

IMAGE MORPHING BASED ON MUTUAL INFORMATION AND OPTIMAL MASS TRANSPORT

Lei Zhu, Yan Yang, Allen Tannenbaum

Steven Haker

Georgia Institute of Technology
Department of Biomedical Engineering
313 Ferst Drive, Atlanta, GA 30332, USA

Brigham and Women's Hospital and Harvard
Medical School, Surgical Planning Laboratory
Boston, MA 02115, USA

ABSTRACT

Time domain image interpolation, or image morphing, refers to a class of techniques for generating a series of smoothly changing intermediate images between two given related images. In this note, we present a novel approach based on the theory of optimal mass transport, using mutual information (MI) as the similarity measurement. The potential applications also include image registration, compression and coding.

Keywords: Image interpolation, Morphing, Optimal mass transport, Mutual information.

1. INTRODUCTION

Image morphing, which is sometimes called *image interpolation* (in time domain), is a class of techniques that deal with the metamorphosis of one image into another [1]. Given two related images, these techniques generate a sequence of intermediate images in which an image gradually changes into another over time. Image morphing first appeared in the movie *Willow* in 1988. From then on, it has been widely used in creating special effects for television commercials, movies and music videos. In this note, we propose a novel approach for image morphing based on optimal mass transport theory, using mutual information as a similarity measure.

There have been a number of different algorithms proposed for image morphing. The main difference between these algorithms is the way of finding a reasonable warping function between the two given images. Among them, the mesh warping method [2] shows good distortion behavior, but has a critical drawback in that it requires the specification of features on the control mesh, which is time consuming and may lead to an arbitrary structure. Field morphing, as described in [3], gives an easy-to-use and explicit method for selecting features. However, undesired distortions referred to as “ghosts” sometimes appear, which means that a part of the initial image may show up in some unrelated region of the intermediate images. Lee *et al.* [1] proposed

an energy minimizing algorithm, which guarantees the one-to-one property of the warping function and thus prevents the warped image from folding back upon itself. However, there is still a need for specifying features. For a general review of image morphing methods, interested readers may refer to [4] and the references therein.

Our previous work [5] was based on the theory of optimal mass transport. The Sum of squared differences (SSD) between aligned images was used as the comparison term. An energy functional was then minimized iteratively by finding the best tradeoff between minimizing the intensity difference and minimizing the transportation cost of moving the mass, given the constraint that total mass was preserved in the process. In this paper, we revise the energy functional by substituting SSD with a mutual information (MI) measurement. The resulting in-between images look more natural than those from the SSD measurement. Our approach does not require the specification of feature points and is thus a “blind” method. It is especially suitable for images where features are hard to be selected, as shown in the examples.

We now outline the contents of this paper. In Section 2, we give a brief review of the optimal mass transport problem. In Section 3, we propose a new functional for image morphing and describe the corresponding method for solving it. In Section 4, we give some examples and compare them with our previous result using SSD. Finally, in Section 5, we summarize the contribution of this paper and discuss some possible future research directions.

2. OPTIMAL MASS TRANSPORT

The optimal mass transport problem was first formulated by a French mathematician Gaspar Monge in 1781 and is also known as the Monge-Kantorovich Problem (MKP). The problem concerns the optimal way, in the sense of minimal transportation cost, of moving a pile of soil from one configuration into another. The total amount of mass is required to be constant in the process. Assume there are two domains Ω_0

and Ω_1 in \mathbf{R}^d , with smooth boundaries, each with a positive density function, μ_0 and μ_1 , respectively, and

$$\int_{\Omega_0} \mu_0 = \int_{\Omega_1} \mu_1, \quad (1)$$

so that the same total mass is associated with both domains. A mapping u from (Ω_0, μ_0) to (Ω_1, μ_1) is said to be mass preserving (MP) if

$$\mu_0 = |Du| \mu_1 \circ u. \quad (2)$$

Here $|Du|$ is the determinant of Jacobian of the mapping function, and \circ represents composition of functions. If a function u satisfies (2), we will write $u \in MP$. Note that equation (2) implies that if a small region in Ω_0 is mapped to a larger region in Ω_1 , then there must be a corresponding decrease in density in order for the mass to be preserved.

The L^p Kantorovich-Wasserstein metric is defined as

$$d_p(\mu_0, \mu_1)^p := \inf_{u \in MP} \int \|u(x) - x\|^p \mu_0(x) dx \quad (3)$$

which defines the distance between two mass densities, by computing the ‘‘cheapest’’ way to transport the mass from one domain to the other with respect to (3). This ‘‘cheapest’’ way is the optimal mapping function \tilde{u} for the Monge-Kantorovich problem.

