MCMC for selective metamorphosis with applications to landmarks

Andreas, Alexis and Colin

November 22, 2018

Contents

1 Introduction

In this paper we consider a generalisation of the classical metamorphosis problem. The purpose of this work is to allow singular control where metamorphosis occurs in space to allow practitioners in e.g. medical imaging to qualitatively assess different growth patterns. Owing to this observation we build a stochastic model on spatial domain and use develop an MCMC algorithm in order to infer where in space growth is the most likely.

This paper is organised as follows. We review some preliminaries in section 2. Next, in section 3 we analyse the problem with an *a priori* observable function ν and derive the corresponding Euler-Lagrange equations for the associated matching functional. Next, section 4 formalises the Bayesian framework for this problem. This is then evaluated in section 5.

2 Background

In this paper we are concerned with diffeomorphometric approaches to image and shape matching. One of the central pillars on which this field relies is the observation made by Arnold '66 (cite) that, under certain conditions, a time-dependent velocity field u occupying some Hilbert space $u_t \in V$ induces a curve on a subgroup G_V of diffeomorphisms [4] via the following ordinary differential equation:

$$\dot{\varphi}_t = u_t \circ \varphi_t \tag{1a}$$

$$\varphi_0 = \mathrm{id} \tag{1b}$$

with $\varphi_t \in G_V$, $\forall t$. This allows for the mathematical (Lagrangian) description of flows on shape and shape spaces with a rich literature of applications (cite theory + applications). A unifying property of these is to use (1) in defining an optimisation problem:

$$\min_{t \mapsto \varphi_t} \int_0^1 \frac{1}{2}\|u_t\|_V^2 \mathrm{d}t \ + \frac{1}{2} S(q_0 \circ \varphi_1, q_1)$$

where S denotes a similarity measure and q_j, $j = 0, 1$.

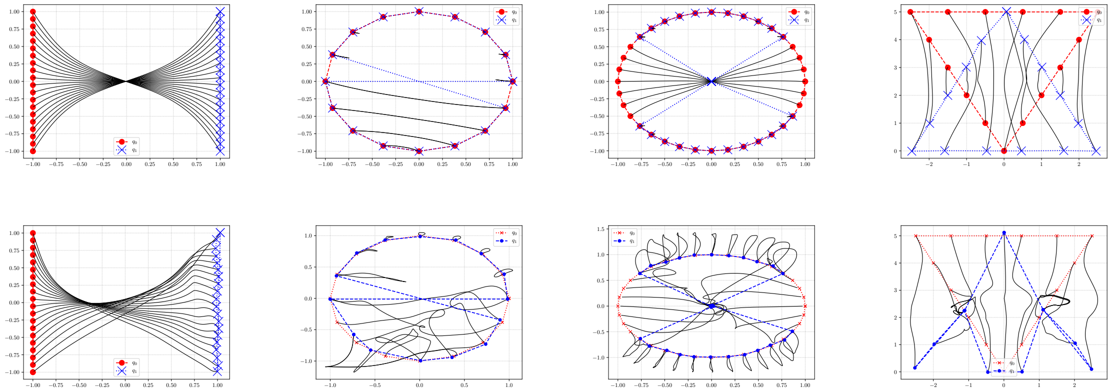

Figure 1: Metamorphosis (top) and LDDMM (bottom).

However deformations of manifolds q_j by members of G_B must obey a certain smoothness, which is ill-suited for arbitrary manifolds e.g. topologically different ones.
Moreover, metamorphosis is a generalisation thereof and provides more robust model for performing shape deformation.
$\longrightarrow$ **Insert more background here**

