Estimating Nonrigid Shape Deformation Using Moments

Wei Liu   Eraldo Ribeiro

Computer Vision and Bio-Inspired Computing Laboratory
Florida Institute of Technology,
Melbourne, FL 32901, USA
`eribeiro@cs.fit.edu`

Abstract

Image moments have been widely used for designing robust shape descriptors that are invariant to rigid transformations. In this work, we address the problem of estimating non-rigid deformation fields based on image moment variations. By using a single family of polynomials to both parameterize the deformation field and to define image moments, we can represent image moments variation as a system of quadratic functions, and solve for the deformation parameters. As a result, we can recover the deformation field between two images without solving the correspondence problem. Additionally, our method is highly robust to image noise. The method was tested on both synthetically deformed MPEG-7 shapes and cardiac MRI sequences.

1. Introduction

Moments have been widely used in computer vision to achieve invariance to transformations [7, 13, 14, 9]. Key applications of moments include the design of robust shape descriptors [11], and the estimation of local image deformation [12, 10]. In this paper, we address the problem of calculating non-rigid deformation fields between image pairs. We propose a moment-based method that recovers dense deformation fields between shapes. Unlike some previous approaches [14, 9], that rely on feature matching, our method estimates nonrigid motion without any feature correspondence.

Our main contribution is based on the observation that moments are integral transforms of an image function $f(x, y)$ with bivariate polynomials $x^p y^q$. By parameterizing image deformation using a single family of polynomials, i.e., $x^p y^q$, changes in image moments can be approximated as a quadratic function of the deformation parameters. Solving this quadratic function

results in deformation parameters that best describe the changes in the moments. Furthermore, our method does not rely on prior information about object shape. Finally, image moments accomplish a robust global representation that do not rely on differential image properties such as intensity or gradients. The proposed method recovers accurate deformation fields from highly noisy images, paving the way for potential applications such as motion tracking in MRI or ultrasound images.

Moments have been mainly used as invariant local descriptors [14, 9] for affine and rigid transformation groups. However, few approaches address the problem of recovering image deformation directly from changes in images moments. For example, the idea of estimating dense deformation fields from texture moments was explored by Sato *et al.* [12]. Recently, Domokos *et al.* [5] proposed an affine registration algorithm using polynomial equations built from shape moments. Our work differs from [12] and [5] in three ways. First, we extend the affine-deformation model to a parameterized polynomial model. Secondly, our method is based on a simple numerical approximation, that leads to accurate and robust registration results. Thirdly, the method in [5] is restricted to binary images, while our method works for both binary and grayscale images. Finally, we define both the deformation field and moments using the same family of basis polynomials, leading to a simplified computation procedure. Polynomials are effective in modeling nonrigid motion fields [8, 6], but to our knowledge, this is the first work studying the interaction between polynomial models and image moments.

2. Image Deformation Model

Given a continuous function $f(x, y)$, the moment of order $(p+q)$ is usually defined by the following integral

185

IEEE
computer
society

transform of the polynomial kernel function $x^p y^q$:

$$M_{p,q} = \iint_\Omega x^p y^q f(x,y) dx dy, \qquad (1)$$

where $p, q \geq 0$, and the integration takes place over the whole support $\Omega \subseteq R^2$ of $f(x,y)$. As a result, $M_{p,q}$ is influenced by all values in Ω, making image moments quite robust to image noise.

Estimating the non-rigid deformation field between two images, $f(x,y)$ and $f'(x',y')$, is an ill-posed problem. The ill-posedness is usually alleviated by casting deformation field estimation as a parametrized model-fitting problem. For this, we can write the coordinate transform of the local image deformation as:

$$x' = x + u(x,y) \quad \text{and} \quad y' = y + v(x,y), \qquad (2)$$

where $[u(x,y), v(x,y)]^\top$ is the deformation field. By assuming a continuous deformation field, analytic to the order N, we can parameterize it using $x^p y^q$ as in (1):

$$u(x,y) = \sum_{s,t=0}^{s,t=N} a_{s,t} x^s y^t \quad \text{and} \quad v(x,y) = \sum_{s,t=0}^{s,t=N} b_{s,t} x^s y^t. \quad (3)$$

Here, when $N = 1$, we have the usual affine model. Other choices of polynomial kernels exist including the Zernike polynomials [13] that produce Zernike moments. In principle, our method is independent on the choice of specific polynomials, provided that the same family of polynomials are used for deformation-field parameterization and to moments definition.

