

A New Framework of Quantifying Differences Between Images by Matching Gradient Fields and Its Application to Image Blending

Wei-Hsun Liao, Chin-Lung Yu, Marvin Bergsneider, Luminita Vese, and Sung-Cheng Huang

Abstract--This paper presents a new method for quantifying the differences between images. The proposed model is based on matching the gradient fields of two images. We first define new image spaces in which images are considered equivalent under a similarity group actions and the difference between two image classes is then defined by employing the Cauchy-Schwarz inequality to the gradient fields. The advantage of our approach is that images are identified by their relative contrasts and thus is scale free. Using this approach, we are able to achieve image blending in a novel way. By modifying the group actions, we extend our basic model to more general equivalence classes. The variational problems and the corresponding Euler-Lagrange equations associated to these models are proposed and the gradient descent time dependent partial differential equations are derived. Fast and efficient solvers employing the Additive Operator Splitting scheme are also presented. We tested our models on simulation images as well as real brain MRI and PET images from normal control subjects.

I. INTRODUCTION

THE mathematical formulation and theories of objects and shapes have been studied extensively in computer vision, statistics, geology, biomedical science and image processing literature. A lot of efforts have been made to mathematically quantify the difference and to define a distance between shapes or objects. In this paper, we turn to the problem of defining differences between images. By doing so, we could achieve image blending in a novel way by minimizing the distance between two images.

Manuscript received November 8, 2002. This work was supported in part by NIH Grant No. NS30308 and DoE Grant No. DE-FC0387ER60615.

Wei-Hsun Liao (e-mail: feuillet@ucla.edu) and Sung-Cheng Huang (e-mail: hhuang@mednet.ucla.edu) are with the Department of Biomathematics and Department of Molecular and Medical Pharmacology at UCLA, Los Angeles, CA 90095 USA.

Chin-Lung Yu is with the Department of Molecular and Medical Pharmacology at UCLA, Los Angeles, CA 90095 USA (e-mail: clyu@ucla.edu).

Luminita Vese is with the Department of Mathematics at UCLA, Los Angeles, CA 90095 USA (e-mail: lvese@math.ucla.edu).

Marvin Bergsneider is with the Brain Injury Research Center of David Geffen School of Medicine at UCLA, Los Angeles, CA 90095 USA (e-mail: mbergsneider@mednet.ucla.edu).

Image blending [1] and image morphing have wide applications in computer vision, image processing and biomedical imaging. For a more complete treatment of this subject we refer the reader to [2, 3]. Image morphing usually consists mainly of two steps. The first one is warping, which involves a coordinate transformation to align user-defined landmarks. The next step is to apply image blending, which is usually done by linear interpolation (or cross dissolving). Linear interpolation is based on the sum of least squares of two images f and g :

$$\text{dist}(f, g) \equiv \frac{1}{2} \int (f - g)^2 dx.$$

Minimizing this distance with respect to f simply gives us a linear interpolation between images f and g .

Although it is the most commonly used, it is not necessarily the best. First, it is not scale invariant, if we rescale one of the images then we obtain a different distance. Moreover, two images that only differ in scaling are considered different under this distance measure. In this paper, we try to interpret images in a different way and avoid these disadvantages. In order to accomplish this, we first redefine the meaning of identical or equivalent images. We turn to the idea of “shape spaces” to look for solutions.

II. RELATED WORK AND MOTIVATIONS

In this section, we closely follow the discussion in [4]. The study of shapes has been proposed in [5-9] as a novel environment in statistics for comparing shapes where a metric and a probabilistic measure could be defined. It has been successfully used in comparing shapes where distinct landmarks could be found. In this case, shapes are defined as the equivalence classes of M points in R^N where the equivalence class is defined by a similarity group including rotation, translation and scaling. The space of all shapes could thus be viewed as a collection of fiber bundles where motions in the space could be “along the fiber” or “across the fiber”. Motions along the fiber are simply motions between equivalent classes, while motions across the fiber correspond to “deformations” from one class to another class.

Another approach that also has motivated our work is the success of variational methods and the use of PDEs to model

shapes and images. For example, in [10-13], a general framework of deformations and warping was proposed and the application of this framework to object comparison has shown to be very successful. In this framework, deformations could be rigorously defined in terms of group actions and the notion of distance could be made precise using these theories. For contributions from other authors, see, for example [14-16].

