FAST NO GROUND TRUTH IMAGE REGISTRATION ACCURACY EVALUATION: COMPARISON OF BOOTSTRAP AND HESSIAN APPROACHES

Jan Kybic

Center for Machine Perception, Czech Technical University, Prague, Czech Republic

kybic@fel.cvut.cz, http://cmp.felk.cvut.cz/~kybic

ABSTRACT

Image registration algorithms provide a displacement field between two images. We consider the problem of estimating accuracy of the calculated displacement field from the input images only and without assuming any specific model for the deformation. We compare two algorithms: the first is based on bootstrap resampling, the second, new method, uses an estimate of the criterion Hessian matrix. We also present a block matching strategy using multiple window sizes where the final result is obtained by fusion of partial results controlled by the accuracy estimates for the blocks involved. Both accuracy estimation methods and the new registration strategy are experimentally compared on synthetic as well as real medical ultrasound data.

***Index Terms*—** image registration, accuracy estimation, bootstrap

1. INTRODUCTION

Image registration is an essential tool for many computer vision tasks such as 3D reconstruction from stereo or motion, motion analysis, motion compensation, and video compression [1]. In medical imaging, it is used for intra-subject, inter-subject, and inter-modality analysis; change, growth, and motion detection and quantification [2].

The accuracy of an image registration procedure can be determined using ground truth data [3, 4, 5], a posteriori from the residuals for low-rank transformations [6], or heuristically [7, 8], often based on the correlation value. For Gaussian noise and simple transformations, the performance limits can be found theoretically [9, 10].

We consider the problem of estimating the local accuracy of the recovered displacement field in the case when no ground truth is available and without assuming any specific model for the deformation. It is useful for any subsequent processing to determine to what extent the recovered displacement information can be trusted and thus the weight a particular region should be given. We are especially interested in the problem of elastography [11] where the tissue mechanical properties are estimated from the displacement.

Here we propose a new estimation method named FRAE (Fast Registration Accuracy Evaluation) based on evaluating the uncertainty of the image similarity criterion and its local quadratic approximation. It requires essentially no computational overhead. We compare it with a previously proposed bootstrap based method [12]. For simplicity, we consider only a block matching approach to image registration. However, both methods can be applied to almost any image registration algorithm based on minimization of an image similarity criterion.

This work was sponsored by the Czech Ministry of Education, Project MSM6840770012.

2. BLOCK MATCHING AND BOOTSTRAP

Block matching registration ([13] and many others) partitions the image into a set of overlapping rectangular blocks. For each block an optimal parameter $\hat{\theta} \in \mathbb{R}^d$ of the transformation T_θ is determined by minimizing a criterion J. In our case an SSD criterion is used and T_θ is a translation $T_\theta(\mathbf{x}) = \mathbf{x} + \theta$:

$$\hat{\theta} = \arg\min_\theta J(\theta) \tag{1}$$

$$J = \sum_{\mathbf{x}\in\Omega} \underbrace{\big(f(\mathbf{x}) - g\big(T_\theta(\mathbf{x})\big)\big)^2}_{e(\mathbf{x})} \tag{2}$$

with f and g being images to register and Ω the set of pixels in a block. We attempt to determine the statistical properties of $\hat{\theta}$, first by the bootstrap method [12], which we recall here briefly: A bootstrap resample $\Omega^{(b)}$ is obtained by selecting $N = |\Omega|$ pixels from Ω with replacement. For each of the M multisets $\Omega^{(b)}$, an estimate $\theta^{(b)}$ is obtained by minimization (1). The set of M vectors $\theta^{(b)}$ is used to calculate the bootstrap estimates of the mean and covariance, $\mu_{\hat{\theta}} \approx \operatorname{mean} \theta^{(b)}$, $\mathbf{C}_{\hat{\theta}} \approx \operatorname{var} \theta^{(b)}$, and the geometrical error ε^2

$$\varepsilon^2 = \mathrm{E}\left[\operatorname*{mean}_{\mathbf{x}\in\Omega} \|T_{\hat{\theta}}(\mathbf{x}) - T_{\theta^*}(\mathbf{x})\|^2\right] = \operatorname{tr} \mathbf{C}_{\hat{\theta}} \tag{3}$$

where θ^* is the true parameter value.