In the following we let H denote a vector space taking values in $\mathbb{R}^d$ to be determined later. Observe the classical landmark metamorphosis problem:

$$\inf_{\substack{u \in L^1([0,1],V) \\ q,\, z \in C^1([0,1],H)}} S = \int_0^1 l(u_t, q_t, z_t)\mathrm{d}t \tag{2a}$$

$$\dot{q}_t = u_t \circ q_t + z_t \tag{2b}$$

$$q_0,\, q_1 \text{ fixed} \tag{2c}$$

with $l(u,q,z) = \frac{1}{2}\|u\|_V^2 + \frac{\sigma^2}{2}\|z\|_2^2$ and where the boundary conditions $q_0,\, q_1 \in \mathbb{R}^{d\times M}$ denotes the landmarks for time $j = 0,\, 1$. σ is a parameter controlling the trade-off between the diffeometric and metamorphosis terms. In LDDMM, the equations are the same but the boundary conditions are different. In particular, q_1 is not fixed, but an initial momentum, p_0, is imposed to close the system. As we see in figure 1, these frameworks lead to different results. Our goal here is to provide an approach that allows selective delivery of a mix between the purely diffeomorphic LDDMM approach and the metamorphic matching.
For landmarks, the velocity field is explicitly expressed in terms of the momentum and landmarks positions, resulting in a boundary value problem (in time):

$$\dot{q}_t = u(q_t) \tag{3a}$$

$$\dot{p}_t = -\nabla u(q_t)^T p_t \tag{3b}$$

$$q_0,\, q_1 \text{ fixed} \tag{3c}$$

$\longrightarrow$ **Insert segue and motivation for next section here**

3 Selective Metamorphosis

This work aims to place a Bayesian framework on this model in such a way to allow control of *where* in the domain $\mathbb{R}^d$ to allow transgression of the usual diffeomorphic constraint $\dot{q}_t = u \circ q_t$. If $\nu = 0$, we recover standard landmark dynamics, if $\nu(x) = \sigma^2$ is a constant, we recover the classic landmark metamorphosis. That is to say, augmenting the Lagrangian to constrain metamorphic transformations only in areas of the domain that correspond to e.g. anatomically interesting places. This can be expressed mathematically by changing the Lagrangian above to:

$$\hat{l}(u,q,p) = \frac{1}{2}\|u\|_V^2 + \frac{1}{2}\sum_i \nu(q_i)|p_i|^2$$

where $\mathbb{R}^d \ni x \mapsto \nu(x)$ takes non-negative values in $\mathbb{R}$. The goal of this work is to infer the unknown function ν so as to gather information about where in the spatial domain we are most likely to observe metamorphic behaviour via a Bayesian framework as in [2].

Remark 1 (Choice of the form of ν) *Taking e.g. $\nu(\cdot) = \sum_k \exp(-h(\cdot))$ for some function $h > 0$ we explicitly state at which locations in space we allow non-diffeomorphic evolution, because when ν goes to zero we observe classical smooth dynamics of the landmarks. It could be useful to develop an algorithm that explores the entire space and infers at which location in space (if we had to choose only one such location!) it would be best to have metamorphosis. This could provide a first-order exploratory tool for physicians, to see if the development of a biological feature stems from a few violations of diffeomorphic evolution, starting with e.g. an inferred, least-cost, constellation of ν.*

On the other hand, if we take $\nu(\cdot) = 1 - \sum_k \alpha \exp(-h(\cdot)) > 0$, we could infer at which location(s) x in space we require $z_t(x) > 0$ in order for the metamorphosis problem to be solved. In contrast, this choice is natural as a diagnostic tool instead to infer where in the domain we have growth, which health experts can attribute biological meaning. This kind of choice allows us to infer where it is best (i.e. cheapest for the functional) physically.

$$\inf_{\substack{u \in L^1([0,1],V) \\ q,p \in C^1([0,1],\mathbb{R}^{d\times M})}} \quad S = \int_0^1 \hat{l}(u_t, q_t, p_t)\mathrm{d}t \tag{4a}$$

$$\dot{q}_t^i = u(q_t^i) + \nu(q_t^i)p_t^i \tag{4b}$$

$$q_0,\, q_1 \text{ fixed} \tag{4c}$$

In this section we analyse the problem described in (4) for an a priori known function ν which we denote *selective metamorphosis problem.* The first task is to show that the functional attains its infimum, i.e. identification of a minimiser. We begin with the following assumption:

Assumption 1 $\|\nu\|_\infty \triangleq \sigma^2 < \infty$ *and ν is bounded from below away from zero by $\nu_{inf} \in \mathbb{R}$.*

