Vehicle X-Ray Scans Registration: A One-Dimensional Optimization Problem

Abraham Marciano[1, 2] *, Laurent D. Cohen[1], Najib Gadi[2]

[1] Université Paris-Dauphine, PSL Research University, CNRS, UMR 7534, CEREMADE, 75016 Paris, France
[2] Smiths Detection, 94405 Vitry-sur-Seine, France
marciano@ceremade.dauphine.fr

Abstract. Over the years, image registration has been largely employed in medical applications, robotics and geophysics. More recently, it has increasingly drawn attention of security and defense industries, particularly aiming at threat detection automation. This paper first introduces a short overview of mathematical methods for image registration, with a focus on variational approaches. In a second part, a specific registration task is presented: the optimal alignment between X-ray scans of an inspected vehicle and an empty reference of the same car model. Indeed, while being scanned by dedicated imaging systems, the car speed is not necessarily constant which may entail non-rigid deformations in the resulting image. The paper simply addresses this issue by applying a rigid transform on the reference image before using the variational framework solved in one dimension. For convergence and speed purposes, line-search techniques and a multiscale approach are used.

Keywords: Image registration, Variational approach, One-dimensional optimization

1 Introduction

Due to the recent global context, security has increasingly become a top priority for governments and agencies all around the world. Meanwhile, larger amounts of data have to be processed in more limited periods of time. Thus, the switch to automation of threat targeting constitutes a crucial step.

The vision industry has significantly been impacted by this move and must address new problems, often by resorting to widely used techniques such as image registration. As defined by Modersitzki et al. [1,2,3], it consists in finding a "reasonable" transformation applied on a template image $\boldsymbol{T}$ to get an optimal alignment with a reference image $\boldsymbol{R}$.

In this paper, we present a challenging registration task: aligning top-view scans of a car under inspection and an empty reference of the same model. The obvious purpose of this process is to automatically identify added objects in a vehicle and target potential

adfa, p. 1, 2011.

security threats or contraband goods. Though, for X-ray imaging systems such as Smiths Detection HCVL, the car is trailed by a mechanism pushing its rear wheels so that the speed may not be constant over time. This imperfection often gives non-rigid deformations in the resulting image (see **Fig. 1.**) due to "over-sampling" (car deceleration) or "sub-sampling" (car acceleration) of the scanned object.

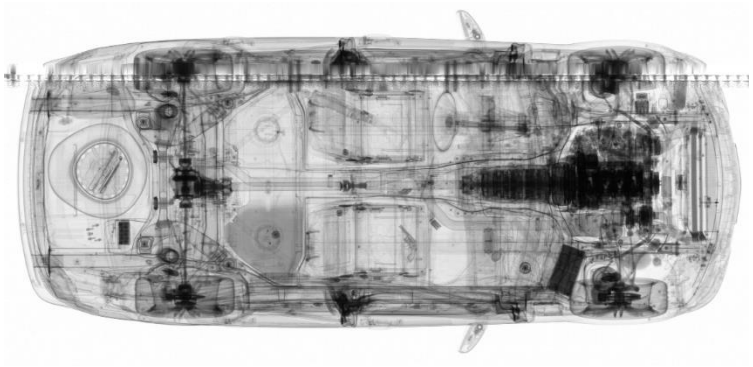

Fig. 1. Top-view scan of a vehicle. The spare wheel shows a non-rigid deformation as a result of the car slowdown during the scanning process.

Note that deforming potential threats in the image is not desirable for further recognition tasks. Thus, we consider the current inspection scan as the static image $\boldsymbol{R}$ and the empty reference image from our database as the moving template $\boldsymbol{T}$ (such that $\boldsymbol{R}$ does not undergo any non-linear warping).

The paper first gives an overview of the general non-rigid registration problem formulated in the variational framework. Parametric and non-parametric methods are detailed along with their numerical resolution. We will also show that, when applied to our registration problem, these schemes fail to yield a consistent transform with respect to the reference image or by violating the car's intrinsic rigidity.

