

3D MOTION FLOW ESTIMATION USING LOCAL ALL-PASS FILTERS

Christopher Gilliam¹, Thomas Küstner^{2,3} and Thierry Blu¹

¹Department of Electronic Engineering, The Chinese University of Hong Kong, Shatin, Hong Kong

²Institute of Signal Processing and System Theory, University of Stuttgart, Stuttgart, Germany

³Department of Radiology, University of Tübingen, Tübingen, Germany

ABSTRACT

Fast and accurate motion estimation is an important tool in biomedical imaging applications such as motion compensation and image registration. In this paper, we present a novel algorithm to estimate motion in volumetric images based on the recently developed Local All-Pass (LAP) optical flow framework. The framework is built upon the idea that any motion can be regarded as a local rigid displacement and is hence equivalent to all-pass filtering. Accordingly, our algorithm aims to relate two images, on a local level, using a 3D all-pass filter and then extract the local motion flow from the filter. As this process is based on filtering, it can be efficiently repeated over the whole image volume allowing fast estimation of a dense 3D motion. We demonstrate the effectiveness of this algorithm on both synthetic motion flows and *in-vivo* MRI data involving respiratory motion. In particular, the algorithm obtains greater accuracy for significantly reduced computation time when compared to competing approaches.

Index Terms— Motion Estimation, All-Pass Filters, Optical Flow, Approximation, MR Images

1. INTRODUCTION

The estimation of a 3D velocity field that describes the motion between two volumetric images is a problem that has many applications in biological and medical imaging. For example, this motion field, also known as the optical flow [1], finds application in image registration [2], cardiac analysis in 3D cine CT images [3] and cell dynamics in confocal microscopy [4]. In this paper, we approach this motion estimation problem from an optical flow point of view.

The dominant ideas in optical flow estimation stem from the seminal works of Horn and Schunck [1], and Lucas and Kanade [5]. Their approaches both start by assuming brightness consistency—the intensity of a point remains constant as it flows from one image to another [6]. Thus, in 3D, two volumetric images, I_1 and I_2 , are related by

$$I_2(\vec{x} + \vec{u}(\vec{x})) = I_1(\vec{x}) \quad (1)$$

where $\vec{u} = (u_x(\vec{x}), u_y(\vec{x}), u_z(\vec{x}))^T$ is the 3D motion flow (i.e. optical flow) and $\vec{x} = (x, y, z)^T$ is the voxel coordinates. Then, under the assumption that the flow's displacement is small, both sets of authors linearise (1) using a first order Taylor approximation. The result is the *Optical Flow Equation*, which in 3D is:

$$I_2(\vec{x}) - I_1(\vec{x}) - \vec{u}^T \nabla I_1(\vec{x}) = 0, \quad (2)$$

where $\nabla = (\partial/\partial x, \partial/\partial y, \partial/\partial z)^T$. Their approaches differ, however, when faced with the issue that (2) offers only one constraint for multiple unknowns, i.e. ill-posed.

¹This work was supported in part by a grant #CUHK14200114 of the Hong Kong Research Grant Council.

To overcome this issue, Horn and Schunck [1] proposed a global L_2 regularization; they minimised the Euclidean norm of the optical flow equation under a quadratic constraint that favoured smoothly varying flows. In contrast, Lucas and Kanade [5] opted to assume the flow is constant over a local region and solved the optical flow equation within that region. Since their publication, these competing approaches have been subject to constant developments, for a complete review of the state-of-the-art in 2D optical flow estimation see [6]. In the case of 3D imaging, 3D versions of Horn-Schunck and Lucas-Kanade were proposed in [7] and the amount of smoothing required for estimation was studied in [8]. Odille *et al* [9] applied optical flow techniques to motion compensation in MRI data. Along side optical flow algorithms, motion estimation can also be achieved using diffusion based/parametric image registration techniques, such as Demons [10] and fast elastic registration [11]. For a review of the state-of-the-art in deformable image registration see [2].

