

Normalized Total Gradient: A New Measure for Multispectral Image Registration

Shu-Jie Chen and Hui-Liang Shen, *Member, IEEE*

Abstract—Image registration is a fundamental issue in multispectral image processing. In filter wheel based multispectral imaging systems, the non-coplanar placement of the filters always causes the misalignment of multiple channel images. The selective characteristic of spectral response in multispectral imaging raises two challenges to image registration. First, the intensity levels of a local region may be different in individual channel images. Second, the local intensity may vary rapidly in some channel images while keeps stationary in others. Conventional multimodal measures, such as mutual information, correlation coefficient, and correlation ratio, can register images with different regional intensity levels, but will fail in the circumstance of severe local intensity variation. In this paper, a new measure, namely normalized total gradient (NTG), is proposed for multispectral image registration. The NTG is applied on the difference between two channel images. This measure is based on the key assumption (observation) that the gradient of difference image between two aligned channel images is sparser than that between two misaligned ones. A registration framework, which incorporates image pyramid and global/local optimization, is further introduced for rigid transform. Experimental results validate that the proposed method is effective for multispectral image registration and performs better than conventional methods.

Index Terms—Multispectral image, multimodal image, filter wheel, image registration, image alignment, similarity measure, total gradient, sparsity, rigid transform, affine transform, global optimization, local optimization, intensity variation.

I. INTRODUCTION

Multispectral color imaging has attracted intensive interest in recent years as it can acquire more spectral information than traditional RGB cameras. A multispectral imaging system can be set up by using a camera and an optical device that splits the visible spectrum reflected from the imaged object. Typical light-splitting devices include electronically controlled tunable filters [1], [2] or mechanically controlled filter wheel [3]. Compare with tunable filters, the filter wheel has the flexibility in filter selection and can achieve high spectral transmittance in the whole visible spectrum range. These two attributes are essential to general-purpose high-quality multispectral imaging.

In this work, the multispectral images are acquired using a filter wheel based imaging system. As illustrated in Fig. 1, the system consists of a monochrome digital camera and a filter wheel installed with a series of optical narrowband filters.

This work was supported by the National Natural Science Foundation of China under Grant 61371160, and in part by the Hong Kong Research Institute of Textiles and Apparel (HKRITA) under Grant ITP/048/13TP. (Corresponding author: Hui-Liang Shen.)

The authors are with the College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China (e-mail: incredible77@zju.edu.cn; shenhl@zju.edu.cn).

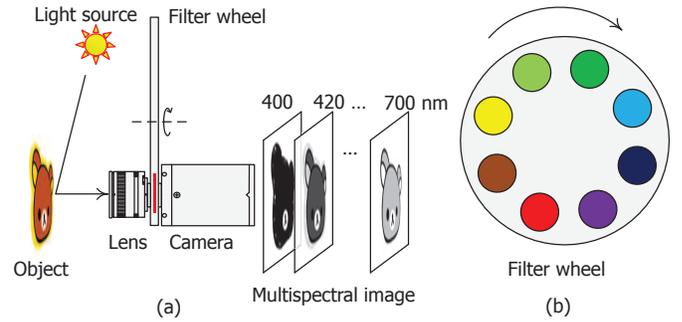


Fig. 1. The multispectral imaging system (a) and filter wheel (b) used in this work. Note that in our system totally 16 filters are installed on the wheel, but only 8 filters are shown for illustration purpose.

These filters are of central wavelengths 400, 420, ..., 700 nm and full width at half maximum (FWHM) 10 nm. The channel images are sequentially captured by positioning the filters between the lens and camera. Note that for illustration purpose only 8 filters are shown in Fig. 1.

The filter wheel based multispectral imaging system usually suffers from image misalignment and out-of-focus blur. The misalignment between channel images is caused by the non-coplanar placement of filters [3]. It severely precludes the direct use of multispectral images. As shown in Fig. 9, the misalignment causes chromatic aberration in the multispectral image (displayed in RGB). Though calibration-based technique for multispectral image alignment is available [3], the system must be re-calibrated when the imaging distance changes. Hence, it is more convenient and favored to explore a calibration-free multispectral image registration method for practical applications. On the other hand, the out-of-focus blur is caused by the different effective focal lengths at individual channels, which originates from the fact that the refractive indices of the lens are wavelength dependent [4]. To deal with this problem, an autofocus method [5] has been proposed to compute the focus positions of lens using a focusing device (step motor). Alternatively, when the channel images are well aligned, the blur can also be computationally removed by multispectral image deblurring [4]. In this work, we will show in Section VII that multispectral images can be restored based on the proposed registration framework and the deblurring method introduced in [4]. In this way, neither the calibration board nor the focusing device is further needed in the multispectral imaging systems.

Image registration aims to find correspondences between two images by maximizing the defined similarity measures.

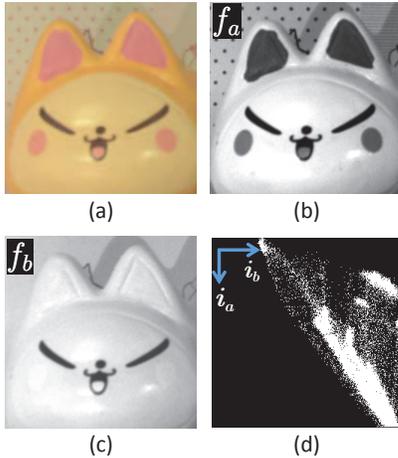


Fig. 2. Local intensity variation of an aligned multispectral image. (a) Multispectral image displayed in RGB. (b) Image f_a at channel 9 (560 nm). (c) Image f_b at channel 16 (700 nm). (d) Joint histogram of the two channel images. Notice that the pink spots on the cat head are sensed in image f_a but invisible in image f_b .

Extensive surveys of registration techniques can be found in [6], [7]. The techniques can be classified into two categories according to the descriptors used to formulate the similarity measures [7], i.e., the feature-based one and intensity-based one. The feature-based methods usually involve feature detection, feature matching, transform model estimation, and resampling. Generally these methods are more efficient but less accurate compared to the intensity-based methods. As one of the most successful feature-based methods, scale invariant feature transform (SIFT) has been widely used in a variety of matching tasks in computer vision [8]. The intensity-based registration methods have been developed and applied in medical imaging [6]. Similarity measures, including sum-of-squared-differences (SSD), correlation coefficient (CC), correlation ratio (CR), and mutual information (MI), have been widely employed in image registration. However, these measures are not robust when registering two images with spatially-varying local intensities [9].

Multispectral image challenges the conventional intensity-based registration methods at two aspects. First, the intensity levels of the corresponding local region in individual channel images can be essentially different. Second, the local intensity may vary in some channel images while keeps stationary in other channel images. Figure 2 illustrates such an example. The contrast between the pink regions and the background is well sensed in the 9th channel image f_a , but vanishes in the 16th channel image f_b . Consequently, there does not exist a functional mapping between the gray levels i_a and i_b in the joint histogram illustrated in Fig. 2(d).

This paper proposes a new intensity-based method, which can resolve the mentioned challenges, for multispectral image registration. The normalized total gradient (NTG), which is applied on the difference between two channel images (referred as *difference image* hereafter), is introduced to measure the alignment of two images. The employment of NTG is based on the key assumption that the gradient of difference

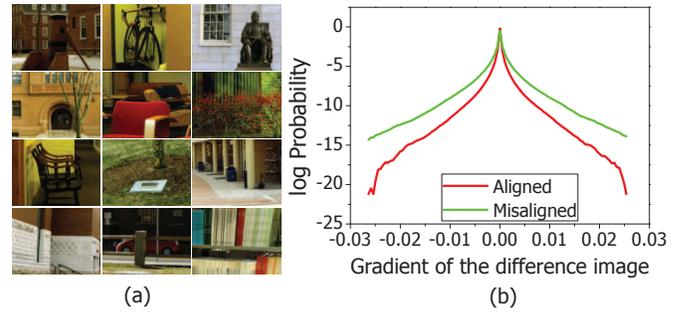


Fig. 3. Multispectral images of real-world scenes displayed in RGB (a) and distribution of gradient of difference image between channel images (b).

image between two aligned images is sparser than that between two misaligned images. This assumption is validated by the statistical investigation on a large number of multispectral images of real-scenes (see Fig. 3). Based on the NTG, an image registration framework, which incorporates image pyramid and global/local optimization, is introduced for general affine transform. During the minimization, an upside down image pyramid is first constructed for computation acceleration. Differential evolution [10], a powerful global optimization algorithm, is then used to find the initial optimal point at the bottom layer of the pyramid. Finally Newton's method is implemented to successively improve the solution at each layers. Using such a framework, the global optimal affine transform can be found.