In section **3**, we will demonstrate that the composition of a rigid transform followed by a one-dimensional optimization scheme applied on $\boldsymbol{T}$ gives a valid solution. An explicit formulation of the problem is given in detail as well as its numerical framework.

2 Problem Setup and Non-Linear Methods Overview

2.1 Mathematical Problem Setup

As mentioned, given $\boldsymbol{R}$ and $\boldsymbol{T}$, respectively a reference image and a template image, we aim at finding a transformation on $\boldsymbol{T}$ such that the warped template matches the reference as closely as possible [1,2,3]. Generally, $\boldsymbol{R}$ and $\boldsymbol{T}$ are defined as d-dimensional images $\boldsymbol{R}, \boldsymbol{T} : \Omega \rightarrow \mathbb{R},\ \Omega \in]0,1[^d$ (normalized coordinates). The transformation is

expressed by $\boldsymbol{u}: \Omega \rightarrow \mathbb{R}^d$, commonly referred to as the displacement field applied on each position $\boldsymbol{x} \in \Omega$ in $\boldsymbol{T}$.

Hence, for a particular $\boldsymbol{x} \in \Omega$, $\boldsymbol{T}(\boldsymbol{x})$ corresponds to the grayscale value of $\boldsymbol{T}$ at $\boldsymbol{x}$ and we wish to find $\boldsymbol{u}$ so that $\boldsymbol{T}(\boldsymbol{x} + \boldsymbol{u})$ is similar to $\boldsymbol{R}(\boldsymbol{x})$ (see [4,5] e.g.). Since our images are 2D, we take $d = 2$.

The unified non-linear registration framework is formulated as follows [2]: find a transformation $\boldsymbol{u}$ minimizing the joint energy:

$$J(\boldsymbol{u}) = D(\boldsymbol{T}(\boldsymbol{u}), \boldsymbol{R}) + \alpha S(\boldsymbol{u}) \quad (1)$$

Where $\boldsymbol{T}(\boldsymbol{u})$ denotes the transformed image with value $\boldsymbol{T}(\boldsymbol{x} + \boldsymbol{u})$ at location $\boldsymbol{x}$, and D represents the data fitting term, quantifying the similarity between $\boldsymbol{T}(\boldsymbol{u})$ and $\boldsymbol{R}$ (external force). Since optimizing D turns to an ill-posed problem, a regularizer term S has been introduced. It refers directly to the "reasonable" aspect of the transformation as defined in section **1**. The smoothing parameter α controls the strength of the displacement field $\boldsymbol{u}$ regularization, ensuring its smoothness during deformation (internal force) [2,3,4,5]. Several methods have been developed to estimate a proper α (see [2], [6]), but in most situations it is conditional upon the type of application and the operator assessment.

Different distance/similarity measures are employed for this task. The most popular is the *sum of differences* distance (SSD) whereas the *normalized-cross-correlation* (NCC) is also widely used for monomodal registration problems. The *mutual information* distance (MI) is dedicated to the registration of images obtained with different modalities (e.g. CT and MRI) [3], [7]. More sophisticated measures have also been developed in the last decade, such as the *normalized gradient fields* (NGF) (see [7,8]). In this paper the SSD distance is preferred for its simplicity and efficiency to solve our monomodal registration issue.

Likewise, several regularizers are outlined in the state of the art literature. The next sub-section gives more details about the different smoothing techniques. Note that the main constraint in the selection of similarity measures or regularizers is the existence of a Gâteaux derivative for the optimization scheme to be used later [4].

2.2 Methods Overview

Landmark-based registration is widely adopted in medical applications for instance ([9]). The displacement field is first evaluated on identified points (landmarks), followed by a TPS (*thin-plate-spline*) interpolation yielding a dense estimation of $\boldsymbol{u}$ (see [3]).