3. Variation of Image Moments

We begin by defining the moments of transformed image $f'(x',y')$ as [7]:

$$M'_{p,q} = \iint x'^p y'^q f'(x',y') dx' dy'$$
$$= \iint (x+u)^p (y+v)^q f(x,y) |J| dx dy, \quad (4)$$

where $|J|$ is the determinant of the Jacobian matrix which is given by:

$$J = \begin{bmatrix} \frac{\partial x'}{\partial x} & \frac{\partial x'}{\partial y} \\ \frac{\partial y'}{\partial x} & \frac{\partial y'}{\partial y} \end{bmatrix} = \begin{bmatrix} 1 + u_x & u_y \\ v_x & 1 + v_y \end{bmatrix}. \qquad (5)$$

By assuming that (u,v) is both small and continuous compared to the object's scale, we can normalize the coordinates (x,y) by the image size such that $x, y \in [0,1]$ for $x, y \in \Omega$, and $u, v \ll 1$. Additionally, since

(u,v) is analytic, u_x and u_y are also $\ll 1$. As a result, $|J|$ can be approximated as:

$$|J| = 1 + u_x + v_y + u_x\,v_y - u_y\,v_x$$
$$\approx 1 + u_x + v_y$$
$$= 1 + \mathbf{div}(u,v). \qquad (6)$$

Additionally, the following approximation is also valid:

$$(x+u)^p \approx x^p + px^{p-1}\,u$$
$$(y+v)^q \approx y^q + qy^{q-1}\,v \qquad (7)$$

and:

$$(x+u)^p(y+v)^q \approx (x^p + px^{p-1}\,u)(y^q + qy^{q-1}\,v)$$
$$\approx x^p y^q + y^q x^{p-1}\,p\,u + x^p y^{q-1}\,q\,v. \quad (8)$$

Here, we can drop the product term of u, v. Substituting (6) and (8) into (4) and expanding, we have:

$$M'_{p,q} \approx \overbrace{\iint x^p y^q f(x,y) dx dy}^{M_{p,q}}$$
$$+ \iint (y^q x^{p-1}\,p\,u + x^p y^{q-1}\,q\,v) f(x,y) dx dy$$
$$+ \iint x^p y^p f(x,y)\mathbf{div}(u,v) dx dy$$
$$+ \iint \underbrace{(y^q x^{p-1}\,p\,u + x^p y^{q-1}\,q\,v)}_{\text{shape variation}} f(x,y) \underbrace{\mathbf{div}(u,v)}_{\text{area variation}} dx dy. \quad (9)$$

Notice that $\mathbf{div}(u,v)$ measures the infinitesimal area-change ratio of the deformation field. Equation 9 shows that the moments after transformation can be approximated by four components: the original moment $M_{p,q}$, the change caused by shape variation (i.e., the second term), the change caused by the transform' stretching (or shrinking) effect (i.e., the third term), and the change caused by the combination of these two factors (i.e., the last term). Next, we show how to recover the deformation field given both $M'_{p,q}$ and $M_{p,q}$.

4. Deformation Field Recovery

Since we parameterized the deformation (u,v) using polynomials $x^p y^q$, Equation 9 can be further simplified. For example, its second term can be expressed as:

$$\iint (y^q x^{p-1}\,p\,u + x^p y^{q-1}\,q\,v) f(x,y) dx dy$$
$$= \iint p \sum_{s,t=0}^{s,t=N} a_{s,t} x^{s+p-1} y^{t+q} f(x,y) dx dy$$
$$+ \iint q \sum_{s,t=0}^{s,t=N} b_{s,t} x^{s+p} y^{t+q-1} f(x,y) dx dy$$
$$= p \sum_{s,t=0}^{s,t=N} a_{s,t} M_{s+p-1,t+q} + q \sum_{s,t=0}^{s,t=N} b_{s,t} M_{s+p,t+q-1} \quad (10)$$
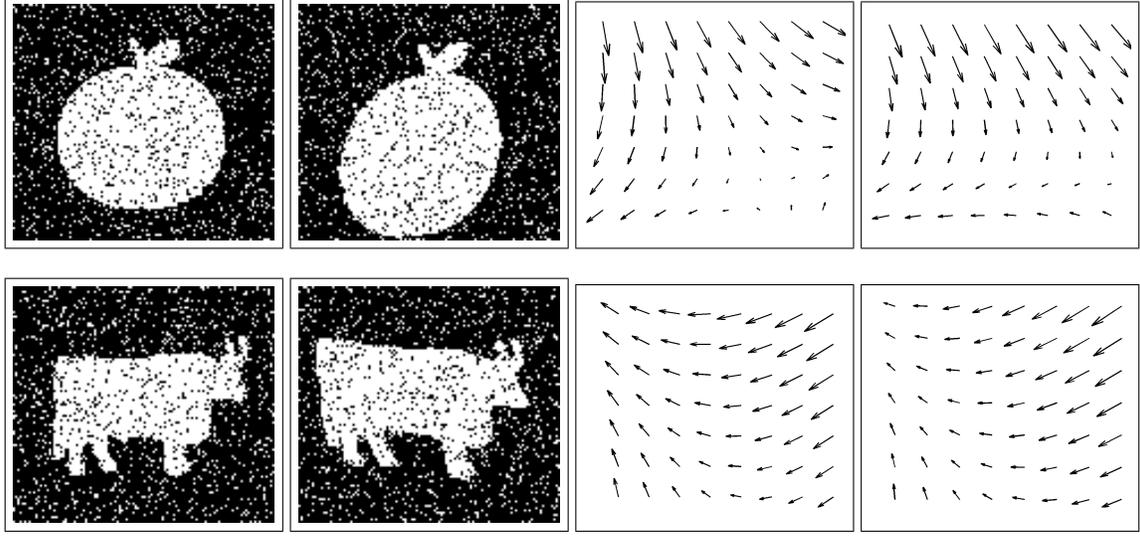
Figure 1. Deformation fields (synthetic images). First and second columns show the image and distorted images corrupted by salt-and-pepper noise (0.2 density). The third column shows the ground truth deformation field, and the last column shows the reconstructed results.