Our work is also closely related to nonlinear image smoothing techniques. Anisotropic diffusion [17, 18] is one of the first nonlinear smoothing techniques. The original idea of anisotropic diffusion is to smooth selectively within more homogeneous areas while preserving edges when a large gradient is detected. The PDE that governs anisotropic diffusion can be written in the following form:

$$\partial I / \partial t = \text{div}(c(|\nabla I|)\nabla I)$$

where I is an image and c is the diffusion coefficient. It has been noted in [19] that this PDE could also be derived by minimizing the following energy:

$$\min_I E(I) = \int_{\Omega} p(|\nabla I|) dx.$$

The function p is the penalty function. The PDE that minimizes the energy has the form:

$$\partial I / \partial t = \text{div}(p'(|\nabla I|)\nabla I / |\nabla I|).$$

Thus the diffusion coefficient c in the original model can be written as a function of p : $c(\xi) = p'(\xi)/\xi$.

In developing our models, we were originally looking for ways to extend the anisotropic diffusion, which could be viewed as a “prior-less” smoothing as opposed to a “prior-based” smoothing. The connection between anisotropic diffusion and our work resides in the fact that, by forming a suitable variational problem, our work could be incorporated into a “prior-based” smoothing when two images being compared contain similar information with one (the template) of the two images having a better image quality. For more details, please refer to the results and discussion sections.

III. DESCRIPTION OF THE MODEL

A. A natural way of defining equivalent images

Inspired by the idea of shape spaces, we want to define differences between images in a way that the difference is defined in terms of equivalence classes of images. Difference between two images will only be nonzero when they are not equivalent. But what similarity group action should be chosen to define equivalent classes?

Human perception of an image depends mostly by the contrast of the image. Areas with high contrast are perceived as boundaries of objects while areas with more homogeneous intensities are often viewed as inside an object. So our goal is to construct a new distance measure that only takes into account the relative change in intensity of the images. This

means we should consider two images I_1 and I_2 to be equivalent if there exist real numbers $k_1 > 0$ and k_2 such that the following holds:

$$I_1(x) = k_1 I_2(x) + k_2, \quad \forall x \in \Omega. \quad (1)$$

So our similarity group should include translation (k_2) and rescaling (k_1) of the intensities. Later on, we will relax the restrictions on k_1 by allowing negativity of k_1 .

This concept shares many features with the shape spaces. For example, images, like shape spaces, are organized in fiber bundles where a fiber bundle is just the collection of images that are equivalent under the similarity group actions we just defined. Also, motions from one image to another also fall into two categories: across the bundles and within the bundles by defining them in a similar way.

Thus, we have a new space on which images must be identified by their gradient fields, regardless of the scaling. Now we need a proper distance measure which is always non-negative and is zero only between equivalent images. We will derive distance measure between images f and g based on the following inequality, which is derived from the Cauchy-Schwarz inequality

$$\sqrt{\int |\nabla f|^2 dx} \sqrt{\int |\nabla g|^2 dx} - \int (\nabla f \cdot \nabla g) dx \geq 0$$

B. The models

We will present three different distance measures based on the above inequality. The first model is related to measuring differences of images where equivalent images are defined as in (1) without allowing negative k_1 , while the second model allows any real value of k_1 . The last one is a further generalization. Each one of them has slightly different properties and leads to its own PDE by employing gradient descent to solve the Euler-Lagrange equations.

For images f and g , we can look at the following minimization problems to match the gradient field of f to the gradient field of g :

model a

$$\min_f E(f) = \sqrt{\int |\nabla f|^2 dx} \sqrt{\int |\nabla g|^2 dx} - \int (\nabla f \cdot \nabla g) dx.$$

model b

$$\min_f E(f) = \int |\nabla f|^2 dx \int |\nabla g|^2 dx - \left(\int \nabla f \cdot \nabla g dx \right)^2.$$

Model B is different from model a in that the negative image of g (up to a scalar) is also considered to be equivalent to g .

