Non-rigid Registration of Images with Geometric and Photometric Deformation by Using Local Affine Fourier-Moment Matching

Hong-Ren Su

Institute of Information Systems and Applications

National Tsing Hua University

Hsinchu,Taiwan

suhongren@gmail.com

Shang-Hong Lai

Department of Computer Science

National Tsing Hua University

Hsinchu,Taiwan

lai@cs.nthu.edu.tw

Abstract

Registration between images taken with different cameras, from different viewpoints or under different lighting conditions is a challenging problem. It needs to solve not only the geometric registration problem but also the photometric matching problem. In this paper, we propose to estimate the integrated geometric and photometric transformations between two images based on a local affine Fourier-moment matching framework, which is developed to achieve deformable registration. We combine the local Fourier moment constraints with the smoothness constraints to determine the local affine transforms in a hierarchal block model. Our experimental results on registering some real images related by large color and geometric transformations show the proposed registration algorithm provides superior image registration results compared to the state-of-the-art image registration methods.

1. Introduction

Image registration is a very fundamental problem in many fields [5][27], such as computer vision, image processing, computer graphics, and so on. However, images could be acquired from different sensors, with different camera acquisition settings, under different illumination conditions, or from different viewpoints. Therefore, the intensities and colors may not have the same distributions between the images to be registered [18], which makes the traditional geometric registration [8][12] methods easy to fail. Figure 1 shows this kind of image registration problems with very different illumination conditions. For example, some images capture the same scene at different time, such as different mountain views acquired in summer or winter, and the city view changes with differen traffic flows. Partial matching is another challenging problem for image registration.
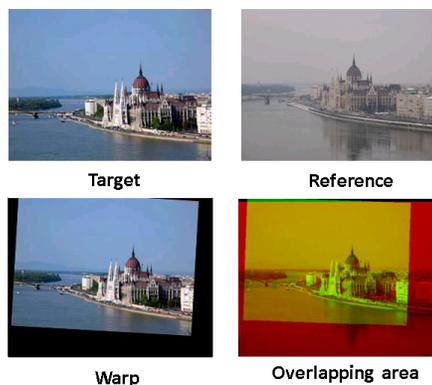

Figure 1. Reference, target images and its warp and overlapping area

An available strategy is to model and correct the appearance variations under different imaging conditions before registering images. Color optical flow [2][23], Eigenflows [19] and NRDC [25] were proposed to transform a target image to similar color distribution as that of the reference image for color transforms and photometrically invariant image representations. However, modeling all possibilities of appearance variations is impossible, thus the registration problem become very challenging when the color images contain large exposure differences or there are few corresponding points between two images.

Hence most previous research works dealing with the image registration taken under different camera and environment conditions are based on feature point correspondences. The SIFT-based feature matching approach [17], such as SURF [4], A-SIFT [20], SIFT flow [16] and CSIFT [1], is robust against lighting changes and can be considered a weak form of photometric invariance [15]. Although the SIFT-based feature matching is robust against illumination changes, the accuracy of feature point matching degrades as the appearance variation increases. In fact, the accuracy

for the registration of images acquired from different sensors or lighting conditions is usually quite limited [3]. A recent method proposed by Lin et al. [15], similar to SIFT flow, can deal with both photometric and geometric registration problems. However, their method is also based on the feature matching, and it is quite complicated and computationally expensive.

Another common approach for image registration [8][12] is based on minimizing cost functions defined by specific image distance measures, such as gradient-based method [6], histogram color mapping [24][13] and non-linear matching on tone mapping (MTM) [9]. The image distance measure is very crucial to the accuracy of image registration. MTM was proposed and proved to perform quite well for image registration with tone mapping. The performance of image registration is limited by the assumption implied in the cost function to be optimized. Shen et al. [21] proposed a new measure, called robust selective normalized cross correlation (RSNCC), which is combined with optical flow to register different types of images, and their method can deal with this problem well. Another idea is to resolve the geometric and photometric registration problems in the same framework, such as Kovalsky et al. [14]. They used a non-linear method to correct the color transform and employed a global affine model for geometric registration.

In this paper, we propose a weighted overlapping local affine image registration algorithm based on the Fourier moment matching (FMM) [22]. The proposed algorithm can not only deal with the image photometric and geometric transforms concurrently, but also deal with partial matching by using the weighted local affine model. The transform parameters can be solved by the proposed hierarchical approach. In addition, our method is more efficient than the previous methods in our experiments.

2. Main Contributions

FMM [22] is based on the fact that the affine transform between two images corresponds to a related affine transform between their Fourier spectrums, whose energies are normally concentrated around the origin in the frequency domain. Then, the affine transform can be solved by minimizing the affine relationship between the moments for the Fourier spectrums of the two images. FMM can be applied to 2D and 3D images easily. Hence, the photometric transform between two images can be considered as three-dimensional color histogram registration with L, a, and b color axis and the intensity from their corresponding frequency of Lab color in each image pixel. By integrating the 2D geometric image registration and 3D photometric alignment based on the Fourier moment matching, we can resolve the affine image registration problem under different illumination.