Still, an accurate automatic detection of landmarks remains a challenging task (see [9] e.g.). In this paper, the focus is therefore given on intensity-based methods which fall into two categories:

- Parametric techniques: the transformation is restricted to a known parametric model. The optimization process aims at automatically identifying the optimal parameters of $\boldsymbol{u}$. Non-linear techniques resort to a linear combination of a small set of basis functions (B-spline, *free-form deformations* – FFD methods). See [3], [7] for a complete overview of PIR methods (Parametric Image Registration). For these techniques, the parametrization itself implicitly integrates a strong regularizer, especially for low dimensional transformation spaces ($S(\boldsymbol{u}) = 0$). Yet, a Tychonov smoothing term may also be employed [3].

- Non-parametric techniques: The transformation is no longer parametrized. The optimization processes over the displacement field $\boldsymbol{u}$ itself making the regularization term $S(\boldsymbol{u})$ essential. *L2 norm*-based models such as curvature [5], fluid [10] and elastic [11] regularizers are the most widespread (see [3] or [7] for a complete overview). For NPIR (Non-Parametric Image Registration) methods, discretization is regularizer-dependent. Different schemes can be adopted such as nodal, centered or staggered grids. See ([3], [7]) for more details about discretization and interpolation issues.
  Thirion's demons popular method [12] derives from optical flow techniques with a diffusion-like regularizer ([1], [5]). It is especially useful for high dimensional and computationally demanding non-linear problems [5].
  Models with rigidity or volume-preserving constraints have also proven their efficiency. Yet, they are beyond the scope of this paper. See [13,14,15] for further details.

2.3 Numerical Resolution

Modersitzki ([3], [7]) proposes a unified numerical optimization framework to minimize (1). Both NPIR and PIR are solved by a quasi-Gauss-Newton descent method ([16]). The gradient $\nabla J(\boldsymbol{u})$ and Hessian $H(\boldsymbol{u})$ of $J(\boldsymbol{u})$ are computed (the Hessian is often approximated). The descent direction $dd = H(\boldsymbol{u})^{-1}\nabla J(\boldsymbol{u})$ is calculated at each iteration. Also, Armijo's line-search backtracking method is employed to get an optimal step size for each update of $\boldsymbol{u}$ ([3], [16]). In general, minimization is stopped whenever the variation of the transformation $\|\boldsymbol{\nabla u}\|_2$ falls below a pre-set tolerance threshold ϵ. See [3] for more details about the stopping criterion.

A multi-level strategy is also used. It yields a smoother objective function at coarser levels, hence easier to minimize. Thus, the resulting displacement field constitutes a good initialization for the finer level and local minima issues can be overcome [2,3].

Alternative optimization paradigms have also been proposed for computationally expensive cases. The *l*-BFGS approximates the inverse of the Hessian using an initial guess H_0 and a sequence composed of descent directions and gradients [16]. Trust-region methods are quite popular as well since they can achieve fast quadratic convergence [16]. See [3] for further details and [17] for a more advanced overview of variational methods and their numerical frameworks.

2.4 Tests on Our Images

In this example, the inspection of the front part of a given vehicle is considered (**Fig. 2.** (a)). A corresponding empty reference scan from the same model is used for registration (from a different car, **Fig. 2.** (b)). Major visual differences are easily identified at the rear-view mirrors, front wheels, the gearbox, the steering wheel, a few liquid tanks and the battery. **Fig. 2.** represents the reference and template images after pose estimation (rigid registration, see paragraph **3.4** for further details). Two methods are tested in this part: B-spline registration and Thirion's demons method.

Fig. 3. shows some results obtained via the demons method with two different smoothing parameters and a multi-level approach. Note that green and pink colors depict respectively the differences originating from the moving template $\boldsymbol{T}$ and image $\boldsymbol{R}$ under inspection (as defined in section **1**).

