Deformable Image Registration

Adrien Bartoli
ALCoV – ISIT
Université d’Auvergne
Clermont-Ferrand, France
The Equations of Registration
Deformable Image Registration

To relate the content of at least two images

ROI $\Omega \in \mathbb{N}^2$

The warp $\varphi_i : \Omega \to \mathbb{R}^2$ is a function that transfers pixels between images

Warp visualization grid
Difficulty: Variation of the Imaged Appearance

Scene geometry
- External occlusions
- Self-occlusions
- Wrinkles and foldings
- Temporal continuity
- Extensibility
- Etc.

Surface appearance
- Lighting
- Lack of / repeating texture
- Reflectance
- Specularities
- Transparency
- Etc.

Imaging conditions
- Pose
- Field of view
- Affine vs perspective
- Optical and motion blur
- Auto-exposure, saturation
- Etc.
Semantic Matching
The warp function $\varphi$:

$\varphi : \Omega \rightarrow \mathbb{R}^2$ is continuous and piecewise ‘smooth’.

3D deformation $\Psi \in C^0$
Image Pair Versus Video Registration

Video registration

Image pair registration

A variational optimization problem

$$\min_{\varphi \in C^0} \mathcal{E}[\varphi]$$

$\mathcal{E}$ is the cost functional
A warp $\varphi$ is an interpolant between pairs of fixed/moving driving features.
Computational Warp Representation

Interpolation between \( l \) driving features:
\[
\mathbf{u}_k \leftrightarrow \mathbf{u}'_k \quad \text{with} \quad k = 1, \ldots, l
\]

The \( \mathbf{u}_k \) are known

The \( \mathbf{u}'_k \) are unknown

Linear Basis Expansion warps:
\[
\mathbf{q}' = \varphi(\mathbf{q}) = \sum_{k=1}^{l} \nu_k(\mathbf{q}) \mathbf{u}'_k
\]

Encompasses the Flow-Field, Mesh Interpolation, Radial Basis Functions (and so the Thin-Plate Spline), Tensors Product (and so the Cubic B-Spline Free-Form Deformation), etc.

[Brunet et al., IJCV’11 ; Bartoli et al., IJCV’10]
The Cost Functional

Data term
- Relates the warp to the image content
- Bayesian likelihood

Regularization weight
- Hyperparameter
- Trades-off data-regularization
- $\approx$ Bayesian noise variance

Regularization term
- Measures closeness to prior knowledge
- Bayesian prior

$$\min_{\varphi \in C^0} E_d[\varphi] + \lambda E_s[\varphi]$$

Small $\lambda$          Large $\lambda$

<table>
<thead>
<tr>
<th></th>
<th>Overfitting</th>
<th>About fine</th>
<th>Underfitting</th>
</tr>
</thead>
</table>

Measuring Unsmoothness

\[
\min_{\varphi \in C^0} \mathcal{E}_d[\varphi] + \lambda \mathcal{E}_s[\varphi]
\]

Partial derivatives of order 1 and 2 are commonly used:

\[
\mathcal{E}_s[\varphi] \overset{\text{def}}{=} \int_{\mathcal{F}} \left\| \frac{\partial^r \varphi}{\partial q^r} \right\|^2 dq \quad \text{with} \quad r \geq 1
\]

This energy is Linear Least Squares (and so convex) for Linear Basis Expansion warps

[Bookstein, PAMI’89]
Measuring Data Fitting

$$\min_{\varphi \in C^0} \mathcal{E}_d[\varphi] + \lambda \mathcal{E}_s[\varphi]$$

1 – Feature-based registration
- Wide-baseline
- Keeps only geometric features
- Less accurate than pixel-based

2 – Pixel-based registration
- Small-baseline
- Uses the whole image content
- More accurate than feature-based
Warp Estimation from Keypoint Matches

Keypoint matches $q_j \leftrightarrow q'_j$ with $j = 1, \ldots, m$

$$\min_{\varphi \in C^0} \sum_{j=1}^{m} \left\| q'_j - \varphi(q_j) \right\|_2^2 + \lambda \mathcal{E}_s[\varphi]$$

Property: for Linear Basis Expansion warps the above cost function is **Linear Least Squares**
Keypoint Detection

• Locations where the image changes significantly
  – Repeatable (stable over appearance variation)
  – Accurately localised
  – Discriminable