In order to make the algorithm more applicable for different viewpoint change, image distortion or image stitching, we propose a two-step hierarchal framework for weighted overlapping local affine registration based on FMM. The first step is to find the coarse corresponding area between two images by using the image block voting method on their corresponding edge maps, followed by estimating the global affine parameters in the photometric and geometric transforms by using the FMM. The second step is to find the accurate location for each pixel in images by partitioning the images into small overlapping blocks, and each block is associated with its own local affine parameters, which are determined by minimizing an energy function consisting of the FMM constraints and spatial smoothness constraints for the local affine parameters. By using a hierarchal scheme to decompose the image into blocks from a coarse scale to a fine scale, we can estimate the local affine parameters efficiently.

In order to increase the registration accuracy, we also use the edge information in the proposed method. We consider the photometric and geometric registration both in the intensity and edge information. We also add a weight function in each overlapping local affine to increase the accuracy especially for partial matching. In addition, local affine models in a hierarchal framework can increase the accuracy but need more time complexity. We can speed up the algorithm based on the property of the hierarchal framework. Our contributions in this paper can be summarized as follows:

1. We extend the FMM-based global affine registration to the hierarchical local affine model for non-rigid registration to deal with both the photometric and geometric registration.

2. We propose a two-step hierarchal estimation framework to estimate the local affine parameeters.

3. In each block of FMM constraints, we adaptively combine the edge information and color information for estimating each local affine model to improve the registration accuracy.

The rest of this paper is organized as follows: in the next section, the proposed FMM framework for image registration is introduced. Subsequently, we extend it for non-rigid registration by using a weighted overlapping local affine model in section 3. Experimental results are given in section 4 to demonstrate the proposed image registration algorithm. Finally, we conclude this paper in section 5.

3. Affine FMM in photometric and geometric registration

In this section, we describe the proposed algorithm in detail. The proposed framework for photometric and geometric image registration is based on using the FMM [22].

FMM is applied on an affine model and can be extended from 2D to 3D space. For image registration, geometric registration can be estimated by the 2D FMM, and photometric registration can be regarded as matching between two 3D color histograms, which can also be estimated by the 3D FMM. In addition, we convert the images from RGB to Lab color system, which has better linearity than other color spaces [7]. Thus we approximate the photometric transformation between images with an affine transform in the Lab space. In this framework, we integrate the 2D and 3D FMM into one system for simultaneous geometric and photometric registration.

3.1. Affine Relationship between Images and their corresponding Fourier Spectrums

Consider two images $f_1(\mathbf{x})$ and $f_2(\mathbf{y})$, $\mathbf{x}, \mathbf{y} \in R^n$ and $n = 2$ or 3, which are related by an affine transformation, $f_1(\mathbf{x}) = f_2(\mathbf{y})$, with

$$\mathbf{y} = A\mathbf{x} + \mathbf{t} \tag{1}$$

where A is an $n \times n$ invertible matrix and corresponds to the linear part of the affine transformation in eq.1. t is a $n \times 1$ translation vector. Assume the Fourier transforms of $f_1(\mathbf{x})$ and $f_2(\mathbf{y})$ are denoted by $F_1(\xi)$ and $F_2(\eta)$, respectively. Then, we can derive the following affine relationship between the Fourier transforms $F_1(\xi)$ and $F_2(\eta)$

$$|F_1(\xi)| = \frac{1}{|\triangle|}|F_2(\eta)| \tag{2}$$

where $\xi = A^T \eta$ and $|\triangle|$ is the determinant of the inverse of the transpose of A.

The matrix A can be solved by the Fourier moment matching technique [22]. From eq. (2), we can see the Fourier spectrums of two images related by an affine transform are also related by the corresponding affine transform and the translation vector t does not appear in eq. (2). Hence, after the invertible transform matrix A is estimated, we can obtain the transformed image $f_2'(\mathbf{x})$ and $f_2(A\mathbf{x})$. By minimizing the sum of square error (SSE) between two images, we can estimate the remaining translation vector $\mathbf{t}$ as follows:

$$\mathbf{t} = arg \min_{\mathbf{t}} ||f_2'(\mathbf{x} + \delta\mathbf{x}) - f_1(\mathbf{x})||^2 \tag{3}$$

3.2. 2D FMM for geometric registration

For 2D image registration, i.e. $n = 2$, we denote the invertible matrix A and translation vector t as follows:

$$A = \begin{bmatrix} a & d \\ b & e \end{bmatrix} \quad , \mathbf{t} = \begin{bmatrix} c \\ f \end{bmatrix} \tag{4}$$

The $(\alpha + \beta)$-th moment for the Fourier spectrum $|F_k(u, v)|$, $k = 1$ or 2, is defined as

$$m_{\alpha,\beta}^k = \int \int u^\alpha v^\beta |F_k(u, v)| du dv \tag{5}$$

With coordinate substitution, we can derive the following relationship for the first-order moments.

$$M_1^1 = \begin{bmatrix} m_{1,0}^1 \\ m_{0,1}^1 \end{bmatrix}, M_1^2 = \begin{bmatrix} m_{1,0}^2 \\ m_{0,1}^2 \end{bmatrix}, A_1 = \begin{bmatrix} a & b \\ d & e \end{bmatrix}, M_1^1 = A_1 M_1^2 \tag{6}$$

