

Similarity Metrics and Efficient Optimization for Simultaneous Registration

Christian Wachinger, Nassir Navab

Computer Aided Medical Procedures (CAMP)

Technische Universität München

Munich, Germany



Motivation

- Analysis of a group of images requires their alignment
 - Alignment of 2D face images, 3D tomographic images, Ultrasound Mosaicing
- Pairwise registration methods with a priori chosen prior lead to bias introduction
- Simultaneous registration presents an interesting alternative but leads to a higher computational complexity because of
 - multivariate similarity measures
 - optimization in higher dimensional parameter space
- Direct estimation of multivariate measures is prohibitive because the dimension grows with the number of images
- We present multivariate similarity metrics and efficient optimization strategies

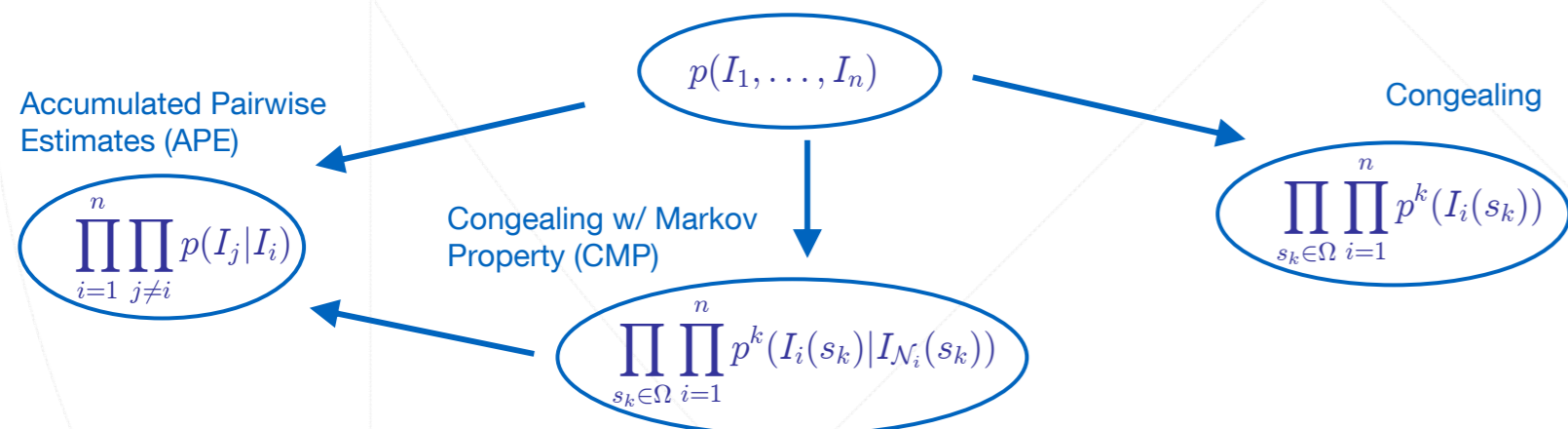
Multivariate Similarity Measures

- Formulate intensity-based registration as maximum likelihood estimation

$$\hat{\mathbf{x}} = \arg \max_{\mathbf{x}} \log p(I_1, \dots, I_n; \mathbf{x})$$

I_i : Images
 \mathbf{x} : Transformation

- Approximations to avoid the estimation of the high-dimensional PDF



- APE and CMP are classes of similarity metrics, while Congealing is a specific one
- APE and CMP need the selection of a noise model

Deduction of APE:

$$I. \quad p(I_1, \dots, I_n) \stackrel{\text{Prod.Rule}}{=} p(I_1, \dots, I_{n-1} | I_n) \cdot p(I_n)$$

$$\stackrel{\text{Cond.Indep.}}{=} \prod_{i=1}^{n-1} p(I_i | I_n) \cdot p(I_n)$$

$$II. \quad p(I_1, \dots, I_n)^n = \prod_{i=1}^n p(I_i) \cdot \prod_{i=1}^n \prod_{j \neq i} p(I_j | I_i)$$

$$\log p(I_1, \dots, I_n) = \frac{1}{n} \sum_{i=1}^n \log p(I_i) + \frac{1}{n} \sum_{i=1}^n \sum_{j \neq i} \log p(I_j | I_i)$$

$$III. \quad \approx \sum_{i=1}^n \sum_{j \neq i} \log p(I_j | I_i)$$

Congealing:

$$p(I_1, \dots, I_n) \approx \prod_{s_k \in \Omega} p^k(I_1(s_k), \dots, I_n(s_k))$$