In this paper, we present a novel algorithm for 3D motion flow estimation using local all-pass filters. Instead of assuming small displacement and using (2), we adapt the 2D optical flow estimation framework outlined recently in [12]. This framework assumes the flow is locally constant and relating local changes in one image to another image using all-pass filters. The optical flow is then extracted from the filters. In contrast to [5], this approach was shown in [12] to produce consistent and accurate 2D flows whilst requiring low computation time. Accordingly, we present a fast filter-based method for estimating smoothly varying motion in 3D images, which we term the 3D Local All-Pass (LAP) algorithm. We then evaluate this algorithm on synthetic data, which exactly satisfies (1), and show that it outperforms two standard non-rigid image registration algorithms in terms of accuracy and speed. Finally, we present initial results for respiratory motion estimation in MR images.

Note that the LAP is not related to Fleet and Jepson's algorithm [13] which relies on the time variation of the spatio-temporal Fourier phase of a *sequence of images*: only spatial filters, between two images, are involved in the LAP.

2. ALL-PASS FILTERING FRAMEWORK

In this section, we introduce and extend the all-pass filtering framework presented in [12] to 3D motion estimation between two volumetric images.

2.1. Relating Shifting and All-Pass Filtering

The central concept in this framework is that a constant motion is equivalent to filtering with an all-pass filter. To observe this equivalence, consider two images, I_1 and I_2 , related by a constant motion $\vec{u}_c = (u_x, u_y, u_z)^T$. Under the assumption of brightness consistency, (1), the relationship linking the images is a simple shifting

operation, $I_2(\vec{x}) = I_1(\vec{x} - \vec{u}_c)$, which in frequency is equivalent to:

$$\hat{I}_2(\vec{\omega}) = \hat{I}_1(\vec{\omega}) e^{-j\vec{u}_c^T \vec{\omega}} \quad (3)$$

where “ $\hat{\cdot}$ ” denotes the Fourier transform of an object and $\vec{\omega} = (\omega_x, \omega_y, \omega_z)^T$ is the frequency coordinates. Accordingly, if we define

$$\hat{h}(\vec{\omega}) = e^{-j\vec{u}_c^T \vec{\omega}},$$

then I_2 is a filtered version of I_1 and the filter, h , is all-pass in nature (i.e. $|\hat{h}(\vec{\omega})| = 1$). Thus, the procedure for determining the motion \vec{u}_c consists of first estimating the all-pass filter h and then extracting the motion information from the filter.

2.2. The All-Pass Filtering Equation

To estimate the filter h , we use the key algorithmic idea proposed in [12]: the all-pass filtering relation between the two images can be expressed linearly in terms of forward and backward filtering. In more detail, assuming ideal sampling with a sinc kernel, we have a digital version of the all-pass filter h . Now, importantly, the $(2\pi, 2\pi, 2\pi)$ -periodic frequency response of any digital all-pass filter can always be expressed as

$$\hat{h}(\vec{\omega}) = \frac{\hat{p}(e^{j\vec{\omega}})}{\hat{p}(e^{-j\vec{\omega}})} \stackrel{\text{def}}{=} \frac{\hat{p}(e^{j\omega_x}, e^{j\omega_y}, e^{j\omega_z})}{\hat{p}(e^{-j\omega_x}, e^{-j\omega_y}, e^{-j\omega_z})}, \quad (4)$$

where $\hat{p}(e^{j\vec{\omega}})$ is the forward and $\hat{p}(e^{-j\vec{\omega}})$ the backward version of a real digital filter p . As a consequence, the filtering operation performed by h can be expressed linearly as a function of p :

$$I_2[\vec{k}] = h[\vec{k}] * I_1[\vec{k}] \Leftrightarrow p[-\vec{k}] * I_2[\vec{k}] = p[\vec{k}] * I_1[\vec{k}] \quad (5)$$

where $\vec{k} = [k, l, m]^T$ is the discrete voxel coordinates. Thus, estimating the forward filter p is equivalent to estimating an all-pass filter that approximates h .