Similarly, we can derive the following relationship for the second-order moments:

$$M_2^1 = \begin{bmatrix} m_{2,0}^1 \\ m_{1,1}^1 \\ m_{0,2}^1 \end{bmatrix}, M_2^2 = \begin{bmatrix} m_{2,0}^2 \\ m_{1,1}^2 \\ m_{0,2}^2 \end{bmatrix}, A_2 = \begin{bmatrix} a^2 & 2ad & d^2 \\ ab & ae + bd & de \\ b^2 & 2be & e^2 \end{bmatrix}$$

$$M_2^1 = A_2 M_2^2 \tag{7}$$

The relationship of the first-order and second-order Fourier moments, given in eq. (6) and (7), respectively, can be used for the least squares estimation [37] of the above four 2D affine parameters

$$E(A) = \sum_{d=1}^2 w_d ||M_d^{(1)} - A_d M_d^{(2)}||^2 \tag{8}$$

where d is the degree of moment, and w_d is the weight associated with the Fourier moment constraint for the d-th degree. From eq. (8), we can estimate the four affine parameters (a,b,d,e). According to eq. (3), we add the translation term into eq. (8) to estimate the translation vector t as follows:

$$E(A, t) = \sum_{d=1}^2 w_d ||M_d^{(1)} - A_d M_d^{(2)}||^2 + ||f_2(A, \mathbf{t}) - f_1||^2 \tag{9}$$

Although the above image registration algorithm was derived for registering grayscale images, it can be easily extended for registering color images. For color images, the FMM can be applied to each color (L, a, b) channel to obtain the cost function, eq. (9), for each channel from the corresponding Fourier moments. Thus, the geometric transform between two color images can be estimated by minimizing the total cost function $E_g(A, t)$, which is the sum of the three cost functions (E_L, E_a and E_b) for channel L, a and b, respectively, with the same geometric transform, given as follows:

$$E_g(A, t) = E_L(A, t) + E_a(A, t) + E_b(A, t) \tag{10}$$

3.3. 3D FMM in photometric registration

Images may be taken at different time, with different sensors, or with different camera settings, and so on. In this situation, images taken at similar views may have very different color distributions. We use the affine transform to model the variation of color distributions. The Lab color histogram of an input image, denoted by a 3D array h, can be used to represent the image color distribution. The affine Lab color mapping can be considered as a typical 3D affine image registration problem, which can be solved by the 3D FMM described in section 3.1. Fig. 1 depicts an example of image registration between two color images captured under different photometric settings. The 3-channel Lab image is transformed into the 3D color histogram image h and the color mapping problem is transformed into a 3D geometric registration problem, which can be solved by the FMM algorithm [22]. Let the Fourier spectrums of the 3D histogram functions $h_1(x,y,z)$ and $h_2(x,y,z)$ be denoted by $H_1(u,v,w)$ and $H_2(u,v,w)$, respectively. If $h_1(x,y,z)$ and $h_2(x,y,z)$ are related by an affine transformation, i.e. $h_1(x) = h_2(Cx + \mathbf{p})$, then their Fourier spectrums are also related by the corresponding affine transform, i.e. $|H_1(u,v,w)| = |H_2(u',v',w')|/\triangle$ with the associated affine relationship between (u,v,w) and (u',v',w') given in eq. (2). Let the affine matrix C and translation vector p be denoted by

$$C = \begin{bmatrix} a & b & c \\ e & f & g \\ i & j & k \end{bmatrix} \quad , \mathbf{p} = \begin{bmatrix} d \\ h \\ l \end{bmatrix} \quad (11)$$

To determine these affine parameters, we employ the moment matching technique to the Fourier spectrums $|H_1(u,v,w)|$ and $|H_2(u,v,w)|$. The $(\alpha+\beta+\gamma)$-th moment for the Fourier spectrum $|F_n(u,v,w)|$ is defined as

$$m^k_{\alpha,\beta,\gamma} = \int\int\int u^\alpha v^\beta w^\gamma |F_k(u,v,w)| du dv dw \quad (12)$$

With coordinate substitution, we can derive the following relationship for the first-order moments.

$$N_1^1 = \begin{bmatrix} m^1_{1,0,0} \\ m^1_{0,1,0} \\ m^1_{0,0,1} \end{bmatrix}, N_1^2 = \begin{bmatrix} m^2_{1,0,0} \\ m^2_{0,1,0} \\ m^2_{0,0,1} \end{bmatrix}, C_1 = \begin{bmatrix} a & e & i \\ b & f & j \\ c & g & k \end{bmatrix}$$

$$N_1^1 = C_1 N_1^2 \quad (13)$$

Similarly, we can derive the following relationship for the second-order moments.

$$C_2 = \begin{bmatrix} a^1 & e^2 & i^2 & 2ae & 2ai & 2ei \\ b^1 & f^2 & j^2 & 2bf & 2bj & 2fj \\ c^1 & g^2 & k^2 & 2cg & 2ck & 2gk \\ ab & ef & ij & af+be & aj+bi & ej+fi \\ ac & eg & ik & ag+ce & ak+ci & ek+gi \\ bc & fg & jk & bg+cf & bk+cj & fk+gj \end{bmatrix}$$

$$N_2^1 = \begin{bmatrix} m^1_{2,0,0} \\ m^1_{0,2,0} \\ m^1_{0,0,2} \\ m^1_{1,1,0} \\ m^1_{1,0,1} \\ m^1_{0,1,1} \end{bmatrix}, N_2^2 = \begin{bmatrix} m^2_{2,0,0} \\ m^2_{0,2,0} \\ m^2_{0,0,2} \\ m^2_{1,1,0} \\ m^2_{1,0,1} \\ m^2_{0,1,1} \end{bmatrix}, N_2^1 = C_2 N_2^2 \quad (14)$$

