Parallel Scales for More Accurate Displacement Estimation in Phase-Based Image Registration

Daniel Forsberg^{*†‡}, Mats Andersson^{*†} and Hans Knutsson^{*†}

^{*}Department of Biomedical Engineering Linköping University, Sweden

[†]Center for Medical Image Science and Visualization (CMIV), Linköping University, Sweden

[‡]Sectra Imtec, Linköping, Sweden

Email: {dafor, matsa, knutte}@imt.liu.se

Abstract—Phase-based methods are commonly applied in image registration. When working with phase-difference methods only a single scale is employed, although the algorithms are normally iterated over multiple scales, whereas phase-congruency methods utilize the phase from multiple scales simultaneously. This paper presents an extension to phase-difference methods employing parallel scales to achieve more accurate displacements. Results are also presented clearly favouring the use of parallel scales over single scale in more than 95% of the 120 tested cases.

I. INTRODUCTION

Phase-difference methods are commonly used in image registration algorithms [1], [2]. The reasons for using phase-difference methods for image registration are manifold. Two of the most frequently ascribed characteristics are sub-pixel accuracy and stability of the phase, the latter extensively investigated in [3]. Most phase-difference methods are based on similar principles, e.g. filter the source image I_S and the target image I_T with a set of quadrature filters in order to estimate the local phase (and the local frequency) of the images. Use the phase-difference (and the frequency) to estimate a displacement field $\mathbf{d}$ to deform the source image, $I_D(\mathbf{x}) = I_S(\mathbf{x} + \mathbf{d}(\mathbf{x}))$. This procedure is often iterated over several scales moving from coarse to finer scales until $I_T(\mathbf{x}) = I_D(\mathbf{x})$.

Related to phase-difference is the concept of phase-congruency [4]. Phase-congruency can be used to create an alternative representation of the images to be registered [5]. A major difference between phase-difference and phase-congruency is that the latter utilizes multiple scales simultaneously whereas the phase-difference methods employ one scale at the time although iterated over multiple scales.

It is well known that the phase is dependent on the scale (center frequency) of the filters used to estimate the phase. However, the change of the phase over different scales is not likely to be sudden but instead it can be argued that visually important features will affect the phase over a larger scale interval than visually unimportant features [6]. To be noted

is that this argumentation is similar to the concept of phase-congruency [4].

The use of multiple/parallel scales and the stability of the phase over a larger scale interval for visually important features are the founding principles for this paper. The aim of this paper is to present an extension to image registration methods based on phase-difference for achieving more accurate displacement estimations. The extension is based upon suggestions both for improving the estimation of the local frequency but also for providing a more robust certainty for the estimated displacement by employing parallel scales instead of a single scale when estimating the phase-difference.

II. METHOD DEVELOPMENT

In this paper we will use the Morphon [7] as a starting point to describe our suggested extension using parallel scales.

A. The Morphon

The Morphon algorithm consists of the following three steps: *local displacement estimation*, *deformation field accumulation*, *deformation*, whereas only the local displacement estimation is of interest to us. For a full description of the Morphon see [7].

The local displacement estimation is based upon a quadrature phase-difference estimation. The source and the target images are filtered with a set of quadrature filters f_k with K different orientations $\hat{\mathbf{n}}_k$. The filter outputs $q_{S_k} = I_S * f_k$ and $q_{T_k} = I_T * f_k$ are used to compute the local phase-difference between the source and the target images based upon the conjugate products of the filter responses $Q_k = q_{S_k} q_{T_k}^*$.

$$d_k = \arg(Q_k) \quad (1)$$

$$c_k = |Q_k|^{1/2} \cos^2\left(\frac{\arg(Q_k)}{2}\right) \quad (2)$$

The phase differences d_k and the coupled certainties c_k are then used to compute an incremental displacement field $\mathbf{d}_i$. This is done by finding the solution to a least square problem

defined as:

$$\min_{\mathbf{d}} \sum_{k=1}^K [c_k \mathbf{T}_S (d_k \hat{\mathbf{n}}_k - \mathbf{d})]^2 \quad (3)$$

where $\mathbf{d}$ is the estimated incremental displacement and $\mathbf{T}_S$ is the local structure tensor of $I_S = \sum_k |q_{S_k}| \mathbf{M}_k$ and $\mathbf{M}_k$ is a projection tensor associated with $\hat{\mathbf{n}}_k$. Solving the least square problem is reduced to solving a linear equation system, $A\mathbf{d} = \mathbf{b} \Leftrightarrow \mathbf{d} = A^{-1}\mathbf{b} = \mathbf{d}_i$.