2.3. Basis Representation of the All-Pass Filter

The final step to obtaining the all-pass filter h uses a standard signal processing technique whereby the filter p is expressed as a linear combination of a few fixed, known real filters p_n . In other words, the filter p is approximated using a filter basis representation:

$$p_{\text{app}}[\vec{k}] = \sum_{n=0}^{N-1} c_n p_n[\vec{k}], \quad (6)$$

where N denotes the number of filters in the basis and c_n are coefficients of the filters. The estimation of the all-pass filter h is thus reduced to determining the coefficients c_n using the filtering scheme in (5). A straightforward solution is a mean square minimisation of the difference between the left and right hand sides of (5) when p satisfies (6), which is equivalent to solving a linear system of equations.

Now, an important question is what type of filters should be used in equation (6)? A simple choice would be to use the canonical representation of a finite impulse response (FIR) filter, supported in a cube of side $2R + 1$. However, as pointed out in [12], R acts as an upper bound on the displacement of the motion that can be estimated. Consequently, such a filter basis would be unsuitable for motions with larger displacement (i.e. $N \approx 4\pi R^3/3$). Instead, the answer comes from the analysis presented in [14]. Here, the authors analysed the quality of approximation obtained when using the all-pass filtering framework in optical flow estimation. They showed that an approximation order of r can be achieved if the basis spans the derivatives of an isotropic function up to order $r/2$. Accordingly, we opt for a filter basis that span the derivatives of a Gaussian function and limit the derivative order to 1 (i.e. $r = 2$). In the discrete

domain, this filter basis is

$$p_0[\vec{k}] = \exp\left(-\frac{k^2 + l^2 + m^2}{2\sigma^2}\right), \quad (7)$$

$$p_1[\vec{k}] = k p_0[\vec{k}], \quad p_2[\vec{k}] = l p_0[\vec{k}], \quad p_3[\vec{k}] = m p_0[\vec{k}],$$

where $\sigma = (R + 2)/4$ and the size of the filters are $(2R + 1)^3$ voxels. Importantly, these filters are completely scalable thus allowing estimation of arbitrary sized motion.

2.4. Obtaining the Motion Estimate

Once the all-pass filter has been obtained, we expect its frequency response to be close to $e^{-j\vec{u}_c^T \vec{\omega}}$. Consequently, the following formula can be used to retrieve the estimate of the motion

$$\vec{u}_{x,y,z} = j \left. \frac{\partial \log(h_{\text{est}}(e^{j\omega_x}, e^{j\omega_y}, e^{j\omega_z}))}{\partial \omega_{x,y,z}} \right|_{\omega_x=\omega_y=\omega_z=0},$$

where h_{est} is the estimate of the all-pass filter. Using the all-pass structure defined in (4), this formula can be expressed in terms of the impulse response of the filter p_{app} as follows:

$$\vec{u}_{\text{est}} = 2 \left(\frac{\sum_{\vec{k}} k p_{\text{app}}[\vec{k}]}{\sum_{\vec{k}} p_{\text{app}}[\vec{k}]}, \frac{\sum_{\vec{k}} l p_{\text{app}}[\vec{k}]}{\sum_{\vec{k}} p_{\text{app}}[\vec{k}]}, \frac{\sum_{\vec{k}} m p_{\text{app}}[\vec{k}]}{\sum_{\vec{k}} p_{\text{app}}[\vec{k}]} \right)^T \quad (8)$$

3. ESTIMATING SMOOTHLY VARYING MOTION

We now relax the assumption that the motion is globally constant and consider the estimation of a smoothly varying 3D motion flow.