Since the limit as $\nu \to 0$ is well-understood these assumptions do not constitute large concessions. Clearly the functional in (4a) is not convex, so we work with a reformulation to ensure the required weak lower semi-continuity. This comes at the cost of showing weak continuity of the constraint variables. Specifically, we define a variable $z_t^i = p_t^i\sqrt{\nu(q_t^i)}$ in the problem:

$$\inf_{\substack{u \in L^2([0,1],V) \\ q,z \in L^2([0,1],\mathbb{R}^{d\times M})}} \quad S = \int_0^1 \frac{1}{2}\|u_t\|_V^2 + \frac{1}{2}\sum_i |z_t^i|^2 \mathrm{d}t \tag{5a}$$

$$\dot{q}_t^i = u(q_t^i) + \sqrt{\nu(q_t^i)}z_t^i \tag{5b}$$

$$q_0,\, q_1 \text{ fixed} \tag{5c}$$

We now show existence of a minimiser to this problem. First, note that owing to the constraint effectively being a boundary value problem, we cannot always find a q for arbitrary pairs of $(u,\, z)$. We define an bounded operator $(q, u) \mapsto \frac{\dot{q}-u(q)}{\sqrt{\nu(q)}} \triangleq z$:

$$\begin{aligned} \|z\|_2^2 &= \|\frac{\dot{q}-u(q)}{\sqrt{\nu(q)}}\|_2^2 \\ &\leq \nu_{\inf}^{-1}\Big(\|\dot{q}\|_2^2 + \|u(q)\|_2^2\Big) \end{aligned}$$

which is clearly bounded since $q \in C^1(0,1,H)$ and since by continuity of u. From this we generate a minimising sequence (q^n, u^n, z^n) admissible to (5).

$\longrightarrow$ **Show weak continuity of the constraint equation for q (the stuff below)**

Passing to subsequences where necessary we can by classic results [4] extract bounded subsequences converging to weak limits where necessary. By convexity of S we have therefore proved the following:

Theorem 1 *Assumption 1 implies that the problem* (5) *attains its unique infimum.* □

We now derive the optimality conditions of (5). Using p_t^i as a Lagrange multiplier and extremising J leads to the following equations:

$$\dot{p_t^i} + \nabla u(q_t^i)p_t^i + \frac{1}{2}\nabla\nu(q_t^i)|p_t^i|^2 = 0 \tag{6a}$$

$$\dot{q_t^i} = u(q_t^i) + \nu(q_t^i)p_t^i \tag{6b}$$

$$q_0,\ q_1 \text{ fixed} \tag{6c}$$

In this paper we are concerned only with landmarks, so we can write the Lagrangian in terms of the state variables $\hat{l}(q,z) = \hat{l}(u,q,z)$ [4]:

$$\|u\|_V^2 = \sum_{i,j=0}^{M} K(q^i,q^j)p^i \cdot p^j$$

where supercripts denote the landmark index.
Extremising the action S we obtain the equations:

$$\dot{p}_i = \frac{1}{2}\sum_j \nabla K(q_i,q_j)p_i \cdot p_j + \frac{1}{2}|p_i|^2\nabla\nu(q_i) \tag{7}$$

$$\dot{q}_i = \sum_j K(q_i,q_j)p_i + \nu(q_i)p_i \tag{8}$$

$$q_0,\ q_1 \text{ fixed} \tag{9}$$

Theorem 2 (Optimality Conditions) *There exists a unique (?)solution u,q,z to this system.*
This will give more conditions on ν, like $\nu(x) \in C^1(\mathbb{R}^2)$.

4 Bayesian Framework

We now place a stochastic model on ν, the purpose of which depends on its definition as mentioned in remark (1). The goal is to develop an algorithm to infer ν, passing via the deterministic problem seen above.

First we present some preliminaries on the Bayesian approach to inverse problems in section 4.1, essentially quoting results from [2]. See also [1] for an exposition of algorithmic aspects of function space MCMC. Section 4.2 then describes how we apply this stochastic approach to inverse problems to ν by a finite-dimensional family of parameterisations.

