THERMOGRAPHY SPATIAL RESOLUTION ENHANCEMENT BY NON-RIGID REGISTRATION WITH VISIBLE IMAGERY

Gerald Mwangi †, Paul Fieguth⋆, Christoph S. Garbe†

†: University of Heidelberg, Germany, ⋆: University of Waterloo, Canada

ABSTRACT

Over the last decade Carbon-Fiber reinforced polymers (CFRP) have become increasingly important in industry and consumer oriented products due to their physical robustness and shape adaptivity. Defect localization and characterization is key to assuring the quality and durability of CFRPs.

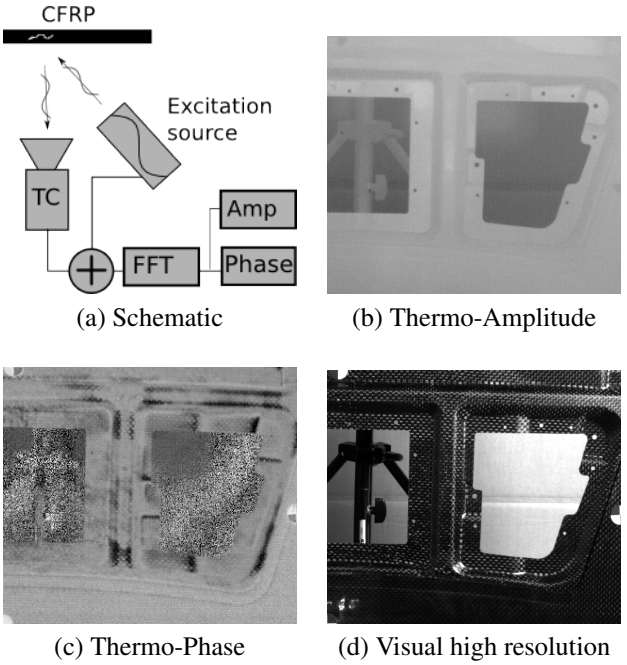
Active thermography has become an important non-invasive technology for the detection of hidden defects within CFRP materials. However current active thermographers exhibit resolutions of only up to 640×512 pixels, whereas visible spectrum imagers routinely have resolutions of 10 megapixels and higher. We therefore propose a fusion of active thermography and visible imagery data, capable of producing higher-resolution thermography images based on the physical boundaries in the visible domain. Our approach handles the disparity between the two imagers *without* the need for stereo or individual calibration.

1. INTRODUCTION

Carbon-Fiber reinforced polymer (CFRP) materials are becoming increasingly widespread in automotive and aerospace industries, but also in consumer goods, due to their adaptivity to different shapes, good rigidity and high strength-to-weight ratio. Improved fabrication techniques are reducing the production costs and time to manufacture. The properties of CFRP strongly depend on the processing of the material, thus defect detection and characterization are indispensable, especially for safety-relevant parts.

Active thermal measurement methods have become vital for the assessment of the quality of CFRP materials. These methods are based on the evaluation of a previously excited heat flow in the tested component and its disturbance by hidden defects, illustrated in Figure 1. The heat flow is generated with a heat pulse or through sinusoidal modulation, observed with a thermography camera, followed by a pixel-wise computation of the complex phase between the excitation signal and the reflected infrared signal. This phase information encodes the heat-loss within a penetration depth δ of the probed material, with depths of 1mm to 2mm typical for CFRPs.

Current state-of-the-art thermography imagers possess resolutions of only 640×512 pixels and a noise equivalent temperature difference of 20mK. Nevertheless, these cameras are very expensive, and the CFRP application domain requires the detection of defects at the noise limit. On the other hand, cameras in the visible spectrum are inexpensive and easily deliver images of 10 megapixels per frame with very little noise. Therefore our strategy is to enhance the resolution of thermography images by utilizing high spatial frequency information from a visible spectrum camera in an image fusion model which takes the disparity between the imagers into account without requiring stereo calibration.



(a) Schematic (b) Thermo-Amplitude

(c) Thermo-Phase (d) Visual high resolution

Fig. 1: A sinusoidal excitation source emits a thermal wave onto a CFRP target. The target may contain a defect not visible on the surface, but which can be detected in the infrared (b,c). However a visual image, (d), can be acquired at a much higher resolution.

The contribution of this paper is a model that is capable of aligning visual and thermal images without the need for camera calibration. The alignment is achieved through the estimation of a non-rigid transformation $\mathbf{d}$ by maximizing correlation, based on the correlation prior $p(s|I)$ from [1]. The results show a significant improvement in image resolution

and defect detectability.