3. FAST REGISTRATION ACCURACY ESTIMATION

Our new method (FRAE) analyzes the uncertainty of the criterion value J and uses the shape of the criterion function $J(\theta)$ to translate it into the uncertainty on $\hat{\theta}$. The simplifying assumption is that the criterion function and the error in parameters are approximately normally distributed. This leads to a simple and computationally efficient algorithm.

3.1. Uncertainty of the criterion

We quantify the uncertainty of J for a fixed θ by determining a confidence interval around the observed value J, where J^* is the ideal noiseless value and $\alpha = 0.05$ the confidence level:

$$\mathrm{P}\big[J^* - \gamma \le J \le J^* + \gamma\big] = 1 - \alpha \tag{4}$$
$$\gamma = \Phi^{-1}(1 - \alpha/2)\,\sigma_J \approx 1.96\,\sigma_J$$

where we assume the normality of J, Φ^{-1} is the inverse normal cumulative distribution function and $\sigma_J^2 = \mathrm{Var}\big[J\big]$. Many similarity criteria, including SSD (2), can be expressed as a sum of approximately independent and identically distributed (i.i.d) terms $e(\mathbf{x})$,

provided that the pixel noise is also i.i.d. Then σ_J can be estimated as

$$\sigma_J^2 = N \operatorname{Var}[e] \approx \sum_{\mathbf{x}\in\Omega} \Big(e(\mathbf{x}) - \frac{1}{N}\sum_{\mathbf{x}\in\Omega} e(\mathbf{x})\Big)^2 \tag{5}$$

To improve the estimate in low amplitude regions, we can add a noise due to quantization:

$$\sigma_J^2 = N \operatorname{Var}[e] + N\sigma_Q^2 \tag{6}$$

where $\sigma_Q^2 = \delta^2/12$ is a variance of a uniformly distributed noise in interval $[-\delta/2, \delta/2]$.

3.2. Uncertainty of the parameter $\hat{\theta}$

The second step is to convert the uncertainty in J into the uncertainty in $\hat{\theta}$. Consider a small neighborhood around $\hat{\theta}$. Assume that the true minimum is reached for θ^*, so $J^*(\theta^*) \leq J^*(\hat{\theta})$. From (4) we have with probability $1-\alpha$

$$J(\theta^*) - \gamma \leq J^*(\theta^*) \quad \text{and} \quad J^*(\hat{\theta}) \leq J(\hat{\theta}) + \gamma$$

which yields a condition for θ^*, with probability $(1-\alpha)^2$

$$J(\theta^*) \leq J(\hat{\theta}) + 2\gamma \tag{7}$$

We approximate the criterion function $J(\theta)$ quadratically

$$J(\theta) = J(\hat{\theta}) + \frac{1}{2}(\theta - \hat{\theta})^T \mathbf{H} (\theta - \hat{\theta}) \tag{8}$$

Luckily, this approximation is available to us for free, since the minimization of (1) is performed by a quasi-Newton optimizer that iteratively updates the estimate of $\mathbf{H}^{-1}$ using the BFGS strategy [12], such that $\mathbf{H}$ is positive definite by construction. From (7) and (8) we have

$$\mathrm{P}\big[(\theta^* - \hat{\theta})^T \mathbf{H} (\theta^* - \hat{\theta}) \leq 4\gamma\big] = (1-\alpha)^2 \tag{9}$$

We find a covariance $\mathbf{C}_{\hat{\theta}}$ that a normally distributed $\hat{\theta}$ would have, given (9). If a zero mean normally distributed random variable $\mathbf{z}$ has a covariance $\mathbf{C}$ then $\mathbf{z}^T\mathbf{C}^{-1}\mathbf{z}$ is χ_d^2 distributed with $d = \dim \mathbf{z}$ and its confidence region is

$$\mathrm{P}\big[\mathbf{z}^T\mathbf{C}^{-1}\mathbf{z} \leq F^{-1}(\beta; d)\big] = \beta \tag{10}$$

where F^{-1} is the inverse cumulative χ_d^2 distribution function. Comparing (9) and (10) we get

$$\mathbf{C}_{\hat{\theta}}^{\mathrm{FRAE}} = \frac{4\gamma}{F^{-1}\big((1-\alpha)^2, d\big)} \mathbf{H}^{-1} = \lambda \sigma_J \mathbf{H}^{-1} \tag{11}$$

As an example, for $\alpha = 0.05$ and $d = 2$ we get $\lambda \approx 1.68$.