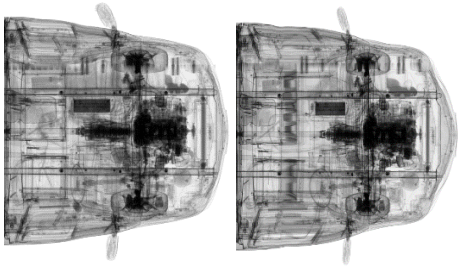

Fig. 2. (a) Empty reference of a vehicle front part; (b) Front part under inspection after pose estimation

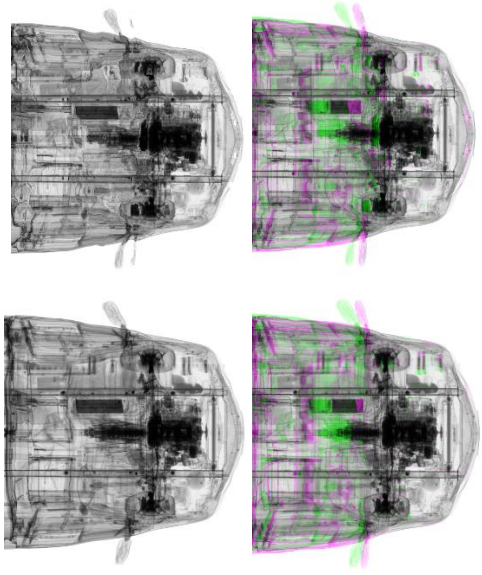

Fig. 3. Demons registration results (resulting $\boldsymbol{T_{final}}$ and overlay of $\boldsymbol{T_{final}}$ and $\boldsymbol{R}$) with a smoothing parameter of 0.5 (top) and 2 (bottom)

B-spline registration is also tested with a multiscale approach and a "thin sheet of metal" smoothness penalty set to 0.01 and 0.1 (**Fig. 4.**).

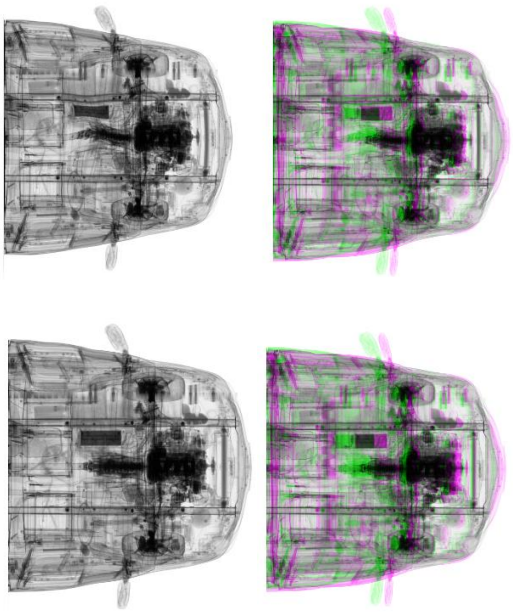

Fig. 4. B-spline registration results (resulting $\boldsymbol{T_{final}}$ and overlay of $\boldsymbol{T_{final}}$ and $\boldsymbol{R}$) with a smoothing parameter of 0.01 (top) and 0.1 (bottom)

Besides computational cost considerations (B-spline registration can be significantly time-consuming), these standard methods do not achieve accurate registration for the particular case addressed here. See **Fig. 10.** for the resulting SSD distances of each method. In fact, the vehicle's rigidity along with the mechanical process behind the non-rigid deformation have to be taken into account. The idea is to simplify the optimization problem by moving from a 2-dimensional to a one-dimensional scheme. More particularly, the displacement field has to be constrained to the longitudinal direction and must remain uniform along the car's width. The next section outlines the motivations and numerical aspects of our method.

3 Registering with a One-Dimensional Optimization Scheme

3.1 Introduction to Our Registration Problem

In HCVL X-ray imaging systems, an inspected vehicle is trailed by a conveyor facility via rollers pushing on its back wheels (**Fig. 5.**). The conveyor speed is meant to be constant and fixed to 12 m/min. Though, the car often rolls off the trailer equipment as a result of shocks between the conveyor rollers and the wheels. The vehicle speed undergoes disturbances affecting the scanning process: a slowdown (resp. acceleration)

implies a local "over-sampling" (resp. "sub-sampling") in specific regions of the car (see **Fig. 1.**).