The 3D affine matrix C and translation vector p can be estimated by minimizing the cost function $E_p(C, \mathbf{p})$ associated with the constraints in eq. (13), (14) and (3), given as follows:

$$E_p(C,p) = \sum_{d=1}^{2} w_d \|N_d^{(1)} - C_d N_d^{(2)}\|^2 + \|h_2(C,\mathbf{p}) - h_1\|^2 \quad (15)$$

3.4. Integration of photometric and geometric registration

In order to register the images captured under different lighting conditions, we combine energy functions associated with the photometric parameter estimation in eq. (15) and geometric parameter estimation in eq. (10) as follows:

$$E(A, \mathbf{t}, C, \mathbf{p}) = E_g(A, \mathbf{t}) + E_p(C, \mathbf{p}) \quad (16)$$

Note that the energy function can be minimized alternativly to obtain estimation for the affine geometric and photometric parameters between two images. In our implementation, we follow the order C, p, A, t to alternatively minimize the energy function. Satisfactory registration results can be obtained when the relationship between the two images can be well modeled by affine transformations in both the photometric and geometric spaces.

4. Local affine transformation

The above FMM approach is restricted by the geometric and photometric affine transform assumptions between the images for registration. Sometimes, the two images for registration may be taken at different time, from different viewpoints, under different lighting conditions, or with different camera settings, and affine transforms may not be satisfactory for modeling their relationship. In fact, the global affine model is limited for application to real images. In this section, we propose to extend the above FMM approach to estimate the non-rigid geometric and photometric transformations between two images by estimating the local affine transform for each block in the image. The local affine transforms for the photometric and geometric registration are estimated by minimizing the local FMM constraints for the image and color histogram spaces as well as the spatial smoothness constraints for the affine parameters.
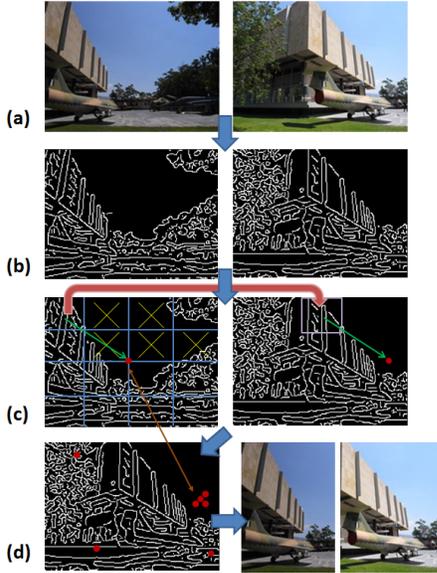
Figure 2. (a) Reference (left) and target (right) images, (b) edge maps of (a), (c) patch matching and sliding window, (d) left: target center voting in edge maps, right: image cropping with the same center.

The proposed local affine registration algorithm consists of two steps. The first step is to find a rough translation between two images in order to provide a good initial transformation for the subsequent refinement step. The second step is to estimate the local affine transformations by minimizing the associated energy function.

4.1. Initial global affine estimation

In order to overcome the limitation of affine transform between two images in the above photometric and geometric registration, we propose to first partition the edge map of reference image, as depicted in Fig. 2(a) and (b), and apply the above 2D geometric registration for each of the partitioned block to the edge map of target image in Fig. 2(c). The average edge amplitude for each block is used to check if the content in this block is rich enough for determining the alignment. The yellow cross in Fig. 2(c) means that the average edge amplitude is too small to be used for alignment. If a block is used, the above geometric registration is applied to estimate the affine transform from a sliding window in the target image in Fig. 2(c). The cross power spectrum $cps(x)$ of a block $q(x)$ and target image $r(x)$ from a sliding window is defined as

$$cps(x) = IFT((Q^{'}(u)R^*(u))/(|Q^{'}(u)R^*(u)|))$$

$$mv = \max real(cps(x)), t = \arg\max_x real(cps(x)) \quad (17)$$

where $Q^{'}(u)$ is the Fourier spectrum of $q^{'}(x)$ which is transformed by $q^{'}(x) = q(Ax)$. $R^*(u)$ is the conjugate

Fourier spectrum of $r(x)$, IFT means the inverse Fourier transform, $real$ means the real part of the complex number in the Fourier domain, mv is the max value of $cps(x)$ and t is the location of the max value mv in the $cps(x)$. In the cross power spectrum theory [12], t measures the displacement of $r(x)$ and $q(x)$ and mv means the similarity of $r(x)$ and $q(x)$. If the mv value is large, the estimation t is accurate.

