circ (-\frac{1}{2}\delta\mu)$	$\mu_0 \circ (\frac{1}{2}\delta\mu) \circ (-\frac{1}{2}\delta\mu)^{-1}$	3.6
B	BCD*	$\mu_0 \circ \delta\mu_I$	$\mu_{\text{ref}} \circ \delta\mu_T^{-1}$	$\mu_0 \circ \delta\mu_I \circ \delta\mu_T$	3.7
	BCO*	$\mu_0 \circ \delta\mu_I$	$\mu_{\text{ref}} \circ (-\delta\mu_T)$	$\mu_0 \circ \delta\mu_I \circ (-\delta\mu_T)^{-1}$	3.7

Table 2. Classification of various existing methods. Category: see Sect. 2.3. Meta-parametrization: see Table 3. Optimization: GN=Gauss Newton, N=Newton, LE=Linear Estimator, O3=Third Order. New insights are obtained for the methods indicated with a symbol *, which are discussed in their respective sections.

C.	Method	M.-Param	Optim.	Sect.
F	Forwards Additive [4]	FA	GN	3.2
	Forwards Compositional [10]	FC	GN	3.2
	Third-order Gradient Method [8]	FA	O3	3.2
I	Inverse Compositional [1] [11]	ICR	GN, N...	3.3
	Hyperplane Approximation [9]	IAR	LE	3.4
	Inverse Additive [3]	IAD	GN	3.5*
D/S	Efficient Second-order Minimization [2]	SCE	GN	3.6*
	Symmetric Gradient Method [5]	SCO	GN	3.6*
	Symmetric third-order Gradient Method [8]	SCO	O3	3.6*
B	Bi-directional Gradient Method [5]	BCO	GN	3.7*

The Jacobian J is specific for each approach. It can be expressed as the concatenation of the gradients $J(\mathbf{x}_i)$ of the pixelwise errors $\mathbf{e}_i$, where:

$$J(\mathbf{x}_i) = \underbrace{\left.\frac{\partial \mathbf{I}(\mathbf{W}(\mu_I(\delta\mu), \mathbf{x}_i))}{\partial \delta\mu}\right|_0}_{J_{\mathbf{I}}(\mathbf{x}_i)} - \underbrace{\left.\frac{\partial \mathbf{T}(\mathbf{W}(\mu_T(\delta\mu), \mathbf{x}_i))}{\partial \delta\mu}\right|_0}_{J_{\mathbf{T}}(\mathbf{x}_i)} \tag{11}$$

Due to a lack of space, only the key equations will be given for each approach. The intermediate steps can be obtained by replacing, in equations (8) and (7), μ_I and μ_T by their expression with respect to $\delta\mu$ from Table 3, and deriving from these equations.

3.2 Forwards Additive (FA) and Forwards Compositional (FC)

The forwards approaches, such as *Forwards Additive* (FA) and *Forwards Compositional* (FC) fit naturally into the bi-directional framework by setting $\mu_T = \mu_{\mathrm{ref}}$. This generalizes the formulation shown in [1] where $\mu_{\mathrm{ref}} = 0$. The two are equivalent by replacing $\mathbf{T}$ with $\mathbf{T}_{\mathrm{ref}} = \mathbf{T}(\mathbf{W}(\mu_{\mathrm{ref}}, \cdot))$. This approach was combined with Newton-type optimization in [1], and a third-order gradient method in [8].

For the FA approach, $\mu_I = \mu_0 + \delta\mu^{FA}$, which yields the following Jacobian:

$$J^{FA}(\mathbf{x}_i) = J_{\mathbf{I}}^{FA}(\mathbf{x}_i) = \nabla \mathbf{I}(\mathbf{W}(\mu_0, \mathbf{x}_i)) \left.\frac{\partial \mathbf{W}(\mu, \mathbf{x}_i)}{\partial \mu}\right|_{\mu_0} \tag{12}$$

For the FC approach, $\mu_I = \mu_0 \circ \delta\mu^{FC}$, which yields

$$J^{FC}(\mathbf{x}_i) = J_{\mathbf{I}}^{FC}(\mathbf{x}_i) = \nabla \mathbf{I}(\mathbf{W}(\mu_0, \mathbf{x}_i)) \left.\frac{\partial \mathbf{W}(\mu_0, \mathbf{x})}{\partial \mathbf{x}}\right|_{\mathbf{x}_i} \left.\frac{\partial \mathbf{W}(\mu, \mathbf{x}_i)}{\partial \mu}\right|_0 \tag{13}$$

An explicit equivalence relationship between the FA and the FC approaches that extends the equivalence proof proposed in [1] is discussed in Sect. 4.2.