From (3), we obtain:

$$u_x = \sum_{s=1,t=0}^{s,t=N} s\, a_{s,t} x^{s-1} y^t \quad \text{and} \quad v_y = \sum_{s=0,t=1}^{s,t=N} t\, b_{s,t} x^s y^{t-1}. \quad (11)$$

As a result, $\mathbf{div}(u,v)$ can also be expressed as a polynomial. Following the same argument for the third and fourth term, Equation 9 can be rewritten as:

$$M'_{p,q} = M_{p,q} + \sum_{s,t=0}^{s,t=N} [a_{s,t}(p+s)M_{s+p-1,t+q}]$$

$$+ \sum_{s,t=0}^{s,t=N} [b_{s,t}(q+t)M_{s+p,t+q-1}] + Q(a_{s,t}, b_{s,t}), \quad (12)$$

where $Q(a_{s,t}, b_{s,t})$, $s,t = 0, \ldots, N$ is the last term in (9), and $Q(a_{s,t}, b_{s,t})$ is quadratic. If $X = [a_{0,0}, b_{0,0}, \ldots, a_{N,N}, b_{N,N}]^{\mathsf{T}}$ are the unknown parameters, we can rewrite Equation 12 as:

$$\Delta M_{p,q} = R_{p,q}^{\mathsf{T}} X + Q(X), \quad (13)$$

where $\Delta M_{p,q} = M'_{p,q} - M_{p,q}$, and $R_{p,q}$ is the coefficient matrix for the linear term. For each pair of (p,q), we obtain a quadratic function. Then, many numeric schemes can be used to solve the overdetermined system. For example, we could use the analytic form of $Q(X) = X^{\mathsf{T}} \widetilde{Q} X$ where $\widetilde{Q}$ is the symmetric and semi-definite matrix of the quadratic form. Then, algorithms for solving general quadratic equations could

be deployed. In this work, we adopt the fixed-point iteration algorithm [4], in which the previous solution X^t is updated to X^{t+1} by solving the linear equations:

$$\epsilon_{p,q}^t = \Delta M_{p,q} - Q(X^t) = R_{p,q}^{\mathsf{T}} X^{t+1}, \quad (14)$$

with X^0 initialized to zeros. Additionally, since higher-order image moments tend to be less reliable, we weight Equation 14 based on its order $W(p+q) = e^{-(p+q)}$.

Two issues should be mentioned. First, Equation 14 works for small deformations only. But large deformations can be handled by incremental warping. Secondly, deforming object parts might enter and exit the image region, causing undesired variation in image moments. We are currently studying ways to address these issues.

5. Experiments

Synthetic Deformations. We randomly selected 15 shapes from MPEG-7 database [2]. For each shape, we used the polynomial model to synthesize 10 different deformation fields, that were in turn used to warp the original images. To show the algorithm's robustness, we added independent salt-and-pepper noise to all images. The deformation model was set to have order $N = 2$ (i.e., affine model). Results were compared with Domoko's method [1] [5]. Figure 1 shows the synthesized images and the estimated deformation fields.