The success of the NTG measure in multispectral image registration is due to two reasons. First, the gradient operation weakens the influence of the local intensity that is slowly varying. Second, the sparseness measure allows for rapidly varying local intensity. In fact, the rapid intensity variation contributes rather few additional edges in the gradient map of the difference image, and consequently the sparseness of the gradient map is rarely affected.

To summarize, the main contribution of this work is twofold. First, the NTG is introduced as a registration measure based on the assumption that the gradient of difference image is most sparsely distributed when the two channel images are aligned. Second, a framework, which comprises image pyramid and global/local optimization, is proposed for multispectral image registration. Experimental results show that the proposed method can efficiently register both multispectral and multimodal images.

The remainder of this paper is organized as follows. Section II reviews the related work on multispectral/multimodal image registration. The NTG is introduced in Section III and the image registration framework is described in Section IV. The computation details of optimization is presented in Section V. Section VI shows the experimental results and Section VII illustrates the applications of the proposed image registration method. Finally, Section VIII concludes the paper.

II. RELATED WORK

Intensity-based image registration can be further divided into unimodal and multimodal image registration based on the

criterion whether corresponding pixels (voxels) have similar intensity values. It is clear that multispectral image registration belongs to the latter one. In the following only the most relevant registration methods are reviewed.

A. Mutual Information

Mutual information (MI) was first introduced for medical image registration in [11] and was extensively investigated in [12]. MI is defined as

$$I(A, B) = H(A) + H(B) - H(A, B), \quad (1)$$

where $H(\cdot)$ denotes the entropy of the corresponding image, and $H(A, B)$ denotes the entropy of the joint histogram of images A and B . When two images are correctly registered, corresponding regions should overlap and the joint histogram will show certain clusters for the intensities of those regions. This consequently results in a least joint entropy $H(A, B)$ [12]. Thus maximizing mutual information can be viewed as registering images A and B such that, in the overlap, the information ($H(A)$ and $H(B)$) provided by the images is large and the local regions are in good match. To further increase immunity against overlap, normalized mutual information has been investigated [13].

As a variant of MI, regionalized mutual information was introduced to cope with the local intensity variation [14]. By assuming that the local intensities are slowly varying across the image, a global mutual information is constructed from local regional distributions and hence the registration accuracy improves. Additionally, the limitation on slowly varying local intensity can also be avoided by computing MI only in the regions with large spatial variations (e.g., corner and edge). This idea is implemented in [15] and [16] by augmenting MI with the Harris operator or saliency measure. Such methods perform relatively well on images with slow intensity variation at the cost of high computation complexity.

B. Correlation Ratio

While MI describes the clustering property of joint histogram by entropy, correlation ratio (CR) characterizes such property by variance [17]. CR stems from the law of total variance

$$\text{var}(A) = \mathbb{E}_B(\text{var}(A|B)) + \text{var}_B(\mathbb{E}(A|B)). \quad (2)$$

When the joint histogram is well clustered, for every single intensity which represents some local region in B , the variance of intensity in the same local region of A should be small. This leads to a small value of $\text{var}(A|B)$. For all intensities of B in the region of overlap, the mean variance of A denoted by $\mathbb{E}_B(\text{var}(A|B))$ are hence the smallest when registered. Accordingly, the proportion of the second term of the right hand side of (2) arrives its maximum, which indicates the correlation ratio

$$CR = \frac{\text{var}_B(\mathbb{E}(A|B))}{\text{var}(A)} \quad (3)$$

is maximized.

Correlation ratio was thrown light upon again in the view of pattern matching recently. The work [18] starts from the discussion of matching by tone mapping (MTM) and ends up with a variant of CR. According to [18], the pattern to window distance in MTM is computed as

$$D(\mathbf{q}, \mathbf{w}) = \frac{1}{|\mathbf{w}| \text{var}(\mathbf{w})} \left(\|\mathbf{w}\|^2 - \sum_j \frac{1}{|\mathbf{q}^j|} (\mathbf{q}^j \cdot \mathbf{w})^2 \right), \quad (4)$$

where \mathbf{w} is the window patch, \mathbf{q}^j is the j th slice in pattern \mathbf{q} , $|\mathbf{q}^j|$ is the pixel number in slice \mathbf{q}^j , and $|\mathbf{w}|$ is the pixel number of the window patch. Note that $\sum_j |\mathbf{q}^j| = |\mathbf{w}|$. The distance (4) can be reformulated as

$$D(\mathbf{q}, \mathbf{w}) = \frac{1}{|\mathbf{w}| \text{var}(\mathbf{w})} \sum_j |\mathbf{q}^j| \left(\frac{\sum_{i \in \mathbf{q}^j} w_{ji}^2}{|\mathbf{q}^j|} - \left(\frac{\sum_{i \in \mathbf{q}^j} w_{ji}}{|\mathbf{q}^j|} \right)^2 \right), \quad (5)$$

where w_{ji} denotes the intensity of the i th pixel in the j th slice of the window patch specified by the j th slice of the pattern. Equation (5) can be further transformed to

$$D(\mathbf{q}, \mathbf{w}) = \frac{1}{|\mathbf{w}| \text{var}(\mathbf{w})} \sum_j |\mathbf{q}^j| \text{var}(\mathbf{w}^j), \quad (6)$$

where \mathbf{w}^j is the j th slice of the window patch with respect to pattern \mathbf{q}^j . With some transformation, it can be validated that the distance (6) equals to $(1 - CR)$. Hence MTM and CR are actually equivalent.

It should be noted that both MI and CR describe the regional correspondence of two different modal images. When the intensity of two images can be well approximated by functional mapping, which allows for a small intensity variation of a single local region, both MI and CR can achieve outstanding registration performance. MI is more robust than CR with respect to intensity variation because entropy is more accurate in measuring the clustering property than variance. Entropy measures the sparseness of the joint histogram and allows for a wider range of distributions of intensity while variance is only best suited to intensity that follows the normal distribution. However, such functional correspondence does not always exist in multispectral images. Figure 2 shows an example of aligned patch of channel 9 (560 nm) and channel 16 (700 nm). The joint histogram of these two patches cannot be approximated as a functional mapping from the intensity of f_a to that of f_b , but the joint distribution still clusters.

C. Local Normalized Correlation Coefficient

While MI and CR are global intensity-based registration methods, local normalized correlation coefficient (NCC) [19] has been introduced to measure the correspondence between two modal images locally. Based on the assumption that the directional derivative energy maps are locally linearly correlated, normalized cross correlation can be used to measure the similarity between two multimodal patches. However, since not all patches satisfy this assumption, an outlier rejection strategy is needed when constructing a global cost function. To

weaken the local linearity assumption, a more robust variant of NCC called robust selective normalized cross correlation (RSNCC) criterion [20] was recently proposed to solve the dense matching problem of both multispectral and natural images. As stated in [20], though structure inconsistency and strong outliers caused by shadow and highlight are well handled in the registration framework, large errors may occur on regions that do not contain informative features.

The local linearity assumption is a strong functional dependence of descriptors in regions specified by user-defined windows. Such descriptors can be the intensity directly or the gradient. Either explicit outlier rejection strategy or implicit robust selective function aims to reduce the risk of registration inaccuracy, which results from those regions that violate such assumption. However, the procedure of selection is empirical and increases the computation complexity.

D. Residual Complexity

Our proposed method is closely related to the work in [9] in the sense of measuring the complexity of the difference image. In [9] the difference of two unimodal images with local intensity variation, referred as the residual image, was used. The author showed that the complexity of the residual image (RC) reduced to its minimum when two images were aligned. RC can be simply represented by using sparse formulation of discrete cosine transform (DCT) coefficients.

However, RC is less suitable for multispectral image registration. When two images are of different modality, the sparseness of the DCT coefficients of the difference image is no longer guaranteed. The reason is that the intensities of two corresponding regions do not counteract while registered. Compared to RC, the proposed method measures the sparseness of the gradient of difference image. By taking the gradient, the residual intensity counteracts with its neighbors so that multi-modality impacts least on the sparseness.