3.3 Inverse Compositional (IC) Variants

The inverse compositional approach proposed in [1] also fits naturally in the framework using parameters shown in Table 3. We further classify it as *Inverse Compositional Reverse* (ICR) because of the presence of $\delta\mu^{-1}$ in its update rule.

The Jacobian is expressed w.r.t. gradients of the reference template $\mathbf{T}_{\text{ref}}$, which allows for their pre-computation, thus improving the online performance:

$$J^{ICR}(\mathbf{x}_i) = -J^{IC}_{\mathbf{T}}(\mathbf{x}_i) = -\underbrace{\nabla\mathbf{T}(\mathbf{W}(\mu_{\text{ref}}, \mathbf{x}_i)) \left.\frac{\partial \mathbf{W}(\mu_{\text{ref}}, \mathbf{x})}{\partial \mathbf{x}}\right|_{\mathbf{x}_i}}_{\nabla\mathbf{T}_{\text{ref}}(\mathbf{x}_i)} \left.\frac{\partial \mathbf{W}(\mu, \mathbf{x}_i)}{\partial \mu}\right|_0 \quad (14)$$

By replacing $\delta\mu$ with $\delta\mu^{-1}$ in μ_T, we can define a new *Inverse Compositional Direct* (ICD) approach, which has a simpler update rule. It has the same complexity as the ICR approach for the estimation of $\delta\mu$ in GN optimization since $J^{ICD} = -J^{ICR}$.

3.4 Inverse Additive Reverse (IAR)

The *Inverse Additive Reverse* (IAR) approach is the dual of the FA approach, by *reversing* the roles of $\mathbf{I}$ and $\mathbf{T}$. This meta-parametrization (shown in Table 3) was used in [9] with a Linear Estimator (LE): the estimation is based on $\delta\mu = A\mathbf{e}(\mu_0, \mu_{\text{ref}})$ where A is a learned matrix. The disturbances $\delta\mu^k$ used for learning A are applied on the template through the parameters ${\mu_T}^k = \mu_{\text{ref}} + \delta\mu^k$, which makes it an IAR approach. The advantages over gradient-based approaches are that A is computed off-line, thus decreasing the online computationnal cost, and that it can handle larger motion amplitude. The use of learning based optimization is one specific advantage of the Inverse approach compared to the others.

3.5 Inverse Additive Direct (IAD): New Insights on [3]

The algorithm introduced in [3] was called Inverse Additive in the survey [1]. In the sequel, we will refer to it as the IA algorithm. The previous justification for this algorithm is based on a FA parametrization, but where the roles of $\mathbf{I}$ and $\mathbf{T}$ are swapped for the computation of the Jacobian, by assuming that the two images are identical up to motion compensation (equation (4)). This allows the authors to derive an efficient algorithm if the factorization of (15) is possible. Because of this strong assumption on the relative content of the two images, the minimized error function can not be expressed in closed form.

We propose a new meta-parametrization (see Table 3), called *Inverse Additive Direct* (IAD), which has an additive update rule. The GN optimization of its closed-form error function then leads naturally to the IA algorithm. Indeed the associated Jacobian matches the one used in the IA algorithm:

$$J^{IAD}(\mathbf{x}_i) = -J^{IAD}_{\mathbf{T}}(\mathbf{x}_i) = -\nabla\mathbf{T}_{\text{ref}}(\mathbf{x}_i) \underbrace{\left(\left.\frac{\partial \mathbf{W}(\mu_0, \mathbf{x})}{\partial \mathbf{x}}\right|_{\mathbf{x}_i}\right)^{-1} \left.\frac{\partial \mathbf{W}(\mu, \mathbf{x}_i)}{\partial \mu_0}\right|_{\mu_0}}_{\text{factored as } \Gamma(\mathbf{x}_i)\Sigma(\mu_0)} \quad (15)$$

where $\nabla \mathbf{T}_{\text{ref}}(\mathbf{x}_i)$ was defined in (14). The previous equation, when plugged into (10) corresponds to one iteration of the IA algorithm [3,1], thus allowing us to integrate this algorithm as an instance of the proposed framework.