model c

We propose a third model that is even more nonlinear by applying the absolute function to the integrand of the last term on the right hand side of model b. It also turns out to be the most interesting of the three models (see results). By equivalence class in this space we mean that the following energy between two images is 0:

$$\min_f E(f) = \int |\nabla f|^2 dx \int |\nabla g|^2 dx - \left(\int \nabla f \cdot \nabla g dx \right)^2.$$

C. Modified models

An undesirable property of these models is that, since they are not properly normalized, images are collapsed to constant images when time goes to infinity. In order to avoid this problem, we could normalize the energy to be minimized so that it is always between 0 and 1. Thus, we look at the following modified versions of our models:

model a'

$$\min_f E(f) = 1 - \frac{\int \nabla f \cdot \nabla g dx}{\left(\int |\nabla f|^2 dx \int |\nabla g|^2 dx \right)^{1/2}},$$

and the gradient descent PDE is:

$$\frac{\partial f}{\partial t} = \frac{\frac{\int \nabla f \cdot \nabla g dx}{\int |\nabla f|^2 dx} \Delta f - \Delta g}{\sqrt{\int |\nabla f|^2 dx \int |\nabla g|^2 dx}}.$$

model b'

$$\min_f E(f) = 1 - \frac{\left(\int \nabla f \cdot \nabla g dx \right)^2}{\int |\nabla f|^2 dx \int |\nabla g|^2 dx},$$

and the gradient descent PDE is:

$$\frac{\partial f}{\partial t} = \frac{\frac{\left(\int \nabla f \cdot \nabla g dx \right)^2}{\int |\nabla f|^2 dx} \Delta f - \left(\int \nabla f \cdot \nabla g dx \right) \Delta g}{\int |\nabla f|^2 dx \int |\nabla g|^2 dx}.$$

model c'

$$\min_f E(f) = 1 - \frac{\int |\nabla f \cdot \nabla g| dx}{\left(\int |\nabla f|^2 dx \int |\nabla g|^2 dx \right)^{1/2}},$$

and the gradient descent PDE is:

$$\frac{\partial f}{\partial t} = \frac{\frac{\int |\nabla f \cdot \nabla g| dx}{\int |\nabla f|^2 dx} \Delta f - \text{div}(\text{sgn}(\nabla f \cdot \nabla g) \nabla g)}{\sqrt{\int |\nabla f|^2 dx \int |\nabla g|^2 dx}}$$

IV. IMPLEMENTATION

Since the PDE we need to solve is an inhomogeneous heat equation, let us write down the PDE in the following form:

$$\frac{\partial f}{\partial t} = F + \alpha \Delta f$$

The semi-implicit scheme is used to solve this PDE, which now reads:

$$\bar{f}^{n+1} = \left(I - \Delta t \alpha^n \sum_{l=1}^d A_l \right)^{-1} (\bar{f}^n + \Delta t \bar{F}^n).$$

Furthermore, by using Additive Operator Splitting (AOS) scheme [20, 21], we update f in the following way:

$$\bar{f}^{n+1} = \frac{1}{d} \sum_{l=1}^d \left(I - d \Delta t \alpha^n A_l \right)^{-1} (\bar{f}^n + \Delta t \bar{F}^n).$$

Here d is the dimension of the problem and A is the one-dimensional discretized Laplacian operator along l th axis. The splitting is of order one in time and order two in space. Moreover it is unconditionally stable. By splitting the operator into a coordinate-by-coordinate fashion we now only need to invert a tri-diagonal matrix along each coordinate and this allows an $O(m^2)$ implementation by the Thomas algorithm.

V. RESULTS

For results presented in this section all calculation were performed on 128 by 128 grid points and the spatial step is 1.

A. Test of the basic model

We only present the numerical results for model a. The goal is to blend a circular shape (Fig. 1(a)) to an oval shape. The intermediate images during blending process are shown in Fig. 1(b)-(d). We notice that along the blending process the image intensities outside the circles also change with time, which does not happen in simple interpolation technique. This phenomenon generates a flow-like motion that gives novel visual effects. These effects might be closely related the resemblance of the governing PDE to the heat equation.