The case $p = 2$ has been widely studied and will be the one used in our paper for image morphing algorithm. Theoretical results [6] show that there is a unique optimal solution $\tilde{u} \in MP$ for MKP, which is characterized as being the gradient of a convex function w , i.e., $\tilde{u} = \nabla w$. Note that from the equation (1) we see that w satisfies

$$\mu_0 = |Hw| \mu_1 \circ (\nabla w) \quad (4)$$

where $|Hw|$ is the determinant of the Hessian of w . Equation (4) is known as *Monge – Ampère* equation. The optimal solution $\tilde{u} = \nabla w$ gives the distance $d_2(\mu_0, \mu_1)$ between μ_0 and μ_1 , in the L^2 Kantorovich-Wasserstein sense via Equation (3).

3. OPTIMAL IMAGE INTERPOLATION

In our pervious work [5], SSD was used as the comparison term, penalizing the intensity difference between the two images. However, SSD implicitly uses constant intensity (or mass density) assumption, which contradicts the mass preserving constraint. The in-between images thus may have some unnatural effects. We now propose the use of mutual information and optimal mass transport to formulate a new method for image morphing. It belongs to the category of energy minimizing algorithms. The idea is to minimize the

following functional over mass-preserving (MP) mappings $u : \Omega_0 \rightarrow \Omega_1$:

$$M_\alpha : = - \int_{i_0 \times i_1} p_u^{I_0, I_1 \circ u}(i_0, i_1) \log \frac{p_u^{I_0, I_1 \circ u}(i_0, i_1)}{p^{I_0}(i_0) p^{I_1 \circ u}(i_1)} di_0 di_1 + \alpha \int_{\Omega_0} \|u(x) - x\|^2 \mu_0(x) dx, \quad (5)$$

for a positive parameter $\alpha \in \mathbf{R}$, whose value can be decided according to subsection 3.2. Here the first term is the negative value of mutual information (MI) measurement. It controls the ‘‘goodness of fit’’ between the two (intensity) images $I_0 : \Omega_0 \rightarrow \mathbf{R}$ and $I_1 : \Omega_1 \rightarrow \mathbf{R}$. The integral taken in the first term is on a 2D domain of $i_0 \times i_1$, where i_0 and i_1 are the intensity values of I_0 and I_1 , respectively. The second term is the transportation cost from L^2 MKP, in which the function $\mu_0(x)$ is the mass density of the source image defined on Ω_0 , and could be the same as I_0 or a smoothed version of I_0 . Similarly, μ_1 is assumed to be the mass density of the target image defined on Ω_1 .

The joint intensity distribution $p_u^{I_0, I_1 \circ u}$ can be estimated from I_0 and I_1 by a non-parametric Parzen-Rozenblatt density model [7]:

$$\begin{aligned} p_u^{I_0, I_1 \circ u}(i_0, i_1) &= \frac{1}{V} \int_{\Omega} \psi(I_0(x) - i_0, I_1(u(x)) - i_1) dx \\ &= \frac{1}{V} \int_{\Omega} \frac{1}{\mu_0} \psi(I_0(x) - i_0, I_1(u(x)) - i_1) \mu_0 dx \end{aligned}$$

For ψ , we use the Gaussian window:

$$\psi(\alpha, \beta) = \frac{1}{2\pi |\Sigma|^{1/2}} e^{-\frac{1}{2} [\alpha, \beta] \Sigma^{-1} [\alpha, \beta]^T} \quad (6)$$

where the covariance matrix Σ of the Gaussian window is chosen to be 10% of the covariance matrix estimated from the aligned images, and V is the volume of samples for estimation. In a discrete case, V is the number of samples used.

We will describe here only the algorithm for finding the optimal mapping \tilde{u} . Mathematical details for general minimization problems under mass-preserving condition can be found in [8] and method for taking derivative of mutual information can be found in [9]. The basic idea for finding the optimal warping function is first to find an initial MP mapping u^0 and update it iteratively to decrease energy functional. When the pseudo time t goes to ∞ , the optimal u will be found, which is \tilde{u} . Basically there are two steps:

3.1. Computing the Initial Mapping

A general method for finding an initial mapping for irregular domains can be found in the work of Moser [10]. Since we are working with images here, we may assume that both

source and target domains are rectangular. Hence, we can use the following simple method [8].

