Appeared in the proceedings of the 2005 VMV Workshop

Feature-based estimation of radial basis mappings for non-rigid registration

Vincent Charvillat, Adrien Bartoli

IRIT-ENSEEIHT-UMR CNRS 5505 & LASMEA - UMR CNRS 6602 - UBP
2, rue Camichel - 31071 Toulouse - France & UBP, Les Cezeaux, 63000 Clermont-Fd - France
Email: Vincent.Charvillat@enseeiht.fr, Adrien.Bartoli@univ-bpclermont.fr

Abstract

We study the challenging problem of registering images of a non-rigid surface by estimating a Radial Basis Mapping from feature matches. We cast the problem as a Maximum Likelihood Estimation coupled with nested model selection. We propose an algorithm based on dynamically inserting centres and refining the transformation parameters under the control of a selection model criterion. We validate the algorithm using extensive simulations and by building on recent feature extraction and matching techniques, we report convincing results on real data.

1 Introduction

The problem of registering images of a non-rigid surface arises in several domains such as augmented reality, medical imaging and video footage processing, for instance changing the texture of a deforming garment. In this context, deriving a physically-based motion model is difficult. Another solution is to compute a parametrized image transformation. *Radial Basis Mappings* (RBMs) have been shown to be effective in representing various non-rigid image deformations. RBMs are non-linear functions defined by centres and coefficients, see e.g. [3, 7, 9, 10].

There exist two main categories of estimation methods: direct and feature-based methods. In direct methods, images are registered by minimizing the intensity difference between the aligned images, *e.g.* [7, 10]. Feature-based methods match features extracted from the images, such as corners or contours, or use landmarks to estimate the transformation. Most of existing methods are defeated by strong deformations, *i.e.* when change in pose or surface deformation yield very distorted images, and by varying appearance, due *e.g.* to change in lightning. Recent work on feature extraction and matching [12] suggests that feature-based methods are more appropriate in the strong deformation case.

The main difficulty in estimating an RBM is due to the varying number of parameters depending, via the number of centers, on the extent of non-rigidities between the images. Computing the right number of centres and their location is crucial to get a good RBM. A possible approach is to use a predefined set of centres such as the nodes of a regular grid or the extracted features. However this either drastically constrain the registration *e.g.* if the grid is too small, or is too flexible, badly extrapolating to the rest of the image.

We tackle the problem of estimating an RBM from feature correspondences. We bring two main contributions. First, §3, we cast the problem as a Maximum Likelihood Estimation (MLE) in a nested model selection framework. We derive a cost function with a data and a complexity term. Second, §4, we propose an algorithm to effectively compute the transformation. The main idea is to iteratively insert centres and jointly minimize the ML cost function over the centres and the transformation coefficients. Inserting centres causes the data term to decrease and the complexity term to increase. An optimization refining the centres and the coefficients is performed at each iteration, see §6. We found this step crucial for the algorithm to effectively converge to the right solution. The dynamic centre insertion procedure is outlined in §5. Experimental results are reported in §7 and our conclusions in §8 The next section gives some notation and preliminaries.

2 Notation and preliminaries

We briefly overview the Radial Basis Mappings (RBMs). For more details, see *e.g.* [3, 7, 10]. A $\mathbb{R}^2 \rightarrow \mathbb{R}^2$ RBM is defined by a basis function ϕ, *e.g.* the Thin-Plate Spline kernel $\phi(\eta) = \eta^2 log(\eta)$, a set of l centres $\mathbf{q}_k$ and a $(l+3) \times 2$ coefficient matrix $\mathsf{H}^\mathsf{T} = \begin{pmatrix} \mathsf{W}^\mathsf{T} & \mathsf{A} \end{pmatrix}$ as:

$$m(\mathbf{x}) = \bar{\mathsf{A}}\mathbf{x} + \mathbf{t} + \sum_{k=1}^{l} \mathbf{w}_k \phi(\|\mathbf{x} - \mathbf{q}_k\|). \quad (1)$$

Parameters $\mathsf{A}_{2\times 3} = \begin{pmatrix} \bar{\mathsf{A}} & \mathbf{t} \end{pmatrix}$ control the affine part. The non-rigid part is a sum of l weighted terms with coefficients $\mathbf{w}_k^\mathsf{T} = \begin{pmatrix} w_k^x & w_k^y \end{pmatrix}$, of the basis function applied to the distance between $\mathbf{x}$ and the centres $\mathbf{q}_k$.