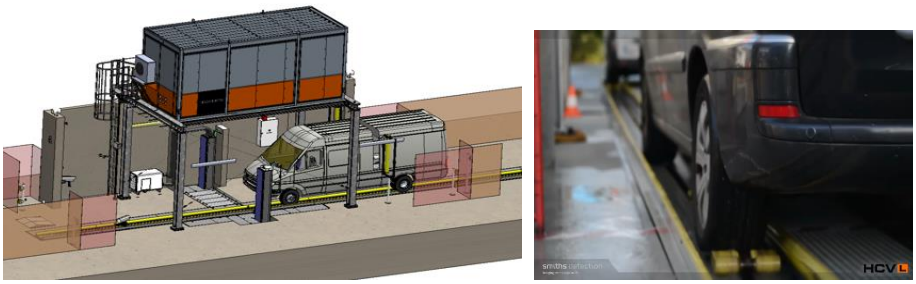

Fig. 5. (a) HCVL scanning system; (b) The conveyor rollers (in yellow) applying a pushing force on the rear wheels

Let's consider two X-ray images of the same car model (not necessarily the same vehicle): we assume that pose estimation has already been performed such that both images were linearly registered through a rigid transform. We formulate two strong hypothesis:

- **Hypothesis 1**: A *columnwise-constant deformation*. We make the reasonable assumption that the field of displacement is uniform along each column. In fact, the car is scanned with a constant sampling rate while entering the X-ray beam line so each column of the resulting image corresponds to a lateral cut of the vehicle scan at a given time (**Fig. 6.**). Thus, any speed disturbance would affect each separate column in a uniform fashion.

- **Hypothesis 2**: The deformation direction is perpendicular to the X-ray beam line (the vertical component of $\boldsymbol{u}$ is null). This assumption is an accurate approximation since the car can hardly strive from the conveyor during scanning (**Fig. 6.**).

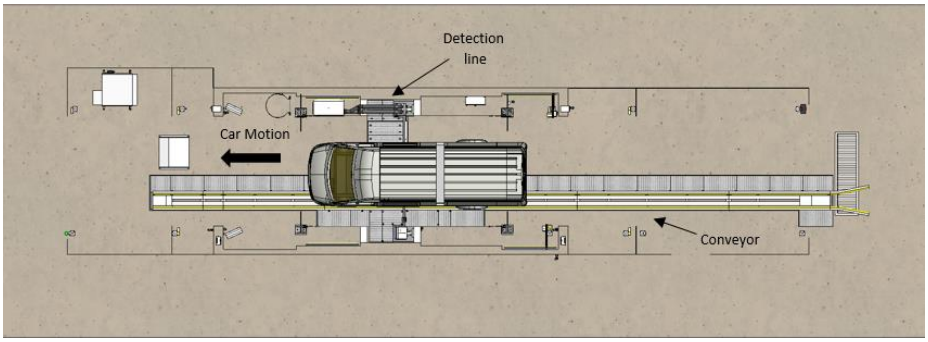

Fig. 6. Top-view description of the HCVL system

3.2 Method Outline

Let's formulate the 1D optimization problem. Given two images $\boldsymbol{R}$ and $\boldsymbol{T}$, pose estimation is performed by applying a rigid transform on $\boldsymbol{T}$. For notation simplification, we keep using $\boldsymbol{T}$ to designate the moving image following this pre-processing warping.