3.1. 3D Local All-Pass Algorithm

To estimate such a motion, the framework defined in the previous section is adapted to a local level; assuming the motion is locally constant, a region in image I_1 can be related to the same corresponding region in I_2 using a local all-pass filter. This relationship is defined by the *local all-pass equation*:

$$p_{\text{app}}[\vec{k}] * I_1[\vec{k}] - p_{\text{app}}[-\vec{k}] * I_2[\vec{k}] = 0, \quad \text{where } \vec{k} \in \mathcal{W}, \quad (9)$$

and \mathcal{W} is a cubic window with $2R + 1$ sides. Accordingly, the method for determining the motion flow consists of solving the above equation to obtain the local filter—corresponding to the central voxel in \mathcal{W} —then shifting the window and estimating a new all-pass filter. In other words, we have a 3D version of the Local All-Pass (LAP) algorithm proposed in [12].

The central aspect of the 3D LAP algorithm is the following minimisation to obtain the local filter $p_{\text{app}}[\vec{k}]$:

$$\min_{\{c_n\}} \sum_{\vec{k} \in \mathcal{W}} \left| p_{\text{app}}[\vec{k}] * I_1[\vec{k}] - p_{\text{app}}[-\vec{k}] * I_2[\vec{k}] \right|^2 \quad (10)$$

where $p_{\text{app}}[\vec{k}] = p_0[\vec{k}] + \sum_{n=1}^3 c_n p_n[\vec{k}]$.

Now, as we set $c_0 = 1$, this minimisation is equivalent to solving a linear system of equations with 3 unknowns, which can be implemented efficiently using convolutions and pointwise multiplication. Note that, in practise, this linear system of equations could be singular for certain voxels. However, as in [12], we shall assume these cases are rare and deal with any erroneous estimates in the framework outlined in Section 3.2. Finally, once a local all-pass filter has been obtained for all the voxel in the image, we then use the formula in (8) to retrieve a dense estimate of the 3D motion flow.

Importantly, although similar to the formulation of the 3D Lucas and Kanade algorithm [7], the 3D LAP is not restricted to estimating small displacement motion flows, and has been shown in 2D optical flow estimation to be much more consistent [12] and achieve greater accuracy in terms of approximation order [14].

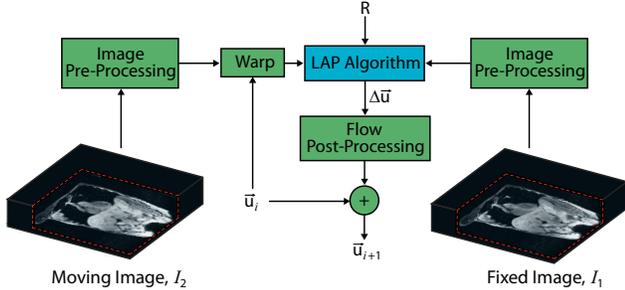


Fig. 1. Diagram illustrating the poly-filter framework for the 3D Local All-Pass (LAP) algorithm. Note that R is the half-support of the LAP filters, \vec{u}_i is the estimate of the motion at the i th iteration and $\Delta\vec{u}$ is the motion increment obtained from the LAP.

3.2. Poly-Filter Framework

Although the 3D LAP is capable of estimating large motion displacement, it requires a filter basis with a large support to do so—the half-support of the filters, R , is the upper bound on the displacement that can be estimated. Thus, to estimate large motion displacements, we must assume the motion flow is varying very slowly, which may not be the case. Accordingly, we use an iterative framework to estimate a broader range of motion flows. However, instead of an iterative refinement involving image downsampling, the framework keeps the full image resolution and changes the support of the filters using the parameter R ; initially, large values of R are used to estimate the large slowly varying parts of the motion flow, then smaller values of R are used for the smaller faster variations in the motion. Thus, we fit the LAP algorithm into a poly-filter framework as shown in Fig. 1.