• Literature is massive (Harris, SIFT, SURF, etc.)
Basic Keypoint-Based Workflow

1. Keypoint detection
2. Putative keypoint matching
3. Warp estimation
Putative Keypoint Matching

There are always mismatches
Robust Keypoint-Based Workflow

Putative keypoint matching → Robust keypoint matching → Warp estimation
Robust Estimation

- Global methods (RANSAC, M-estimators, etc.) are not adapted
- Local methods do much better

Local consistency $\rightarrow$ strong evidence of correctness

1 – Keep all locally consistent matches as strong correct matches
2 – Add matches locally consistent with strong correct matches

[Pizarro et al, IJCV, to appear]
Robust Estimation Results
Robust Estimation Results

Inaccuracies are (partly) due to self-occlusions
Self-Occlusions

- Self-occlusion boundary
- Self-occlusion
- Visible

Current warp profile

Self-occlusions

Desired warp profile
Self-Occlusions

Let $\eta: \mathbb{R}^2 \to \mathbb{R}$ give the warp’s Jacobian ($\eta = |\frac{\partial \varphi}{\partial \mathbf{q}}|$)

Differential properties:

$\eta(\mathbf{q}) \leq 0$ means point $\mathbf{q}$ is self-occluded

$\eta(\mathbf{q}) > 0$ is not conclusive
The ‘Oversmoothing’ Algorithm

Unfortunately $\eta \geq 0$ is a non convex constraint

1 - Estimate $\varphi$ ignoring self-occlusions

2 - Detect partial self-occlusions using $\eta \leq 0$

3 - Oversmooth $\varphi$ in the self-occlusion

$$\tilde{E}_s[\varphi] \overset{\text{def}}{=} \int \kappa \left\| \frac{\partial^r \varphi}{\partial q^r} \right\|_F^2 dq$$

$\kappa(q) = K$ if $q$ if self-occluded and 1 otherwise

The problem stays Linear Least Squares

[Pizarro et al, IJCV, to appear]
Self-Occlusion Estimation Results

1. Initial warp
2. Partial self-occlusion map
3. Self-occlusion resistant warp
Measuring Data Fitting

\[
\min_{\varphi \in C^0} \mathcal{E}_d[\varphi] + \lambda \mathcal{E}_s[\varphi]
\]

1 – Feature-based registration

• Wide-baseline

• Keeps only geometric features

• Less accurate than pixel-based

2 – Pixel-based registration

• Small-baseline

• Uses the whole image content

• More accurate than feature-based
Warp Estimation from the Brightness Constancy Assumption

Matching pixels have the same color

\( \mathcal{T}(q) \approx \mathcal{I}(\varphi(q)) \)
Warp Estimation from the Brightness Constancy Assumption

\[
\min_{\varphi} \int \| T - I \circ \varphi \|_2^2 \, dq + \lambda \mathcal{E}_s[\varphi]
\]

The above cost function is **Nonlinear Least Squares**
\[\rightarrow\text{Iterative local minimization (gradient descent-like)}\]
Color Change

\[
\min_{\varphi, \mathcal{P}} \int \left\| \mathcal{T} - \mathcal{P} \circ \mathcal{I} \circ \varphi \right\|_2^2 \, dq + \lambda \mathcal{E}_s[\varphi]
\]

\[
\mathcal{T}(q) \approx \mathcal{P}(\mathcal{I}(\varphi(q)))
\]
Explicit Photometric Transformation

Video: no photometric model

Video: affine photometric model

[Bartoli, PAMI’08]
Light-Invariant Registration

Template

Video: no photometric model

Light-invariant registration

Image

[Pizarro et al, SCIA’07, CVPR’08]
Registration Results

Video: visualization grid, retexturing

Video: retexturing
Videos: original, visualization grid and retexturing

Videos: original, retargetting and retexturing
What About External Occlusions?

Video: failure in case of external occlusion
Robustification

Idea: inhibitate the data terms that get ‘too large’

The square loss function is replaced by an M-estimator ρ

\[
\min_{\varphi, \mathcal{P}} \int \rho(\mathcal{T} - \mathcal{P} \circ \mathcal{I} \circ \varphi) \, dq + \lambda \mathcal{E}_s[\varphi]
\]
Robust Registration Results

Videos: visualization grid and template with outlying pixels in blue
What About Self-Occlusions?