4.1 General Framework

Our setting is the following. Given an inverse problem
min $\quad J$

we formally define a centered Gaussian distribution, denoted $N(0,B)$ over this space of functions defined up to a constant by the density

$$\mu_0 = e^{-J}$$

In general, for inverse problems on function space, several key properties documented in [**?**] must be verified before the inverse-problem is well-posed. Beyond showing existence of the MAP estimator extremising J above, the infinite-dimensional version of Bayes' rule must also be check i.e. the Radon-Nikodym derivative of the prior with respect to the posterior must exist and be absolutely continuous. Finally, we request continuity of the posterior distribution w.r.t the initial data corresponding in a sense to Hadamard well-posedness in a stochastic framework. Proceeding bona fide we therefore have the key ingredients necessary to define a pCN MCMC algorithm:

1. Sample a function.

2. $f^{k+1} = \alpha f^k + (1-\alpha) f^k$

3. accept/reject..

$\longrightarrow$ **Finish**

4.2 Finite-dimensional Parameterisation

We now introduce the main problem of this paper in the setting above. We consider ν as a random variable, indeed a random *function*, occupying some space yet to be determined. This is an appropriate framework because if the lack of uniqueness allows for a qualitative evaluation of a solution to selective metamorphosis. In this paper we consider the case where ν is given by a sum of exponentials:

$$\nu_h(x) \triangleq \sum_k e^{-\|h_k - x\|^2_{\mathbb{R}^d}} \tag{10}$$

Here the finite family h_k of centroids in $\mathbb{R}^d$ fully determine ν_h, thus greatly reducing the complexity of sampling.

$$\Phi(a_k) = S(\nu, q_0, q_1) + |\nu|^2_V \tag{11}$$

where, for example $V = L_2(\mathcal{D})$.

Algorithm 1 MCMC on ν

procedure QUASIMCMC(N)
  $k \leftarrow 0$
  $\nu^k \leftarrow$ some initial guess
  while $k < N$ **do**
    $\nu \leftarrow$ SAMPLECENTROID()   **# sample a new centroid**
    $(q, z, u) \leftarrow$
    if RANDOMUNIT() $<$ ACCEPTANCEPROBABILITY(q, z) **then**
      do update
    else
      don't do update
    $k \leftarrow k + 1$
  return q^k, z^k, u^k

Note that the *solveMetamorphosis*() operation, in general, depends on how the spatial domain is treated e.g. whether we use landmarks or treat full images. In this paper we are only concerned with the former, so we use a standard and easily implemented procedure (cite) below

$\longrightarrow$ **Describe shooting/solution method**

Algorithm 2 Metamorphosis for fixed ν

procedure SOLVEMETAMORPHOSIS
  something either shooting or Firedrake

$\longrightarrow$ **Something about adjoints here**

5 Selective Metamorphosis for Landmarks

This penultimate section displays some numerical results for our method. First, as a proof of concept
As we have seen in figure **??**, classic metamorphosis is an appropriate model for non-diffeomorphic matching. We now evaluate our stochastic model described in algorithm 1 to the landmark configurations in **??**.

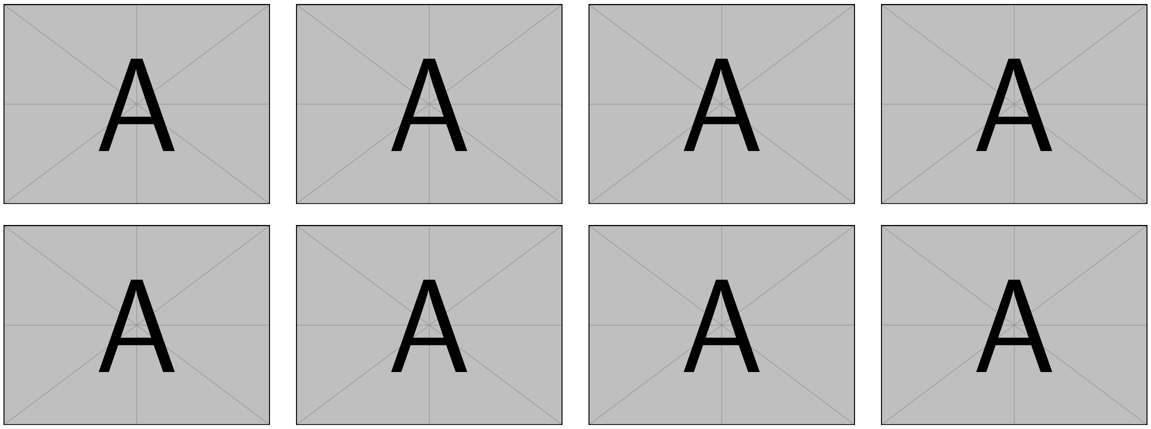

Figure 2: Selective metamorphosis for the pinch example. The top row displays the geodesics corresponding a sample realisation from the chain, centroid heat map with four MAP estimators, functional histogram and autocorrelation plot. The bottom row shows the geodesics for four MAP estimators.