III. NORMALIZED TOTAL GRADIENT (NTG)

Statistics of image features are of significant use in image processing and computer vision. The sparseness of gradients of natural images has been prevalently applied in single image deblurring [21], denoising [22], and inpainting [23]. Statistics of multispectral images have also been explored for the efficient representation of multispectral images [24]. In this work, the gradient of difference image is studied and the sparseness of such feature is employed in multispectral image registration.

A. Sparseness

To validate the sparseness assumption of the gradient of difference image, a multispectral image database with 77 real-world scenes from [24] is employed. Some images are shown in Fig. 3(a). Each multispectral image consists of 31 channels and covers the visible spectrum from 420 to 720 nm at an interval of 10 nm. Without loss of generality, image intensities are normalized to the $[0, 1]$ range. All multispectral images are originally aligned. Random displacements are imposed on all

the channels, except for the reference channel, to synthesize the misaligned multispectral images. The 16th channel is chosen to serve as the reference. The *difference images* are then computed as the differences between the reference and the other channel images.

The distributions of the gradient of difference image in cases of alignment and misalignment are shown in Fig. 3(b). It is observed that the distribution corresponding to the aligned channel images is more heavy-tailed than that corresponding to the misaligned ones. More zero gradients and less large gradients of the difference image are obtained by registration. In other words, the gradient of difference image becomes sparser when the multispectral image is aligned. This can be mathematically formulated as

$$\sum_l |\nabla_l \{f - f_R\}| \leq \sum_l |\nabla_l \{\tilde{f} - f_R\}|, \quad (7)$$

where f_R represents the reference channel image, f and \tilde{f} denote the aligned and misaligned channel images, respectively. The operator ∇_l , with $l \in \{x, y\}$, denotes the gradient computation along the direction l . Notice that $|\cdot|$ is the L_1 norm, the absolute gradient of difference image $f - f_R$ can be explicitly expanded as

$$|\nabla_l \{f - f_R\}| = \sum_{\mathbf{x}} |\nabla_l \{f(\mathbf{x}) - f_R(\mathbf{x})\}|,$$

where $\mathbf{x} = (x, y)^T$ denotes the spatial location.

The explanation of the sparsest distribution of gradient in circumstance of well alignment is straightforward. It is expected that, when two images are aligned, the gradients of these images generally overlap and hence the number of nonzero gradients of the difference image reduces. A rigorous paradigm that characterizes such number should be the L_0 norm. In this work we employ the L_1 norm instead since it is computationally efficient and performs comparatively to the L_0 norm [25].

In what follows, we refer to the L_1 norm of the gradient of the image along all directions (x and y in this work) as the total gradient (TG) of the image, i.e.,

$$TG(f) = \sum_l |\nabla_l f|.$$

Note that this definition is different to total variation (TV) [26]. With the definition, the expression (7) can be interpreted that the TG of difference image is minimized when two images are registered. Hence TG functions as a simple measure for multispectral image registration.

B. Normalization

Region of overlap between two images is composed by those pixels whose intensities or features are used to compute the measure. A measure is of limited use if it is sensitive to the overlap variation. The work [13] presents an extensive discussion on the overlap, and proposes the normalized mutual information (NMI) to reduce the effect of varying overlap on entropy measure.

The TG of the difference image is sensitive to the overlap because the summation of absolute gradients is defined on the

region of overlap. The TG reduces when the region of overlap shrinks. By considering the L_1 norm as a kind of energy measure, we propose the normalized total gradient (NTG) of difference image as follows,

$$NTG(f, f_R) = \frac{\sum_l |\nabla_l \{f - f_R\}|}{\sum_l |\nabla_l f| + \sum_l |\nabla_l f_R|}. \quad (8)$$

In (8), the numerator is the TG of the difference image and the denominator is the total energy that normalizes the TG measure. It can be easily verified that $0 \leq NTG \leq 1$. Minimizing NTG is equivalent to registering two images while increasing their total energy in the region of overlap.

IV. IMAGE REGISTRATION USING NTG

The proposed measure NTG is minimized when two images are registered. Though NTG can be applied to a variety of registration tasks, we narrow down to image registration with global affine transform model. The affine transform model is frequently used as the pre-processing of medical image registration [27]. Besides, it has been established that the multispectral imaging registration in the filter wheel based imaging system can be well approximated by the affine transform [28].

A. Affine Transform

The aim of image registration is to find a transform \mathbf{p} such that the transformed floating image $g(\mathbf{x}, \mathbf{p}) = f(u(\mathbf{x}, \mathbf{p}), v(\mathbf{x}, \mathbf{p}))$ matches the reference image $f_R(\mathbf{x})$. The affine transform can be formulated as

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} p_1 & p_2 & p_3 \\ p_4 & p_5 & p_6 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \mathbf{P} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}, \quad (9)$$

where $\mathbf{p} = (p_1, p_2, \dots, p_6)^T$ is the vector form of \mathbf{P} . According to the discussion above we know that when the transformed floating image g and the reference image f_R are well matched, the $NTG(g, f_R)$ measure is minimized. Hence registering the floating image f and the reference image f_R is equivalent to solving the following optimization problem,

$$\begin{aligned} \min_{\mathbf{p}} J(\mathbf{p}) &\equiv NTG(g, f_R) \\ &= \frac{\sum_l \sum_{\mathbf{x} \in \Omega(\mathbf{p})} |\nabla_l \{g(\mathbf{x}, \mathbf{p}) - f_R(\mathbf{x})\}|}{\sum_l \sum_{\mathbf{x} \in \Omega(\mathbf{p})} (|\nabla_l g(\mathbf{x}, \mathbf{p})| + |\nabla_l f_R(\mathbf{x})|)}, \end{aligned} \quad (10)$$

where $g(\mathbf{x}, \mathbf{p}) = f(u(\mathbf{x}, \mathbf{p}), v(\mathbf{x}, \mathbf{p}))$, $(u, v)^T = \mathbf{P}(x, y, 1)^T$, $l \in \{x, y\}$, and $\Omega(\mathbf{p})$ denotes the region of overlap.

B. Image Registration Framework

While the problem specified in (10) is a global optimization problem, local optimization methods can result in a transform that corresponds to a local minimum of the objective function, despite their simple and fast computation. On the other hand, global optimization methods are not advisable to be directly employed due to their unacceptable computation time when images are large. Hence we employ a hybrid strategy to balance computation efficiency and accuracy.

Images are first subsampled into K layers to form an upside down image pyramid whose lower layer is the subsampled

version of the upper one. Then global optimization is imposed on the bottom layer to furnish a good initial estimate for the successive optimization. The initial point is transferred to the upper layer and is refined by local optimization until the top layer is reached. In this work the differential evolution [10] and the well-known Newton's method are used as the global and local optimizers, respectively.

Differential evolution (DE) is a genuinely useful global optimization algorithm and has earned a reputation as a very effective and reliable global optimizer. Similar to most evolution algorithms, three steps are needed after choosing the initial population, namely, mutation, crossover, and selection. The initial vector population is chosen randomly and covers the constrained parameter space. New parameter vectors are generated by adding the weighted difference between two population vectors to a third one. The mutated vector's parameters are then mixed with the parameters of another predetermined vector (the target vector), yielding the trial vector. The target vector is replaced if the trial vector has a lower cost value. This has been done over all population vectors so that competitions take place in each generation [29].

In the global optimization, we use the DeMat library [10] to search the initial point at the bottom layer. A population of 30 individuals and 200 generations are usually adequate for deriving a good initial point. With a four-layered image pyramid, the searching space of $\mathbf{p} = (p_1, \dots, p_6)^T$ is constrained from $[0.95, -0.05, -10, -0.05, 0.95, -10]$ to $[1.05, 0.05, 10, 0.05, 1.05, 10]$ since no severe scaling and rotation take place in the captured multispectral images. Other DE parameters are set by default [10].

In the local optimization, the parameters at each layer is updated as

$$\mathbf{p}^{t+1} = \mathbf{p}^t - (\mathbf{H}_{\mathbf{p}}^J)^{-1} J_{\mathbf{p}}|_{\mathbf{p}=\mathbf{p}^t}, \quad (11)$$

where $\mathbf{H}_{\mathbf{p}}^J$ and $J_{\mathbf{p}}$ are, respectively, the Hessian matrix and gradient of the objective function $J(\mathbf{p})$ with respect to \mathbf{p} . The superscribe t represents the t th iteration. When the initial point locates near the optimal solution, which is guaranteed by global optimization at the bottom layer, 6 iterations are usually sufficient for the convergence of Newton's method [30].