3.3. Experiment 1 — Error with respect to noise

We take a 512^2 pixel 8 bit grayscale Lena image, shift it by a uniformly distributed random vector with a maximum amplitude of 0.5 pixels in two opposite directions, add a Gaussian noise with standard deviation σ to both shifted images, and register a 60^2 pixel center part by minimizing (1,2). The geometrical error ε (3) was evaluated using the bootstrap and FRAE (11) estimates of the covariance $\mathbf{C}$. The experiment is repeated 100 times for each σ (Figure 1). For noise levels up to $\sigma = 10$ (SNR=28 dB), FRAE performs very well, while the bootstrap underestimates. For higher noise levels bootstrap accuracy improves significantly, while FRAE suffers from underestimation. This is because at high noise levels the criterion function is rough and has high local curvature, leading to an overestimation of the Hessian.

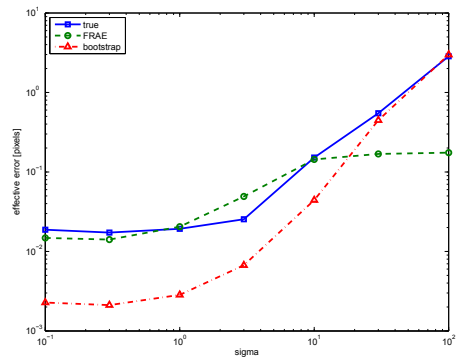

Fig. 1. Mean effective error ε as a function of the noise standard deviation σ — the true value compared with the bootstrap and FRAE estimates.

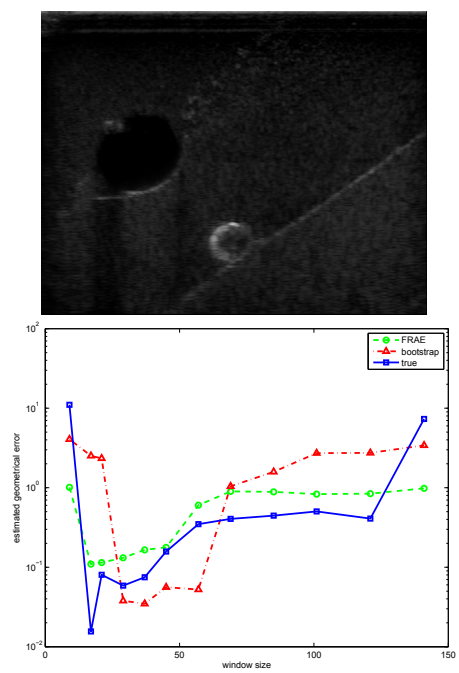

Fig. 2. An ultrasound phantom image (top). Block size versus the true and estimated geometrical error, estimated by bootstrap and the FRAE methods (bottom).

3.4. Experiment 3 — Block size

In Figure 2 we show how registration accuracy for the block matching depends on the block size for an ultrasound phantom image deformed by a known deformation [12]. We see that the registration performance is low for very small block sizes (9 pixels), reaches optimum for width equal to 17 pixels and then the gradually deteriorates again for increasing block sizes because of the inhomogeneity of the deformation field. Both bootstrap and FRAE capture this behavior well, FRAE being better in localizing the minimum, while bootstrap predicts better its value.

4. MULTIPLE BLOCK SIZE MATCHING

We perform the matching for several block sizes and then obtain the final result by combining partial results based on the uncertainty information. We start with a block corresponding to the whole image and progressively subdivide it, maintaining a half-size overlap for blocks at the same level. The subdivision can be terminated when the estimated geometric error ε of a block fails to improve with respect to its parent. The pruning accelerates the algorithm about 3 times but occasionally leads to premature termination. We use a multiresolution strategy for each block to improve robustness and speed.