Independence \swarrow \searrow Markov Property

$$\prod_{s_k \in \Omega} \prod_{i=1}^n p^k(I_i(s_k)) \quad \prod_{s_k \in \Omega} \prod_{i=1}^n p^k(I_i(s_k) | I_{N_i}(s_k))$$

CMP like measure: Sum of Voxel Variances

$$\log p(I_1, \dots, I_n) \approx - \sum_{s_k \in \Omega} \sum_{i=1}^n (I_i(s_k) - \mu_k)^2$$

$$\mu_k = \frac{1}{n} \sum_{l=1}^n I_l(s_k)$$

Efficient Non-Linear Optimization

Transformation Parameterization

Rigid Transformations do not form vector spaces but Lie Groups. We perform a geometric optimization where the group structure is taken care of intrinsically. Special Euclidean Group:

$$\mathbf{x}_i = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} \in \text{SE}(3) \quad \mathbf{x} = [\mathbf{x}_1, \dots, \mathbf{x}_n]$$

The tangent space is the Lie algebra

$$\mathfrak{se}(3) = \left\{ \begin{bmatrix} \Omega & \mathbf{v} \\ \mathbf{0} & 0 \end{bmatrix} \mid \Omega \in \mathbb{R}^{3 \times 3}, \mathbf{v} \in \mathbb{R}^3, \Omega^\top = -\Omega \right\}$$

$$\mathbf{h}_j \in \mathfrak{se}(3) \quad \mathbf{h}_j = \sum_{i=1}^6 h_{i,j} \mathbf{l}_i \quad \mathbf{h} = [\mathbf{h}_1, \dots, \mathbf{h}_n]$$

Exponential map: Lie Algebra to Lie Group

$$\exp : \mathfrak{se}(3) \rightarrow \text{SE}(3)$$

Update step:

$$\mathbf{y} = \mathbf{x} + \mathbf{h}$$

$$\mathbf{y} = \mathbf{x} \circ \exp(\mathbf{h}) = \mathbf{x} \circ \exp\left(\sum_{i=1}^6 h_{i,j} \mathbf{l}_i\right)$$

Optimization Methods

Cost function

$$E(\mathbf{x}) = \sum_{i \neq j} F_{i,j}(\mathbf{x}) = \sum_{i \neq j} \frac{1}{2} \|\mathbf{f}_{i,j}(\mathbf{x})\|^2$$

Taylor series

$$E(\mathbf{x} \circ \exp(\mathbf{h})) \approx E(\mathbf{x}) + \mathbf{J}_E(\mathbf{x}) \cdot \mathbf{h} + \frac{1}{2} \mathbf{h}^\top \cdot \mathbf{H}_E(\mathbf{x}) \cdot \mathbf{h}$$

- Newton-Raphson

$$\mathbf{H}_{F_{i,j}} \mathbf{h}_{i,j}^{\text{NR}} = -\mathbf{J}_{F_{i,j}}^\top \quad \mathbf{x} \leftarrow \mathbf{x} \circ \exp(\mathbf{h}^{\text{NR}})$$

$$\mathbf{h}^{\text{NR}} = \left[\sum_i \mathbf{h}_{i,1}^{\text{NR}}, \dots, \sum_i \mathbf{h}_{i,n}^{\text{NR}} \right]$$

- Steepest Descent

$$\alpha \cdot \mathbf{h}^{\text{SD}} = -\mathbf{J}_E^\top(\mathbf{x})$$

- Gauß-Newton

$$\left(\mathbf{J}_{F_{i,j}}^\top \mathbf{J}_{F_{i,j}} \right) \mathbf{h}_{i,j}^{\text{GN}} = -\mathbf{J}_{F_{i,j}}^\top \mathbf{f}_{i,j}$$

- ESM

$$\left(\mathbf{J}_{F_{i,j}}^{\text{ESM}^\top} \mathbf{J}_{F_{i,j}}^{\text{ESM}} \right) \mathbf{h}_{i,j}^{\text{ESM}} = -\mathbf{J}_{F_{i,j}}^{\text{ESM}^\top} \mathbf{f}_{i,j}$$