2. BACKGROUND

Thermal cameras are widely used in fields such as remote sensing [2], the agricultural and food industries [3, 4] and building inspection [5], but have recently attracted increasing interest in computer vision. Applications range from using thermal imagery as a segmentation prior [6], in augmented reality [7] and in surveillance [8, 9].

Multi-Modal image fusion has seen many applications, in remote sensing, time of flight imaging [10], and medical image registration [11]. Image registration can also be formulated as optical flow [12, 13] or particle image velocimetry [14].

The data acquisition apparatus consists of a visible spectrum camera (VSC) mounted on top of a thermography camera (TC). The resolution of the VSC is 1226×1028 pixels while that of the TC is 640×512 pixels, both cameras with a focal length of 25 mm. We used a sinusoidal excitation source with a frequency of 0.1 Hz, which corresponds to a penetration depth of approximately 1.3 mm in the CFRP.

3. IMAGE FUSION

The rationale behind image fusion is to combine the desired features of all images, while avoiding limitations, here based on spectrum and resolution.

The image data do not appear well-suited for image registration, due to their strong structural differences and noise. Therefore we propose to apply the relatively straightforward correlation ratio of [15, 1], instead of resorting to more complex approaches using landmark based thin-plate spline methods or mutual information [16, 17]. Furthermore, we avoided notoriously unreliable stereo calibration methods [18] by applying ideas from variational regularization techniques regularly used in optical flow [13, 19].

The method of [1] combines a global correlation prior and an image deconvolution likelihood

$$p\left(s|y,I\right) \sim p\left(y|s\right) \cdot p\left(s|I\right) \tag{1}$$

The likelihood $p\left(y|s\right)$ models the low resolution thermographic image, y, as the result of a blurring and subsequent down-sampling of the unknown high resolution thermographic image s with additive white noise n:

$$y = W_\sigma s + n \quad n \sim \mathcal{N}\left(0|C_n\right) \tag{2}$$

The actions of Gaussian blurring and down-sampling are encoded into the filter W_σ. The prior $p\left(s|I\right)$ extracts the high frequency structural information from the visible spectrum image I and adds it to s. We assume that both s and I are drawn from an approximately Gaussian distribution $p\left(s,I\right)$, an assumption which is at least loosely valid for our data since, from Figure 2b, the bulk of the probability mass in the joint probability distribution function $p\left(y,I\right)$ is approximately Gaussian. Then $p\left(s|I\right) = \frac{p(s,I)}{p(I)}$ must also both be Gaussian

$$\begin{aligned} p\left(s|I\right) &\sim \mathcal{N}\left(\mu_{s|I}|C_{s|I}\right) \\ C_{s|I} &= C_{s,s} - C_{s,I}^2 \cdot C_{I,I}^{-1} \\ \mu_{s|I}\left(\mathbf{x}\right) &= \mu_s + C_{s,I} \cdot C_{I,I}^{-1}\left(I\left(\mathbf{x}\right) - \mu_I\right) \end{aligned} \tag{3}$$

where $C_{\cdot,\cdot}$ is the covariance

$$C_{u,v} = \frac{1}{|\Omega| - 1} \int_\Omega \left(u\left(x\right) - \mathbb{E}\left(u\right)\right) \cdot \left(v\left(x\right) - \mathbb{E}\left(v\right)\right)$$

The conditional mean $\mu_{s|I}$ is linear in the VSC image I and thus encodes the assumption regarding the linearity between the TC and the VSC data. $C_{s|I}$ is an estimator for the functional relationship between s and I

$$\begin{aligned} C_{s|I} = 0 &\rightarrow s = \mu_{s|I} \\ C_{s|I} = C_{s,s} &\rightarrow s \perp I \end{aligned}$$

and is associated with the correlation ratio η (see [20]) as

$$C_{s|I} = C_{s,s}\left(1 - \eta\left(s,I\right)\right)$$

The maximization of (1) with respect to s gives the analytical solution [1]

$$\hat{s} = \mu_{s|I} + C_{s|I} \cdot W_\sigma^T \left(W_\sigma \cdot C_{s|I} \cdot W_\sigma^T + C_n\right)^{-1}\left(y - W_\sigma \mu_{s|I}\right) \tag{4}$$

(4) is intractable to compute due to the dense operator W_σ and the matrix-inverse operation. In [1] a computationally tractable approximation was introduced

$$\begin{aligned} \hat{s} &= \mu_{s|I} + C_{\tilde{s}|\tilde{I}} \cdot \left(C_{\tilde{s}|\tilde{I}} + C_n\right)^{-1}\left(y - \tilde{\mu}_{s|I}\right) \\ \tilde{I} &= W_\sigma I \quad \tilde{s} = W_\sigma s \approx y \end{aligned}$$

On the basis of the test, shown in Figure 2, we set the scaling parameter $\sigma = 4$ to minimize $C_{\hat{s}|\hat{I}}$.

