A Duality Based Algorithm for TV-L^1-Optical-Flow Image Registration

Thomas Pock[1], Martin Urschler[1], Christopher Zach[2], Reinhard Beichel[3], and Horst Bischof[1]

[1] Institute for Computer Graphics & Vision, Graz University of Technology, Austria,
{pock,urschler,bischof}@icg.tu-graz.ac.at
[2] VRVis Research Centre, Graz, Austria,
zach@vrvis.at
[3] Electrical & Computer Engineering and Internal Medicine,
The University of Iowa, USA,
reinhard-beichel@uiowa.edu

Abstract. Nonlinear image registration is a challenging task in the field of medical image analysis. Intensity based registration is based on the entire image information and generates a dense displacement field. In many applications discontinuities may be present in the displacement field, and intensity variations may occur. In this work we utilize an energy functional which is based on Total Variation regularization and a robust data term. We propose a novel, fast and stable numerical scheme to find the minimizer of this energy. Our approach combines a fixed-point procedure derived from duality principles combined with a fast thresholding step. We show experimental results on synthetic and clinical CT lung data sets at different breathing states as well as registration results on inter-subject brain MRIs.

1 Introduction

A large number of medical image analysis applications require nonlinear (deformable) registration of data sets acquired at different points in time or from different subjects. The deformation of soft tissue organs, such as the lung or the liver often requires the compensation of breathing motion. Surveys on nonlinear registration methods in medical imaging can be found in Maintz and Viergever [12] or Crum et al. [9]. In the literature a distinction is also made between feature based and intensity based methods. However, the majority of publications utilizes the intensity based methods [2, 8, 15] mainly because all available image information is used for registration. The main drawback of those methods is the required computational effort.

The optical-flow based approach is very popular for intra-modality registration tasks [10, 15, 16]. The variational formulation of the optical flow framework typically consists of quadratic data and regularization terms. This approach has two disadvantages. At first, only smooth displacement fields can be recovered.

In some medical volume registration applications it is however necessary to allow for discontinuities in the displacement field. A typical example is breathing motion, where the diaphragm undergoes a heavy deformation, whereas the rib cage remains almost rigid. Therefore smooth displacement fields are not sufficient to model such complex motions. In addition, quadratic error norms are not robust against outliers in the intensity values. Image differences due to contrast agent application or inter-subject registration tasks frequently violate the intensity-constancy assumption.

In this work we present a novel, numerical algorithm for an optical flow based method which allows for discontinuities in the displacement field and is robust with respect to varying intensities. In Section 2 we introduce variational optical flow methods and derive the TV-L^1-optical-flow model. For minimization we derive a novel numerical scheme by combining a duality based formulation of the variational energy and a fast thresholding scheme. In Section 3 we evaluate our algorithm using synthetically transformed and clinical data sets. In the last section we give some conclusions and suggest possible directions for future investigations.

2 Optical Flow

The recovery of motion from images is a major task of biological and artificial vision systems. In their seminal work, Horn and Schunck [11] studied the so-called *optical flow*, which relates the image intensity at a point and given time to the motion of an intensity pattern.

2.1 Model of Horn and Schunck

The classical optical flow model of Horn and Schunck (HS) for 2D images is given by the minimizer of the following energy:

$$\min_{\boldsymbol{u}} \left\{ E_{\mathrm{HS}} = \int_{\Omega} |\nabla u_1|^2 + |\nabla u_2|^2 \,\mathrm{d}\Omega + \lambda \int_{\Omega} \left(I_1(\boldsymbol{x} + \boldsymbol{u}(\boldsymbol{x})) - I_0(\boldsymbol{x})\right)^2 \mathrm{d}\Omega \right\} . \tag{1}$$

I_0 and I_1 is the image pair, $\boldsymbol{u} = (u_1(\boldsymbol{x}), u_2(\boldsymbol{x}))^T$ is the two-dimensional displacement field and λ is a free parameter. The first term (regularization term) penalizes for high variations in $\boldsymbol{u}$ to obtain smooth displacement fields. The second term (data term) is basically the optical flow constraint, which assumes that the intensity values of $I_0(\boldsymbol{x})$ do not change during its motion to $I_1(\boldsymbol{x} + \boldsymbol{u}(\boldsymbol{x}))$.

Since the HS model penalizes deviations in a quadratic way, it has two major limitations. It does not allow for discontinuities in the displacement field and it does not allow for outliers in the data term. To overcome these limitations, several models including robust error norms and higher order data terms have been proposed [1, 3, 13].