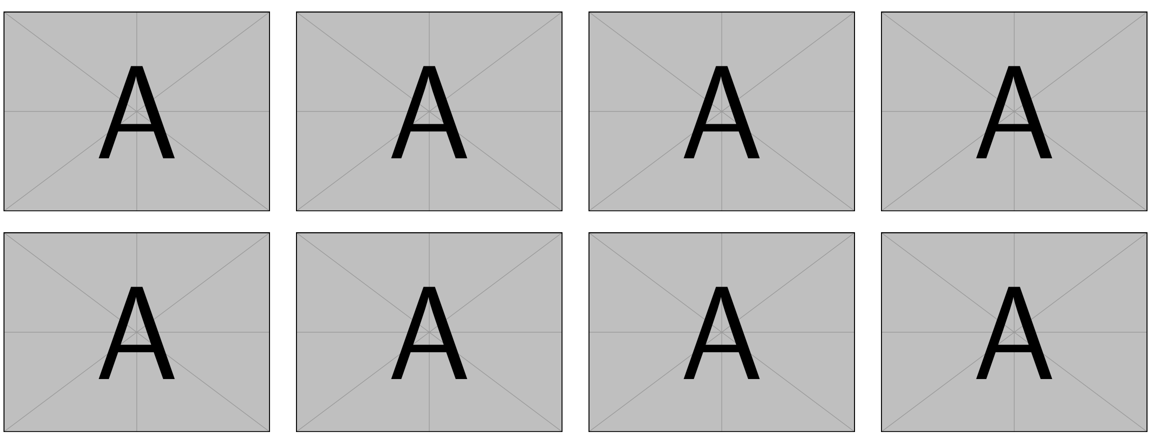

Figure 3: Selective metamorphosis for the inverted landmarks example. The top row displays the geodesics corresponding a sample realisation from the chain, centroid heat map with four MAP estimators, functional histogram and autocorrelation plot. The bottom row shows the geodesics for four MAP estimators.

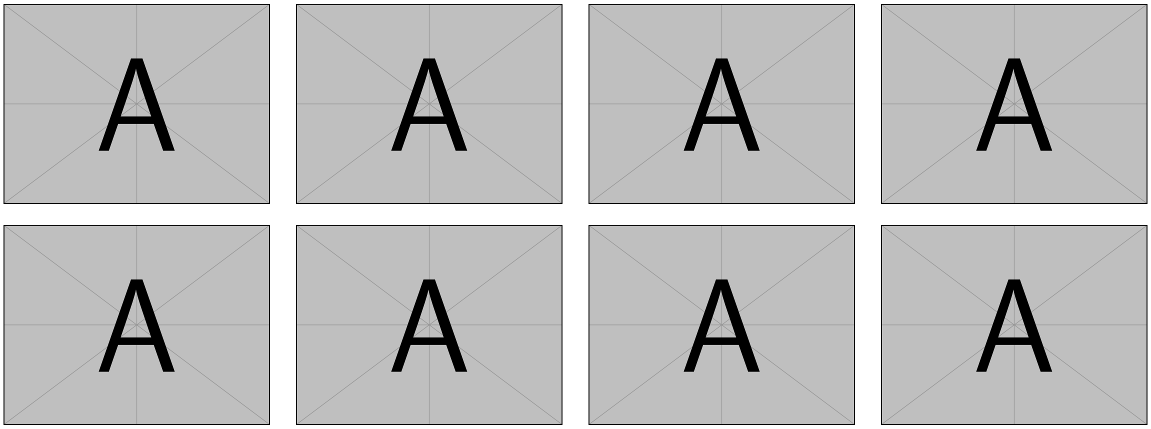

Figure 4: Selective metamorphosis for the inverted landmarks example The top row displays the geodesics corresponding a sample realisation from the chain, centroid heat map with four MAP estimators, functional histogram and autocorrelation plot. The bottom row shows the geodesics for four MAP estimators.