4.1. Combining results from different blocks

For each point $\mathbf{x} = (x, y)$ in the image we have a set of estimates θ_i with associated covariances $\mathbf{C}_i$ and geometrical errors ε_i (3), corresponding to all blocks (of different sizes) with centers (x_i, y_i) that include $\mathbf{x}$. The simplest strategy to obtain the final estimate $\theta(\mathbf{x})$ is 'winner-takes-all':

$$\theta(\mathbf{x}) = \theta_i \qquad \text{where} \qquad i = \arg\min_j \varepsilon_j \tag{12}$$

Averaging of θ_i by inverse covariances corresponds to an ML estimate for normally distributed measurements:

$$\theta(\mathbf{x}) = \big(\sum_i \mathbf{C}_i^{-1}\big)^{-1} \sum_i \mathbf{C}_i^{-1}\theta_i \tag{13}$$

Further smoothing is obtained by windowing:

$$\theta(\mathbf{x}) = \big(\sum_i w_i(\mathbf{x})\, \mathbf{C}_i^{-1}\big)^{-1} \sum_i w_i(\mathbf{x})\, \mathbf{C}_i^{-1}\theta_i \tag{14}$$

$$w_i(\mathbf{x}) = \left(1 - \kappa\Big(\frac{x - x_i}{h_x/2}\Big)^2\right)\left(1 - \kappa\Big(\frac{y - y_i}{h_y/2}\Big)^2\right) \tag{15}$$

where $h_x \times h_y$ is the block size and $\kappa = 0.99$. A simplification is to consider only the diagonal terms of $\mathbf{C}_i^{-1}$.

4.2. Experiment 4 — Multiple block size approach

Figure 3 shows that multiple block size matching outperforms the fixed block size approach when applied to an artificially deformed ultrasound image (Figure 2) [12] . FRAE results are slightly better than bootstrap; windowing (14) produces smoother fields than the winner-takes-all strategy (12) at the expense of introducing errors at the movement discontinuity boundary (not shown).

4.3. Experiment 5 — Real images

We register an ultrasound image from Figure 2 with an image from the same sequence, acquired after pressure was applied from the top (Figure 4). The block-matching approach produces many artifacts around the cyst. The bootstrap-based multilevel approach produces the smoothest looking results. However, the registration is incorrect inside the cyst, where the bootstrap is too confident. The FRAE-based multilevel result is similar, except that the displacement field is interpolated also in the cyst region. The FRAE registration took about 15 minutes (5 minutes with pruning) while the bootstrap took over 2 hours.

0

5. CONCLUSIONS

We have presented a new, fast and practical method to estimate image registration accuracy, called FRAE. We found experimentally that FRAE often outperforms the bootstrap method, even though it is less general in theory. It is not yet possible to estimate the absolute registration error reliably. However, relative comparison of registration errors within the same image or to study the effect of registration parameters starts to be feasible.

References

[1] Barbara Zitová and Jan Flusser, "Image registration methods: a survey," *Image and Vision Computing*, , no. 21, pp. 977–1000, 2003.

[2] J.B. Maintz and M. A. Viergever, "A survey of medical image registration," *Medical Image Analysis*, vol. 2, no. 1, pp. 1–36, 1998.

[3] M. A. Snyder, "The precision of 3-D parameters in correspondence based techniques: the case of uniform translational motion in rigid environment," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 5, no. 11, pp. 523–528, 1998.

[4] R. M. Haralick, H. Joo, C. N. Lee, X. Zhuang, V. G. Vaidya, and M. B. Kim, "Pose estimation from corresponding point data," *IEEE Trans. Systems, Man and Cybernetics*, vol. 6, no. 19, pp. 1426–1446, 1989.

[5] Jay West et al., "Comparison and evaluation of retrospective intermodality brain image registration techniques," *Journal of Computer Assisted Tomography*, vol. 21, no. 4, pp. 554–568, 1997.

[6] X. Pennec and J.-P. Thirion, "A framework for uncertainty and validation of 3d registration methods based on points and frames," *International Journal of Computer Vision*, vol. 25, no. 3, pp. 203–229, 1997.