The key issue is that this method requires both modalities, I and y, to be co-registered, which is simply not the case for our TC and VSC data. It is for this reason that a registration / disparity strategy is required.

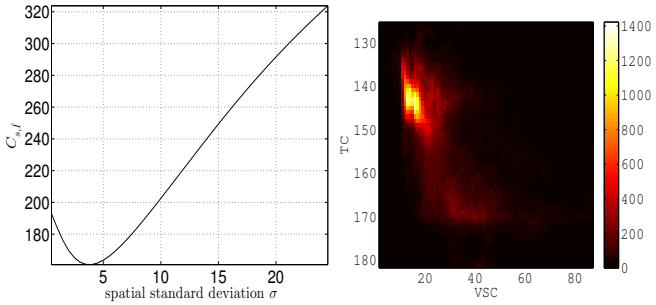
4. DISPARITY

We model the disparity between the TC image y and the VSC image I by assuming y to be co-registered with $I_{\mathbf{d}}$,

$$I_{\mathbf{d}}\left(\mathbf{x}\right) = I\left(\mathbf{x} + \mathbf{d}\left(\mathbf{x}\right)\right)$$

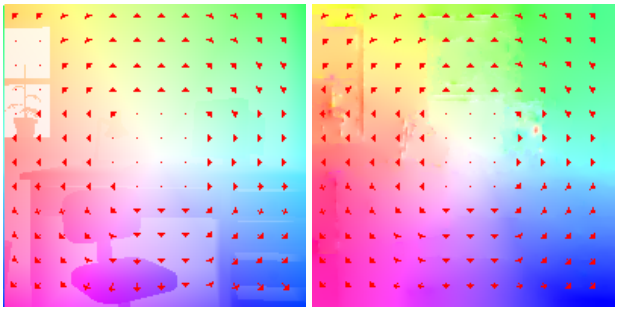
where $\mathbf{d}\left(\mathbf{x}\right)$ is an unknown disparity field. Substituting into (1) and following, we obtain the likelihood

$$p\left(s|y,I,\mathbf{d}\right) = p\left(s|y,I_{\mathbf{d}}\right) \tag{5}$$



(a) (b)

Fig. 2: 2a: Dependence of $C_{\tilde{s}|\tilde{I}}$ on the scaling parameter σ. 2b: Joint Histogram $p(y, I)$ of the TC and smoothed VSC image pair y and $\tilde{I}$ at the optimum $\sigma^{\star} = 4$, the scale at which y and $\tilde{I}$ are maximally linear.



(a) (b)

Fig. 3: Synthetic test based on the office sequence from the middleburry dataset. The ground truth 3a and estimated 3b flow fields are shown, with excellent match. The mean Endpoint error (EPE) is 0.12.

The maximization of (5) with respect to $\mathbf{d}(\mathbf{x})$ is an ill-posed problem. For regions $\mathcal{V}_I$ of I and $\mathcal{V}_y$ of y in which I and y contain little structural information, a pixel $\mathbf{x}_I \in \mathcal{V}_I$ can be mapped to any pixel in $\mathcal{V}_y$, meaning that $\mathbf{d}(\mathbf{x}_I)$ is ambiguous. We therefore apply a smoothness prior for $\mathbf{d}(\mathbf{x})$. A common choice for such a prior would be the total variation (TV) norm, [21, 22], . However our tests showed that the TV prior caused sharp discontinuities from the texture in the VSC image I to lead to discontinuities in the disparity field $\mathbf{d}$. Instead we propose an alternate prior

$$p(\mathbf{d}) = \exp\left(-\frac{\alpha}{2}\int_{\Omega} \|\nabla \mathbf{d}\|^2\right) \tag{6}$$

The prior (6) isotropically penalizes variations of $\mathbf{d}$ across boundaries induced by texture in the VSC image. $\mathbf{d}$ is thus not susceptible to such distortions.