Assume source domain $\Omega_0 = [0, A_0] \times [0, B_0]$ and $\Omega_1 = [0, A_1] \times [0, B_1]$. Our initial mapping will be of the form $u^0(x, y) = (a(x), b(x, y))$. Since both μ_0 and μ_1 are defined to be positive everywhere, we may define $u^0 = (a(x), b(x, y))$ be the following:

$$\begin{aligned} \int_0^{a(x)} \int_0^{B_1} \mu_1(\eta, y) dy d\eta &= \int_0^x \int_0^{B_0} \mu_0(\eta, y) dy d\eta \\ a'(x) \int_0^{b(x, y)} \mu_1(a(x), \rho) d\rho &= \int_0^y \mu_0(x, \rho) d\rho \end{aligned} \quad (7)$$

It can be easily proved that $u^0 = (a(x), b(x, y))$ satisfies MP property [8], i.e.

$$|Du^0|_{\mu_1 \circ u^0} = \mu_0.$$

The construction of u^0 can be explained as finding a 1D MKP in the x direction and then finding a family of 1D MKPs in the y direction.

3.2. Gradient Descent Method for Optimization

Here we only consider the gradient descent method in \mathbf{R}^2 . For a general \mathbf{R}^d problem, please refer to [8] and the references therein. In 2D case, we define \perp be the rotation by 90 degrees. For example, for a scalar function h , $\nabla^\perp h = (-h_y, h_x)$, which is a divergence free vector field. By taking the derivative of (5) respect to the pseudo time t , doing change of variable twice and applying the gradient descent algorithm, it can be seen that the update of the MP mapping u should have the following form:

$$\Delta f = -\operatorname{div}(P^\perp), \quad (8)$$

$$f = 0 \text{ on } \partial\Omega_0, \quad (9)$$

$$u_t = -\frac{1}{\mu_0} Du \nabla^\perp f \quad (10)$$

where P is defined as:

$$\begin{aligned} P = & \frac{1}{V} \left\{ \left[(1 + \log \frac{p_u^{I_0, I_1 \circ u}}{p^{I_0} p^{I_1 \circ u}}) * \psi_\alpha \right] (I_0(x), I_1 \circ u(x)) \frac{\nabla I_0(x)}{\mu_0(x)} \right. \\ & + \left. \left[(1 + \log \frac{p_u^{I_0, I_1 \circ u}}{p^{I_0} p^{I_1 \circ u}}) * \psi \right] (I_0(x), I_1 \circ u(x)) \frac{\nabla \mu_0(x)}{\mu_0^2(x)} \right\} \\ & + 2\alpha(u(x) - \underline{\text{id}}) \end{aligned} \quad (11)$$

in which $*$ stands for convolution, V is the number of samples used in estimating joint density function (or more specifically the number of pixels in image I_0), ψ is the Gaussian window (6), ψ_α is the derivative of ψ with respect to its first variable, and $\underline{\text{id}}$ is the identity map. The first term of

P comes from the (negative mutual information) comparison term and the second term of P comes from the transport cost of L^2 MKP. The two terms in equation (5) can be solved individually. Thus the flow $\nabla^\perp f$ can be divided into two parts: one for the comparison term and the other for the L^2 MKP transport cost term. The parameter α is then chosen to make the initial flows (i.e. when $t = 0$) from the two terms comparable.

4. EXPERIMENTAL RESULTS

We use standard techniques to solve Equations (8) to (10). In particular we have employed an unwinding scheme when computing $Du \nabla^\perp f$, and the FFT when inverting the Laplacian on a rectangular grid. Centered differences are used for other spatial derivatives. Once we numerically solve for the right hand side of (10), we use the result to update u . The optimal map is obtained as $t \rightarrow \infty$. In practice, we iterate until convergence with respect to a specified tolerance.

We demonstrate now our image interpolation method with two examples. The intermediate images are generated by a standard *cross-dissolving* method.

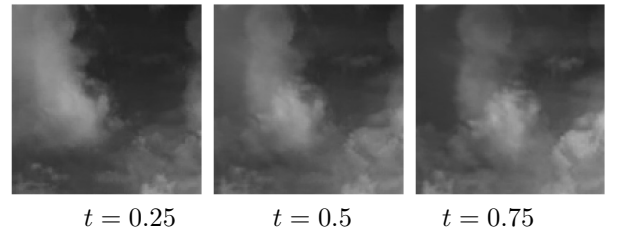
The first example is an interpolation between two *cloud* images. We take Figure 1(a) and Figure 1(b) to be the starting and ending images, respectively. Figures 2 shows the intermediate images generated by our MI morphing algorithm, at times $t=0.25, 0.5$, and 0.75 , respectively.



(a): source image

(b): target image

Fig. 1. The source and target *cloud* images.