We define *x* and *y* as the horizontal and vertical coordinates of $\boldsymbol{x} \in \Omega$:

$$\boldsymbol{x} = (x, y) \tag{2}$$

Similarly, the displacement field:

$$\boldsymbol{u} = \big(u_x(x,y), u_y(x,y)\big) \tag{3}$$

Hypothesis 1 implies that

$$u_x(x,y) = u_x(x) \tag{4}$$

And **Hypothesis 2** yields

$$u_y(x,y) = 0 \tag{5}$$

Eventually, combining (5) and (4) gives:

$$\boldsymbol{u} = (u_x(x), 0). \tag{6}$$

With an *SSD* distance and a regularizer of the form: $S(\boldsymbol{u}) = \frac{1}{2}\|\nabla u\|_2^2$, (1) becomes:

$$\text{Find } \boldsymbol{u} \text{ minimizing } J(\boldsymbol{u}) = \frac{1}{2}\|\boldsymbol{T}(\boldsymbol{u}) - \boldsymbol{R}\|_2^2 + \frac{1}{2}\alpha\|\nabla \boldsymbol{u}\|_2^2 \tag{7}$$

On the basis of (6), (7) turns to a one-dimensional optimization problem that can be solved via simple descent techniques.

3.3 Numerical Resolution

For a given column index *x*, (7) is equivalent to:

$$\text{Find } u_x(x) \text{ minimizing}$$

$$J(u_x(x)) = \frac{1}{2}\big\|\boldsymbol{T}\big(u_x(x)\big) - \boldsymbol{R}\big\|_2^2 + \frac{1}{2}\alpha\left\|\frac{\partial u_x(x)}{\partial x}\right\|_2^2 \tag{8}$$

A first-degree descent scheme would be relevant for this 1D optimization problem (low complexity). We resort to the gradient descent method, in combination Armijo's backtracking line search method [16]. We evolve the following equation with Dirichlet boundary conditions:

$$\frac{\partial u_x(x)}{\partial t} = -\nabla J\big(u_x(x)\big) \; for \; x \in \Omega \setminus \partial\Omega \; with \; u_x(x) = 0 \; for \; x \in \partial\Omega \tag{9}$$

The gradient of $J(u_x(x))$ is computed from (8):

$$\nabla J\big(u_x(x)\big) = \frac{\partial \boldsymbol{T}(x+u_x(x))}{\partial u_x}\Big(\boldsymbol{T}\big(x+u_x(x)\big) - \boldsymbol{R}(x)\Big) - \alpha\frac{\partial^2 u_x(x)}{\partial x^2} \tag{10}$$

By abuse of notation: $\boldsymbol{T}\big(x+u_x(x)\big)$ corresponds to the column of $\boldsymbol{T}$at $x+u_x(x)$. We note $X = x+u_x(x)$ such that $\frac{\partial \boldsymbol{T}(x+u_x(x))}{\partial u_x} = \frac{\partial \boldsymbol{T}(x+u_x(x))}{\partial \mathrm{X}}$, referring to the gradient of $\boldsymbol{T}$ at $x+u_x(x)$. We get the final expression:

$$\nabla J\big(u_x(x)\big) = \frac{\partial \boldsymbol{T}(x+u_x(x))}{\partial X}\Big(\boldsymbol{T}\big(x+u_x(x)\big) - \boldsymbol{R}(x)\Big) - \alpha\frac{\partial^2 u_x(x)}{\partial x^2} \tag{11}$$

A multi-scale approach is used to speed up the registration process and in order to avoid convergence at local minima. Thereby, the displacement ${u_x}^{(p)}$ obtained at level *p* gives a strong initialization for the displacement estimate ${u_x}^{(p-1)}$ for the finer level *p – 1*. The number of levels is fixed to $l = 4$.

Empirically, a smoothing parameter set to $\alpha = 0.5$ for all scales yields accurate and stable results. The stopping criterion is analogous to the method described in section **2.3** with a chosen threshold $\epsilon = 1$.

At the end, an estimate of $\boldsymbol{T}$ transformed via the optimal displacement field $\boldsymbol{u}^*$ is calculated using linear or cubic interpolation (both techniques give similar results in our case).