Combining (5) and (6) we obtain the posterior

$$p(s, \mathbf{d}|y, I) = p(s|y, I, \mathbf{d}) \cdot p(\mathbf{d}) \tag{7}$$

The global optimum of the posterior (7) with respect to s is

$$\hat{s}_{I,\mathbf{d}} = \mu_{s|I_{\mathbf{d}}} = \mu_s + C_{s,I_{\mathbf{d}}} \cdot C^{-1}_{I_{\mathbf{d}},I_{\mathbf{d}}} (I_{\mathbf{d}} - \mu_I) \tag{8}$$

Inserting $\hat{s}_{I,\mathbf{d}}$ into posterior (7), the (locally) optimal disparity map $\hat{\mathbf{d}}$ is defined by

$$\hat{\mathbf{d}} = \underset{\mathbf{d}}{\operatorname{argmax}} \{p(\hat{s}_{I,\mathbf{d}}, \mathbf{d}|y, I)\} \tag{9}$$

Figures 4c and 4d depict the resulting flow $\hat{\mathbf{d}}$ and high resolution thermography image $s_{I,\hat{\mathbf{d}}}$. The alignment of $s_{I,\hat{\mathbf{d}}}$ and I at the physical boundaries is well achieved. However the textured regions of I get in-painted into $s_{I,\hat{\mathbf{d}}}$ since they also belong to the high-frequency domain. To resolve this problem we deploy a texture decomposition method based on a TV norm to separate the image $I_{\mathbf{d}}$ into a cartoon part u and a texture part v

$$I_{\mathbf{d}} = u + v$$

The cartoon image u is modeled with a total variation model (TV) defined by the posterior $p(u|I, \mathbf{d})$

$$-\ln p(u|I, \mathbf{d}) = \frac{1}{2}\int_{\Omega} \left\{(u - I_{\mathbf{d}})^2 + \alpha |\nabla u|\right\} d\mathbf{x} \tag{10}$$

removing small high-frequency variations from the image u while preserving sharp discontinuities. We combine $p(u|I, \mathbf{d})$ together with our image alignment posterior (7) to form the joint pdf for s, u and $\mathbf{d}$:

$$p(s, u, \mathbf{d}|I, y) = p(s|y, u) \cdot p(u|I, \mathbf{d}) \cdot p(\mathbf{d}) \tag{11}$$

5. RESULTS

Synthetic Data

We have tested our approach on the office sequence provided in [23] (Figure 3). The first image of the sequence was inverted and noise ($\sigma = 0.2$) added while the second image was left untouched. Figure 3 shows the result of estimating the disparity, $\hat{\mathbf{d}}$, between the frames. The endpoint error [24]

$$\text{EPE} = \left|\hat{\mathbf{d}} - \text{GT}\right|$$

is the basis on which we assess the learned disparity. The mean EPE is 0.12 while the mean GT length is 0.53 meaning that $\hat{\mathbf{d}}$ has a measurement uncertainty of approximately 20%; the EPE is larger in the floor region of the image, since that region has is largely structureless and has a low signal to noise ratio.