3.6 Symmetric Compositional (SC): New Insights on [2], [8] and [5]

The *Symmetric Compositional Midway* (SCM) approach is defined by compensating both **I** and **T** towards each other using two transformations that are inverse one from the other (see Table 3). In that case the common compensated coordinate frame lies exactly midways from both images coordinate frames from a compositional point of view. Its update rule is $\hat{\mu} \leftarrow \mu_0 \circ (\frac{1}{2}\delta\mu) \circ (\frac{1}{2}\delta\mu)$.

By reusing the notations from (13) and (14), its Jacobian is equal to:

$$J^{SCM}(\mathbf{x}_i) = \frac{1}{2}\left(J^{FC}_{\mathbf{I}}(\mathbf{x}_i) + J^{IC}_{\mathbf{T}}(\mathbf{x}_i)\right) \tag{16}$$

In a similar way as the IA algorithm discussed before, the justification of the ESM algorithm [2] relies on assumption (4). We now propose to instantiate this algorithm into the bi-directional framework by adapting the SCM approach.

The update rule $\hat{\mu} \leftarrow \mu_0 \circ \mu(\delta v)$ used in [2] is slightly different from the SCM rule. The update step $\delta\mu = \mu(\delta v)$ is indeed further parametrized around the identity by a vector δv using an exponential map (associated with a Lie Group on projective transformations matrices denoted by $\mathbf{G}(\mathbf{x})$ in [2]). The optimization is then done with respect to δv instead of $\delta\mu$.

This parametrization has two interesting properties : $\mu(2v) = \mu(v) \circ \mu(v)$, $\mu(-v) = \mu(v)^{-1}$. We can therefore split δv into two symmetrical parts, to obtain the *Symmetrical Compositional Exponential map* (SCE) approach:

$$\mu_I = \mu_0 \circ \mu\Big(\frac{1}{2}\delta v\Big) \quad \text{and} \quad \mu_T = \mu_{\text{ref}} \circ \mu\Big(-\frac{1}{2}\delta v\Big) = \mu_{\text{ref}} \circ \left(\mu\Big(\frac{1}{2}\delta v\Big)\right)^{-1} \tag{17}$$

Because additionnally $\delta\mu = \mu(\delta v) \approx \mu(0) + \delta v$ to the first order around the identity, due to the exponential map properties, the associated Jacobian is identical to (16). We can therefore conclude that the ESM algorithm corresponds to the GN optimization with respect to δv of the closed form SCE error function.

A very similar approach was proposed for GN optimization [5] and higher-order optimization [8]. By taking into account the compensations that occur in the described algorithms explicitely (in a similar way as in Sect. 3.2), this translates into a compositional meta-parametrization, which we call *Symmetric Compositional Opposite* (SCO), the O corresponding to the term $-\frac{1}{2}\delta\mu$, that appears in μ_T (see Table 3). According to (9), the associated update rule should be $\hat{\mu} \leftarrow \mu_0 \circ (\frac{1}{2}\delta\mu) \circ (-\frac{1}{2}\delta\mu)^{-1}$. It is different from the rule $\hat{\mu} \leftarrow \mu_0 + \delta\mu$ used in [5]. This issue is discussed in more detail in Sect. 4.2.

One of the advantages of such symmetrical approaches is that the estimation is precise up to the second-order in $\delta\mu$, even when using only first-order approximation, provided (4) is satisfied and the motion increment to estimate is small enough (shown for SCE in [2] and SCO in [5]).

3.7 Bi-directional Compositional (BC) Variants

The bi-directional approach was first proposed for Gradient Methods in [5]. As in the previous subsection, we propose to reformulate their approach by taking into account explicitely the initial parameters μ_0 and μ_{ref}, which leads to the *Bi-directional Compositional Opposite* (BCO) meta-parametrization (see Table 3). The corresponding update rule can be approximated by $\hat{\mu} \leftarrow \mu_0 \circ (\delta\mu_I + \delta\mu_T)$ for a small $\delta\mu$, which is different from rule $\hat{\mu} \leftarrow \mu_0 + (\delta\mu_I + \delta\mu_T)$ used in [5]. The consequence of this difference is discussed in Sect. 4.2.