2.2 TV-L^1-Optical-Flow

In this paper we utilize the non-quadratic error norm $\rho(s) = |s|$ for both the regularization term and data term. The major advantage is that it allows for outliers while still being convex. Using this error norm and extending Eq. (1) to N dimensions, the robust optical flow model is given by

$$\min_{\boldsymbol{u}} \left\{ E_{\mathrm{TV}-L^1} = \int_\Omega \sum_{d=1}^{N} |\nabla u_d| \, \mathrm{d}\Omega + \lambda \int_\Omega |I_1(\boldsymbol{x} + \boldsymbol{u}(\boldsymbol{x})) - I_0(\boldsymbol{x})| \, \mathrm{d}\Omega \right\}, \quad (2)$$

where $\boldsymbol{u} = (u_1, u_2, \ldots, u_N)^T$ is the N-dimensional displacement field. Although this model seems to be simple and the modifications compared to Eq. (1) are minor, it offers some desirable improvements. At first, the regularization term allows for discontinuities. We note that this term is the well known Total Variation (TV) regularizer which has been proposed by Rudin Osher and Fatemi (ROF) for image denoising [14]. Secondly, the data term uses the robust L^1 norm and is therefore less sensitive to intensity variations.

Besides its clear advantages, the TV-L^1-optical-flow model also leads to some computational difficulties for minimization. The main reason is that both the regularization term and the data term are non-differentiable at zero. One approach would be to replace $|s|$ by some differentiable approximation e.g. $\sqrt{s^2 + \varepsilon^2}$ [4, 6]. However, for small ε this approach usually shows slow convergence and, on the other hand, using large ε leads to blurred displacement fields.

2.3 Solution of the TV-L^1-Optical-Flow Model

Based on the influential work of Chambolle [5], we introduce an additional variable $\boldsymbol{v} = (v_1, v_2, \ldots, v_N)^T$ and propose to minimize the following convex approximation of Eq. (2).

$$\min_{\boldsymbol{u},\boldsymbol{v}} \left\{ E_{\mathrm{TV}-L^1} = \int_\Omega \sum_{d=1}^{N} |\nabla u_d| + \frac{1}{2\theta} (u_d - v_d)^2 \, \mathrm{d}\Omega + \lambda \int_\Omega |\rho(\boldsymbol{x})| \, \mathrm{d}\Omega \right\}, \quad (3)$$

where the parameter θ is small so that we almost have $\boldsymbol{u} \approx \boldsymbol{v}$ and $\rho(\boldsymbol{x}) = I_1(\boldsymbol{x}+\boldsymbol{v}_0)+(\nabla I_1(\boldsymbol{x}+\boldsymbol{v}_0))^T(\boldsymbol{v}-\boldsymbol{v}_0)-I_0(\boldsymbol{x})$ is a first order Taylor approximation of the image residual. This formulation has several advantages: The minimization with respect to $\boldsymbol{u}$ can be performed using Chambolle's algorithm. It is based on a dual formulation of the total variation energy and does not suffer from any approximation error [5]. The minimization with respect to $\boldsymbol{v}$ reduces to a simple 1-D minimization problem which can be solved by an efficient thresholding scheme. Thus, to solve the optimization problem Eq. (3), we propose an iterative algorithm by alternating the following two steps:

1. For every d and fixed v_d, solve

$$\min_{u_d} \left\{ \int_\Omega \left\{ |\nabla u_d| + \frac{1}{2\,\theta} (u_d - v_d)^2 \right\} d\boldsymbol{x} \right\}. \quad (4)$$

2. For fixed $\boldsymbol{u}$, solve

$$\min_{\boldsymbol{v}} \left\{ \sum_d \frac{1}{2\,\theta} (u_d - v_d)^2 + \lambda \, |\rho(\boldsymbol{x})| \right\}. \tag{5}$$

The solution of these two optimization steps is given by the following propositions:

Proposition 1 *The solution of Eq. (4) is given by* $u_d = v_d - \theta \, \mathbf{div}\, \boldsymbol{p}$*, where* $\boldsymbol{p} = (p^1, p^2)$ *fulfills* $\nabla(\theta \, \mathbf{div}\, \boldsymbol{p} - v_d) = |\nabla(\theta \, \mathbf{div}\, \boldsymbol{p} - v_d)| \, \boldsymbol{p}$*, which can be solved by the following iterative fixed-point scheme:*

$$\boldsymbol{p}^{k+1} = \frac{\boldsymbol{p}^k + \tau \nabla(\mathbf{div}\, \boldsymbol{p}^k - v_d/\theta)}{1 + \tau |\nabla(\mathbf{div}\, \boldsymbol{p}^k - v_d/\theta)|}\,, \tag{6}$$

where $\boldsymbol{p}^0 = \mathbf{0}$ *and the time step* $\tau \leq 2^{(N+1)}$*.*