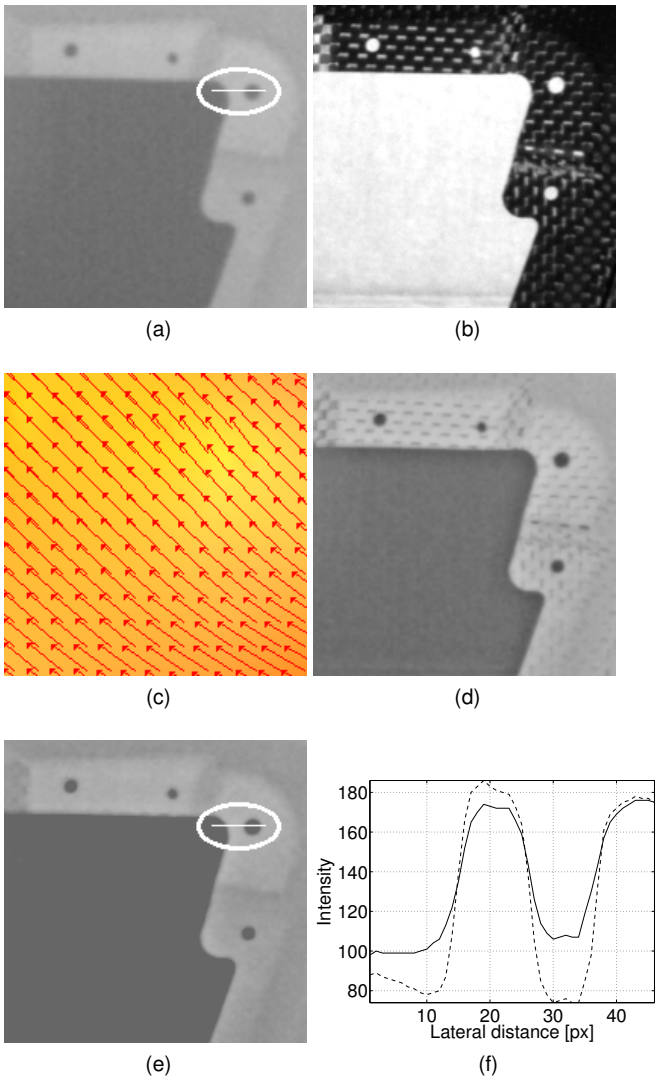


Fig. 4: Based on the amplitude image of the TC 4a and the high-resolution VSC 4b, the disparity 4c was found. 4d shows the estimated high resolution thermographic image $s_{I,\hat{\mathbf{d}}}$, which possesses undesirable high frequency texture from the visible domain. The TV-Prior of (10) effectively removes the high frequency texture, so that $s_{u,\hat{\mathbf{d}}}$ in 4e inherits only the sharp physical boundaries from the VSC image 4b. 4f shows the intensity along a line in the encircled regions in 4a (solid) and 4e (dashed). It shows that the gradients in 4e are steeper as in 4a

Real Data

Figure 4 shows the proposed method applied to visual and thermographic data. The recorded object is a carbon-fiber section of the hood of a car. The images recorded are not co-aligned, since the cameras naturally have a displacement on the camera rig. The estimated disparity, estimated per (9), is shown in Figure 4c, followed by (8) to estimate the high resolution thermographic image $\hat{s}_{I,\hat{\mathbf{d}}}$. The disparity map $\hat{\mathbf{d}}$ in Figure 4c clearly aligns the images, such that the physical edges of the CFRP in the thermographic image (Figure 4a) are enhanced with the edge information of the visible spectrum imager (Figure 4b). Since the car hood is in thermal equilibrium its intensity distribution is spatially uniform in the thermal domain, whereas in the visible domain the carbon fiber structure is visible as a texture which gets in-painted in $s_{I,\hat{\mathbf{d}}}$ (Figure 4d). This texture is clearly an undesired effect, as the aim of our approach is to enhance only the physical boundaries. Consequently, the TV prior for the cartoon image u resolves this problem by removing the texture while preserving the sharp edge discontinuity, shown in Figure 4e. To illustrate the edge sharpness more clearly, Figure 4f plots a sample line scan of interest in Figures 4a (solid line) and 4e (dashed line), showing the improved sharpness of the discontinuities in resolution-enhanced image compared to those in the TC image.

Conclusion

In this paper we introduced active thermography imaging as a means for quality assurance of Carbon-Fiber Reinforced Polymers (CFRP). We tackled the problem of low resolution of current active thermographers by proposing a combination of active thermographer and higher resolution visible spectrum imager. Our proposed model jointly computes the physical disparity between both imagers and a higher resolution thermography image. With this approach we can compute the disparity without the need of camera calibration, as was shown in tests on both synthetic and real data with compelling results.

6. REFERENCES

[1] Y. Zhang, "Spatial resolution enhancement for hyperspectral image based on wavelet baysian fusion," *4th International Congress on Image and Signal Processing*, 2011. 1, 3, 3, 3

[2] G. Hong and Y. Zhang, "Comparison and improvement of wavelet-based image fusion," *International Journal of Remote Sensing*, vol. 29, no. 3, pp. 673–691, 2008. 2

[3] R. Vadivambal and D. S Jayas, "Applications of thermal imaging in agriculture and food industry - a review," *Food and Bioprocess Technology*, vol. 4, no. 2, pp. 186–199, 2011. 2

[4] S. Cubero, N. Aleixos, E. Moltó, J. Gómez-Sanchis, and J. Blasco, "Advances in machine vision applications for automatic inspection and quality evaluation of fruits and

vegetables," *Food and Bioprocess Technology*, vol. 4, no. 4, pp. 487–504, 2011. 2