---

[1] http://www.inf.u-szeged.hu/~kato/
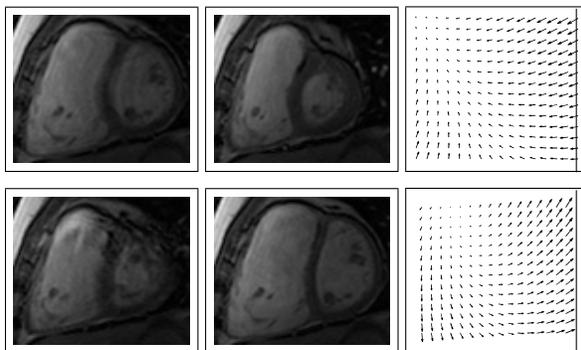
Figure 3. Cardiac MRI sequence. First row: two frames of a cardiac contraction cycle and the estimated deformation field. Second row: relaxation cycle.

To measure the reconstruction quality quantitatively, we used the concept of Average End-Point Error (APE) from optical flow [3]. In Figure 5, we plot the APE mean and variance as a function of noise energy level. In comparison with Domoko's method, our method has lower estimation error and smaller estimation variance.
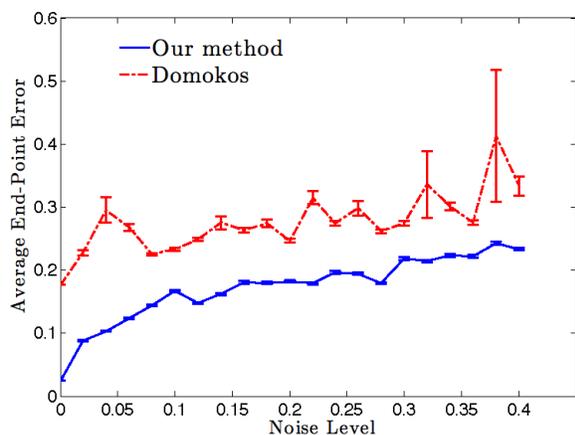
Figure 2. APE in terms of noise. Reconstruction error of deformation increases moderately as the noise level increases.

Cardiac MRI Sequence. We tested our algorithm on a 95×80-pixel Cardiac MRI sequence [1]. We used deformation model order $N = 3$, that is more flexible than the affine model. Figure 3 shows the deformation field estimated from the cardiac contraction and relaxation cycles. Unlike [1] and many existing works, our algorithm does not need prior shape or appearance models.

6. Conclusion

We presented a image-deformation estimation method that uses a polynomial deformation model and image moments. Future work includes the use of different basis functions as well as integrating the approach into spline-based registration methods.

References

[1] A. Andreopoulos and J. K. Tsotsos. Efficient and generalizable statistical models of shape and appearance for analysis of cardiac MRI. *Medical Image Analysis*, 12(3):335 – 357, 2008.

[2] X. Bai, X. Yang, L. J. Latecki, W. Liu, and Z. Tu. Learning context sensitive shape similarity by graph transduction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 99, 2009.

[3] S. Baker, S. Roth, D. Scharstein, M. Black, J. Lewis, and R. Szeliski. A database and evaluation methodology for optical flow. In *ICCV*, pages 1–8, 2007.

[4] R. L. Burden and J. D. Faires. *Numerical Analysis*. Brooks Cole, 2000.

[5] C. Domokos and Z. Kato. Parametric estimation of affine deformations of planar shapes. *Pattern Recogn.*, 43(3):569–578, 2010.

[6] J. Hoey and J. J. Little. Bayesian clustering of optical flow fields. *ICCV*, 2:1086, 2003.

[7] M. K. Hu. Visual pattern recognition by moment invariants. *IEEE Transactions on Information Theory*, IT-8:179–187, February 1962.

[8] O. Kihl, B. Tremblais, and B. Augereau. Multivariate orthogonal polynomials to extract singular points. In *ICIP*, pages 857–860, 2008.

[9] J. Maintz and M. Viergever. A survey of medical image registration. *Medical Image Analysis*, 2(1):1–36, 1998.

[10] A. Makadia and K. Daniilidis. Rotation recovery from spherical images without correspondences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28:1170–1175, 2006.

[11] K. Mikolajczyk and C. Schmid. A performance evaluation of local descriptors. *IEEE Trans. Pattern Anal. Mach. Intell.*, 27(10):1615–1630, 2005.

[12] J. Sato and N. Hollinghurst. Image registration using multi-scale texture moments. *Image and Vision Computing*, 13(5):496–513, 1995.

[13] M. R. Teague. Image analysis via the general theory of moments. *Journal of the Optical Society of America (1917-1983)*, 70:920–930, August 1980.

[14] B. Zitova. Image registration methods: a survey. *Image and Vision Computing*, 21(11):977–1000, October 2003.