Lehrstuhl für Bildverarbeitung
Institute of Imaging & Computer Vision

Multispectral Filter-Wheel Cameras: Geometric Distortion Model and Compensation Algorithms

Johannes Brauers and Nils Schulte and Til Aach
Institute of Imaging and Computer Vision
RWTH Aachen University, 52056 Aachen, Germany
tel: +49 241 80 27860, fax: +49 241 80 22200
web: www.lfb.rwth-aachen.de

in: IEEE Transactions on Image Processing. See also BIBTEX entry below.

BIBTEX:

```
@article{Brauers2008d,
  author     = {Johannes Brauers and Nils Schulte and Til Aach},
  title      = {Multispectral Filter-Wheel Cameras:
                Geometric Distortion Model and Compensation Algorithms},
  journal    = {IEEE Transactions on Image Processing},
  publisher  = {IEEE},
  year       = {2008},
  month      = {Dec},
  volume     = {17},
  pages      = {2368--2380},
  number     = {12},
}
```


document created on: November 28, 2008
created from file: main.tex
cover page automatically created with CoverPage.sty
(available at your favourite CTAN mirror)

RWTH AACHEN
RHEINISCH-WESTFÄLISCHE TECHNISCHE HOCHSCHULE AACHEN

Multispectral Filter-Wheel Cameras: Geometric Distortion Model and Compensation Algorithms

Johannes Brauers*, Nils Schulte and Til Aach, *Senior Member, IEEE*

***Abstract*—Multispectral image acquisition considerably improves color accuracy in comparison to RGB technology. A common multispectral camera design concept features a filter-wheel consisting of six or more optical bandpass filters. By shifting the filters sequentially into the optical path, the electromagnetic spectrum is acquired through the channels, thus making an approximate reconstruction of the spectrum feasible. However, since the optical filters exhibit different thicknesses, refraction indices and may not be aligned in a perfectly coplanar manner, geometric distortions occur in each spectral channel: The reconstructed RGB images thus show rainbow-like color fringes. To compensate for these, we analyze the optical path and derive a mathematical model of the distortions. Based on this model we present two different algorithms for compensation and show that the color fringes vanish completely after application of our algorithms. We also evaluate our compensation algorithms in terms of accuracy and execution time.**

***Index Terms*—multispectral imaging model, geometric distortion, multispectral image processing, registration, chromatic aberration, affine transform**

I. Introduction

CONSUMER cameras in mobile phones or for digital photography as well as many professional cameras use mass-produced RGB sensors for color image acquisition. These outperform multispectral cameras in terms of low costs, robustness and ease of use but not in terms of color fidelity [1]; RGB cameras provide a "preferred reproduction [...], rather than accurate one" [2], [3], which is perfectly acceptable for most users.

However, RGB cameras produce a systematic color error caused by violating the Luther rule [4], which states that for faithful color reproduction, the camera's spectral sensitivity curves have to be a linear combination of the CIE observer's ones. Another drawback of the most frequently used 1-chip RGB camera type is its need to interpolate missing color information: The sensor of this camera type is covered with a color filter array (CFA) which maps three spectral sensitivities to the spatial domain, resulting in a downsampled image of each color channel. An interpolation algorithm called demosaicking [5], [6], [7] is required to estimate the missing color information, but typically does not achieve a perfect reconstruction of the original. Remaining interpolation artifacts reduce the image quality and fidelity. A birefringent blur filter [8] brought into the optical path may reduce aliasing by lowpass filtering prior to the acquisition, but considerable systematic limitations still remain. Another disadvantage of the CFA is the induced shift variance [9], which causes the reconstructed image to be dependent on spatial shifts. In contrast to this, multispectral cameras mostly fulfill the Luther rule and are not afflicted by the problems caused by a color filter array. This makes them suitable for high quality color inspection tasks in industry, acquisition of art paintings [10] or imaging applications in medicine requiring high color fidelity [11], [12].