V. IMPLEMENTATION DETAILS OF NEWTON'S METHOD

To minimize the objective function $J(\mathbf{p})$ in (10) we first denote $m(\mathbf{p})$ and $n(\mathbf{p})$ as the numerator and denominator of $J(\mathbf{p})$, respectively, as follows,

$$\begin{aligned} m(\mathbf{p}) &= \sum_{\mathbf{x} \in \Omega(\mathbf{p})} (|g_x(\mathbf{x}, \mathbf{p}) - f_{R,x}(\mathbf{x})| + |g_y(\mathbf{x}, \mathbf{p}) - f_{R,y}(\mathbf{x})|), \\ n(\mathbf{p}) &= \sum_{\mathbf{u} \in \Omega(\mathbf{p})} (|g_x(\mathbf{x}, \mathbf{p})| + |g_y(\mathbf{x}, \mathbf{p})| + |f_{R,x}(\mathbf{x})| + |f_{R,y}(\mathbf{x})|), \end{aligned}$$

where the subscripts x and y denote the partial derivatives along the x and y directions. With these notations, the objective function (10) is represented as $J(\mathbf{p}) = \frac{m(\mathbf{p})}{n(\mathbf{p})}$. By taking the derivative with respect to \mathbf{p} we have

$$J_{\mathbf{p}} = n^{-2}(nm_{\mathbf{p}} - mn_{\mathbf{p}}), \quad (12)$$

where $m_{\mathbf{p}}$ and $n_{\mathbf{p}}$ denote the gradient of m and n with respect to \mathbf{p} . Note that in (12) the parameter \mathbf{p} in parenthesis is omitted for notation simplification. The (i, j) th entry in Hessian matrix $\mathbf{H}_{\mathbf{p}}^J$ is

$$\begin{aligned} \mathbf{H}_{\mathbf{p}}^J(i, j) &= J_{p_i, p_j} \\ &= n^{-2}(nm_{p_i, p_j} - mn_{p_i, p_j} + 2mn^{-1}n_{p_i}n_{p_j} \\ &\quad - m_{p_i}n_{p_j} - n_{p_i}m_{p_j}). \end{aligned}$$

Hence the Hessian matrix $\mathbf{H}_{\mathbf{p}}^J$ becomes

$$\begin{aligned} \mathbf{H}_{\mathbf{p}}^J &= n^{-2}(n\mathbf{H}_{\mathbf{p}}^m - m\mathbf{H}_{\mathbf{p}}^n + 2mn^{-1}n_{\mathbf{p}}n_{\mathbf{p}}^{\top} \\ &\quad - m_{\mathbf{p}}n_{\mathbf{p}}^{\top} - n_{\mathbf{p}}m_{\mathbf{p}}^{\top}), \end{aligned} \quad (13)$$

where $\mathbf{H}_{\mathbf{p}}^m$ and $\mathbf{H}_{\mathbf{p}}^n$ denote the Hessian matrices of m and n with respect to \mathbf{p} .

To this end, in order to derive the gradient $J_{\mathbf{p}}$ and the Hessian matrix $\mathbf{H}_{\mathbf{p}}^J$ we resort to computing $m_{\mathbf{p}}$, $n_{\mathbf{p}}$, $\mathbf{H}_{\mathbf{p}}^m$, and $\mathbf{H}_{\mathbf{p}}^n$, which will be elaborated below.

A. Computing $m_{\mathbf{p}}$ and $n_{\mathbf{p}}$

Let $d(\mathbf{x}, \mathbf{p})$ denote the difference image, i.e.

$$d(\mathbf{x}, \mathbf{p}) = g(\mathbf{x}, \mathbf{p}) - f_R(\mathbf{x}),$$

thus it is easy to derive that

$$\begin{aligned} d_x &= g_x - f_{R,x}, \quad d_y = g_y - f_{R,y}, \\ d_{x,\mathbf{p}} &= g_{x,\mathbf{p}}, \quad d_{y,\mathbf{p}} = g_{y,\mathbf{p}}. \end{aligned} \quad (14)$$

Let $\rho_1(\cdot)$ denote the first order derivative of the absolute function $|\cdot|$ (see Subsection C). Since the overlap changes slightly in the two consecutive updates of \mathbf{p} , i.e. $\Omega_{\mathbf{p}}(\mathbf{p}^t) \approx \mathbf{0}$, by taking the derivatives of m and n with respect to \mathbf{p} we have

$$m_{\mathbf{p}} \approx \sum_{\mathbf{x} \in \Omega(\mathbf{p})} (\rho_1(d_x)g_{x,\mathbf{p}} + \rho_1(d_y)g_{y,\mathbf{p}}) \quad (15)$$

and

$$n_{\mathbf{p}} \approx \sum_{\mathbf{x} \in \Omega(\mathbf{p})} (\rho_1(g_x)g_{x,\mathbf{p}} + \rho_1(g_y)g_{y,\mathbf{p}}). \quad (16)$$

We narrow our focus down to the computation of $g_{x,\mathbf{p}}$ and $g_{y,\mathbf{p}}$ since all the other terms are known given the current transform parameter \mathbf{p}^t . By implementing $g(\mathbf{x}, \mathbf{p}) = f(u(\mathbf{x}, \mathbf{p}), v(\mathbf{x}, \mathbf{p}))$, we have

$$\begin{pmatrix} g_x \\ g_y \end{pmatrix} = \begin{pmatrix} u_x & v_x \\ u_y & v_y \end{pmatrix} \begin{pmatrix} f_u \\ f_v \end{pmatrix}.$$

Consequently it can be derived that

$$\begin{aligned} \begin{pmatrix} g_{x,p_i} \\ g_{y,p_i} \end{pmatrix} &= \begin{pmatrix} u_x & v_x \\ u_y & v_y \end{pmatrix} \begin{pmatrix} f_{uu} & f_{uv} \\ f_{vu} & f_{vv} \end{pmatrix} \begin{pmatrix} u_{p_i} \\ v_{p_i} \end{pmatrix} \\ &\quad + \begin{pmatrix} u_{x,p_i} & v_{x,p_i} \\ u_{y,p_i} & v_{y,p_i} \end{pmatrix} \begin{pmatrix} f_u \\ f_v \end{pmatrix}, \end{aligned} \quad (17)$$

where $p_i \in \{p_1, p_2, \dots, p_6\}$. In (17), the gradient of the objective function is finally decomposed to the transformed first order and second order gradients of the floating image, and the gradients of the transform model.

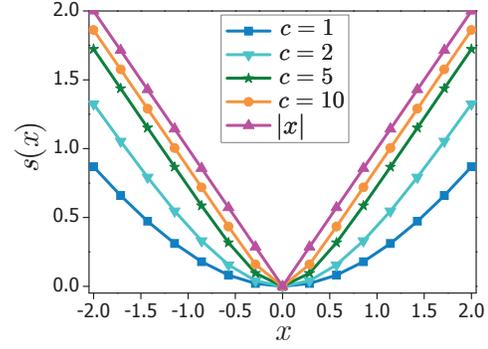


Fig. 4. Function $s(x)$ with respect to different values of c . The value of x is in the range $[-2, 2]$. The absolute function $|x|$ can be well approximated by $s(x)$ when $c = 10$.

According to the affine transform model (9), we have

$$\begin{pmatrix} u_x & v_x \\ u_y & v_y \end{pmatrix} = \begin{pmatrix} p_1 & p_4 \\ p_2 & p_5 \end{pmatrix},$$

$$\begin{aligned} u_{\mathbf{p}} &= (x, y, 1, 0, 0, 0)^{\top}, \quad v_{\mathbf{p}} = (0, 0, 0, x, y, 1)^{\top}, \\ u_{x,\mathbf{p}} &= (1, 0, 0, 0, 0, 0)^{\top}, \quad u_{y,\mathbf{p}} = (0, 1, 0, 0, 0, 0)^{\top}, \\ v_{x,\mathbf{p}} &= (0, 0, 0, 1, 0, 0)^{\top}, \quad v_{y,\mathbf{p}} = (0, 0, 0, 0, 1, 0)^{\top}. \end{aligned}$$

To this end, the gradient of the objective function can be computed by inserting (15), (16), and (17) into (12).