[7] M.H. Chan, Y.B. Yu, and A.G. Constantinides, "Variable size block matching motion compensation with applications to video coding," *IEE Proceedings*, , no. 4, Aug. 1990.

[8] G. R. Martin, R. A. Packwood, and I. Rhee, "Variable size block matching motion estimation with minimal error," in *SPIE Proceedings*, Feb. 1996, pp. 324–333.

[9] D. Robinson and P. Milanfar, "Fundamental performance limits in image registration," *IEEE Transactions on Image Processing*, vol. 13, no. 9, pp. 1185–1199, 2004.

[10] I. S. Yetik and Arye Nehorai, "Performance bounds on image registration," *IEEE Transactions on Signal Processing*, vol. 54, no. 5, pp. 1737–1748, May 2006.

[11] J. Ophir, S.K. Alam, B. Garra, F. Kallel, E. Konofagou, T. Krouskop, and T. Varghese, "Elastography: ultrasonic estimation and imaging of the elastic properties of tissues," *Proc. Instn Mech Engrs*, vol. 213, pp. 203–233, 1999.

[12] Jan Kybic and Daniel Smutek, "Image registration accuracy estimation without ground truth using bootstrap," in *CVAMIA: Computer Vision Approaches to Medical Image Analysis*, R. Beichel and M. Sonka, Eds. May 2006, number 3117 in Lecture Notes in Computer Science, pp. 61–72, Springer.

[13] J.R. Jain and A.K. Jain, "Displacement measurement and its application in interframe image coding," *IEEE Trans. Commun.*, pp. 1799–1808, Dec. 1981.