According to the value mv, a similarity measure fm can be defined as follows:

$$fm = \frac{\max real(IFT((Q^{'}(u)R^*(u))/|Q^{'}(u)R^*(u)|))}{\max real(IFT((R(u)R^*(u))/|R(u)R^*(u)|))} \quad (18)$$

We compute the score fm for each block and a fixed threshold is used to determine if it is a match for the voting in each block. After the alignment for all the blocks in the sliding windows, the locations of the reference point are determined from the voting map for the target image computed from the alignment of the partitioned blocks in the target image. Fig 2 (d) depicts an example of the voting map. The aligned locations of all valid sub-blocks are included into the voting on the target image for all matched sliding windows to produce the voting map, which indicates the possibility of a reference point appearing at each location in the target image. The larger value at a location in voting map indicates the higher possibility of a reference point located at this location. Finally, according to the location, we can crop the reference point and target image into the same region and apply the photometric and geometric registration on the same region both in the reference and target image to determine the initial global affine parameters for the second step.

4.2. Overlapping local affine registration

After the initial affine transform is estimated from the first step, we partition the image uniformly into M-by-N blocks with 50% overlapping as depicted in Fig. 4 and estimate the corresponding local affine transforms for each block to achieve non-rigid image registration. We combine the local Fourier moment constraints for each block in the reference and target images and the spatial smoothness constraints for the affine transforms in the geometric and photometric space. For each block in the image, it is associated with a geometric 2-by-2 affine matrix $A_{i,j}$, a 2D translation vector $t_{i,j}$, a photometric 3-by-3 affine matrix $C_{i,j}$, and a 3D translation vector $p_{i,j}$. The whole parameter set of an image are $D = \{A_{i,j}C_{i,j}|i = 1, \ldots, n; j = 1, \ldots, m\}$ and $T = \{t_{i,j}^T p_{i,j}^T\}$. Thus, the local affine estimation can be extended from eq. (16) to minimize the following energy function:

$$E(D,T) = E_{data}(D,T) + \lambda E_s(D,T) \quad (19)$$

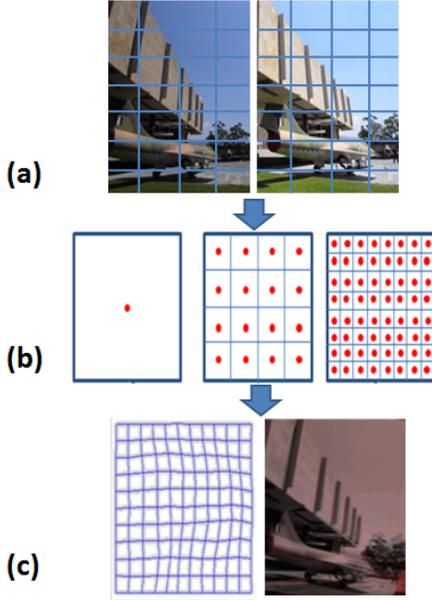where λ is the weight for the smoothness constraint, and the

Figure 3. (a) Initial global affine estimation and image partitioning. (b) Hierarchal block partitioning for coarse-to-fine local affine estimation. (c) Image registration result.

data term $E_{data}(D, T)$ is defined as

$$E_{data}(D, T) = \sum_{i=1}^{n} \sum_{j=1}^{m} [E_g(A_{i,j}, t_{i,j}) + E_p(C_{i,j}, p_{i,j})$$

$$+ \alpha E_e(A_{i,j}, t_{i,j})] e^{-(\beta_1(1-Sg_{i,j}) + \beta_2(1-Sc_{i,j})) ln(1+(n-1)/\beta_3)}$$

(20)

where $E_e(A_{i,j}, t_{i,j})$ is the Fourier moment cost function for the (i, j)-th block in the corresponding edge image, α is the weighting factor for the edge constraint, and $Sg_{i,j}$ and $Sc_{i,j}$ are the fm value estimated by eq. (18) showing similarity of the corresponding local affine blocks. $Sg_{i,j}$ is the geometric similarity measure and $Sc_{i,j}$ is the photometric similarity measure. Note that n is the hierarchal block partition level. β_3 is larger than 1. if n equals to 1, there is no weighting for all local affine blocks. When n is larger than 1, the weighting in each local affine blocks is more important. There is the same importance in the traditional local affine blocks, but there may be no corresponding blocks in different types of images when doing local affine. Hence we add the similarity weighting in each local affine block at different level n. This term can deal with photometric and geometric registration with occlusion well. $E_e(A_{i,j}, t_{i,j})$ is quite important if a block contains more edge information, and it can help the algorithm to increase the registration accuracy.

The smoothness term $E_s(D, T)$ is defined as

$$E_s(D, T) = E_{s_1}(D) + E_{s_2}(T)$$

$$E_{s_1}(D) = \sum [\|A_{i,j} - A_{i+1,j}\|_F^2 + \|A_{i,j} - A_{i-1,j}\|_F^2$$
$$+ \|A_{i,j} - A_{i,j+1}\|_F^2 + \|A_{i,j} - A_{i,j-1}\|_F^2$$
$$+ \|C_{i,j} - C_{i+1,j}\|_F^2 + \|C_{i,j} - C_{i-1,j}\|_F^2$$
$$+ \|C_{i,j} - C_{i,j+1}\|_F^2 + \|C_{i,j} - C_{i,j-1}\|_F^2]$$

$$E_{s_2}(T) = \sum [\|t_{i,j} - t_{i+1,j}\|_F^2 + \|t_{i,j} - t_{i-1,j}\|_F^2$$
$$+ \|t_{i,j} - t_{i,j+1}\|_F^2 + \|t_{i,j} - t_{i,j-1}\|_F^2$$
$$+ \|p_{i,j} - p_{i+1,j}\|_F^2 + \|p_{i,j} - p_{i-1,j}\|_F^2$$
$$+ \|p_{i,j} - p_{i,j+1}\|_F^2 + \|p_{i,j} - p_{i,j-1}\|_F^2]$$

(21)

where $\|\|_F$ is the Frobenius norm.