B. Computing $\mathbf{H}_{\mathbf{p}}^m$ and $\mathbf{H}_{\mathbf{p}}^n$

Let $\rho_2(\cdot)$ be the second order derivative of the absolute function $|\cdot|$. By using (14) and (15), the (i, j) th entry in $\mathbf{H}_{\mathbf{p}}^m$ can be computed as

$$\begin{aligned} \mathbf{H}_{\mathbf{p}}^m(i, j) &= m_{p_i, p_j} \\ &\approx \sum_{\mathbf{x} \in \Omega(\mathbf{p})} (\rho_1(d_x)g_{x,p_i,p_j} + \rho_2(d_x)g_{x,p_i}g_{x,p_j} \\ &\quad + \rho_1(d_y)g_{y,p_i,p_j} + \rho_2(d_y)g_{y,p_i}g_{y,p_j}). \end{aligned}$$

Hence the Hessian matrix $\mathbf{H}_{\mathbf{p}}^m$ becomes

$$\begin{aligned} \mathbf{H}_{\mathbf{p}}^m &\approx \sum_{\mathbf{x} \in \Omega(\mathbf{p})} (\rho_1(d_x)\mathbf{H}_{\mathbf{p}}^{g_x} + \rho_2(d_x)g_{x,\mathbf{p}}g_{x,\mathbf{p}}^{\top} \\ &\quad + \rho_1(d_y)\mathbf{H}_{\mathbf{p}}^{g_y} + \rho_2(d_y)g_{y,\mathbf{p}}g_{y,\mathbf{p}}^{\top}), \end{aligned} \quad (18)$$

where $\mathbf{H}_{\mathbf{p}}^{g_x}$ and $\mathbf{H}_{\mathbf{p}}^{g_y}$ are the Hessian matrices of g_x and g_y . Similarly, we have

$$\begin{aligned} \mathbf{H}_{\mathbf{p}}^n(i, j) &= n_{p_i, p_j} \\ &\approx \sum_{\mathbf{x} \in \Omega(\mathbf{p})} (\rho_1(g_x)g_{x,p_i,p_j} + \rho_2(g_x)g_{x,p_i}g_{x,p_j} \\ &\quad + \rho_1(g_y)g_{y,p_i,p_j} + \rho_2(g_y)g_{y,p_i}g_{y,p_j}), \end{aligned}$$

and thus

$$\begin{aligned} \mathbf{H}_{\mathbf{p}}^n &\approx \sum_{\mathbf{x} \in \Omega(\mathbf{p})} (\rho_1(g_x)\mathbf{H}_{\mathbf{p}}^{g_x} + \rho_2(g_x)g_{x,\mathbf{p}}g_{x,\mathbf{p}}^{\top} \\ &\quad + \rho_1(g_y)\mathbf{H}_{\mathbf{p}}^{g_y} + \rho_2(g_y)g_{y,\mathbf{p}}g_{y,\mathbf{p}}^{\top}). \end{aligned} \quad (19)$$

According to the affine model given in (17), we can derive that

$$\begin{aligned} \begin{pmatrix} g_{x,p_i,p_j} \\ g_{y,p_i,p_j} \end{pmatrix} &= \begin{pmatrix} u_{x,p_j}, v_{x,p_j} \\ u_{y,p_j}, v_{y,p_j} \end{pmatrix} \begin{pmatrix} f_{uu}, f_{uv} \\ f_{vu}, f_{vv} \end{pmatrix} \begin{pmatrix} u_{p_i} \\ v_{p_i} \end{pmatrix} \\ &+ \begin{pmatrix} u_x, v_x \\ u_y, v_y \end{pmatrix} \begin{pmatrix} f_{uu,p_j}, f_{uv,p_j} \\ f_{vu,p_j}, f_{vv,p_j} \end{pmatrix} \begin{pmatrix} u_{p_i} \\ v_{p_i} \end{pmatrix} \\ &+ \begin{pmatrix} u_{x,p_i}, v_{x,p_i} \\ u_{y,p_i}, v_{y,p_i} \end{pmatrix} \begin{pmatrix} f_{u,p_j} \\ f_{v,p_j} \end{pmatrix}, \end{aligned} \quad (20)$$

where

$$\begin{pmatrix} f_{u,p_j} \\ f_{v,p_j} \end{pmatrix} = \begin{pmatrix} f_{uu}, f_{uv} \\ f_{vu}, f_{vv} \end{pmatrix} \begin{pmatrix} u_{p_j} \\ v_{p_j} \end{pmatrix},$$

$$\begin{pmatrix} f_{uu,p_j} \\ f_{vu,p_j} \end{pmatrix} = \begin{pmatrix} f_{uuu}, f_{uuv} \\ f_{vuu}, f_{vvv} \end{pmatrix} \begin{pmatrix} u_{p_j} \\ v_{p_j} \end{pmatrix},$$

and

$$\begin{pmatrix} f_{uv,p_j} \\ f_{vv,p_j} \end{pmatrix} = \begin{pmatrix} f_{uvu}, f_{uvv} \\ f_{vuv}, f_{vvv} \end{pmatrix} \begin{pmatrix} u_{p_j} \\ v_{p_j} \end{pmatrix}.$$

It is clear that the Hessian matrix of the objective function can finally decomposed to the transformed gradients of the floating image up to third orders. To this end, the Hessian matrix can be computed by inserting (18), (19), and (20) into (13).

C. Form of ρ_1 and ρ_2

Since the absolute function $|x|$ is not differentiable at $x = 0$, we use the function

$$s(x) = x + \frac{2}{c} \log\left(\frac{1 + e^{-cx}}{2}\right) \quad (21)$$

to approximate it. Figure 4 illustrates the function $s(x)$ with respect to different values of c . The value of x , which is the gradient of difference image, is in the range $[-2, 2]$. As illustrated, $s(x)$ gradually approximates $|x|$ as c increases. In this work we set $c = 10$. The first and second derivatives of $s(x)$ are, respectively,

$$\rho_1(x) = \frac{1 - e^{-cx}}{1 + e^{-cx}} \quad (22)$$

and

$$\rho_2(x) = \frac{2c}{e^{-cx} + e^{cx} + 2}. \quad (23)$$

VI. EXPERIMENTS

Experiments were conducted on 10 multispectral images of real-scenes which are shown in Fig. 5. These images were captured using the imaging system illustrated in Fig. 1. The performances of MI [11], RC [9], CR [17], RSNCC [20] and the proposed NTG were compared at two aspects. First, the robustness of these measures was investigated on multispectral images with local intensity variation. Second, the registration accuracy was quantitatively evaluated on real multispectral images.



Fig. 5. The multispectral sample images ($S1, S2, \dots, S10$) and pattern image (displayed in RGB) used in the experiments. The pattern image is employed in the quantitative evaluation of registration accuracy.

A. Robustness to Intensity Variation

Figure 6 shows the effect of local intensity variation on different measures. In the experiment, channel 1 (400 nm) was used as the reference and channels 5 (480 nm), 9 (560 nm), 12 (620 nm), 16 (700 nm) were used as the floating ones. The synthetic affine transforms include horizontal and vertical translations, as well as rotation and scaling. The cost function maps of negative MI, negative CR, RSNCC and NTG are shown in Fig. 6 from column 3 to 12. The top half of the figure shows the cost maps of different measures in case of rapid local intensity variation (see the spots on cat's face and ears). The bottom half shows the cost maps in case of additional slow local intensity variation. The slow variation was simulated by dark clouds whose intensity gradually decreases. The centers of the maps correspond to the ground truth parameters, i.e. (0, 0) for the (x, y) translation and (0, 1) for the rotation and scaling. In these cost maps, lower cost values are encoded in blue and higher cost values are encoded in red. The locations of the global minimal costs correspond to the estimated transform parameters.

It is observed that MI, RC, CR, and RSNCC are more or less influenced by intensity variation. The reason is that the assumptions like functional mapping, unimodality, and local linearity are violated when local intensity varies. On the contrary, NTG exhibits an excellent performance. The cost function of NTG is scarcely influenced by the varying local intensity. Its robustness to intensity variation takes advantages from the combination of the gradient operation and L_1 norm sparseness measure.