[5] H. González-Jorge, S. Lagüela, P. Krelling, J. Armesto, and J. Martínez-Sánchez, "Single image rectification of thermal images for geometric studies in façade inspections," *Infrared Physics & Technology*, 2012. 2

[6] J. van Baar, P. Beardsley, M. Pollefeys, and M. Gross, "Sensor fusion for depth estimation, including tof and thermal sensors," in *3D Imaging, Modeling, Processing, Visualization and Transmission (3DIMPVT), 2012 Second International Conference on*. IEEE, 2012, pp. 472–478. 2

[7] M. R. Mine, J. van Baar, A. Grundhofer, D. Rose, and B. Yang, "Projection-based augmented reality in disney theme parks," *Computer*, vol. 45, no. 7, pp. 32–40, 2012. 2

[8] X. Zou and B. Bhanu, "Tracking humans using multi-modal fusion," in *Computer Vision and Pattern Recognition-Workshops, 2005. CVPR Workshops. IEEE Computer Society Conference on*. IEEE, 2005, pp. 4–4. 2

[9] C. Town, "Multi-sensory and multi-modal fusion for sentient computing," *International Journal of Computer Vision*, vol. 71, no. 2, pp. 235–253, 2007. 2

[10] R. Nair, F. Lenzen, S. Meister, H. Schäfer, C. Garbe, and D. Kondermann, "High accuracy tof and stereo sensor fusion at interactive rates," in *Computer Vision–ECCV 2012. Workshops and Demonstrations*, pp. 1–11. Springer Berlin Heidelberg, 2012. 2

[11] B. Zitova and J. Flusser, "Image registration methods: a survey," *Image and vision computing*, vol. 21, no. 11, pp. 977–1000, 2003. 2

[12] S. Baker, D. Scharstein, J. P. Lewis, S. Roth, M. J. Black, and R. Szeliski, "A database and evaluation methodology for optical flow," *International Journal of Computer Vision*, vol. 92, no. 1, pp. 1–31, 2011. 2

[13] B. K.P. Horn and B. G. Schunck, "Determining optical flow," *Artificial Intelligence*, vol. 17, issues 1-3, pp. 185–203, 1981. 2, 3

[14] R. Adrian and J. Westerweel, *Particle image velocimetry*, Cambridge University Press, 2010. 2

[15] W. Zhang, A. Wong, A. Mishra, P. Fieguth, and D. A. Clausi, "Efficient globally optimal registration of remote sensing imagery via quasi-random scale-space structural correlation energy functional," *IEEE GEOSCIENCE AND REMOTE SENSING LETTERS*, vol. 8 No. 5, pp. 997–1001, 2011. 3

[16] A. Biesdorf, S. Wörz, H. Kaiser, C. Stippich, and K. Rohr, "Hybrid spline-based multimodal registration using local measures for joint entropy and mutual information," in *Medical Image Computing and Computer-Assisted Intervention*, pp. 607–615. Springer, 2009. 3

[17] F. Oliveira and J. M. Tavares, "Medical image registration: a review," *Computer Methods in Biomechanics and Biomedical Engineering*, pp. 1–21, 2012. 3

[18] H. Hirschmuller and S. Gehrig, "Stereo matching in the presence of sub-pixel calibration errors," in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*. IEEE, 2009, pp. 437–444. 3

[19] A. Wedel, T. Pock, C. Zach, H. Bischof, and D. Cremers, "An improved algorithm for tv-l 1 optical flow," in *Statistical and Geometrical Approaches to Visual Motion Analysis*, pp. 23–45. Springer, 2009. 3

[20] A. Roche, G. Malandain, X. Pennec, and N. Ayache, "The correlation ratio as a new similarity measure for multimodal image registration," in *Medical Image Computing and Computer-Assisted Interventation*, pp. 1115–1124. Springer, 1998. 3

[21] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin, "An iterative regularization method for total variation-based image restoration," *Multiscale Modeling & Simulation*, vol. 4, no. 2, pp. 460–489, 2005. 4

[22] A. Chambolle, "An algorithm for total variation minimization and applications," *Journal of Mathematical imaging and vision*, vol. 20, no. 1-2, pp. 89–97, 2004. 4

[23] B. McCane, K. Novins, D. Crannitch, and B. Galvin, "On benchmarking optical flow," *Computer Vision and Image Understanding*, vol. 84, no. 1, pp. 126–143, 2001. 5

[24] D. Sun, S. Roth, and M. J. Black, "Secrets of optical flow estimation and their principles," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*. IEEE, 2010, pp. 2432–2439. 5