One particular system used for multispectral image acquisition divides the incoming visible electromagnetic spectrum into several spectral channels by application of spectral bandpass filters [13], [14], [15], [16]. Typically, these filters are mounted in a computer-controlled filter wheel located between lens and gray level sensor. Acquisition of the entire visible spectrum is done by sequentially placing the filters into the optical path and grabbing an image for each filter. Our multispectral camera (see Fig. 1) consists of seven spectral bandpass filters which sample the spectral interval from 400nm to 700nm in discrete steps of 50nm. This makes it possible to differentiate between metamere colors, i.e., colors with different spectra but whose color impressions are the same for a human viewer or an RGB camera. A multispectral camera thus extends the ability to differentiate between different light spectra compared to an RGB camera in a similar way as RGB cameras do in comparison to a gray scale camera, where the latter one cannot differentiate colors which have the same intensity but different hue and saturation values. Another advantage is the ability to correctly generate reproductions from *one* given acquisition for *different* light sources. Although images acquired with three color channels can be adapted to other white points, a realistic reproduction fails due to the insufficient knowledge of the true underlying spectrum.

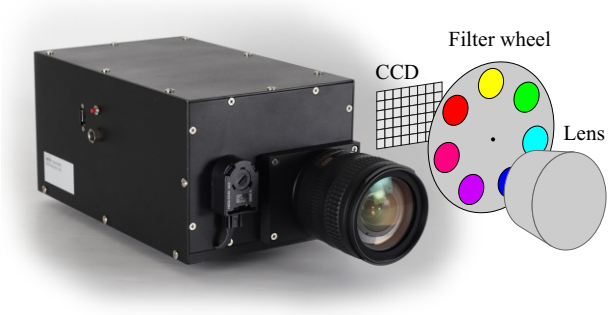


Fig. 1. Left: Our multispectral camera, using 7 optical bandpass filters. Right: Sketch of its internal configuration.

Besides the above filter wheel approach [16], [15], other

Manuscript received November 28, 2008.

The authors are with the Institute of Imaging&Computer Vision, RWTH Aachen University, D-52056 Aachen, Germany (corresponding author to provide phone (J. Brauers): 0049-241-8027866; fax: 0049-241-22200; e-mail: Johannes.Brauers@lfb.rwth-aachen.de).

techniques for multispectral image acquisition exist: In [17], a medium format 1-chip RGB camera in conjunction with a customized spectral filter dividing each RGB color component into two spectral halves enables the acquisition of six spectral channels. Demosaicking is circumvented by applying a piezo-driven pixel shift to the sensor, which ensures that each color component is sampled by the complete sensor matrix. In [18] and [19] this RGB-based concept is improved further by using two 3CCD HDTV cameras with a similar customized spectral filter as described above. The camera pair is able to record multispectral videos and does not exhibit moving mechanical components. Interestingly, a related principle called the "6-Primary Display" [20], [21], [22] exists, which employs two RGB projectors with additional appropriate spectral filters to project images with higher color fidelity. To enable acquisition with a higher dynamic range, a neutral density filter is also attached to one 3CCD camera [19]. Another possibility to avoid mechanical parts is the application of a liquid crystal tunable filter (LCTF) [23], whose filter characteristics can be adjusted by applying different voltages. A line scanning approach splitting the spectrum using a prism is presented in [24]: The spectrum is mapped along one coordinate of the sensor while one spatial coordinate of the scene is mapped along the other coordinate of the sensor. The remaining spatial coordinate is acquired sequentially over time. Despite the advantages of other acquisition techniques, the filter wheel approach is a widely-used, low cost, almost temperature drift independent [1] and mechanically manageable solution to multispectral image acquisition of static scenes.

A drawback of all multispectral camera types using more than one optical filter is the alteration of the optical path by the filters. Since it is practically unfeasible to align these filters in a perfectly coplanar manner, the normal of each filter is tilted by an individual angle with respect to the optical axis, as shown in Fig. 2. As we will show quantitatively later on, even small deviations between the optical axis and the filter normals can cause noticeable distortions in the image. Additionally, the filter thicknesses and refraction indices differ. These differences cause two different types of optical aberrations: The *transversal* aberrations affect the displacement of rays within the sensor plane and result in geometrically distorted color components which, when all spectral channels are combined into one color image, cause color fringes. Our objective in this contribution is to derive a mathematical model explaining these distortions and, based on this model, to develop algorithms for their compensation.

On the other hand, *longitudinal* aberrations as mentioned in [25] describe the divergence of rays through an optical element. Since the rays emerging from a single object point do not coincide on one focus point in the sensor plane, but on a point beyond or before the sensor layer, the color channels are blurred. In this paper, we focus on the transversal aberrations, i.e., geometric distortions.