Note that $\mathbf{d}_i$ has to be multiplied with a scale dependent parameter to translate the estimated phase-difference to a pixel displacement. This corresponds to dividing the phase-difference with the center frequency of the employed filters. A more accurate approach is described in [8] where the filter responses are utilized to estimate the local frequency of the two images and to use the average frequency instead of the center frequency of the employed filters.

B. Extension to parallel scales

The basic idea of our suggested extension is to use parallel scales when filtering the source and the target images, i.e. to use a set of multiple quadrature filters $f_{k,l}$, with different but adjacent center frequencies ρ_l , for each filter orientation $\hat{\mathbf{n}}_k$. Different center frequencies correspond to different scales, hence the term parallel scales.

If using lognormal filters as quadrature filters the ratio between two filter outputs with different center frequencies can be used to estimate the local frequency [9]. However, this local frequency estimation is only valid if the spectrum of the signal falls within the range of the filters (an argument against the use of a single scale to estimate the local frequency as suggested in [8]). To overcome this, [9] also introduces a wide range frequency estimation coupled with an estimation of the variance of the frequency based upon using multiple filters with different center frequencies ρ_l .

Our suggestion is to use this wide range frequency estimation, with minor modifications, to achieve a better displacement estimation but also to let the variance of the spectrum influence the certainty of our displacement. E.g. if the variance of the spectrum is large, the different filters will suggest different displacements for different scales and thus the certainty of the displacement should decrease.

A 2D lognormal filter $F(\mathbf{u})$ in the Fourier domain consists of a radial frequency function $R_l(\rho)$ and a directional function $D_k(\hat{\mathbf{u}})$. The frequency function is defined as $R_l(\rho) = \exp(-\frac{4}{b^2 \ln 2} \ln^2(\rho/\rho_l))$ where ρ_l is the center frequency and b is the 6 dB relative bandwidth in octaves. The directional function is defined as $D_k(\hat{\mathbf{u}}) = (\hat{\mathbf{u}} \cdot \hat{\mathbf{n}}_k)^2$ if $\mathbf{u} \cdot \hat{\mathbf{n}}_k > 0$ and $D_k(\hat{\mathbf{u}}) = 0$ otherwise. A set of L quadrature filters $f_{k,l}$, for each of the K filter orientations $\hat{\mathbf{n}}_k$, is obtained by letting $\rho_l = \rho_0 2^{a_l}$, where a determines the ratio between the center frequencies of the filters.

As before we start off by filtering the image I (for the moment neglecting the subscripts S and T) with the parallel

quadrature filters $q_{k,l} = I * f_{k,l}$. The filter responses are then used to compute a set of local frequency estimates.

$$\rho_{k,l} = \sqrt{\rho_l \rho_{l+1}} \left| \frac{q_{k,l+1}}{q_{k,l}} \right|^{\frac{b^2}{8a}} \quad (4)$$

The set of local frequency estimates are combined to obtain a wide range frequency estimation coupled with a spectrum variance.

$$\tilde{\rho}_k = \left[\sum_{l=1}^{L-1} c_{\rho_{k,l}} \right]^{-1} \sum_{l=1}^{L-1} c_{\rho_{k,l}} \rho_{k,l} \quad (5)$$

$$\tilde{\sigma}_k^2 = \left[\left(\sum_{l=1}^{L-1} c_{\rho_{k,l}}^2 \right) \right]^{-1} \sum_{l=1}^{L-1} c_{\rho_{k,l}}^2 (\rho_{k,l} - \tilde{\rho}_k)^2 \quad (6)$$

$$c_{\rho_{k,l}} = \sqrt{|q_{k,l}| |q_{k,l+1}|} \quad (7)$$

Based upon the spectrum variance estimation we can create a certainty measure for our wide range frequency estimation. This certainty measure accounts both for the variance of the spectrum and for the ratio of the different center frequencies of the filters. The estimated certainty is then used to create a frequency estimate for each scale.

$$c_{\tilde{\sigma}_k^2} = \left(1 + \frac{2\tilde{\sigma}_k^2}{(2^{\frac{a}{\alpha}} - 1)^2} \right)^{-1} \quad (8)$$

$$\tilde{\rho}'_{k,l} = \rho_l^{(1-c_{\tilde{\sigma}_k^2})} \tilde{\rho}_k^{c_{\tilde{\sigma}_k^2}} \quad (9)$$

These computations are applied both to the source I_S and the target I_T images to compute $\tilde{\rho}_{S_k}$, $\tilde{\rho}_{T_k}$, $\tilde{\sigma}_{S_k}^2$, $\tilde{\sigma}_{T_k}^2$, $\tilde{\rho}'_{S_{k,l}}$ and $\tilde{\rho}'_{T_{k,l}}$.