4.3. Hierarchical block partition for coarse-to-fine local affine estimation

Local affine model is sensitive to the initial estimation of the affine transforms. Hence we use the hierarchal block partitioning model to define the local affine transforms, which are solved by minimizing the energy function in eq. (14). For layer 0, we obtain the global affine parameters from the first step described in section 3.1 as the initial estimation for layer 1. The local affine parameters in layer 1 can be solved by minimizing eq. (14) and they are used as initial transforms for layer 2. This coarse-to-fine estimation process is repeated for several iterations until the desired block partitioning is reached. Fig. 3 shows the hierarchical block partitioning for coarse-to-fine local affine estimation. In our implementation, we use five layers to solve the local affine parameters. After the above process, the local affine parameters for each block are obtained. However, using the same affine transform for all pixels in one block is not smooth enough. In addition, there is a continuity problem for the transform across the boundary of adjacent blocks. To overcome this problem, we apply a 2D bilinear interpolation for the local affine transform parameters to compute the affine transform parameters for each pixel in the image, and then we compute the transformed values, photometric and geometric ones, from the associated affine transforms for all pixels to obtain the transformed image after the local affine transformation. Fig. 4 depicts an example of image registration experiment comparing the proposed method, FMM [22], and Ho et al. [11]. Note that the latter two methods correspond to the affine registration methods based on the moment matching in the Fourier domain and spatial domain, respectively. The results of FMM and Ho et al. were obtained by applying them to the edge maps in Fig. 4(b) and 4(d). The results show the proposed local affine registration algorithm provides the best registration result.

4.4. Computational Speedup

In the hierarchal block partition, if the similarity for a pair blocks is very high, we do not estimate its transfor-
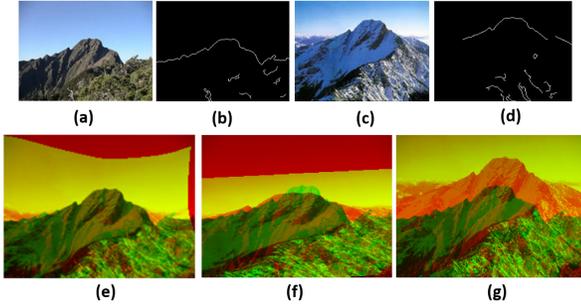
Figure 4. An example of image registration results: (a) reference image and (b) its edge map, (c) target image and (d) its edge map, and registered results by (e) proposed method, (f) FMM [22], and (g) Ho et al. [11]
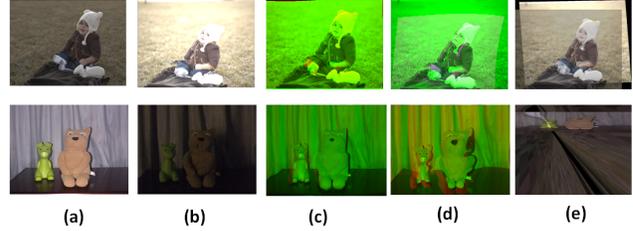

Figure 5. Some examples in the dataset from Shen et al. [21]. Images in the upper row are images of different exposures, images in the lower row are the flash/no flash examples. (a) Reference images, (b) target images, registered results by (c) the proposed algorithm, (d) Lin et al. [15], and (e) Zeragoza et al.[26]
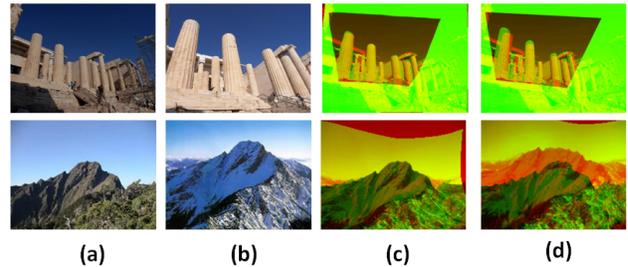

Figure 6. Some examples in our image dataset. (a) Reference images, (b) target images, registered results by the proposed algorithm (c) with and (d) without block weighting.

mation in the next layer of parameter estimation and fix its affine transformation in this layer. For the photometric registration, three layers are enough to describe color transformations. For geometric registration, we use 5 layers for the hierarchical local affine model estimation. By using this strategy of early termination, we can significantly reduce the execution time.