The geometric distortions between multispectral color channels were already observed by other researchers: Helling [16] uses a local region-based registration algorithm to compensate for the geometric inter-channel distortions. However, without an underlying mathematical model, only regions with sufficient structural information for registration can be corrected. Also, the correlation-based similarity measure used in [16] is not ideally suited because the values of different spectral channels are generally related non-linearly. Another registration algorithm using a restricted affine model fitted by global optimization of the model's parameters is presented in [26]. Because of the global optimization, this approach does not provide the ability to validate the underlying mathematical model in a region-specific manner. In [27], a phase correlation based algorithm for subpixel registration is used to perform registration on, among others, multispectral images. Since the phase-based similarity measure only supports translation, the more complex distortions caused by the optical filters may not be compensated satisfactorily. A registration algorithm for the estimation of affine transformations in the frequency domain is presented in [28]. The algorithm is, however, not suited for our application since it requires a manual selection of features, and does not consider the contrast inversions between the spectral channels. Working on satellite images, Kern et al. [29] improve the reliability of registration by using an extended mutual information measure and utilize an affine model. Still, the derivation in this paper is not based on a physical model, and no further details about the search strategy are described. Since the focus in [10] is on other topics, no details on the registration algorithm are given. In [30], registration is performed via the local adaptation of a filter mask. Since the similarity criterion between the registered image and the reference image is the quadratic error, typical brightness variations or contrast inversions between multispectral channels are not accounted for. A feature-based similarity measure in combination with an affine distortion model is used in [11]. Since the correction of geometric distortions is not the main topic of the paper, neither implementation details nor results concerning registration are provided. The template matching algorithm for registering volumetric brain images from a PET brain scanner in [31] exhibits some parallels to our approach: Regions for local registration are first preselected using a gradient measure. Then, template matching by full search is carried out based on a model incorporating translation and rotation. However, [31] has a different imaging background and utilizes a model estimation algorithm, which may not be able to cover stochastic errors in registration. Most important to us is that none of the mentioned papers investigates the physical background of a multispectral filter wheel camera to derive a model.

In this paper, we start by deriving a physical model describing the distortions of optical bandpass filters and complement it with an approximation for the chromatic aberrations caused by the lens. Based on our model, we then develop two registration algorithms in section III. Experimental results with our multispectral camera are provided in section IV. We conclude with a model validation in section V.

Compared to an earlier conference version of this work in [32], we present here a more general derivation of the physical model using vector- and matrix notation. Moreover, we

[1] The temperature drift of the center wavelength of our filters is stated to be about 0.015nm/°C at 400nm and 0.02nm/°C at 700nm.

develop two registration algorithms which are both validated in detail in the results section.

II. Physical Modeling

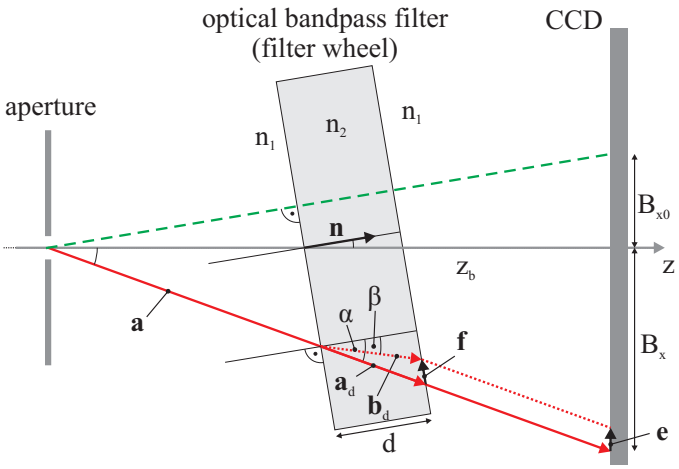


Fig. 2. Optical bandpass filter between CCD sensor and lens (not shown; the aperture is assumed to be very small). The lower solid ray $\mathbf{a}$ is refracted twice by the filter (dotted ray). The upper dashed ray is normal to the filter and thus not refracted. The displacement does not depend on the distance between the filter and the sensor.