It is important to note that the corrective parameter vector $\delta\mu = \begin{pmatrix} \delta\mu_I \\ \delta\mu_T \end{pmatrix}$ belongs to the full bi-directional parameter space $\mathcal{P}^2$. The Jacobian of the error can thus be expressed as the horizontal concatenation of the Jacobians of the simple compositional approaches: $J^{BCO} = \left[J^{FC}, -J^{ICR}\right]$.

In the same spirit as the ICD approach, the *Bi-directional Compositional Direct* (BCD) approach shown in Table 3 has a slightly simpler update rule (see Table 3), and the same estimation cost as BCR, since $J^{BCD} = J^{BCO}$.

4 Discussion

4.1 Integration into an Iterative Scheme

Up to now, we have mostly derived results corresponding to one estimation step. This step is generally included in an iterative algorithm, in order to refine the estimation. The iterative schemes in the reviewed approaches update the initial parameters from iteration n to $n+1$ by finding the equivalent forwards parametrization to the estimated bi-directional parameters based on (9): $\mu_0^{n+1} \leftarrow {\mu_I}^n \circ ({\mu_T}^n)^{-1} \circ {\mu_{\mathrm{ref}}}^n$, ${\mu_{\mathrm{ref}}}^{n+1} \leftarrow {\mu_{\mathrm{ref}}}^n$. This allows to keep R fixed across iterations in order to avoid a drift from the initial interest region.

The region R in (8) represents the region of interest in the common compensated coordinates frame (see Fig. 1). It is generally chosen such that it corresponds to the reference region of interest in the template at initialization: $R = \mathbf{W}(({\mu_{\mathrm{ref}}}^n)^{-1} \circ {\mu_{\mathrm{ref}}}^0, R_{\mathrm{ref}})$. Once defined, R is considered to be fixed for the optimization of (8), in order to avoid the spurious terms in the error derivatives that would reflect the variation of R. These terms have always been neglected in the studied methods. This scheme also leads to fixed derivatives with respect to the template, thus decreasing the computational cost within the (I), (D) and (B) approaches.

Figure 2 illustrates the error function $E(\mu_I, \mu_T)$ in the bi-directional space $\mathcal{P}^2$. It can be noted that it forms a valley along the curve $\mu_I = \bar{\mu} \circ {\mu_{\mathrm{ref}}}^{-1} \circ \mu_T$, which is valid for μ_T close to $\mu_{\mathrm{ref}} = 0$. Indeed, when μ_T is too far away, the interest region R includes elements from the background, which increases the error. The trajectories of the initial $(\mu_0^n, {\mu_{\mathrm{ref}}}^n)$ and the estimated $({\mu_I}^n, {\mu_T}^n)$ parameters are plotted for one approach of each category. The differences in the meta-parametrization are clearly reflected as different types of trajectories in the bi-directional space $\mathcal{P}^2$.