Video: failure in case of self-occlusion
The ‘Shrinker’ Approach: Principle

Remember that $\eta$ is the warp’s Jacobian
Differential properties are simplified since $\eta \geq 0$:
\[
\eta(q) = 0 \text{ means point } q \text{ is self-occluded}
\]
\[
\eta(q) > 0 \text{ means point } q \text{ is not self-occluded}
\]

[Warp profile]

[Gay-Bellile et al., PAMI’10]
The ‘Shrinker’ Approach: Implementation

Repeat

1 - Compute the self-occlusion map $\mathcal{H}$ using $\eta \leq \epsilon$

2 - Constrained minimization of the cost functional

$$\min_{\varphi, \mathcal{P}} \int \mathcal{H}_\rho (I - \mathcal{P} \circ I \circ \varphi) \, dq + \lambda_s \mathcal{E}_s[\varphi] + \lambda_f \mathcal{E}_f[\varphi]$$
Self-Occlusion Detection Results

Videos (with self-occlusion): visualization grid and template with self-occluded pixels in white
Self-Occlusion Detection Results

Videos (with external and self-occlusion): visualization grid and template with self-occluded pixels in white and external occluded pixels in blue
Self-Occlusion Detection Results

Videos (with self-occlusion): visualization grid and template with self-occluded pixels in white
Self-Occlusion Detection Results

Video (with self-occlusion): retexturing
Video (with self-occlusion): retexturing
Registration Workflow for an Image Pair

1. Feature-based initialization
   - Convex optimizations (Linear Least Squares)
   - Wide baseline, less accurate
   - Partial detection of self-occlusions

2. Pixel-based refinement
   - Non-convex optimization
   - Short baseline, more accurate
   - Complete detection of self-occlusions
Registering a Video

The warps $\varphi_1, ..., \varphi_n$ are **temporally coherent**

$$\min_{\varphi_1, ..., \varphi_n} \mathcal{E}_d[\varphi_1, ..., \varphi_n] + \lambda_s \mathcal{E}_s[\varphi_1, ..., \varphi_n] + \lambda_t \mathcal{E}_t[\varphi_1, ..., \varphi_n]$$

- **data**
- **Spatial smoothness**
- **Temporal smoothness**
Subspace Trajectory Constraints

\[ \min_{\phi_1, \ldots, \phi_n} \sum_{i=1}^{n} \epsilon_d[\phi_i] + \lambda_t \sum_{i=1}^{n} |u_i - \phi_i|_2 + \lambda_s \left| \frac{\partial \alpha}{\partial q} \right|_1 \]

- Data (L1 norm)
- Temporal smoothness
- Spatial smoothness

\[ \phi_i \approx u_i \]

\[ u_i(q) = B_i \alpha(q) \]

Fixed 2D trajectory basis
- PCA of sparse tracks
- Low frequency components of DCT
- B-Splines

\[ \Omega \]

[Garg et al, EMMCVPR’11]
Subspace Trajectory Constraints

\[
\min_{\varphi_1, \ldots, \varphi_n} \sum_{i=1}^{n} E_d[\varphi_i] + \lambda_t \sum_{i=1}^{n} |u_i - \varphi_i|^2 + \lambda_s \frac{\partial \alpha}{\partial \mathbf{q}}^1
\]

Data (L1 norm)  \hspace{1cm} Temporal smoothness  \hspace{1cm} Spatial smoothness

Initialize all warps to identity

For \( i = 1, \ldots, n \)

\[
\min_{\varphi_i} E_d[\varphi_i] + \lambda_t |u_i - \varphi_i|^2
\]

Solved point-wise

\[
\min_{\alpha(\mathbf{q}), \mathbf{q} \in \Omega} \lambda_t \sum_{i=1}^{n} |u_i - \varphi_i|^2 + \lambda_s \left| \frac{\partial \alpha}{\partial \mathbf{q}} \right|^1
\]

Can be solved independently for each basis

This is embedded in a coarse-to-fine approach

→ Show « vid_flag_orig.mp4 »

[Garg et al, EMMCVPR’11]
Video Registration Results

From left to right:
- Subspace trajectory constraints, PCA basis [Garg et al, EMMCVPR’11]
- Subspace trajectory constraints, DCT basis [Garg et al, EMMCVPR’11]
- Large Displacement Optical Flow [Brox et al, ECCV’10]
- Baseline method (no motion basis) [Wedel et al, 2009]
- Pairwise B-Spline pixel-based estimation [Pizarro et al, IJCV]
Deformable Image Registration

Adrien Bartoli
ALCoV – ISIT
Université d’Auvergne
Clermont-Ferrand, France