Fig. 2 shows a detailed sketch of the internal configuration of a camera with a filter wheel. The incoming ray $\mathbf{a}$ is refracted twice by the tilted filter glass, namely, when entering the front side and when leaving the rear side of the filter. It then hits the CCD sensor, which is positioned at the image distance z_b behind the projection center. The filter normal $\mathbf{n}$ itself is tilted with respect to the optical axis. The degree of refraction depends on the filter thickness d, the refraction indices n_1 and n_2 and the incident ray angle. As already mentioned in the introduction, the filters vary in their thicknesses, orientations and refraction indices, causing the incoming rays to be refracted differently by each filter. The resulting misalignments between the color channels then cause color fringes (see Fig. 9). Since the undisplaced and the refracted ray are parallel, the distance between the filters and the sensor has no influence on the model.

We define the optical axis of our setup as

$$\mathbf{e_z} = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}^T . \tag{1}$$

A principal ray

$$\mathbf{a} = \begin{pmatrix} B_x & B_y & z_b \end{pmatrix}^T \tag{2}$$

with its physical image coordinates B_x, B_y and the image distance z_b hits the optical bandpass filter with thickness d under an angle

$$\alpha = \arccos\left(\mathbf{a}_n \cdot \mathbf{n}\right) , \quad \mathbf{a}_n = \frac{\mathbf{a}}{||\mathbf{a}||} , \tag{3}$$

with respect to the filter normal $\mathbf{n}$. Here, the dot operator denotes the scalar product of two vectors. After the ray is refracted, its angle β with respect to the normal follows Snell's law

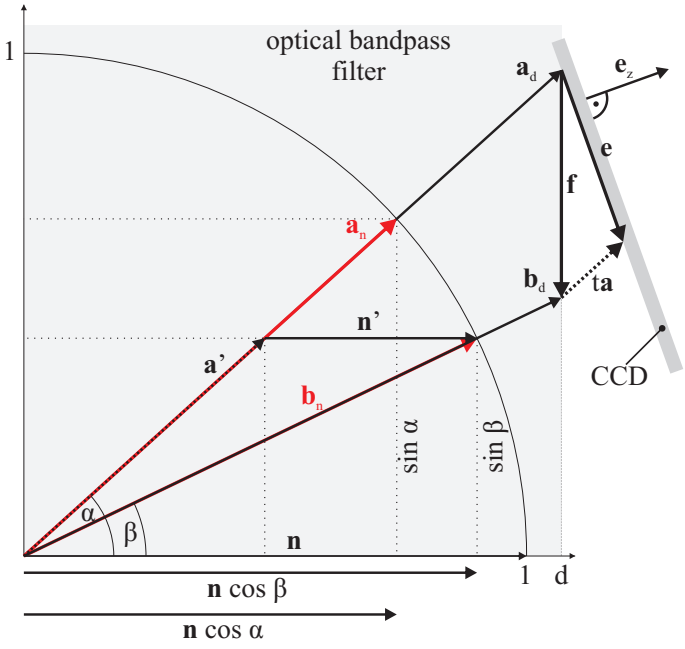


Fig. 3. Sketch of the plane spanned by the vectors $\mathbf{a}_n$ representing the incoming ray, $\mathbf{b}_n$ representing the ray after refraction by the filter, and the filter normal $\mathbf{n}$ (horizontal axis). Obeying Snell's law, these vectors are coplanar. The incoming normalized ray in the direction of $\mathbf{a}_n$, i.e., with the angle α relative to the normal $\mathbf{n}$, is refracted to $\mathbf{b}_n = \mathbf{a}' + \mathbf{n}'$ ($\angle(\mathbf{b}_n, \mathbf{n}) = \beta$). For a filter with thickness d, the displacement between unrefracted and refracted ray on the filter exit surface is $\mathbf{f} = \mathbf{b}_d - \mathbf{a}_d$. The displacement on the CCD surface is $\mathbf{e} = \mathbf{f} + t\mathbf{a}$. The optical axis $\mathbf{e_z}$ is perpendicular to the CCD surface.