$t = 0.25$

$t = 0.5$

$t = 0.75$

Fig. 2. The intermediate *cloud* images.

The second one is a *flame* example. The original images were taken from a video clip by *Arbeats Digital Film Library*. The starting image was the 24th frame in the sequence (the leftmost one in figure 3) and the ending image

was the 29th frame in the sequence (the rightmost one in figure 3). The middle images in figure 3 were generated at time $t=0.25, 0.5,$ and $0.75,$ using our algorithm as described in Section 3. Figure 4 shows the results at $t=0.5$ where SSD and MI were used as similarity measurements, respectively. If we zoom in to the bottom of the fire as shown in Figure 5, we may find out MI performs better than SSD as comparison term, by reducing the undesired “broken” effect. However, this improvement is more dynamic than static. It is strongly recommended for interested readers to view the two video clips from the following link:
<http://www.prism.gatech.edu/~gte538w/icip2004/icip2004.htm>.

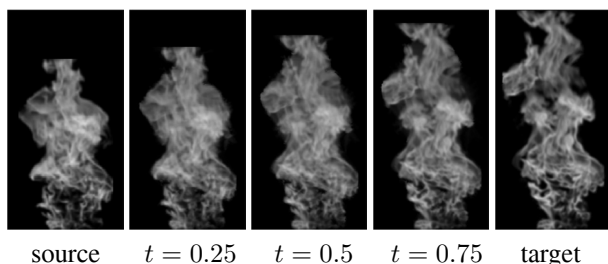


Fig. 3. Interpolated flame images.

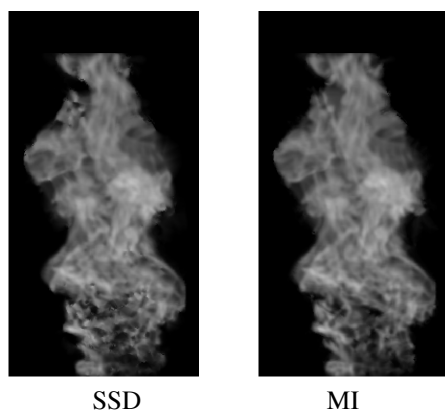


Fig. 4. Comparison of SSD and MI as similarity measurement



Fig. 5. Part of Figure 4

5. CONCLUSIONS

In this note, we presented a new algorithm for image morphing based on the optimal mass transport theory, using mu-

tual information as a similarity measure. We described an algorithm for finding the initial mapping and then iteratively optimizing it using a gradient descent method. We have also showed the advantage of MI over SSD, as a similarity measure in this application. In our future work, we plan to apply other measurements, such as cross-correlation, for the comparison term. We also feel that L^2 penalty is often too severe for measuring the transport cost. One possible solution is to use a L^1 measurement with some smoothness term added. It should also be noted that our algorithm may also be applied to other image processing tasks including image registration, compression and coding.

6. REFERENCES

- [1] S.-Y. Lee, K.-Y. Chwa, J. Hahn, , and S.Y. Shin, “Image morphing using deformation techniques,” *J. Visualization and Computer Animation*, vol. 7, no. 1, pp. 3–23, 1996.
- [2] G. Wolberg, “Recent advances in image morphing,” in *Proc. of Computer Graphics International 1996*, 1996, pp. 64–71.
- [3] T. Beier and S. Neely, “Feature-based image metamorphosis,” in *Computer Graphics (Proc. SIGGRAPH ’92)*, 1992, pp. 35–42.
- [4] G. Wolberg, “Image morphing: A survey,” *Visual Computer*, vol. 14, pp. 360–372, 1998.
- [5] L. Zhu and A. Tannenbaum, “Image interpolation based on optimal mass preserving mappings,” in *Proc. of IEEE International Symposium on Biomedical Imaging*, 2004, pp. 21–24.
- [6] W. Gangbo and R. McCann, “The geometry of optimal transportation,” *Acta Math*, vol. 177, pp. 113–161, 1996.
- [7] R. Duda, P. Hart, and D. Stork, *Pattern Classification*, Wiley-Interscience, 2000.
- [8] S. Angenent, S. Haker, and A. Tannenbaum, “Minimizing flows for the Monge–Kantorovich problem,” *SIAM J. Math. Analysis*, vol. 35, pp. 61–97, 2003.
- [9] G. Hermosillo, C. Chefd’Hotel, and O. Faugeras, “A variational approach to multi-modal image matching,” Tech. Rep. n°4117, INRIA, February 2001.
- [10] J. Moser, “On the volume elements on a manifold,” *Trans. Amer. Math. Soc.*, vol. 120, pp. 286–294, 1965.