Gauß-Newton

Linear approximation of \mathbf{f} : $\mathbf{f}(\mathbf{x} \circ \exp(\mathbf{h})) \approx \mathbf{f}(\mathbf{x}) + \mathbf{J}_f(\mathbf{x}) \cdot \mathbf{h}$ leading to

$$E(\mathbf{x} \circ \exp(\mathbf{h})) \approx \sum_{i \neq j} \frac{1}{2} \|\mathbf{f}_{i,j}(\mathbf{x} \circ \exp(\mathbf{h}))\|^2$$

$$= \sum_{i \neq j} \frac{1}{2} \mathbf{f}_{i,j}(\mathbf{x} \circ \exp(\mathbf{h}))^\top \mathbf{f}_{i,j}(\mathbf{x} \circ \exp(\mathbf{h}))$$

$$= \sum_{i \neq j} \left(F_{i,j}(\mathbf{x}) + \mathbf{h}^\top \mathbf{J}_{F_{i,j}}^\top \mathbf{f}_{i,j} + \frac{1}{2} \mathbf{h}^\top \mathbf{J}_{F_{i,j}}^\top \mathbf{J}_{F_{i,j}} \mathbf{h} \right)$$

ESM (Efficient Second-Order Minimization)

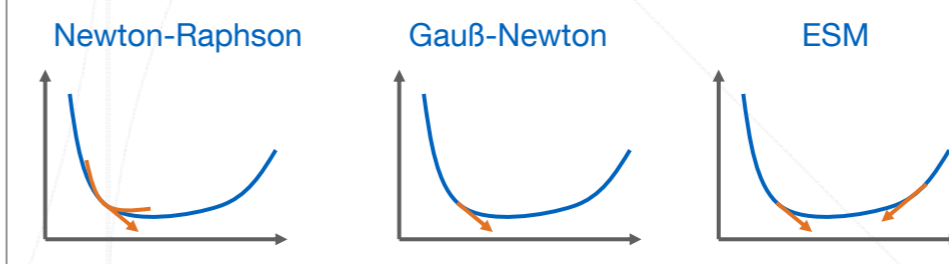
Second-order approximation of $\{$

$$\mathbf{f}(\mathbf{x} \circ \exp(\mathbf{h})) \approx \mathbf{f}(\mathbf{x}) + \mathbf{J}_f(\mathbf{x}) \cdot \mathbf{h} + \frac{1}{2} \mathbf{h}^\top \cdot \mathbf{H}_f(\mathbf{x}) \cdot \mathbf{h}$$

$$\mathbf{J}_f(\mathbf{x} \circ \exp(\mathbf{h})) \approx \mathbf{J}_f(\mathbf{x}) + \mathbf{H}_f(\mathbf{x}) \cdot \mathbf{h}$$

$$\mathbf{f}(\mathbf{x} \circ \exp(\mathbf{h})) \approx \mathbf{f}(\mathbf{x}) + \frac{1}{2} [\mathbf{J}_f(\mathbf{x}) + \mathbf{J}_f(\mathbf{x} \circ \exp(\mathbf{h}))] \mathbf{h}$$

$$\mathbf{J}_f^{\text{ESM}} = \frac{1}{2} (\mathbf{J}_f(\mathbf{x}) + \mathbf{J}_f(\mathbf{x} \circ \exp(\mathbf{h})))$$



Gradient Calculation

Steepest Descent

$$\mathbf{J}_E(\mathbf{x}) = \frac{\partial E(\mathbf{x} \circ \exp(\mathbf{h}))}{\partial \mathbf{h}} = \sum_{i \neq j} \frac{\partial F_{i,j}(\mathbf{x} \circ \exp(\mathbf{h}))}{\partial \mathbf{h}}$$

$$= \sum_{i \neq j} \frac{\partial \text{SM}(I_i(\mathbf{x}), I_j(\mathbf{x} \circ \exp(\mathbf{h})))}{\partial \mathbf{h}} = \sum_{i \neq j} \nabla \text{SM}(I_i, I_j^\dagger) \cdot \nabla I_j^\dagger \cdot \mathbf{J}_w \cdot \Theta(\mathbf{l})$$

Gauß-Newton

Intrinsically non-squared similarity measures in a least-squares optimization framework \Rightarrow squared similarity measures

$$\mathbf{J}_{F_{i,j}}(\mathbf{x}) = \frac{\partial \mathbf{f}_{i,j}(\mathbf{x} \circ \exp(\mathbf{h}))}{\partial \mathbf{h}} \Big|_{\mathbf{h}=0} = \nabla \text{SM}_{i,j} \cdot \nabla I_j^\dagger \cdot \mathbf{J}_w \cdot \Theta(\mathbf{l})$$

ESM

Assumption $\nabla I_j(\mathbf{x} \circ \mathbf{h}^{\text{opt}}) \approx \nabla I_j(\mathbf{x})$ leads to the Jacobian:

$$\mathbf{J}_{F_{i,j}}(\mathbf{x} \circ \exp(\mathbf{h}^{\text{opt}})) \approx \nabla \text{SM}_{i,j} \cdot \nabla I_i \cdot \mathbf{J}_w \cdot \Theta(\mathbf{l})$$