$$\frac{\sin\alpha}{\sin\beta} = \frac{n_2}{n_1} = n_2 , \tag{4}$$

where $n_1 = 1$ is the refraction index of air and n_2 the refraction index of the optical filter. The refracted ray lies in the plane spanned by the filter normal and the incoming ray (see Fig. 3). Using vector notation, the refracted normalized vector $\mathbf{b}_n$ can be expressed as a weighted combination $\mathbf{b}_n = \mathbf{a}' + \mathbf{n}'$ of the filter normal $\mathbf{n}$ and the incident ray $\mathbf{a}$ as shown in Fig. 3. With (4) and the theorem on intersecting lines we derive

$$\mathbf{a}' = \mathbf{a}_n \frac{\sin\beta}{\sin\alpha} = \frac{\mathbf{a}_n}{n_2} \tag{5}$$

for the component $\mathbf{a}'$ of $\mathbf{b}_n$ parallel to $\mathbf{a}_n$. The component

$$\mathbf{n}' = \mathbf{n}\cos\beta - \mathbf{n}\frac{\cos\alpha}{n_2} \tag{6}$$

finally leads to the (normalized) refracted ray vector

$$\mathbf{b}_n = \mathbf{a}' + \mathbf{n}' = \frac{\mathbf{a}_n}{n_2} + \mathbf{n}\left(\cos\beta - \frac{\cos\alpha}{n_2}\right) . \tag{7}$$

By using the theorem on intersecting lines

$$\frac{\mathbf{b}_d}{d} = \frac{\mathbf{b}_n}{\cos\beta} \qquad \frac{\mathbf{a}_d}{d} = \frac{\mathbf{a}_n}{\cos\alpha} \tag{8}$$

the distance vector $\mathbf{f}$ between the unrefracted and refracted points on the filter exit layer is

$$\mathbf{f} = \mathbf{b}_d - \mathbf{a}_d = \left(\frac{\mathbf{b}_n}{\cos\beta} - \frac{\mathbf{a}_n}{\cos\alpha} \right) d \tag{9}$$

where $\mathbf{f}$ is perpendicular to the filter normal $\mathbf{n}$. To compute the final displacement $\mathbf{e}$ on the CCD surface, we project the vector $\mathbf{f}$ onto the xy-plane, i.e., the z-component of $\mathbf{e}$ must be zero ($\mathbf{e} \cdot \mathbf{e_z} = 0$). Since the refraction indices outside the filter glass are identical on both sides of the filter, the second refraction at the filter surface next to the CCD causes the outgoing ray to be parallel to the incoming ray $\mathbf{a}$. With a shift

$$\mathbf{e} = \mathbf{f} + t\mathbf{a} \tag{10}$$

of $\mathbf{f}$ along $\mathbf{a}$ and the condition $\mathbf{e} \cdot \mathbf{e_z} = 0$ (we assume that the CCD sensor is perpendicular to the optical axis) we finally find

$$\mathbf{e} = \mathbf{f} - \frac{\mathbf{f} \cdot \mathbf{e_z}}{\mathbf{a} \cdot \mathbf{e_z}} \mathbf{a} \tag{11}$$

for the displacement on the CCD layer.

Equations (3) to (11) form our reference model for the geometric distortion of *one* optical filter. The trigonometric terms involved in these equations, however, make the necessary computations more complex than it is desired for a fast correction algorithm. In practice, we therefore assume a paraxial optical system and use approximations for small angles. Our filter set consists of seven filters with two refraction indices, being either $n_2 = 1.45$ for the filters with center wavelengths 400nm and 450nm, or $n_2 = 2.05$ for the filters with center wavelengths 500nm, 550nm, 600nm, 650nm and 700nm. The tilt angles for the optical filters are less than $0.5°$. The image distance z_b is greater than 50mm, and the sensor size is 6.4mm $\times$ 4.8mm, yielding a maximum horizontal angle $\theta_x = \pm 3.7°$ and a maximum vertical angle $\theta_y = \pm 2.7°$ for the incident ray $\mathbf{a}$.

Under these conditions, simplifying our model using small-angle approximations is justified, as will be verified in section IV (see Fig. 20). With the approximations for small angles

$$\cos\beta \approx 1 \;, \quad \cos\alpha \approx 1 \tag{12}$$

equation (7) simplifies to

$$\widetilde{\mathbf{b}}_n = \frac{\mathbf{a}_n}{n_2} + \mathbf{n}\left(1 - \frac{1}{n_2}\right) \;. \tag{13}$$

With (13), the displacement $\mathbf{f}$ in (9) on the filter exit surface can be approximated to

$$\begin{aligned} \widetilde{\mathbf{f}} &= d\left(\frac{\mathbf{a}_n}{n_2} + \mathbf{n}\left(1 - \frac{1}{n_2}\right) - \mathbf{a}_n \right) \\ &= d\left(\frac{1}{n_2} - 1 \right) (\mathbf{a}_n - \mathbf{n}) \;. \end{aligned} \tag{14}$$