The framework illustrated in Fig. 1 includes the following processing steps: i) Image Pre-Processing—this step consists of pre-filtering the images with a high pass filter to suppress image noise. ii) Warp—a procedure whereby the moving image, I_2 , is warped closer to the fixed image, I_1 , using the current estimate of the motion flow. This warping is achieved using high quality interpolation, in particular, fast shifted linear interpolation [15]. iii) Flow Post-Processing—this step comprises two elements: Firstly, an inpainting procedure [16] to replace erroneous motion estimates caused when (10) is singular. These estimates are identified if their displacement is greater than the current value of R . Secondly, 3D Gaussian filtering is performed to smooth any errors not previously identified. The filter parameters are $\sigma = 2R$ and a $(4R)^3$ window size.

4. RESULTS

We now compare the performance of the 3D LAP against two types of image registration algorithms: Elastix [17] using a parametric multiresolution cubic B-Spline model - often used in respiratory MR motion estimation [18] - and the non-parametric Demons [19] algorithm, which is comparable to optical flow methods. For the LAP, we use 5 iteration with $R = 16, 8, 4, 2, 1$. All algorithms are run on an Intel Core i7-3770 3.4 GHz with 16 GB RAM.

4.1. Synthetic Evaluation

We start by evaluating the algorithms on synthetic data where the brightness consistency (1) is exactly satisfied – image 1 is generated by warping the image 2 using a known ground truth motion. Under these conditions, we estimate two types of motion flows: a constant flow and a smoothly varying flow. In both cases, the maximum displacement is 8 voxels and we apply these flows to two types of images: Gaussian noise image (i.e. arbitrary image variation) and

a coronal whole-body 3D MR image. For reference, given the original motion flow \vec{u} and its estimate \vec{u}_{est} , performance is measured in terms of the computation time in seconds, the End-point Error, $EE = \|\vec{u} - \vec{u}_{\text{est}}\|_2$, in voxels and the Angular Error (AE), see [6] for more details, in degrees. These errors are then averaged over the whole motion flow (excluding boundaries).

The results of this synthetic evaluation are shown in the Table 1. From the table, we observe that the 3D LAP consistently outperforms the other algorithms when estimating the motion flow in these synthetic conditions; it is roughly 10 times more accurate. Moreover, the table demonstrates that the LAP is significantly faster than the other two; about 2 times faster than Elastix and about 7 times faster than Demons. Importantly, unlike the others, this computation time is achieved using only a Matlab implementation (no C++ code).

4.2. Respiratory Motion Estimation in MRI

Now, we present an initial evaluation of the 3D LAP on real data: respiratory motion estimation on three *in-vivo* MRI datasets. The evaluation is based on the accuracy of registering I_2 to I_1 using the motion estimate obtained from each algorithm. This accuracy is measured in terms of 1) lung segmentation – we perform automatic lung segmentation, using a modified version of method [21], on both I_1 and the registered image I_r , and then measure the overlap between the two segmentations using Dice coefficients [20]. 2) Peak Signal-to-Noise-Ratio (PSNR), in dB, between I_1 and the registered image I_r .

The results of this evaluation, averaged over the three datasets, are shown in the far right handside of Table 1. An example of the estimated respiratory motion and registered image produced by the LAP are shown in Fig. 2. From the table, we observe that the LAP outperforms the other algorithms; it achieves a gain of 0.03, or greater, in lung segmentation accuracy and a PSNR gain of over 1dB. Also, similar to the synthetic case, the LAP requires significantly less computation time. Finally, we observe that the LAP does not estimate spurious motion outside the torso area and exactly identifies the static spine area. Further images illustrating these results and comparing it against elastix and demons can be found at <https://sites.google.com/site/cwsgilliam/3D-LAP>.

5. CONCLUSIONS

In this paper, we have presented a new algorithm for 3D motion estimation in volumetric images. The algorithm is based on a recent filtering framework for optical flow estimation. Specifically, the algorithm uses an all-pass filter to relate a local region in one image to the corresponding region in another image and then extract a local estimate of the motion from the filter. We demonstrated that this algorithm is significantly faster and more accurate than two standard non-rigid image registration algorithms when estimating a constant and a smoothly varying motion flow in synthetic conditions. Finally, we presented initial results for respiratory motion estimation on *in-vivo* MRI data and showed that the flows produced by the LAP enabled more accurate motion compensation than the competition.