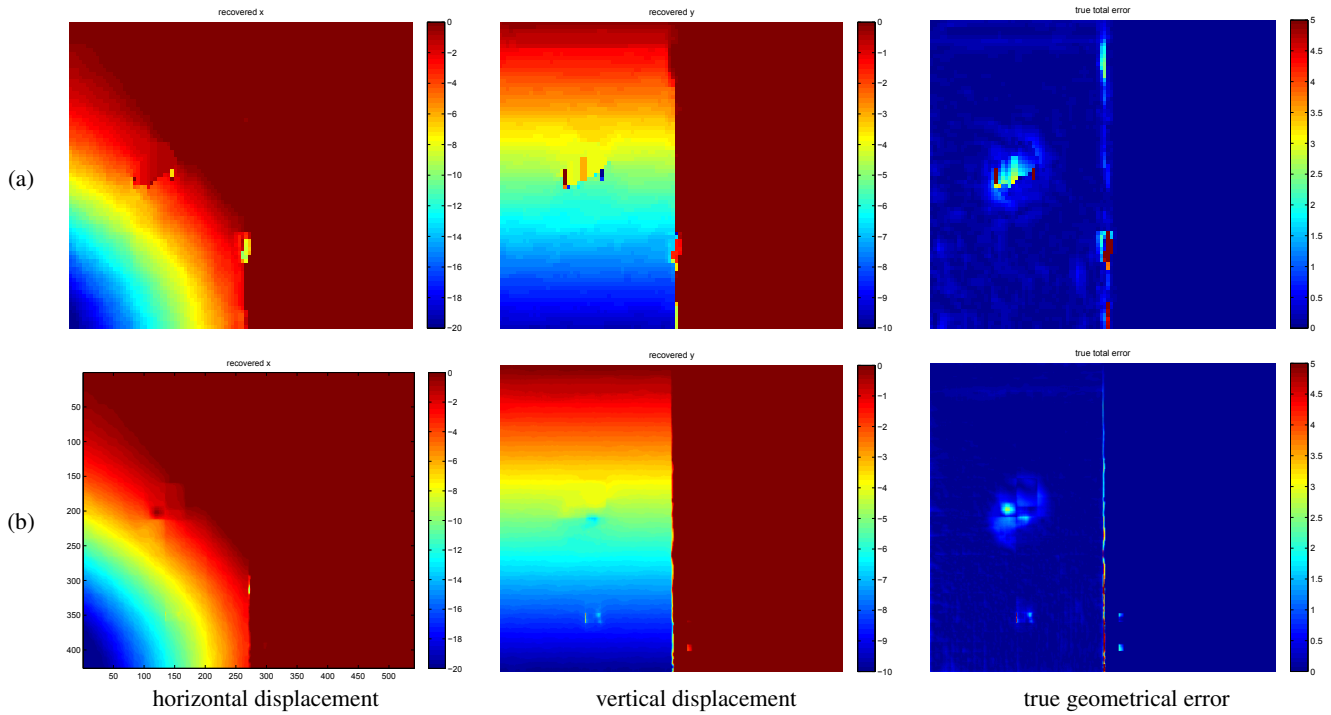

Fig. 3. Recovered deformation and a true geometrical error of the block size registration on an ultrasound image pair with a known synthetic deformation: (a) fixed block-size matching, and (b) multiple block-size matching with a FRAE estimator and windowing. *(We recommend looking at color versions of Figures 3 and 4 available in the electronic version of this paper.)*

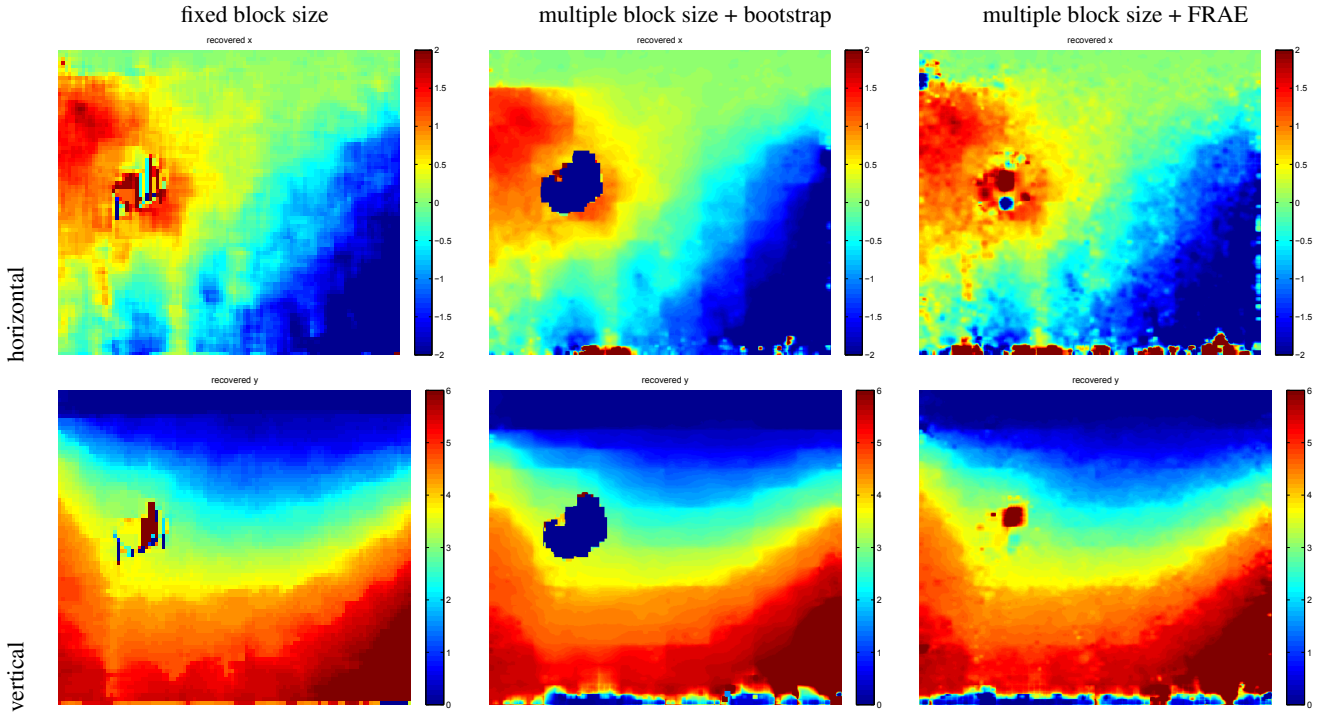

Fig. 4. Horizontal (top row) and vertical (bottom row) displacement from a pair of real ultrasound images. A pressure was applied from the top. Block matching (left), multiple block size with bootstrap (middle) and FRAE (right).