5. Experimental Results

In our experiments, we test the proposed image registration algorithm and some state-of-the-art algorithms on two image datasets. The first is the open dataset provided by Shen et al.[21]. The dataset contains 4 image pairs with different exposures and 3 image pairs with and without flash. The dataset also provides a set of human labeled corresponding points as the ground truth in the 7 image pairs that can be used for the quantitative evaluation of image registration accuracy. Figure 5 (a) and (b) shows an example of the image pairs. The other dataset is our dataset that contains 42 real color image pairs. It can be categorized into three clusters, 16 image pairs with large geometric transformations, 21 image pairs with large illumination variations and 5 image pairs with occlusion. We also select about 50 to 100 corresponding points in each pair of images for quantitative evaluation of image registration accuracy.

Table 1. Average registration errors (unit: pixels) of different methods on the dataset by Shen et al. [21]

Methods	Flash/No-flash	Diff. Exposure
SIFT Flow [16]	8.76	10.03
Hermosillo et al.[10]	16.57	13.24
Shen et al.[21](color+grad.)	4.56	2.25
Lin et al.[15]	6.42	4.75
Zaragoza et al.[26]	6.75	4.32
Zaragoza et al.+NRDC[25]	6.23	4.33
Proposed method without block weighting	2.37	4.02
Proposed method	1.86	2.32

Table 2. Average registration errors (unit: pixels) of different methods on our dataset.

Methods	Large Motion	Lighting Change	Occlusion
Lin et al.[15]	5.76	3.37	8.25
Zaragoza et al.[26]	4.68	3.51	5.87
Zaragoza et al.+NRDC[25]	5.23	3.26	6.17
Proposed method without block weighting	5.27	3.02	6.98
Proposed method	3.51	1.62	3.12

We demonstrate the proposed hierarchical local affine registration algorithm for estimating the geometric and photometric transforms for the image pairs in the two datasets in Fig. 5 and Fig. 6. The first image registration experiment is performed on the open dataset from Shen et al. [21]. We use the image pairs with different exposure and with/without flash in our experiments. We compared with the experimental results in Shen et al. [21] and also compared with Lin et al. [15], Zaragoza et al.[26], and Zaragoza et al. with color correction by using NRDC[25] as a preprocessing step. Table 1 shows the average errors of 7 image pairs. Our proposed algorithm can be used to solve such challenging image registeration problems. Especially adding the block weighting term in our proposed algorithm can significantly improve the registration accuracy. Some of the results are depicted in Fig. 5. It is evident from Table
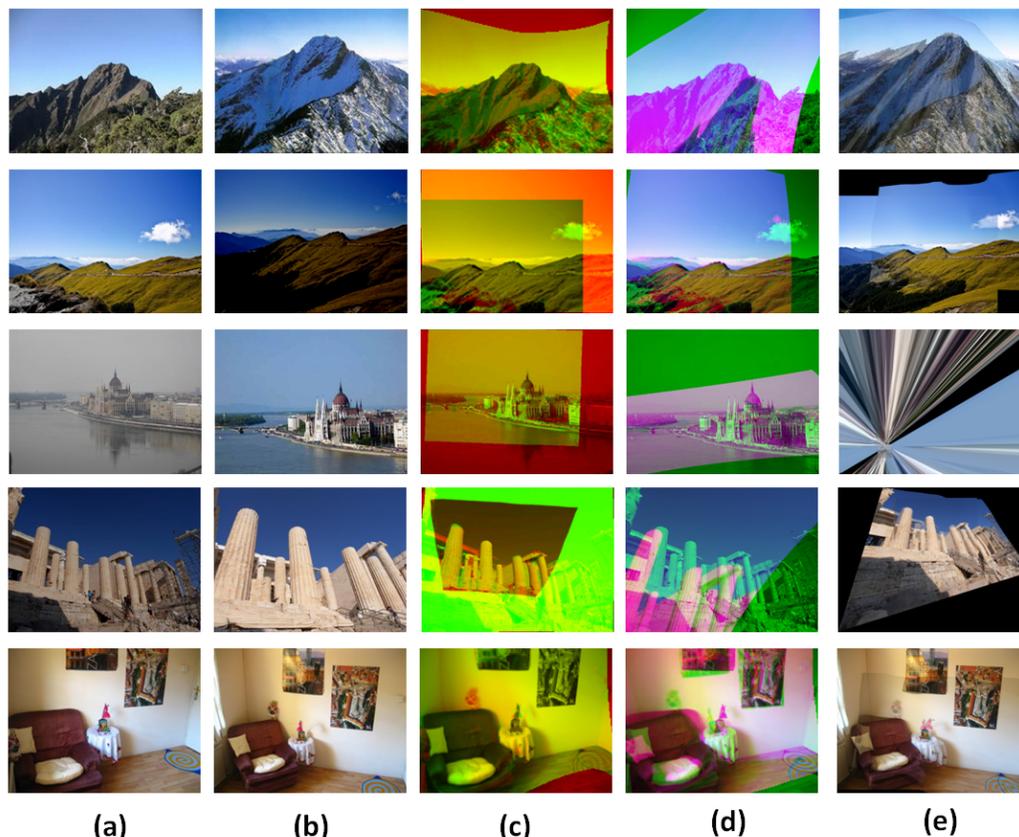
Figure 7. (a) Reference images, (b) target images, and registered results by (c) the proposed algorithm, (d) Lin et al. [15], and (e) Zeragoza et al.[26].

1 that the proposed algorithm outperforms the other methods for the flash/no-flash category and its accuracy on the different exposure category is comparable to the best result by Shen et al. [21].