B. Test of modified model on medical images

Simple modification of model c' allows us to combine features of two images and to generate a hybrid or fusion image. The idea is similar to [18, 22]. We consider denoising an image f_0 by solving f in the following problem:

$$\min E(f) = \frac{1}{2} \int (f - f_0)^2 dx + c \left[1 - \int |\nabla f \cdot \nabla g| dx / \left(\int |\nabla f|^2 dx \int |\nabla g|^2 dx \right)^{1/2} \right]$$

The PDE that governs this minimization problem is:

$$\frac{\partial f}{\partial t} = (f_0 - f) + c \left(\int |\nabla f|^2 dx \int |\nabla g|^2 dx \right)^{-1/2} \times \left[\frac{\int |\nabla f \cdot \nabla g| dx}{\int |\nabla f|^2 dx} \Delta f - \text{div}[\text{sgn}(\nabla f \cdot \nabla g) \nabla g] \right]. \quad (2)$$

The above variational problem could be viewed as a prior-based image denoising. The idea is that f_0 is a corrupted image that we would like to recover. Furthermore, we know that g contains the same information but we do not know exactly how the intensity values relate in these two images. By applying the above model, we seek to find a recovered image that is close to g while constraining the result to be not far away from the initial image f_0 .

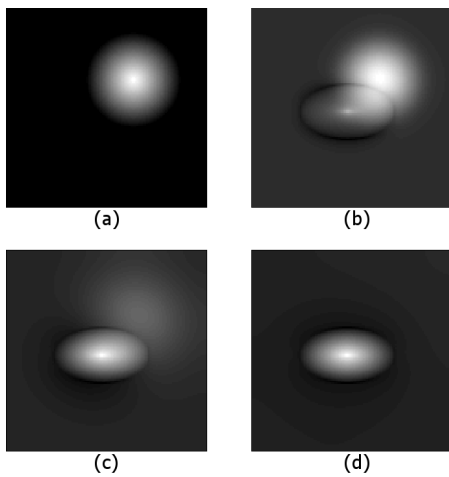


Figure 1. Image blending process described in \mathcal{A} of the results section. (a): The initial image containing an circular object; (b), (c) and (d): Intermediate images. The template image is visually similar to (d).

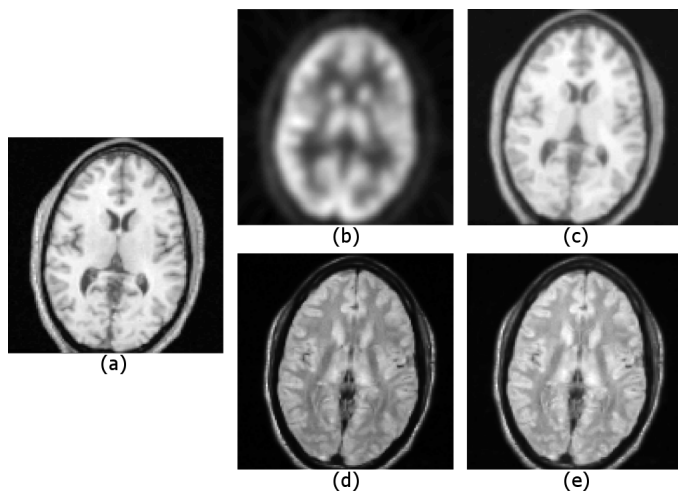


Figure 2. Image blending in brain MRI imaging. (a): SPGR MRI image as the initial image; (b): FDG-PET image as the template; (c): hybrid of image (a) and (b); (d): T2-star MRI image as another template; (e): hybrid image of (a) and (d).

In order to test this modified model, we turn to biomedical imaging for test images. In biomedicine, images of different modalities, such as CT, MRI, and PET, are often obtained for the same subject. This is especially true in brain imaging as different images could provide information on brain structures and functions. We chose to test our model on MRI and PET images.

FDG-PET imaging and MRI imaging of the brain were obtained from a normal subject in the Brain Trauma Project at UCLA. The available sequences in MRI images include SPGR, T2, and T2-star. All MRI images were first co-registered to PET using mutual information [23, 24] as the similarity measure. The mutual information between MRI and PET images were optimized using Powell's multi-dimensional search algorithm. We looked at two specific issues. First, we combined the features in MRI and PET images by solving (2). Second, we examined how different sequences in MRI images

influenced each other under this model. For all computations, the intensity values of all images were normalized to 0 to 256 before computation.