3 A Model Selection Approach

Constructing an automatic non-rigid registration algorithm raises hard steps since the number of centres l as well as their locations $\mathbf{q}_k$ and the coefficients H are unknown. We use a Maximum Likelihood (ML) and nested model selection approach. Let $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$ be the n measured point matches. Assuming that the noise on image point position is Gaussian centred and i.i.d., our goal is to minimize the following cost function:

$$G = \underbrace{\frac{1}{n}\sum_{i=1}^{n} \|m(\hat{\mathbf{x}}_i) - \mathbf{x'}_i\|^2 + \|\hat{\mathbf{x}}_i - \mathbf{x}_i\|^2}_{J_o^2} + \sigma^2 P, \quad (2)$$

over corrected point positions $\hat{\mathbf{x}}_i$ in the first image, the centres $\mathbf{q}_k$ and the coefficient matrix H of the transformation. We define J_o as the *data term* and P as the *complexity term*.

There exist many model selection criteria for balancing the residual and the degrees of freedom of the model. Most of them are based on statistical and information-theoretic criteria. The most widely used criteria are the Akaike's criterion (AIC) and the Minimum Description Length (MDL)[1]:

$$\mathrm{P_{AIC}} = 2k \quad \mathrm{P_{gMDL}} = -k log(\sigma^2),$$

where k is the number of degrees of freedom of the transformation (proportional to l in our case and depending on the effective unknown values: the $\mathbf{q}_k$ may be included as unknowns). The noise level σ^2 cannot be simply estimated from the residuals since an interpolating transformation is reached by increasing the number of centres, causing σ to vanish. We rather estimate σ based on the uncertainty of the feature extractor. The following table shows the criteria compared in our experiments:

[1]We present the geometric MDL (gMDL) due to [8].

1. *Initialization:* compute an initial affine transformation.
2. *Initial centres:* insert 4 centres, see §5. $l \leftarrow 4$.
3. *Refinement:* refine the centre positions and the transformation coefficients, see §6.
4. *Model selection:* if the cost function (2) decreases insert a new centre ($l \leftarrow l+1$, see §5) and loop on 3 else stop.

Table 1: The proposed registration algorithm.

Criterion	Complexity term P	Reference
AIC	$2k$	[1]
CAIC	$k(logn + 1)$	[2]
BIC	$2klog(n)$	[6]
MDL	$\frac{1}{2}klog(n)$	[8]
gMDL	$-klog(\sigma^2)$	[8]

4 Algorithm Overview

We propose an algorithm to minimize the cost function (2) based on *dynamic centre insertion*. The idea is, starting from a rigid affine transformation, to iteratively insert new centres to increase the extent of non-rigidy until it becomes irrelevant. We optimize both the tranformation coefficients and the centre positions at each iteration. The algorithm is summarized in table 1. Described in the next section is the dynamic centre insertion procedure. More details on the computation of the coefficient matrix and the centres, given their number l, are given in §6.

5 Dynamic centre insertion

When inserting a centre, we have to choose its position according to some criterion. One may think of some heuristic procedures. In the context of intensity-based direct methods, [10] proposes to insert the center by considering peaks in the differ-

ence image. Experiments show that this method can not be directly employed in the feature-based case. This is due to the leveraging and averaging effects in the least-squares estimation of the initial affine transformation. Spreading out the centres uniformly in the image is another possible heuristic insertion procedure. Without prior knowledge about the transformation, a uniform distribution is used.