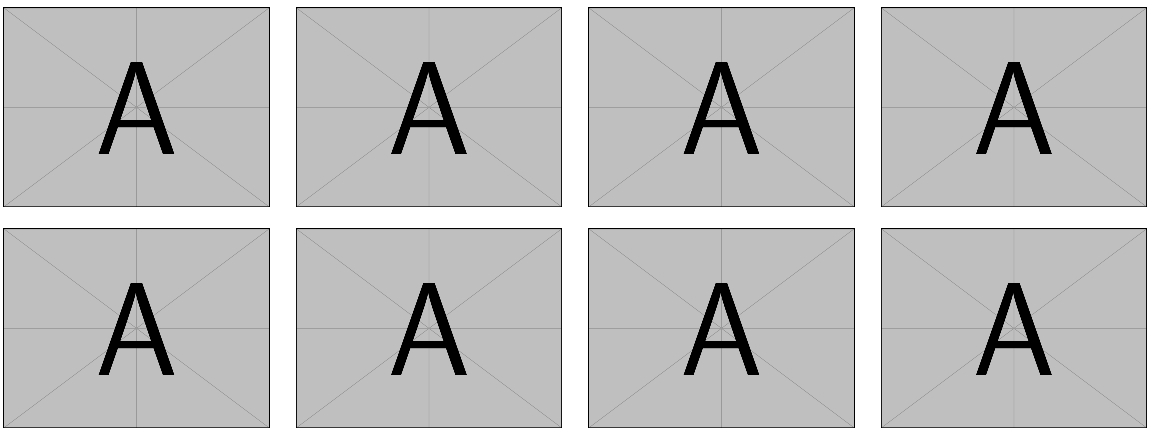

Figure 5: Selective metamorphosis for the inverted landmarks example The top row displays the geodesics corresponding a sample realisation from the chain, centroid heat map with four MAP estimators, functional histogram and autocorrelation plot. The bottom row shows the geodesics for four MAP estimators.

$\longrightarrow$ **Explain results**

We note that since the Euler-Lagrange equations for p and q are time-reversible, the configuration in **??** both lead to a notion of collapse and hole creation for the landmarks.

We note that the source code for all of these experiments is documented here (insert link to github.io).

6 Outlook

Another choice is

$$l(q,p) = \|u\|_B^2 + \sum_i p_i \cdot \nu(q_i)\,, \tag{12}$$

where $\nu : \mathbb{R}^2 \to \mathbb{R}^2$, which will give more freedom on the metamorphosis term, and give equations of motion as

$$\dot{p}_i = \nabla u(q_i) p_i + p_i \cdot \nabla \nu(q_i) \tag{13}$$

$$\dot{q} = u(q_i) + \nu(q_i) \tag{14}$$

Natural extensions of this work include fully determining e.g. the Fourier modes of ν via the stochastic framework.

Future work also includes extending this framework to images e.g. using the kernel framework in [3], or developing a space-time method. Formalising the measure-theory necessary to consider the limit of (**??**) as $k \to \infty$.

Moreover, adding a time-dependency to ν is also to be explored. It is our hope that we can extend the theory developed here to encompass classic metamorphosis; that is to say, to develop the necessary theory in order to place a stochastic model on the state space consisting of velocities and source functions. Being able to sample random pairs (MCMC) of these would permit a novel numerical approach to metamorphosis as well as other problems shape analysis.

References

[1] S. L. Cotter, G. O. Roberts, A. M. Stuart, and D. White. Mcmc methods for functions: modifying old algorithms to make them faster. *Statistical Science*, pages 424–446, 2013.

[2] M. Dashti and A. M. Stuart. The bayesian approach to inverse problems. *Handbook of Uncertainty Quantification*, pages 311–428, 2017.

[3] C. L. Richardson and L. Younes. Metamorphosis of images in reproducing kernel hilbert spaces. *Advances in Computational Mathematics*, 42(3):573–603, 2016.

[4] L. Younes. *Shapes and diffeomorphisms*, volume 171. Springer Science & Business Media, 2010.