Combining the estimates from both the source and the target images we can now estimate a single frequency estimate for each scale coupled with two certainty measures related to the difference of the estimated frequencies and to the difference in the variance of the estimated frequencies:

$$\tilde{\rho}'_{k,l} = \frac{\tilde{\rho}'_{D_{k,l}} + \tilde{\rho}'_{T_{k,l}}}{2} \quad (10)$$

$$c_{\text{freq}} = e^{-\beta(\tilde{\rho}_{D_k} - \tilde{\rho}_{T_k})^2} \quad (11)$$

$$c_{\text{var}} = e^{-\gamma(\tilde{\sigma}_D^2 - \tilde{\sigma}_T^2)^2} \quad (12)$$

The estimation of d_k and c_k are then extended to include the employed parallel scales but also the estimated frequencies and their coupled certainty measures. We begin by extending Q_k and d_k to include the use of parallel scales, i.e. $Q_{k,l} = q_{S_{k,l}} q_{T_{k,l}}^*$ and $d_{k,l} = \arg(Q_{k,l})$. Also c_k is extended to the use of parallel scales but here a product term is included to suppress displacements that are too large for the filters with higher center frequencies in order to avoid the problem of phase wrapping. This assumes that the phase-difference estimated by the filter with the lowest

Figure 1. From Left: Ploop, Lena, F16/Yosemite, Noise and Mix.

center frequency does not result in phase wrapping.

$$c_{k,l} = |Q_{k,l}|^{1/2} \prod_{m=1}^l \cos^{2(\frac{\rho_l}{\rho_m})^2} \left(\frac{\arg(Q_{k,m})}{2} \right) \quad (13)$$

The estimated frequency and the phase-difference of each scale is used to estimate a single displacement estimate for each filter orientation.

$$d'_{k,l} = \frac{d_{k,l}}{\rho'_{k,l}} \quad (14)$$

$$d_k = \frac{\sum_{l=1}^L 2^{\zeta \rho_l} c_{k,l} d'_{k,l}}{\sum_{l=1}^L 2^{\zeta \rho_l} c_{k,l}} \quad (15)$$

A fourth certainty measure is estimated, determining the consistency of the phase-difference for the different scales, before all certainty measures are combined to a single certainty measure.

$$c_{\text{cons}} = \left[\frac{\left| \sum_{l=1}^L |Q_{k,l}|^{\delta} e^{i d_{k,l} 2^{\rho_l / \rho_l}} \right|}{\sum_{l=1}^L |Q_{k,l}|^{\delta}} \right]^{\epsilon} \quad (16)$$

$$c_k = \left(\sum_{l=1}^L c_{k,l} \right) c_{\text{cons}} c_{\text{freq}} c_{\text{var}} \quad (17)$$

The estimated displacements d_k and the coupled certainties c_k are then used to compute the incremental displacement field $\mathbf{d}_i$ according to (3).

III. RESULTS

The suggested extension has been evaluated on 10 image sequences. The images are made up of a single or several images that have been cut into circles and then patched together, where the patched circles can occlude each other. The image sequences consist of five sequences (see Figure 1) in two versions, where the second include a time-varying contrast gradient for each circle. Each image sequence consists of 13 frames where the circles are moving with different velocities.

The performance of the image registration process has been measured using two similarity measures: normalized cross-correlation (NCC) and mutual information (MI). The

Table I
FRACTION OF IMAGE PAIRS PER IMAGE SEQUENCE AND MEASURE WHERE THE USE OF PARALLEL SCALES PERFORMED BETTER THAN USING A SINGLE SCALE.