B. Registration Accuracy

In [3] a pattern board was used to correct the geometric distortion in a multispectral imaging system. Instead, in this work, we employed the pattern board shown in Fig. 5 to evaluate the registration measures. In the imaging procedure,

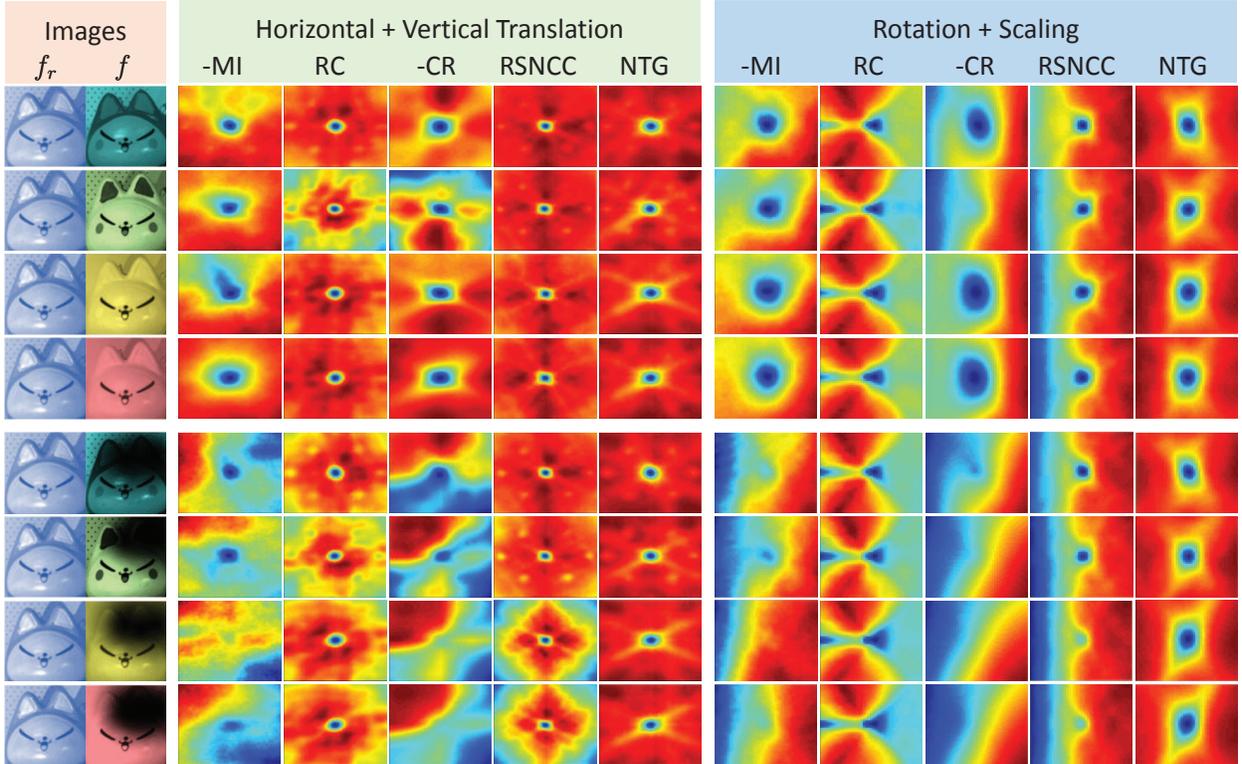


Fig. 6. Effect of local intensity variation on different measures. Image of channel 1 (400 nm) is the reference image and images of channel 5 (480 nm), 9 (560 nm), 12 (620 nm), 16 (700 nm) are the floating images. Cost maps of negative MI, RC, negative CR, RSNCC, and NTG with respect to horizontal and vertical translations (in range $[-20, 20]$ pixels) are shown from column 3 to 7. Cost maps with respect to rotation (in range $[-10^\circ, 10^\circ]$) and scaling (in range $[0.8, 1.2]$) are shown from column 8 to 12. The centers of the maps correspond to the ground truth parameters, i.e. $(0, 0)$ for translations and $(0, 1)$ for rotation and scaling. In the color maps, lower values are in blue and higher values are in red.

we kept the imaging conditions (e.g., imaging distance and focal length) fixed, but changed the objects to be imaged. The 10 different scenes are illustrated in Fig. 5. Note that, as the imaging conditions were fixed, the transform parameters of any individual channel with respect to the reference channel (No. 9) should be identical in all the captured scenes. In the experiment, we computed the rigid transforms from the sample images ($S1, S2, \dots, S10$), and the obtained transforms were then applied on the pattern image. For each channel of the pattern image, the displacements between the circle centers on the registered channel image and those on the reference channel image were computed. Finally, the root mean square error (RMSE) of all these displacements was computed as the registration error of each channel.

Table I lists the median RMSE (in pixels) of the 10 samples produced by different registration measures. It is observed that NTG yields subpixel accuracy and performs comparative to or better than its competitors.

As an example, Fig. 7 shows the distributions of registration errors of the samples $S1$ and $S2$ in different channels. It is observed that the registration errors by the proposed method are generally smaller than those by other methods. Figure 8 further shows the close-up views of image registration results of the samples $S1$ and $S2$. Here channel 2 was chosen to be the floating image and channel 9 was set as the reference image. The two channel images before and after registration are displayed in an overlapping manner with pseudocolors. It

is observed that, for the sample $S1$, small misalignment exists in the registration results by MI, RC, CR, and RSNCC. Rather severe misalignments take place in the registration results of sample $S2$. In comparison, the floating image is robustly and accurately registered by the proposed NTG measure.

VII. APPLICATIONS

It is mentioned in Section I that the acquired multispectral images suffer from misalignment and out-of-focus blur. We discuss the application of the proposed method in two circumstances. In the first circumstance (Subsection A), the multispectral image has a focusing device [5] and hence the acquired multispectral images are well focused. Thus the proposed registration method can be directly applied on the sharp multispectral images. In the second circumstance (Subsection B), as no focusing device is employed, we must deal with both misalignment and out-of-blur. We will further show in Subsection C that, in addition to multispectral image, the proposed method can also be applied on multimodal and other images.

A. Multispectral Image Registration

Figure 9 shows some registration results of well-focused multispectral images of real scenes using the proposed method. The chromatic aberrations caused by image misalignment are obvious in the original acquired multispectral images. By

TABLE I
REGISTRATION ERRORS PRODUCED BY DIFFERENT MEASURES (IN PIXELS). THE SMALLEST ERRORS ARE IN BOLD.

Channel No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
MI	1.15	0.68	0.66	0.61	0.61	0.51	0.58	0.56	-	0.68	0.58	0.66	0.61	0.85	1.47	1.27
RC	2.12	1.34	1.24	1.26	1.17	0.75	0.52	0.58	-	0.99	0.81	0.94	1.23	0.81	1.05	1.02
CR	3.50	1.81	1.95	1.47	1.22	0.82	0.62	0.48	-	0.79	0.60	0.98	0.82	1.20	2.12	2.98
RSNCC	0.38	0.90	0.92	1.28	1.49	1.65	1.89	1.89	-	1.74	1.20	0.85	1.30	0.90	1.56	0.43
Proposed NTG	0.80	0.37	0.43	0.32	0.26	0.35	0.23	0.24	-	0.36	0.32	0.33	0.34	0.27	0.95	0.45

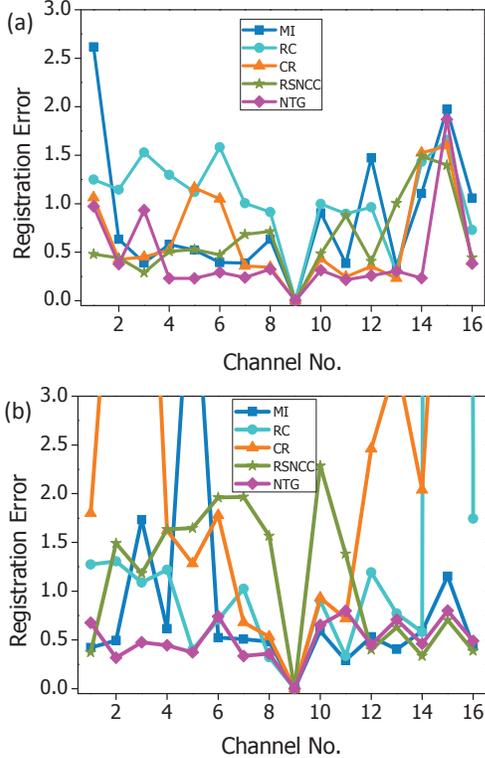


Fig. 7. Registration errors (in pixels) produced by different measures. (a) Sample *S1*. (b) Sample *S2*.

applying image registration, the multispectral images are well aligned and accordingly the chromatic aberrations disappear.