The second image registration experiment is performed on our dataset of image pairs with three different categories; namely, large motion, lighting change and occlusion . Table 2 summarizes the average registration errors of the 42 image pairs under the three categories. In order to compare with the state-of-the-art non-rigid registration methods, we apply the methods by Lin et al. [15] and Zaragoza et al.[26] to the image pairs in this dataset. These two methods are representative recent approaches for image registration using local affine and feature point correspondences. It is obvious from Table 2 that our algorithm that can deal with photometric and geometric deformation simultaneously outperforms the other two competing state-of-the-art methods. Some image registration results from this dataset are depicted in Fig. 7.

In addition, we also compare the registration results of the proposed algorithm with and without using the block weighting in both datasets, and the results show the block weighting significantly imrpoves the registration accuracy for both datasets. Some examples of the registration results are shown in Fig. 6.

6. Conclusion

In this paper, we proposed a novel hierarchical local affine FMM-based algorithm for registering images with deformable geometric and photometric variations. We extend the FMM from global affine transformation to an overlapping local affine model, and adding the edge term and block weighting to improve the registration accuracy. We also use the hierarchial estimation framework to speed up the computational speed. Our experiments show the proposed non-rigid image registration algorithm is very robust and accurate for registering pairs of images related by large geometric and photometric transformations. We demonstrate superior registration accuracy of the proposed algorithm over the previous methods on two image datasets through quantitative assessment.

References

[1] A. E. Abdel-Hakim and A. A. Farag. Csift: A sift descriptor with color invariant characteristics. *In CVPR*, 2006.

[2] A. R. J. Andrews and B. C. Lovell. Color optical flow. *Workshop on Digital Image Computing*, 2003.

[3] A. Baumberg. Reliable feature matching acrosswidely separated views. *In CVPR*, 2000.

[4] H. Bay, T. Tuytelaars, and L. V. Gool. Surf: Speeded up robust features. *In ECCV*, 2006.

[5] L. G. Brown. A survey of image registration techniques. *ACM Computing Survey*, 24(4):325–376, 1992.

[6] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. *In CVPR*, 2005.

[7] M. Dan. Photoshop lab color: The canyon conundrum and other adventures in the most powerful colorspace. *Berkeley, Calif.*, 2006.

[8] J. W. et al. Comparison and evaluation of retrospective intermodality brain image registration techniques. *Journal of Computer Assisted Tomography*, 21(4):554–566, 1997.

[9] Y. Hel-Or, H. Hel-Or, and E. David. Fast template matching in non-linear tone-mapped images. *In ICCV*, 2011.

[10] G. Hermosillo, C. ChefdHotel, and O. Faugeras. Variational methods for multimodal image matching. *IJCV*, 50(3):329–343, 2002.

[11] J. Ho, A. Peter, A. Ranganranjan, and M. Yang. An algebraic approach to affine registration of point sets. *In ICCV*, 2009.

[12] P. W. Josien, J. B. Pluim, M. Antoine, and A. V. Max. Mutual information based registration of medical images: a survey. *IEEE Trans Med Imaging*, 22(8):986–1004, 2003.

[13] S. Kagarlitsky, Y. Moses, and Y. Hel-Or. Piecewise-consistent color mappings of images acquired under various condition. *In CVPR*, 2009.

[14] S. Kovalsky, G. Cohen, R. Hagege, and J. M. Francos. Decoupled linear estimation of affine geometric deformations and non-linear intensity transformations of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 2010.

[15] W.-Y. Lin, L. Liu, Y. Matsushita, K.-L. low, and S. Liu. Aligning images in the wild. *In CVPR*, 2012.

[16] C. Liu, J. Yuen, A. Torralba, J. Sivic, and W. T. Freeman. Sift flow: Dense correspondence across different scenes. *In ECCV*, 2008.

[17] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *IJCV*, 2004.

[18] H. Q. Luong, B. Goossens, A. Pizurica, and W. Philips. Joint photometric and geometric image registration in the total least square sense. *Pattern Recognition Letters*, 32:2061–2067, 2011.

[19] E. G. Miller and K. Tieu. Color eigenflows: Statistical modeling of joint color changes. *In ICCV*, 2001.

[20] J. Morel and G. Y. Asift. A new framework for fully affine invariant image comparison. *SIAM Journal on Imaging Sciences*, 2009.

[21] X. Shen, L. Xu, Q. Zhang, and J. Jia. Multi-modal and multi-spectral registration for natural images. *In ECCV*, 2014.

[22] H. R. Su and S. H. Lai. Ct-mr image registration in 3d k-space based on fourier moment matching. *In PSIVT*, 2011.

[23] B. J. van de Weijer and T. Gevers. Robust optical flow from photometric invariants. *In ICIP*, 2004.

[24] Y. Wei and L. Tao. Efficient histogram-based sliding window. *In CVPR*, 2010.

[25] H. Yoav, S. Eli, B. G. Dan, and L. Dani. Nrdc. non-rigid dense correspondence with applications for image enhancement. *ACM Transactions on Graphics (Proc. SIGGRAPH)*, 2011.

[26] J. Zaragoza, T. J. Chin, M. Brown, and D. Suter. As-projective-as-possible image stitching with moving dlt. *In CVPR*, 2013.

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