6. REFERENCES

- [1] B. Horn and B. Schunck, “Determining optical flow,” *Artificial Intell.*, vol. 17, no. 1, pp. 185–203, 1981.
- [2] A Sotiras, C. Davatzikos, and N. Paragios, “Deformable medical image registration: A survey,” *IEEE Trans. Med. Imag.*, vol. 32, no. 7, pp. 1153–1190, 2013.

Table 1. Comparison for the 3D LAP and two image registration algorithms when performing 3D motion estimation.

	Noise Images ($128 \times 128 \times 64$ voxels)						MR Images ($256 \times 256 \times 72$ voxels)								
	Constant Flow			Smoothly Varying Flow			Constant Flow			Smoothly Varying Flow			Respiratory Motion		
	AEE	AAE	Time	AEE	AAE	Time	AEE	AAE	Time	AEE	AAE	Time	LS-DC	PSNR	Time
3D LAP	0.014	0.065	9.320	0.019	0.319	9.290	0.007	0.038	34.77	0.048	0.771	40.82	0.90 (0.01)	39.93	36.28
Elastix [17]	0.174	0.558	47.20	0.223	4.400	49.80	0.196	0.914	69.42	0.494	7.809	76.00	0.87 (0.02)	37.30	61.55
Demons [19]	0.173	0.784	66.14	0.253	4.853	134.5	0.240	1.070	246.7	0.230	3.070	235.5	0.73 (0.05)	38.23	434.6

* AEE is the mean EE in voxels, AAE is the mean AE in degrees, Time is the computation time in seconds, LS-DC is the mean Dice Coefficient [20] for Lung Segmentation (standard derivation in brackets) and PSNR is the mean error (in dB) between I_1 and the registered version of I_2 using the motion flow.

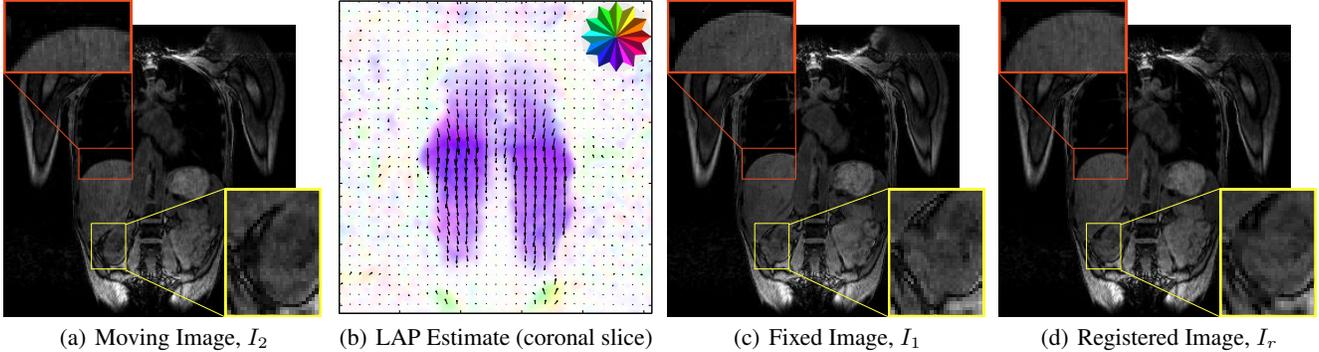


Fig. 2. Example of estimating respiratory motion in MR images using the LAP. For a coronal slice, (a) shows the moving image I_2 , (b) is motion estimate from the LAP, (c) is the fixed image I_1 and (d) is the registered image, I_r , generated using (a) and (b).