3.4 Experimental Results on Scan Images

Our method is applied on the example of section **2.4**. **Fig 7**. displays $\boldsymbol{T}$ and $\boldsymbol{R}$ prior to rigid registration. In a first stage, a rigid transform is automatically computed using SURF [18] feature points matching and RANSAC filtering [19]. The transform applied on $\boldsymbol{T}$ gives a pose estimation between both images. See **Fig. 8.** for a visual description of alignment before and after pose estimation. This pre-processing step is often necessary to rectify slight differences of car positioning or geometry variations between separate X-ray systems (at different sites).

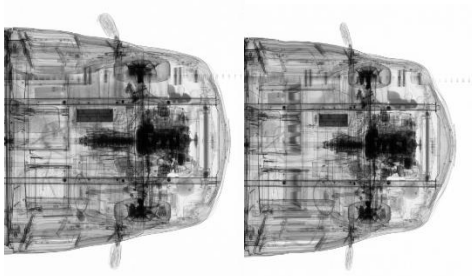

Fig. 7. (a) Empty reference of a vehicle front part; (b) Front part under inspection

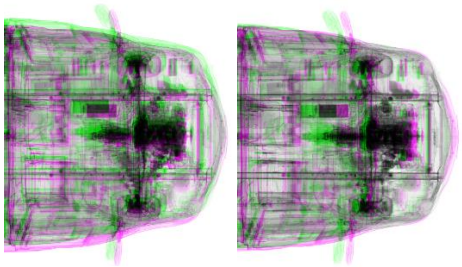

Fig. 8. (a) Overlay of both images before pose estimation; (b) Overlay following pose estimation by rigid registration

Our multiscale one-dimensional minimization scheme yields the following image $\boldsymbol{T_{final}}$ with the corresponding columnwise-constant displacement field $\boldsymbol{u^*}$(**Fig. 9.**).

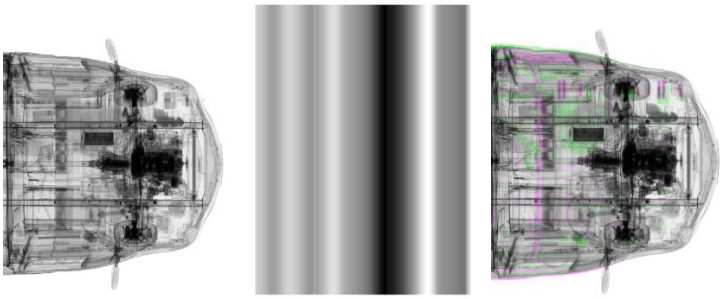

Fig. 9. (a) Resulting image; (b) Displacement field; (c) Overlay of $\boldsymbol{R}$ and $\boldsymbol{T_{final}}$

After reaching convergence to global minima of the objective function (8), the resulting alignment shows the vanishing of initial major differences: the mirrors as well as the battery or the steering wheel align perfectly (**Fig. 9.**). Obviously, our method achieves the lowest SSD cost in comparison with the different methods tested (**Fig. 10.**).

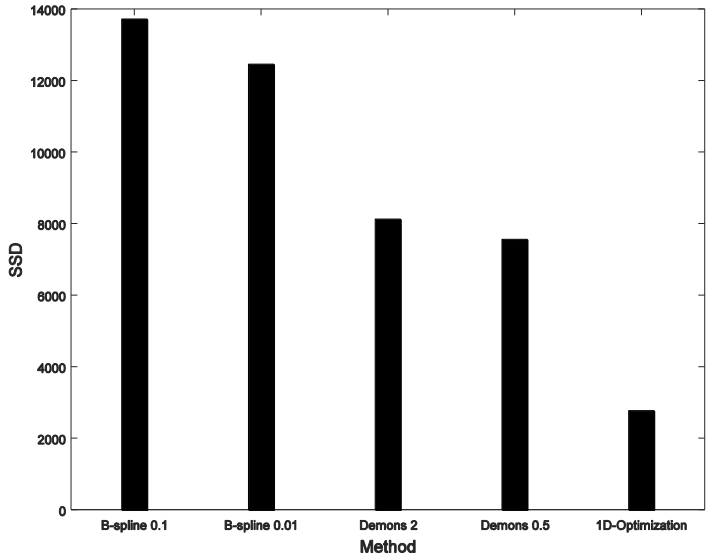

Fig. 10. SSD distances reached for each method (with its corresponding smoothing parameter)