Both previous heuristic solutions tend to overfit the global transformation since the first inserted centers may be badly located. We propose to use a global optimization method proposed by Nelder and Mead in 1965, see *e.g.* [11]. It uses a simplex in the parameter space, for instance a triangle in $\mathbb{R}^2$, to search a globally optimal solution. Each vertex is associated with a value of the criterion to be optimized so that the vertices are sorted. The Nelder-Mead algorithm heuristically improves the worst vertex. Several elementar yet general operations such as reflection, expansion, contraction of the vertices drive the search. This update procedure is repeated until the variance of the cost function or the volume of the simplex fall below a threshold. In our implementation the optimization is performed over the centre positions ($2l$ parameters) and the criterion is the registration error J_t computed with the linear estimate of H for a candidate set of centres. The side-conditions $\mathsf{P}^\mathsf{T}\mathsf{W} = \mathbf{0}$ are enforced using numerical elimination via QR decomposition of matrix W [4].

6 Estimation Given the Number of Centres

When the number l of centres is fixed, the complexity term P of the cost function (2) becomes irrelevant.

6.1 Linear Estimation

We propose to minimize an approximation to the ML data term J_o, over the coefficients H, given by the least-squares transfer error J_t:

$$J_t^2 = \frac{1}{n}\sum_{i=1}^{n} \|m(\mathbf{x}_i) - \mathbf{x}_i'\|^2 = \| \begin{pmatrix} \mathsf{K} & \mathsf{P} \end{pmatrix} \mathsf{H} - \mathsf{P}' \|^2, \tag{3}$$

where the i-th row of P is $\begin{pmatrix} \mathbf{x}_i^\mathsf{T} & 1 \end{pmatrix}$ and the (i,c)-th entry of K is $\phi(\|\mathbf{x}_i - \mathbf{q}_c\|)$. Ensuring the side conditions $\mathsf{P}^\mathsf{T}\mathsf{W} = \mathbf{0}$ is important to get a well-behaved transformation [3]. We employ a QR-based method to solve the resulting constrained linear least squares problem [4].

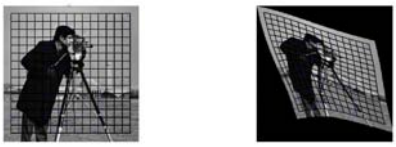

Figure 1: *Simulated data: the reference image and injected transformation represented by the black grid.*

6.2 Maximum Likelihood Refinement

Given the initial solution, we minimize J_o over the centres and coefficients. The Levenberg-Marquardt algorithm is used, adapted to exploit the sparse block structure of the Jacobian matrix. We use the netlib software ODRPACK which implements the Boggs *et al.* algorithm [5].

7 Experimental Results

7.1 Simulated data

The aim of these experiments is to validate the algorithm and determinate conditions under which it converges. As figure 1 shows, we inject a transformation $\hat{m}$ involving $l = 7$ centers to an image. The 15×15 grid illustrate the deformation and serves as ground-truth. Uniformly chosen correspondences $\{\mathbf{x}_i \leftrightarrow \mathbf{x}_i' = \hat{m}(\mathbf{x}_i)\}$ are shown by blue circles, independently corrupted with a gaussian centred noise with standard deviation σ.

Figure 2 shows a correct estimation of the non-rigidities thanks to model selection. At each iteration an estimate with increasing degrees of freedom is provided (affine model at iteration 1, 4 centres inserted at iteration 2 and one more centre for each following step). For each estimate the reference grid is transferred on the second image. The distances between these transferred nodes and the target ones give the ground truth error. The small noise level allows a good superimposition between the J_o blue plot and this ground thruth. Both AIC

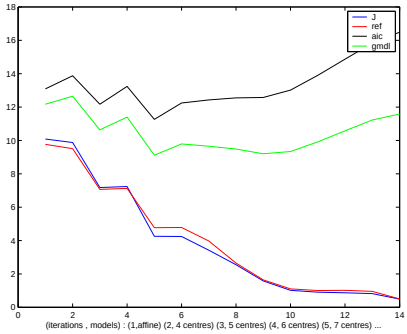

Figure 2: *Simulated data ($\sigma = 0.5$ pixel). transfer error and ground truth are similar (lowest plots) and model selection criteria perform well (hightest plots)*

$\sigma = 0.4$				
Criteria	Underfit	Correct	+1/2	Overfit
AIC	0%	40%	42%	18%
CAIC	2%	45%	46%	7%
BIC	4%	59%	34%	3%
MDL	0%	41%	39%	20%
gMDL	0%	55%	27%	18%

Table 2: Results with $\sigma = 0.4$

and gMDL criteria detect the correct dof $(l = 7)$ at the 5-th iteration.