Proposition 2 *The solution of Eq. (5) is given by the following thresholding scheme:*

$$\boldsymbol{v}(\boldsymbol{x}) = \boldsymbol{u}(\boldsymbol{x}) + \begin{cases} \lambda \, \theta \, \nabla I_1 & \text{if} \quad \rho(\boldsymbol{x}) < -\lambda \, \theta \, |\nabla I_1|^2 \\ -\lambda \, \theta \, \nabla I_1 & \text{if} \quad \rho(\boldsymbol{x}) > \lambda \, \theta \, |\nabla I_1|^2 \\ -\rho(\boldsymbol{x}) \, \nabla I_1 / |\nabla I_1|^2 & \text{if} \quad |\rho(\boldsymbol{x})| \leq \lambda \, \theta \, |\nabla I_1|^2\,, \end{cases} \tag{7}$$

The proof of the n-dimensional case is presented in [17].

2.4 Implementation

Computing optical flow is a non-convex inverse problem, which means, that no globally optimal solution can be computed in general. In order to avoid such physically non-relevant solutions, one typically applies a coarse to fine strategy. Note, that due to the linearization of the image residual, the single-level minimization problem Eq. (2) becomes a convex one. However, the overall problem still remains non-convex.

For this purpose, we build a full Gaussian image pyramid by successively down-sampling the images by a factor of 2. On each level of the image pyramid, we solve the minimization problem Eq. (3), starting with the coarsest level. The solution is then propagated until the finest level has been reached. At the coarsest level, the displacement field is initialized by $\boldsymbol{u} = \mathbf{0}$.

In the following we describe the numerical scheme used to compute the solution at one single level of the image pyramid. For the implementation of the fixed-point iteration scheme Eq. (6) we use backward differences to approximate the discrete divergence operator and forward differences to approximate the discrete gradient operator [5]. For the implementation of the thresholding scheme Eq. (7) we use central differences to approximate the image derivatives. Finally, our iterative procedure consist of alternately computing the solutions of the fixed-point scheme and the thresholding step. In practice only a small number

of iterations $(100 - 200)$ is necessary to reach convergence. Moreover, empirical tests have shown that one iteration of the fixed-point scheme is sufficient to obtain fast convergence of the entire algorithm.

The numerical algorithm has two free parameters. The parameter θ acts as a coupling between the two optimization steps. It basically depends on the range of the intensity values of the input images. Empirically, we have selected $\theta = 0.02$ in all experiments. The parameter λ is used to control the amount of regularization. Typical values of λ are between 10 and 100.

3 Experimental Results

To assess the validity of our approach we performed qualitative and quantitative evaluations using synthetically transformed and clinical thorax CT data sets showing breathing motion and inter-subject brain data sets. All our experiments were performed on an AMD Opteron machine with 2.2 GHz and 16 GB RAM, running a 64-bit Linux system. The run-time of our algorithm is approximately 15 minutes for 256^3 voxel data sets when performing 200 iterations on the finest level of the image pyramid. Note that this performance is superior compared to the majority of intensity based registration algorithms.

3.1 Synthetic Data

Synthetic experiments were performed on a thoracic CT test data set with a size of 256^3 voxels, which was taken from the NLM data collection project [4]. To demonstrate the robustness of our method we generated two data sets. The first one (NLM-SM) contains a simulated breathing motion. For simulation of diaphragm movement we applied a translational force in the axial direction. This force sharply decreases at those locations where the diaphragm is attached to the rib cage, and linearly decreases in the axial direction (see Fig. 1(b)). For the second data set (NLM-SMC) we simulated the application of a contrast agent by increasing the intensity values of pre-segmented lung vessels by 300 Hounsfield Units (HU) (see Fig. 1(c)).