4 Conclusion

This paper addresses the registration problem of top-view X-ray scans from two different vehicles of the same model. It especially aims at dealing with non-linear deformations induced by possible speed variations of the vehicle during scanning.

A simple and intuitive solution is described: in a first stage, a pose estimation between both scans is performed. Then, assuming that the field of displacement is columnwise-constant and parallel to the car's motion, a one-dimensional optimization scheme is formulated. It is solved by well-known descent techniques in combination with Armijo's backtracking method and a multiscale approach. Both visual and numerical results presented in this paper demonstrate the necessity for our method as well as its high performances in terms of registration accuracy.

References

1. Haber, E., Heldmann, S., & Modersitzki, J. (2009). A computational framework for image-based constrained registration. Linear Algebra and its Applications, 431(3), 459-470.
2. Haber, E., & Modersitzki, J. (2006). A multilevel method for image registration. SIAM Journal on Scientific Computing, 27(5), 1594-1607.
3. Modersitzki, J. (2009). FAIR: flexible algorithms for image registration (Vol. 6). SIAM.
4. Fischer, B., & Modersitzki, J. (2003). Fast image registration: a variational approach. In Proceedings of the International Conference on Numerical Analysis & Computational Mathematics, G. Psihoyios (ed.), Wiley (pp. 69-74).
5. Fischer, B., & Modersitzki, J. (2004). A unified approach to fast image registration and a new curvature based registration technique. Linear Algebra and its applications, 380, 107-124.
6. Haber, E., Ascher, U. M., & Oldenburg, D. (2000). On optimization techniques for solving nonlinear inverse problems. Inverse problems, 16(5), 1263.
7. Modersitzki, J. (2004). Numerical methods for image registration. Oxford University Press on Demand.
8. Haber, E., & Modersitzki, J. (2006, October). Intensity gradient based registration and fusion of multi-modal images. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 726-733). Springer Berlin Heidelberg.
9. Crum, W. R., Hartkens, T., & Hill, D. L. G. (2014). Non-rigid image registration: theory and practice. The British Journal of Radiology.
10. Christensen, G. E. (1994). Deformable shape models for anatomy (Doctoral dissertation, Washington University).
11. Broit, C. (1981). Optimal registration of deformed images.
12. Thirion, J. P. (1998). Image matching as a diffusion process: an analogy with Maxwell's demons. Medical image analysis, 2(3), 243-260.
13. Haber, E., & Modersitzki, J. (2004). Numerical methods for volume preserving image registration. Inverse problems, 20(5), 1621.
14. Modersitzki, J. (2008). FLIRT with rigidity—image registration with a local non-rigidity penalty. International Journal of Computer Vision, 76(2), 153-163.
15. Staring, M., Klein, S., & Pluim, J. P. (2006, March). Nonrigid registration using a rigidity constraint. In Medical Imaging (pp. 614413-614413). International Society for Optics and Photonics.
16. Nocedal, J., & Wright, S. (2006). Numerical optimization. Springer Science & Business Media.
17. Chumchob, N., & Chen, K. (2011). A robust multigrid approach for variational image registration models. Journal of Computational and Applied Mathematics, 236(5), 653-674.
18. Bay, H., Tuytelaars, T., & Van Gool, L. (2006, May). Surf: Speeded up robust features. In European conference on computer vision (pp. 404-417). Springer Berlin Heidelberg.
19. Fischler, M. A., & Bolles, R. C. (1981). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM, 24(6), 381-395.