The following tables (tables 2 and 3) report some results about the fitted transformations after extensive simulations (averages over 50 runs). For a moderate noise ($\sigma \leq 0.5$ pixels) most of the criteria perform well.

We notice that all criteria tend to overfit the data. The worst criteria is this case are AIC, MDL and gMDL.If we increase the noise level ($\sigma \geq 1$ pixel) most criteria become very unstable: AIC and MDL underfit and overfit, CAIC and BIC underfit and gMDL largely overfit the data. In the light of these results, we recommend CAIC or BIC if the noise variance is subpixel and CAIC if the noise is strong than a pixel.

$\sigma = 1.2$				
Criteria	Underfit	Correct	+1/2	Overfit
AIC	12%	39%	37%	12%
CAIC	21%	48%	30%	1%
BIC	37%	57%	6%	0%
MDL	12%	40%	37%	11%
gMDL	0%	1%	31%	68%

Table 3: Results with $\sigma = 1.2$

Figure 3: *The 'baby carpet' image pair.*

7.2 Real data

We apply our algorithm to the images shown in figures 3 and 6. These images were acquired by crumpling some patterned materials. The extracted features, shown on figure 4, are the affine covariant features of [12]. Visually the registrations, shown on figures 5 and 7, are satisfying. A registration error (obtained from manual landmark selection) has been computed to assess the performance. These errors are below 3 pixels in average and confirm good visual results.

Figure 4: *Extracted affine covariant features.*

Figure 5: *Estimated deformation.*

Figure 6: *The 'kitchen towel' image pair.*

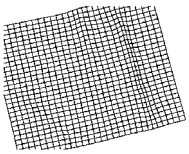

Figure 7: *The 'kitchen towel' deformation.*

8 Conclusions and Further Work

We tackled the problem of registering images of a non-rigid surface. We built on recent feature extraction and matching techniques and state a Maximum Likelihood Estimation problem for Radial Basis Mappings. A framework based on nested model selection allows us to derive an original algorithm using data-driven, dynamic centre insertion, to refine the estimated transformation. This provides a theoretically derived way of assessing convergence. Possible avenues for future work include robustness to outliers and a mixed feature-based and direct approach.

References

[1] H. Akaike "A new look at the statistical model identifi cation" *Trans. on Aut. Ctrl.*, vol. 19, no. 6, 1974

[2] H. Bozdogan "Model selection and Akaike's information criterion" *Psychometrika*, vol. 52, no. 3, 1987

[3] F.L. Bookstein "Principal warps: Thin-plate splines and the decomposition of deformations " *IEEE Trans. on PAMI*, vol. 11, no. 6, 1989

[4] A. Bjorck "Numerical Methods for Least Squares Problems " *SIAM*, Philadelphia (PA), 1996

[5] P.T. Boggs and R.H. Byrd and R.B Schnabel "A stable and effi cient algorithm for non linear orthogonal distance regression" *SIAM Journal on Scientific and Statistical Computing*, 1987

[6] P.H.S. Torr "Geometric motion segmentation and model selection" *Phil. Trans. of the Royal Society of London*, 356, 1998.

[7] H.J Johnson and G.E. Christensen "Consistent landmark and intensity-based image registration" *IEEE Trans. on Medical Imaging*, vol 21, no. 5, 2002.

[8] K. Kanatani "Uncertainty modeling and model selection for geometric inference" *IEEE Trans. on PAMI*, vol 26, no. 4, 2004

[9] M. Orr "Introduction to radial basis function networks" *TR Univ. of Edinburgh*, 1996

[10] A. Bartoli and A. Zisserman "Direct Estimation of Non-Rigid Registrations" *BMVC*, pp. 899-908, 2004.

[11] R. M. Lewis and V. Torczon and M. W. Trosset "Direct search methods: then and now" *J. of. Comp. and App. Math.*, vol 124, pp. 191-207, 2000.

[12] K. Mikolajczyk and C. Schmid, "Scale and Affi ne invariant interest point detectors." *IJCV* 1(60):63-86, 2004