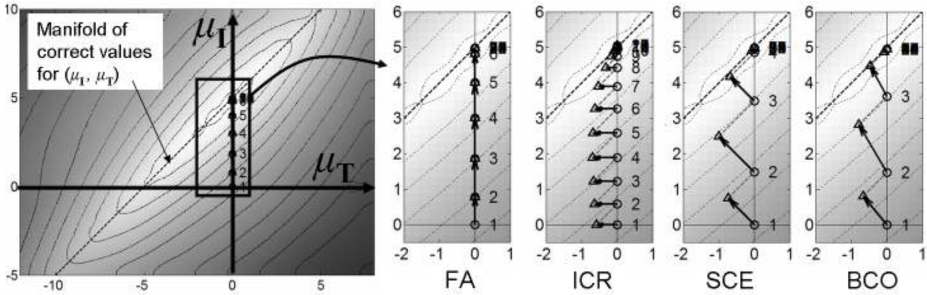

Fig. 2. Error function $E(\mu_I, \mu_T)$ corresponding to the image of Fig. 1 displayed on a $(\mu_{I,1}, \mu_{T,1})$ slice of the bi-directional space $\mathcal{P}^2$, where the index 1 stands for the horizontal translation coefficient. Translation was estimated using GN optimization with one method in each category. Each iteration n is drawn with an arrow, that links $(\mu_0^n, {\mu_{\text{ref}}}^n)$ (numbered $\odot$ bullets) to $({\mu_I}^n, {\mu_T}^n)$ ($\triangle$ bullets). The true deformation is a 5 pixels horizontal translation $\bar{\mu}_1 = 5$, and the initialisation is $(\mu_{0,1}^1, \mu_{\text{ref},1}^1) = (0, 0)$. The dashed line $(\mu_{I,1}, \mu_{T,1}) = (\mu_1, 5 + \mu_1)$ represents the set of correct estimates. The shape of the trajectories reflect the contraints put on the meta-parametrization.

4.2 Equivalences and Incompatibilities in Terms of Update Rules

When using GN optimization, the equivalence between FA and FC approaches was argued in [1]. On the basis of our formulation, we can extend this result by providing the exact relationship between the two update steps $\delta\mu^{FA}$ and $\delta\mu^{FC}$ computed by both approaches. Indeed, thanks to the regularity of the considered functions, we can show that

$$J^{FC} = J^{FA} M_0 \quad \text{where} \quad M_0 = \left.\frac{\partial \mu_0 \circ \delta\mu}{\partial \delta\mu}\right|_0 \tag{18}$$

For GN optimization, (10) then yields $\delta\mu^{FC} = {M_0}^{-1} \delta\mu^{FA}$. From this relationship, we can conclude on the equivalence to the first order:

$$\mu_0 \circ \delta\mu^{FC} \approx \quad \mu_0 + M_0 \delta\mu^{FC} = \quad \mu_0 + \delta\mu^{FA} \tag{19}$$

The same methodology can be used to show that, when using a GN optimization, all proposed approaches are equivalent to the first order within the same category. The equivalences are based on:

$$\delta\mu^{ICR} = -\delta\mu^{ICD}, \quad \delta\mu^{IAD} = M_0 \delta\mu^{ICD}$$

$$\delta\mu^{IAR} = M_{\text{ref}} \delta\mu^{ICR} \quad \text{where} \quad M_{\text{ref}} = \left.\tfrac{\partial \mu_{\text{ref}} \circ \delta\mu}{\partial \delta\mu}\right|_0 \tag{20}$$

$$\delta\mu^{SCM} = \delta\mu^{SCE} = \delta\mu^{SCO} \quad \text{and} \quad \delta\mu^{BCO} = \delta\mu^{BCD} \tag{21}$$

One issue that the previous analysis reveals, is that although the final estimate $\hat{\mu}$ may be approximately equal between related additive and compositional

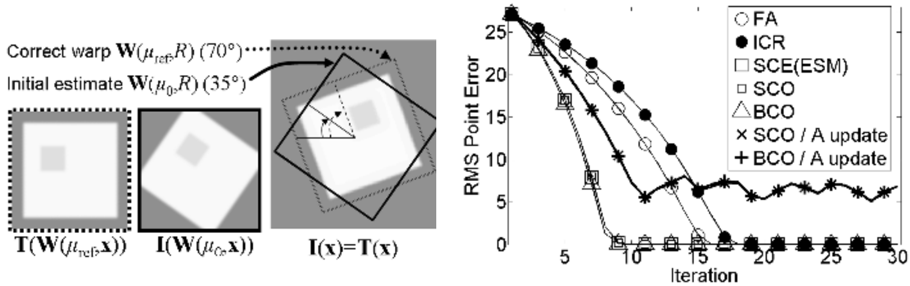

Fig. 3. The equivalence of two approaches does not mean the egality of their corrective parameter $\delta\mu$. This counter-example is based on affine motion estimation using a standard benchmark [1]. The error is expressed in pixels. A rotation $\bar{\mu}$ of 70° around the center of the object is to be estimated, from an initial rotation μ_0 of 35°. All approaches use GN optimization. The hybrid approach composed of an SCO or BCO meta-parametrization combined with an additive update rule [5] converge faster that the F and I approaches when the angle of μ_0 is small, but cannot converge to the correct estimate as explained in the text. Using instead the rule stemming from the framework corrected this problem, to achieve the best results with this type of optimization.

approaches, the corresponding $\delta\mu$ are not equal in the general case, because of the presence of the matrix M. Therefore *the update step should always be consistent with the used meta-parametrization.*

When this is not the case, a correctly estimated $\delta\mu$ can lead to an incorrect estimate $\hat{\mu}$. This causes convergence problems, especially when rotations are involved, which seem to appear for example in some experiments from [5] and [8]. These effects are illustrated in Fig. 3, where the error corresponding to an affine motion estimate oscillates without converging around the correct estimate when μ_0 corresponds to a large rotation. The proposed framework offers a systematic methodology to avoid such problems when designing template based alignment.