- [3] S. Song and R. Leahy, "Computation of 3D velocity fields from 3D cine CT images of a human heart," *IEEE Trans. Med. Imag.*, vol. 10, no. 3, pp. 295–306, 1991.
- [4] J. Delpiano, J. Jara, and J. Scheer *et al.*, "Performance of optical flow techniques for motion analysis of fluorescent point signals in confocal microscopy," *Mach. Vision Applicat.*, vol. 23, no. 4, pp. 675–689, 2012.
- [5] B. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *Proc. Int. Joint Conf. Artificial Intell., Vancouver, Canada*, 1981, vol. 2, pp. 674–679.
- [6] S. Baker, D. Scharstein, and J. Lewis *et al.*, "A database and evaluation methodology for optical flow," *Int. J. Comput. Vision*, vol. 92, no. 1, pp. 1–31, 2011.
- [7] J.L. Barron, "Experience with 3D optical flow on gated MRI cardiac datasets," in *Canadian Conf. Computer Robot Vision*, London, ON, Canada, 2004, pp. 370–377.
- [8] A. Bab-Hadiashar, R. Tennakoon, and M. de Bruijne, "Quantification of smoothing requirement for 3D optic flow calculation of volumetric images," *IEEE Trans. Image Process.*, vol. 22, no. 6, pp. 2128–2137, 2013.
- [9] F. Odille, P. A. Vuissoz, and P. Y. Marie *et al.*, "Generalized reconstruction by inversion of coupled systems (GRICS) applied to free-breathing MRI," *Magn. Reson. Med.*, vol. 60, no. 1, pp. 146–157, 2008.
- [10] J. P. Thirion, "Image matching as a diffusion process: an analogy with Maxwell's demons," *Med. Image Anal.*, vol. 2, no. 3, pp. 243–260, 1998.
- [11] J. Kybic and M. Unser, "Fast parametric elastic image registration," *IEEE Trans. Image Processing*, vol. 12, no. 11, pp. 1427–1442, 2003.
- [12] C. Gilliam and T. Blu, "Local all-pass filters for optical flow estimation," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, Brisbane, Australia, 2015, pp. 1533–1537.
- [13] D. Fleet and A. Jepson, "Computation of component image velocity from local phase information," *Int. J. Comput. Vision*, vol. 5, no. 1, pp. 77–104, 1990.
- [14] T. Blu, P. Moulin, and C. Gilliam, "Approximation order of the LAP optical flow algorithm," in *Proc. IEEE Int. Conf. Image Process.*, Québec City, Canada, 2015.
- [15] T. Blu, P. Thévenaz, and M. Unser, "Linear interpolation revitalized," *IEEE Trans. Image Process.*, vol. 13, no. 5, pp. 710–719, 2004.
- [16] M. Oliveira, B. Bowen, and R. McKenna *et al.*, "Fast digital image inpainting," in *Proc. Int. Conf. Visual. Imag. Image Process. (VIIP)*, Marbella, Spain, 2001, pp. 261–266.
- [17] S. Klein, M. Staring, and K. Murphy *et al.*, "elastix: A toolbox for intensity-based medical image registration," *IEEE Trans. Med. Imag.*, vol. 29, no. 1, pp. 196–205, 2010.
- [18] J. McClelland, D. Hawkes, and T. Schaeffter *et al.*, "Respiratory motion models: a review," *Med. Image Anal.*, vol. 17, no. 1, pp. 19–42, 2013.
- [19] D.-J. Kroon and C. Slump, "MRI modality transformation in demon registration," in *Proc. IEEE Int. Symp. Biomed. Imag.*, Boston, MA, USA, 2009, pp. 963–966.
- [20] L. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.
- [21] P. Kohlmann, J. Strehlow, and B. Jobst *et al.*, "Automatic lung segmentation method for MRI-based lung perfusion studies of patients with chronic obstructive pulmonary disease," *Int. J. Comput Assist Radiol Surg*, vol. 10, no. 4, pp. 403–417, 2015.