We ran our algorithm on these two data sets using $\lambda = 40$ and 200 iterations on the finest scale level. For quantitative evaluation we calculated the mean and the standard deviation of the computed displacement field ($\boldsymbol{u}$) with respect to the known ground-truth displacement field. For comparison we additionally ran the popular Demons algorithm [15] on our synthetic data. The implementation was taken from the ITK [5]. We set up the ITK Demons algorithm in a multi-scale manner with 6 pyramid levels, a total number of 50 iterations at the finest scale level and a displacement field smoothing value of $\sigma = 1$. Table 1 shows the results of our experiments. From this we can see that our algorithm performs slightly worse on NLM-SM, but outperforms the Demons algorithm on NLM-SMC, showing the robustness of our algorithm.

[4] http://nova.nlm.nih.gov/Mayo/
[5] http://www.itk.org

(a) (b) (c)

Fig. 1. (a) Original data set. (b) Simulated diaphragm motion. (c) Simulated contrast agent in lung vessels.

Table 1. Quantitative evaluation of the synthetic experiments.

	NLM-SM			**NLM-SMC**		
	Initial	*Demons*	*TV-L^1*	*Initial*	*Demons*	*TV-L^1*
mean [mm]	3.788	**1.555**	1.919	3.788	2.260	**1.768**
std [mm]	2.997	**2.185**	2.285	2.997	2.441	**2.301**

3.2 Clinical Data

The algorithm was also evaluated using clinical data sets. For clinical data clearly no ground-truth displacement field is available. To still allow a quantitative assessment, we calculated two similarity measures, the decrease of root mean squared error of the intensity values (RMS) and the increase of $[0, 1]$-normalized mutual information (NMI).

The first experiment was breathing motion compensation of two thoracic CT data sets (LCT1, LCT2). Each of them consists of two scans at different breathing states. The size of both data sets is 256^3. The second experiment was intra-subject registration of brain MRI. The brain data base consists of four data sets (BMRI1, BMRI2, BMRI3, BMRI4) provided by the Non-Rigid Image Registration Evaluation Project (NIREP) [7]. The size of each data set again was 256^3. We chose BMRI1 as reference image and registered the remaining images to the reference. For the first experiment we ran our algorithm using $\lambda = 50$ and 200 iterations. For the second experiment we used $\lambda = 10$.

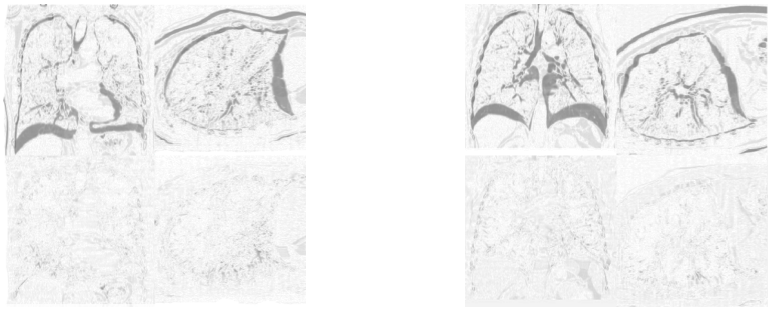
The quantitative results of our experiments are given in Tab. 2. From this we can see a significant improvement in both similarity measures. For visual assessment we also provide qualitative results which are shown in Fig. 2 and Fig. 3.

4 Conclusion

In this work we presented a novel variational approach for nonlinear optical flow based image registration by employing a TV-L^1-optical-flow model. For mini-

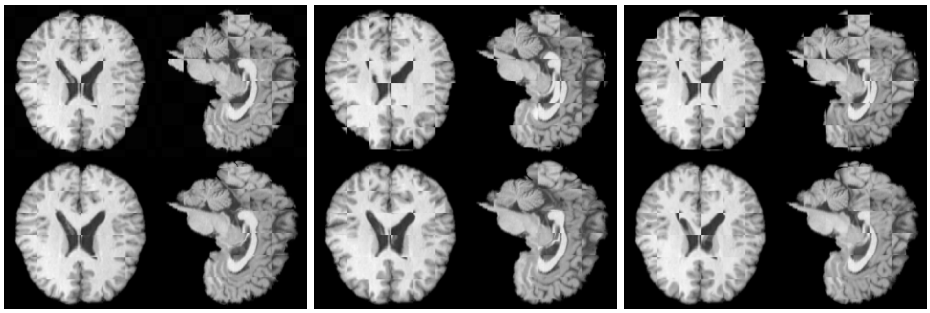
Table 2. Quantitative evaluation of the clinical data sets.