To investigate the first issue, we chose SPGR MRI images as the test images since SPGR images had the best spatial resolution and image quality. After MRI-PET coregistration was performed, an SPGR image slice (Fig. 2(a)) and a PET image slice (Fig. 2(b)) were extracted. We then used the SPGR image as the initial image and the corresponding PET image as the template. The fusion image was calculated by solving (2). The weight c of the regularizer is determined empirically. The final c used in this test was 2×10^6 . The corresponding PDE was solved at $t = 2$ to get the final fusion image. The result is shown in Fig. 2(c). Note that the overall contrast of the fusion images is very much like that of the initial SPGR image while the shape of the gray and white matters resemble that of the corresponding PET image.

For the second issue, we chose T2-star and SPGR images. T2-star images usually have lower resolution and in this subject the extracted slice was noted to have magnetic field inhomogeneity that resulted in shading of the image (frontal lobe has a much higher value than the occipital lobe) as shown in Fig. 2(d). On the other hand, the corresponding SPGR (Fig. 2(a)) was of good quality without shading. We thus used the SPGR image as the template and the T2-star as the initial image and see if we could improve the quality of the T2-star image. The result is shown in Fig. 2(e). The weight c used in this test was 6×10^6 and again the corresponding PDE was solved at $t = 2$. We notice that, although the resolution of the fusion image did not improve significantly, the intensity inhomogeneity almost completely disappeared.

VI. DISCUSSIONS AND CONCLUSIONS

In this paper, we propose a new method of defining the difference of two images. The idea is to look at the gradient field instead of the image itself. This results in a novel approach of image blending and a method of combining or hybridizing two images. The advantage of our approach is that it is scale free; the scale of the image is determined by the PDE that governs the minimization process. The drawback of our approach is that since we are working on the gradient fields we have to restrict ourselves to the space H^1 , which in theory does not allow discontinuous solutions. Though we do point out that numerically images are discretized signals and the gradient could always be viewed as finite.

Experimental results showed that our approach generated flow-like transition between images. This flow-like blending phenomenon is not observed in simple linear interpolation blending and provides a novel perception to human eyes. Although our model generates better results than simple interpolation method, it still has the undesirable overlapping effect of objects in the transition process. Thus, we have to emphasize that our model is not intended to improve or replace

image warping, but instead offers a different view of image blending and of quantifying differences between images.

We also applied our approach to biomedical images. We generated a hybrid MRI-PET image with features from both modalities. We also showed that by incorporating the information from SPGR image we were able to correct the magnetic field inhomogeneity of the corresponding T2-star images.

Another issue is that, in model c we seek to minimize the following energy:

$$\min_f E(f) = 1 - \frac{\int |\nabla f \bullet \nabla g| dx}{\left(\int |\nabla f|^2 dx \int |\nabla g|^2 dx\right)^{1/2}}$$

We can further generalize the model by replacing the absolute value function in the numerator with a continuous function Ψ :

$$\min_f E(f) = 1 - \frac{\int \Psi(\nabla f \bullet \nabla g) dx}{\left(\int |\nabla f|^2 dx \int |\nabla g|^2 dx\right)^{1/2}}$$

If we further require that this function is always non-negative and symmetric and its graph lies below the absolute value function, then our energy is still defined and between 0 and 1.

Moreover, if Ψ is differentiable, then the associated gradient descent PDE becomes:

$$\frac{\partial f}{\partial t} = \frac{\int \Psi(\nabla f \bullet \nabla g) dx}{\int |\nabla f|^2 dx} \Delta f - \frac{\text{div}(\Psi'(\nabla f \bullet \nabla g) \nabla g)}{\sqrt{\int |\nabla f|^2 dx \int |\nabla g|^2 dx}}$$

As a concrete example, the following choice of Ψ satisfies all the requirements:

$$\Psi(\xi) = \sqrt{\varepsilon + \log(1 + \xi^2)}, \text{ where } \varepsilon \text{ is a small number.}$$

By modifying the model this way, one can achieve an anisotropic form of the original model. Work to explore how different choices of Ψ will influence the results is still ongoing in our laboratory.