B. Multispectral Image Restoration

In multispectral restoration, our aim is to remove both misalignment and out-of-focus blur. The process is as follows. First, each channel images were initially deblurred using a single-image-based blind deconvolution method [31]. Second, the proposed method was applied to register the deblurred individual channel images with respect to the reference channel image. Third, the state-of-the-art interchannel-correlation-based deblurring method [4] was applied on the registered blurry images, and the high-quality deblurred multispectral image was finally obtained. Figure 10 shows the multispectral image restoration results using the above process. For illustration, the two blurriest channels (Nos. 1 and 16) and the reference channel (No. 9) are displayed in composite views. It

is observed that quality of the multispectral image (displayed in RGB) are much improved after image restoration. From the composite views in the second and third rows, image registration and deblurring have been successfully accomplished in the 1th and 9th channel images.

C. Other Registration Tasks

The proposed method is also applicable to a variety of registration tasks. We present some typical applications in image processing and computer vision. Figure 11 shows the registration results on medical images, flash/no-flash images [32], RGB/disparity images [33]. In these image pairs, both rotation and translation have been simulated in the floating images. It is observed that, despite the rapidly/slowly intensity variations, the proposed method produces quite satisfactory registration results.

VIII. CONCLUSIONS

This paper proposed a new measure, namely normalized total gradient (NTG), for multispectral image registration. The employment of NTG was based on the observation that the gradient of difference image becomes sparser when the two channel images are well aligned. An registration framework, which consists of image pyramid and global/local optimization, was introduced for general affine transform. Experimental results showed that the proposed NTG measure is well suited to multispectral image registration and outperforms the conventional measures. It was validated that the proposed registration method can be applied in multispectral image restoration (registration and deblurring) and other relevant registration tasks.

A limitation of this work is that the proposed method is only implemented for rigid (affine) image registration, which is only applicable to static and rigid objects. However, nonrigid registration is required when objects are deformable or moving in the image acquisition process. Hence, in the future we plan to incorporate the NTG measure into nonrigid registration methods to further extend its applications.

REFERENCES

- [1] J. Y. Hardeberg, F. Schmitt, and H. Brettel, "Multispectral color image capture using a liquid crystal tunable filter," *Optical Engineering*, vol. 41, no. 10, pp. 2532–2548, 2002.
- [2] J. Katrašnik, F. Pernuš, and B. Likar, "Radiometric calibration and noise estimation of acousto-optic tunable filter hyperspectral imaging systems," *Applied Optics*, vol. 52, no. 15, pp. 3526–3537, 2013.

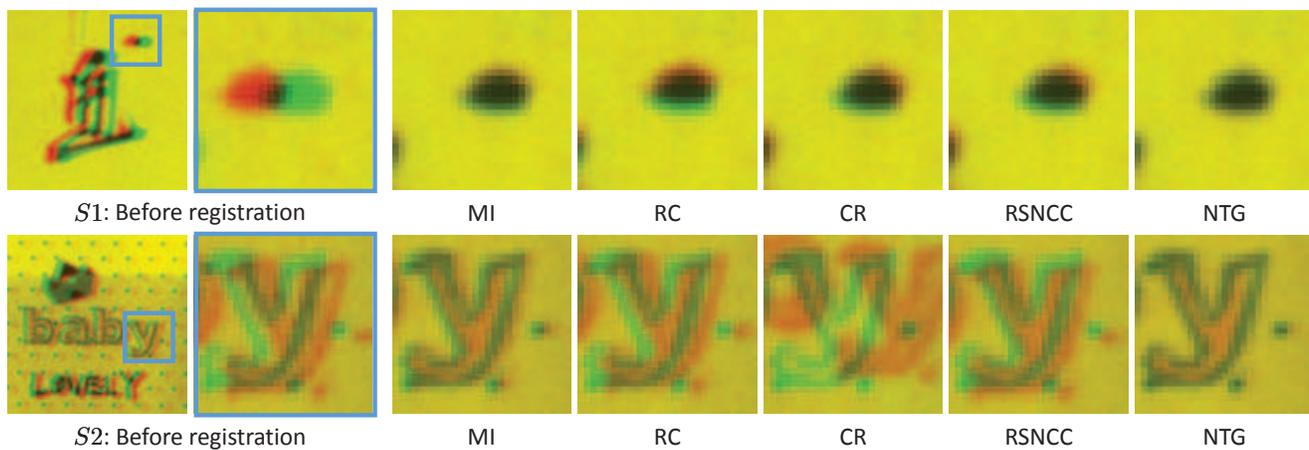


Fig. 8. Close-up views of multispectral image registration results of samples $S1$ and $S2$ using different measures. The two channel images, which are in pseudocolor, are displayed in an overlapping manner.

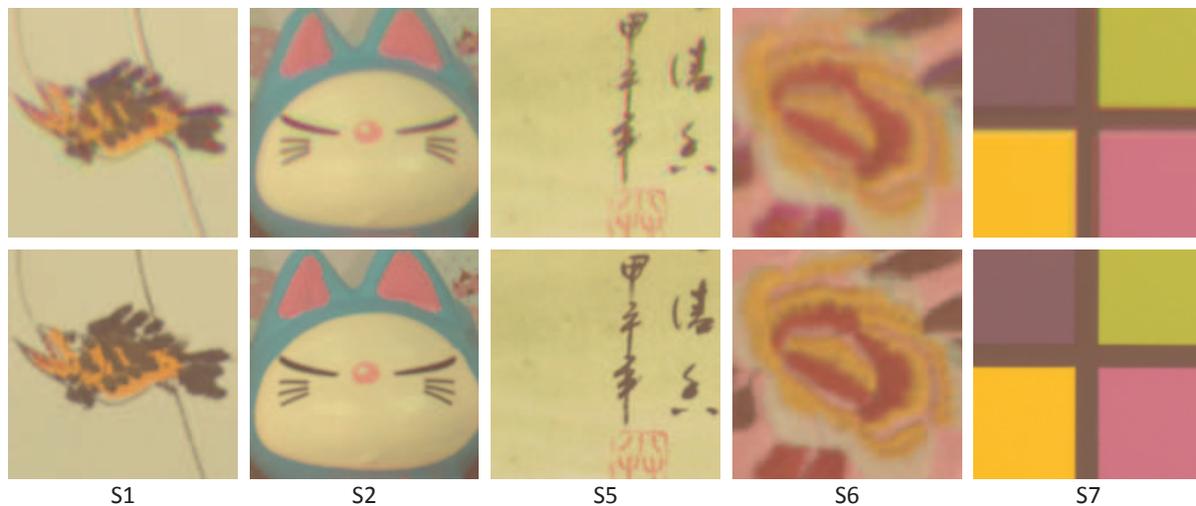


Fig. 9. Examples of multispectral image registration using NTG. The first and second rows show multispectral images before and after registration. All multispectral images are displayed in RGB.

- [3] J. Brauers and T. Aach, "Geometric calibration of lens and filter distortions for multispectral filter-wheel cameras," *IEEE Transactions on Image Processing*, vol. 20, no. 2, pp. 496–505, 2011.
- [4] S.-J. Chen and H.-L. Shen, "Multispectral image out-of-focus deblurring using interchannel correlation," *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 4433–4445, 2015.
- [5] H.-L. Shen, Z.-H. Zheng, W. Wang, X. Du, S.-J. Shao, and J. H. Xin, "Autofocus for multispectral camera using focus symmetry," *Applied Optics*, vol. 51, no. 14, pp. 2616–2623, 2012.
- [6] J. A. Maintz and M. A. Viergever, "A survey of medical image registration," *Medical Image Analysis*, vol. 2, no. 1, pp. 1–36, 1998.
- [7] B. Zitova and J. Flusser, "Image registration methods: a survey," *Image and Vision Computing*, vol. 21, no. 11, pp. 977–1000, 2003.
- [8] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.
- [9] A. Myronenko and X. Song, "Intensity-based image registration by minimizing residual complexity," *IEEE Transactions on Medical Imaging*, vol. 29, no. 11, pp. 1882–1891, 2010.
- [10] K. Price, R. M. Storn, and J. A. Lampinen, *Differential Evolution: A Practical Approach to Global Optimization*. Springer Science and Business Media, 2006.
- [11] 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, no. 2, pp. 187–198, 1997.
- [12] J. P. Pluim, J. A. Maintz, and M. A. Viergever, "Mutual-information-based registration of medical images: a survey," *IEEE Transactions on Medical Imaging*, vol. 22, no. 8, pp. 986–1004, 2003.
- [13] C. Studholme, D. L. Hill, and D. J. Hawkes, "An overlap invariant entropy measure of 3D medical image alignment," *Pattern Recognition*, vol. 32, no. 1, pp. 71–86, 1999.
- [14] C. Studholme, C. Drapaca, B. Iordanova, and V. Cardenas, "Deformation-based mapping of volume change from serial brain mri in the presence of local tissue contrast change," *IEEE Transactions on Medical Imaging*, vol. 25, no. 5, pp. 626–639, 2006.
- [15] J. Woo, M. Stone, and J. L. Prince, "Multimodal registration via mutual information incorporating geometric and spatial context," *IEEE Transactions on Image Processing*, vol. 24, no. 2, pp. 757–769, 2014.
- [16] H. Luan, F. Qi, Z. Xue, L. Chen, and D. Shen, "Multimodality image registration by maximization of quantitative–qualitative measure of mutual information," *Pattern Recognition*, vol. 41, no. 1, pp. 285–298, 2008.
- [17] A. Roche, G. Malandain, X. Pennec, and N. Ayache, "The correlation ratio as a new similarity measure for multimodal image registration," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI'98*. Springer, 1998, pp. 1115–1124.
- [18] Y. Hel-Or, H. Hel-Or, and E. David, "Matching by tone mapping: Photometric invariant template matching," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 2, pp. 317–330, 2014.