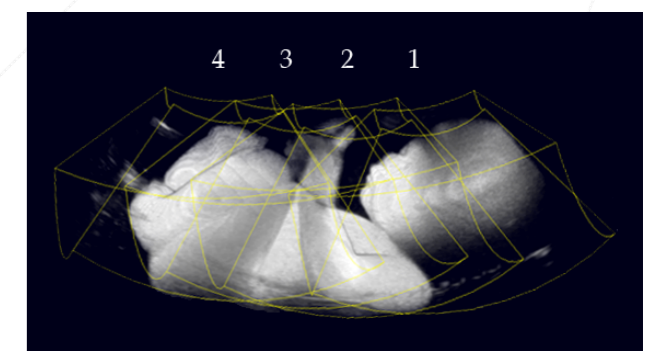
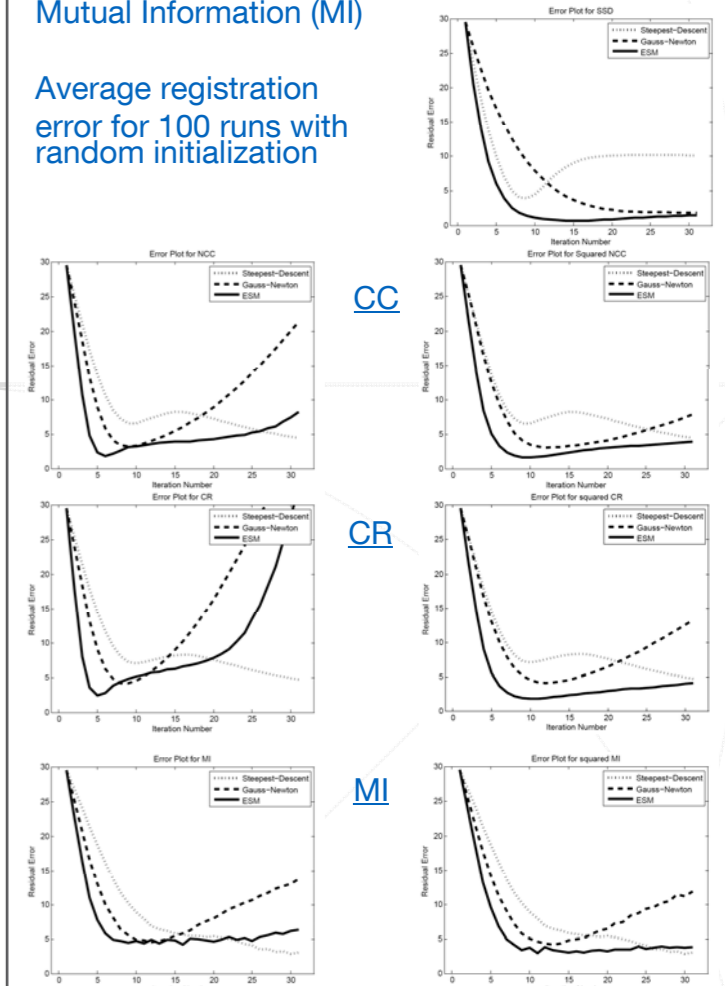
And finally to the Jacobian for ESM:

$$\mathbf{J}_{F_{i,j}}^{\text{ESM}}(\mathbf{x}) = \frac{1}{2} \cdot \nabla \text{SM}_{i,j} \cdot (\nabla I_i + \nabla I_j^\dagger) \cdot \mathbf{J}_w \cdot \Theta(\mathbf{l})$$

Results

Evaluation for Ultrasound Mosaicing. Similarity Measures: Sum of Squared Differences (SSD), Correlation Coeff. (CC), Correlation Ratio (CR), Mutual Information (MI)

Average registration error for 100 runs with random initialization



Conclusion

- Strict mathematical deduction of the new similarity class APE
- Relation of APE to the Congealing framework and extension to CMP
- Derivation of steepest descent, Gauß-Newton, and ESM for APE
- Application of ESM to intrinsically non-squared similarity metrics CC, CR, MI
- Evaluation on 3D ultrasound data

We thank S. Benhimane, D. Zikic, and L. Zöllei for valuable discussions and Siemens Corporate Research for software tools and ultrasound data. This work was partly funded by the European Project "PASSPORT", ref. number 223904.