	LCT1		**LCT2**		**BMRI2**		**BMRI3**		**BMRI4**	
	before	*after*	*before*	*after*	*before*	*after*	*before*	*after*	*before*	*after*
RMS	232.88	50.33	245.86	45.94	25.79	12.31	31.11	13.35	30.80	13.06
NMI	0.233	0.418	0.289	0.454	0.299	0.401	0.265	0.394	0.269	0.395



(a) LCT1 (b) LCT2

Fig. 2. Lung CT breathing motion compensation. The upper row shows differences of sagittal and coronal slides before registration, the lower row the same slides after registration.



(a) BMRI2 to BMRI1 (b) BMRI3 to BMRI1 (c) BMRI4 to BMRI1

Fig. 3. Checkerboard representation of inter-subject brain MRI registration. The upper row shows axial an sagittal slides before registration, the lower row shows the same slides after registration.

mization of the model we proposed a novel fast and stable numerical scheme by combining a dual formulation of the total variation energy and an efficient thresholding scheme. For quantitative and qualitative evaluation we used synthetically deformed lung CT data sets, clinical intra-subject thorax CT images and inter-subject brain images. For future work we see two potential directions. One direction is that we plan to use more advanced similarity measures such mutual information. Since our numerical algorithm can be easily parallelized,

a second direction is to speed up our method using state-of-the-art graphics hardware.

References

1. G. Aubert, R. Deriche, and P. Kornprobst. Computing optical flow via variational techniques. *SIAM J. Appl. Math.*, 60(1):156–182, 1999.
2. R. Bajcsy and S. Kovacic. Multiresolution Elastic Matching. *Computer Vision, Graphics and Image Processing*, 46(1):1–21, 1989.
3. M.J. Black and P. Anandan. A framework for the robust estimation of optical flow. In *ICCV93*, pages 231–236, 1993.
4. T. Brox, A. Bruhn, N. Papenberg, and J. Weickert. High accuracy optical flow estimation based on a theory for warping. In *European Conference on Computer Vision (ECCV)*, pages 25–36, 2004.
5. A. Chambolle. An algorithm for total variation minimization and applications. *J. Mathematical Imaging and Vision*, 20:89–97, 2004.
6. T. F. Chan, G. H. Golub, and P. Mulet. A nonlinear primal-dual method for total variation-based image restoration. In *ICAOS '96 (Paris, 1996),*, volume 219, pages 241–252, 1996.
7. G. E. Christensen, X. Geng, J. G. Kuhl, J. Bruss, T. J. Grabowski, I. A. Pirwani, M. W. Vannier, J. S. Allen, and H. Damasio. Introduction to the non-rigid image registration evaluation project (NIREP). In *WBIR 2006, Lecture Notes in Computer Science (LNCS)*, volume 4057, pages 128–135, 2006.
8. G. E. Christensen, R. D. Rabbitt, and M. I. Miller. Deformable Templates Using Large Deformation Kinematics. *IEEE Trans. Image Processing*, 5(10):1435–1447, 1996.
9. W. R. Crum, T. Hartkens, and D. L. G. Hill. Non-rigid image registration: theory and practice. *The British Journal of Radiology - Imaging Processing Special Issue*, 77:S140–S153, December 2004.
10. P. Hellier, C. Barillot, E. Memin, and P. Perez. Hierarchical Estimation of a Dense Deformation Field for 3D Robust Registration. *IEEE Trans. Med. Imag.*, 20(5):388–402, May 2001.
11. B. Horn and B Schunck. Determining optical flow. *Artificial Intelligence*, 17:185–203, 1981.
12. J. B. A. Maintz and M. A. Viergever. A Survey of Medical Image Registration. *Medical Image Analysis*, 2(1):1–36, 1998.
13. N. Papenberg, A. Bruhn, T. Brox, S. Didas, and J. Weickert. Highly accurate optic flow computation with theoretically justified warping. *Int'l J. Computer Vision*, pages 141–158, 2006.
14. L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.
15. J.-P. Thirion. Image matching as a diffusion process: An analogy with Maxwell's demons. *Medical Image Analysis*, 2(3):243–260, 1998.
16. J. Weickert and C. Schnörr. A Theoretical Framework for Convex Regularizers in PDE-Based Computation of Image Motion. *International Journal of Computer Vision*, 45(3):245–264, 2001.
17. C. Zach, T. Pock, and H. Bischof. A duality based approach for realtime TV-L1 optical flow. In *Patter Recognition (Proc. of the 29th DAGM-Symposium)*, Lecture Notes in Computer Science, to appear. Springer, 2007.