Image sequence	NCC	MI	G-RMS	ADE
Ploop	1.00	1.00	1.00	1.00
Ploop _{grad}	1.00	1.00	1.00	1.00
Lena	1.00	1.00	1.00	1.00
Lena _{grad}	1.00	1.00	1.00	1.00
F16/Yosemite	0.75	0.75	1.00	1.00
F16/Yosemite _{grad}	0.75	0.83	1.00	1.00
Noise	1.00	1.00	1.00	1.00
Noise _{grad}	1.00	1.00	1.00	1.00
Mix	1.00	1.00	1.00	1.00
Mix _{grad}	1.00	1.00	1.00	1.00
All sequences	0.95	0.96	1.00	1.00

overall smoothness of the displacement field has been measured using the gradient root mean square (G-RMS):

$$\sqrt{\frac{1}{N} \sum_{k=1}^N \|\nabla \mathbf{d}_k\|} \quad (18)$$

where a smoother field (i.e. lower G-RMS value) is considered to be better if the similarity measures show no difference. Since the true displacement field is available we have also measured the correctness of the estimated displacement field, using the average displacement error (ADE). These measures have been estimated for the registration of each consecutive image pair in all the image sequences (i.e. 12 image pairs per sequence and in total 120 image pairs). The suggested extension (parallel scales) has been compared to using a single scale and a scale dependent parameter as previously described. To evaluate the two methods we have compared each measure for each image pair and image sequence to see which method performed better.

IV. DISCUSSION

The results in Table I show that the use of parallel scales was favorable in almost all cases and for all measures. To be noted is that the use of single scale was only favorable for 2 of the 10 image sequences (the 2 versions of the F16/Yosemite sequence) and only when measured by NCC or MI.

Although the results in Table I are rather straightforward the validity of the results can be questioned since no statistical tests have been used to evaluate the difference between using parallel or single scales. The reason for not presenting any parametric test results (paired t-test) is due to the estimated measures not being Gaussian when considering each measure for all sequences as a separate data set. Non-parametric tests (Wilcoxon signed rank test) also turned out to be insufficient since the data sets were not symmetric around their medians. However, a paired t-test was applied to each measure for each image sequence and where parallel scales were found to perform significantly better in most cases where the data showed normality.

Note that the implementation of parallel scales can be made very efficiently, e.g. using filter networks [10].

Nevertheless, the obtained results are promising and clearly favours the use of parallel scales. Future work includes to test the use of parallel scales on a larger set of images applicable for image registration and to perform statistical tests. Another area to work further with is to compare parallel scales with the use of single scale but with the frequency estimation described in [8] instead of the currently used scale dependent parameter.

Figure 2. A close-up of the registration result ($I_D - I_T$) for the first image pair in the Ploop sequence. Left: Parallel scales and Right: Single scale

REFERENCES

- [1] M. Hemmendorff, M. Andersson, T. Kronander, and H. Knutsson, "Phase-based multidimensional volume registration," in *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2000.
- [2] T. Gautama and M. M. V. Hulle, "A phase-based approach to the estimation of the optical flow field using spatial filtering," *IEEE Transactions on Neural Networks*, vol. 13, no. 5, pp. 1127–1136, September 2002.
- [3] D. J. Fleet and A. D. Jepson, "Stability of phase information," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 12, pp. 1253–1268, 1993.
- [4] P. Kovcs, "Phase congruency: A low-level image invariant," *Psychological Research*, vol. 64, pp. 136–148, 2000.
- [5] A. Wong and J. Orchard, "Robust multimodal registration using local phase-coherence representations," *J. Signal Process. Syst.*, vol. 54, no. 1-3, pp. 89–100, 2009.
- [6] L. Haglund, H. Knutsson, and G. H. Granlund, "Scale analysis using phase representation," in *The 6th Scandinavian Conference on Image Analysis (SCIA)*, 1989.
- [7] H. Knutsson and M. Andersson, "Morphons: Segmentation using elastic canvas and paint on priors," in *International Conference on Image Processing (ICIP'05)*.
- [8] D. J. Fleet, A. D. Jepson, and M. R. M. Jenkin, "Phase-based disparity measurement," *CVGIP Image Understanding*, vol. 53, no. 2, pp. 198–210, March 1991.
- [9] H. Knutsson, C.-F. Westin, and G. H. Granlund, "Local multiscale frequency and bandwidth estimation," in *International Conference on Image Processing (ICIP'94)*.
- [10] B. Svensson, M. Andersson, O. Smedby, and H. Knutsson, "Efficient 3-D adaptive filtering for medical image enhancement," in *International Symposium on Biomedical Imaging (ISBI)*, 2006.