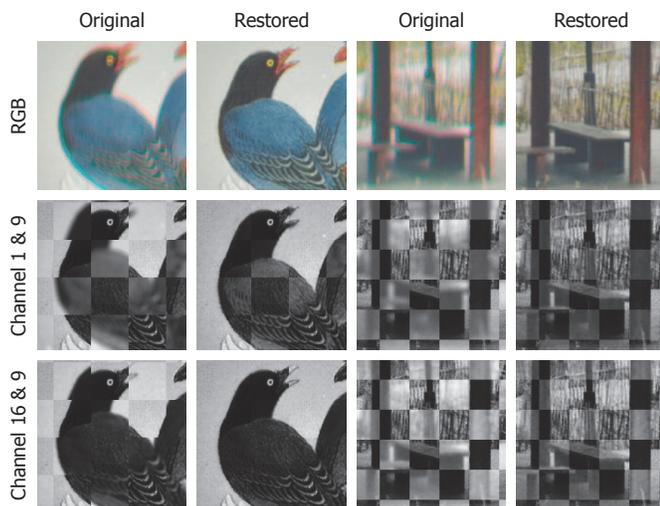


Fig. 10. Multispectral image restoration (registration and deblurring) results. The quality of the multispectral images (displayed in RGB) is much improved after image restoration. The two blurriest channel images (Nos. 1 and 16) are displayed in composite view with respect to the reference channel image (No. 9).

subpixel-accurate ground truth,” in *36th German Conference on Pattern Recognition (GCPR)*. Springer, 2014, pp. 31–42.

- [19] M. Irani and P. Anandan, “Robust multi-sensor image alignment,” in *Sixth International Conference on Computer Vision*. IEEE, 1998, pp. 959–966.
- [20] X. Shen, L. Xu, Q. Zhang, and J. Jia, “Multi-modal and multi-spectral registration for natural images,” in *European Conference on Computer Vision*. Springer, 2014, pp. 309–324.
- [21] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman, “Removing camera shake from a single photograph,” *ACM Transactions on Graphics*, vol. 25, no. 3, pp. 787–794, 2006.
- [22] J. Portilla, V. Strela, M. J. Wainwright, and E. P. Simoncelli, “Image denoising using scale mixtures of Gaussians in the wavelet domain,” *IEEE Transactions on Image Processing*, vol. 12, no. 11, pp. 1338–1351, 2003.
- [23] A. Levin, A. Zomet, and Y. Weiss, “Learning how to inpaint from global image statistics,” in *International Conference on Computer Vision*. IEEE, 2003, pp. 305–312.
- [24] A. Chakrabarti and T. Zickler, “Statistics of real-world hyperspectral images,” in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, 2011, pp. 193–200.
- [25] E. J. Candes and T. Tao, “Near-optimal signal recovery from random projections: Universal encoding strategies?” *IEEE Transactions on Information Theory*, vol. 52, no. 12, pp. 5406–5425, 2006.
- [26] L. Rudin, S. Osher, and E. Fatemi, “Nonlinear total variation based noise removal algorithm,” *Physica D*, vol. 60, no. 1–4, pp. 259–268, 1992.
- [27] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. Hill, M. O. Leach, and D. J. Hawkes, “Nonrigid registration using free-form deformations: application to breast mr images,” *IEEE Transactions on Medical Imaging*, vol. 18, no. 8, pp. 712–721, 1999.
- [28] J. Klein and T. Aach, “Multispectral filter wheel cameras: modeling aberrations for filters in front of lens,” in *IS&T/SPIE Electronic Imaging*. International Society for Optics and Photonics, 2012, pp. 82 990R–82 990R.
- [29] R. Storn and K. Price, “Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces,” *Journal of Global Optimization*, vol. 11, no. 4, pp. 341–359, 1997.
- [30] S. J. Wright and J. Nocedal, *Numerical Optimization*. Springer New York, 1999, vol. 2.
- [31] J. P. Oliveira, M. Figueiredo, and J. M. Bioucas-Dias, “Parametric blur estimation for blind restoration of natural images: Linear motion and out-of-focus,” *IEEE Transactions on Image Processing*, vol. 23, no. 1, pp. 466–477, 2014.
- [32] G. Petschnigg, R. Szeliski, M. Agrawala, M. Cohen, H. Hoppe, and K. Toyama, “Digital photography with flash and no-flash image pairs,” *ACM transactions on graphics (TOG)*, vol. 23, no. 3, pp. 664–672, 2004.
- [33] D. Scharstein, H. Hirschmüller, Y. Kitajima, G. Krathwohl, N. Nešić, X. Wang, and P. Westling, “High-resolution stereo datasets with

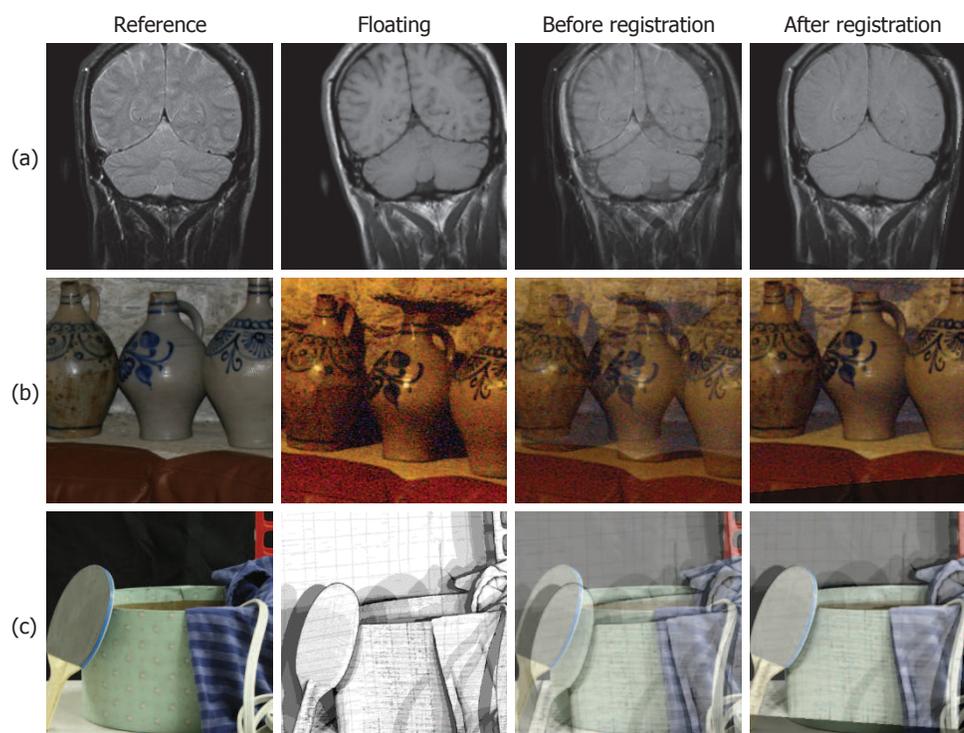


Fig. 11. Other registration tasks using NTG. (a) Medical images. (b) Flash/no-flash images. (c) RGB and disparity images.