VII. ACKNOWLEDGMENTS

This study was partially supported by NIH Grant NS30308 and DoE Grant DE-FC0387-ER60615. We would like to thank Dr. Richard Tsai for the valuable discussions.

VIII. REFERENCES

- [1] R. T. Whitaker, "A level-set approach to image blending," *IEEE Transactions on Image Processing*, vol. 9, pp. 1849-61, 2000.
- [2] C. Watkins, A. Sadun, and S. Marenka, *Modern image processing: warping, morphing, and classical techniques*. Boston: Academic Press, 1993.
- [3] G. Wolberg, *Digital image warping*. Los Alamitos, Calif.: IEEE Computer Society Press, 1992.
- [4] S. Soatto and A. J. Yezzi, "Deformation: Deforming Motion, Shape Average and the Joint Registration and Segmentation of Images," UCLA, Los Angeles, UCLA CAM report 01-32, December 2001.
- [5] T. K. Carne, "The Geometry of Shape Spaces," *Proceedings of the London Mathematical Society*, vol. 61, pp. 407-432, 1990.
- [6] K. V. Mardia and I. L. Dryden, "Shape Distributions for Landmark Data," *Advances in Applied Probability*, vol. 21, pp. 742-755, 1989.
- [7] D. G. Kendall, "Shape Manifolds, Procrustean Metrics, and Complex Projective Spaces," *Bulletin of the London Mathematical Society*, vol. 16, pp. 81-121, 1984.
- [8] D. G. Kendall, *Shape and shape theory*. New York: Wiley, 1999.
- [9] H. L. Le and D. G. Kendall, "The Riemannian Structure of Euclidean Shape Spaces - a Novel Environment for Statistics," *Annals of Statistics*, vol. 21, pp. 1225-1271, 1993.
- [10] L. Younes, "Computable elastic distances between shapes," *Siam Journal on Applied Mathematics*, vol. 58, pp. 565-586, 1998.
- [11] L. Younes, "A distance for elastic matching in object recognition," *Comptes Rendus De L Academie Des Sciences Serie I-Mathematique*, vol. 322, pp. 197-202, 1996.
- [12] L. Younes, "Optimal matching between shapes via elastic deformations," *Image and Vision Computing*, vol. 17, pp. 381-389, 1999.
- [13] M. I. Miller and L. Younes, "Group actions, homeomorphisms, and matching: A general framework," *International Journal of Computer Vision*, vol. 41, pp. 61-84, 2001.
- [14] B. B. Kimia, A. R. Tannenbaum, and S. W. Zucker, "Shapes, Shocks, and Deformations .1. The Components of 2-Dimensional Shape and the Reaction-Diffusion Space," *International Journal of Computer Vision*, vol. 15, pp. 189-224, 1995.
- [15] B. B. Kimia, A. Tannenbaum, and S. W. Zucker, "On the Evolution of Curves Via a Function of Curvature .1. The Classical Case," *Journal of Mathematical Analysis and Applications*, vol. 163, pp. 438-458, 1992.
- [16] B. B. Kimia, A. Tannenbaum, and S. W. Zucker, "Toward a Computational Theory of Shape - an Overview," *Lecture Notes in Computer Science*, vol. 427, pp. 402-407, 1990.
- [17] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, pp. 629-39, 1990.
- [18] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," 1992.
- [19] Y. Yu-Li, X. Wenyuan, A. Tannenbaum, and M. Kaveh, "Behavioral analysis of anisotropic diffusion in image processing," *IEEE Transactions on Image Processing*, vol. 5, pp. 1539-53, 1996.
- [20] J. Weickert, "Recursive separable schemes for nonlinear diffusion filters," 1997.
- [21] J. Weickert, B. M. T. H. Romeny, and M. A. Viergever, "Efficient and reliable schemes for nonlinear diffusion filtering," *IEEE Transactions on Image Processing*, vol. 7, pp. 398-410, 1998.
- [22] G. Aubert and L. Vese, "A variational method in image recovery," *SIAM Journal on Numerical Analysis*, vol. 34, pp. 1948-79, 1997.
- [23] P. Viola and W. M. Wells, III, "Alignment by maximization of mutual information," 1995.
- [24] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Transactions on Medical Imaging*, vol. 16, pp. 187-98, 1997.