4.3 Practical Considerations on the Classification

In order to facilitate the choice of an approach, here is a short summary of the main properties of the respective categories. The FA and FC approaches have been shown to be equivalent to first order in [1]. We have shown that this equivalence is also true within each separate Inverse, Symmetrical, and Bidirectionnal categories. The Inverse approach has the fastest step computation thanks to the offline computation of the Jacobian. This category is also the only one to benefit from a learned parameter estimator [9] which yields a direct estimation in one step. Additionally, when (4) is satisfied, the Symmetrical approaches (SCE [2], SCO [5]) need a lower number of GN iterations to converge than the Forwards and Inverse approaches. According to [5], BCO should not be more performant than SCO when (4) holds, but may outperform it in the more general case.

5 Conclusion

In this paper, we have presented a formal framework for pixel-based image alignment methods, associated to a simple and consistent classification. The proposed criteria have been applied to a wide range of image alignment methods. In particular, this methodology has led to a new formulation of the IA algorithm and the ESM algorithm, based on a closed form error function without any assumption on the content of the images. This unification revealed useful to make an explicit description of the equivalence to the first order between the methods, and to give new insights with respect to the use of a mismatching update rule.

We think such a framework offers a structured formulation of the parametric image alignment problem, which, we hope, will help in understanding, designing and evaluating the performance of alignment algorithms. Perspectives are to extend the formalization to include non purely geometric models such as illumination compensation [12], and study the interaction between model parametrization and meta-parametrization.

References

1. Baker, S., Matthews, I.: Lucas-Kanade 20 years on: A unifying framework. International Journal of Computer Vision 56(3), 221–255 (2004)
2. Benhimane, S., Malis, E.: Real-time image-based tracking of planes using efficient second-order minimization. In: IROS 2004, Sendai, Japan, vol. 1, pp. 943–948 (2004)
3. Hager, G.D., Belhumeur, P.N.: Efficient region tracking with parametric models of geometry and illumination. IEEE Transactions on Pattern Analysis and Machine Intelligence 20(10), 1025–1039 (1998)
4. Lucas, B., Kanade, T.: An iterative image registration technique with an application to stereo vision. In: IJCAI 1981, pp. 674–679 (1981)
5. Keller, Y., Averbuch, A.: Fast motion estimation using bi-directional gradient methods. IEEE Transactions on Image Processing 13(8), 1042–1054 (2004)
6. Baker, S., Gross, R., Matthews, I., Ishikawa, T.: Lucas-Kanade 20 years on: A unifying framework: Part 2. Technical Report CMU-RI-TR-03-01, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA (February 2003)
7. Mégret, R., Mikram, M., Berthoumieu, Y.: Inverse composition for multi-kernel tracking. In: Kalra, P.K., Peleg, S. (eds.) ICVGIP 2006. LNCS, vol. 4338, pp. 480–491. Springer, Heidelberg (2006)
8. Keller, Y., Averbuch, A.: Global parametric image alignment via high-order approximation. Computer Vision and Image Understanding 109(3), 244–259 (2008)
9. Jurie, F., Dhome, M.: Hyperplane approximation for template matching. IEEE Transactions on Pattern Analysis and Machine Intelligence 24(7), 996–1000 (2002)
10. Shum, H.Y., Szeliski, R.: Construction of panoramic image mosaics with global and local alignment. International Journal of Computer Vision 36(2), 101–130 (2000)
11. Matthews, I., Baker, S.: Active appearance models revisited. International Journal of Computer Vision 60(2), 135–164 (2004)
12. Bartoli, A.: Groupwise geometric and photometric direct image registration. IEEE Transactions on Pattern Analysis and Machine Intelligence (2008) preprint (January 14, 2008)