Finally, by insertion of (14) into (11) we find

$$\begin{aligned} \widetilde{\mathbf{e}} &= d\left(\frac{1 - n_2}{n_2} \right) \left(\frac{\mathbf{n} \cdot \mathbf{e_z}}{\mathbf{a}_n \cdot \mathbf{e_z}} \mathbf{a}_n - \mathbf{n} \right) \\ &= \underbrace{d\left(\frac{1 - n_2}{n_2} \right)}_{C} \left(\frac{\mathbf{n} \cdot \mathbf{e_z}}{z_b} \mathbf{a} - \mathbf{n} \right) \end{aligned} \tag{15}$$

with $z_b = \mathbf{a} \cdot \mathbf{e_z}$. Eq. (15) now establishes a linear relation between the displacement $\widetilde{\mathbf{e}}$ in the image plane and the filter parameters d, n_2 and $\mathbf{n}$. When the optical filter and CCD sensor are aligned in a coplanar manner, i.e. $\mathbf{n} = \mathbf{e_z}$, the displacement $\widetilde{\mathbf{e}}$ depends multiplicatively on the image coordinates B_x, B_y, contained in the incident ray $\mathbf{a}$ (see Eq. (2)), and a factor C, consisting of the filter thickness d and the refraction index n_2. The case of an unrefracted ray is considered by setting $\mathbf{a}_n = \mathbf{n}$: The ray then impinges on the filter with the angle $\alpha = 0$ and passes the filter with the displacement $||\mathbf{e}|| = 0$.

Eq. (15) also allows to calculate the influence of the filter orientations alone: for $d = 5$mm, $n_2 = 2.05$, a sensor size of 6.4mm$\times 4.8$mm and a resolution of 1280×960 pixels, each degree of misalignment shifts the color component by 4.5 pixels. In other words, if two filters are each misaligned by only 0.1 degrees in opposite directions with respect to the optical axis, the induced shift between the corresponding color components is already about one pixel.

Since each bandpass filter in the optical path exhibits its own filter normal (or tilt angle) $\mathbf{n}_i$, refraction index $n_{2,i}$ and thickness d_i, it causes individual geometric distortions $\widetilde{\mathbf{e}}_i$ in the image plane. The difference between the displacements $\widetilde{\mathbf{e}}_i$ for two filters with normal vectors $\mathbf{n}_i$ and parameters $C_i = d_i(1 - n_{2,i})/n_{2,i}$, $i = 1, 2$ in the image plane then is

$$\begin{aligned} \Delta\widetilde{\mathbf{e}} &= \widetilde{\mathbf{e}}_2 - \widetilde{\mathbf{e}}_1 \\ &= \underbrace{\frac{1}{z_b} \left[C_2 \mathbf{n}_2 \cdot \mathbf{e_z} - C_1 \mathbf{n}_1 \cdot \mathbf{e_z} \right]}_{g_F} \mathbf{a} + \underbrace{C_1 \mathbf{n}_1 - C_2 \mathbf{n}_2}_{\mathbf{t}_F} \;. \end{aligned} \tag{16}$$

Eq. (16) again yields a linear relationship, which later serves as the basis for our registration algorithm. It shows that the different filter angles, thicknesses and refraction indices generate a global translation $\mathbf{t}_F = \begin{pmatrix} t_{x,F} & t_{y,F} & \epsilon \end{pmatrix}^T$ between color components, which does not depend on the coordinates B_x, B_y in the image plane. In addition, the different filter thicknesses and refraction indices cause a position-dependent displacement $g_F \mathbf{a}$. Moreover, the factor g_F is identical for both x- and y-direction. Writing the differential distortion between two filters in homogeneous coordinates, (16) can be expressed as

$$\Delta\hat{\mathbf{e}}_F^H = \begin{pmatrix} g_F & 0 & t_{x,F} \\ 0 & g_F & t_{y,F} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} B_x \\ B_y \\ 1 \end{pmatrix} = \mathbf{T}_F \begin{pmatrix} B_x \\ B_y \\ 1 \end{pmatrix} \;, \tag{17}$$

thus integrating the additive translation into the matrix. The superscript $()^H$ in $\Delta\hat{\mathbf{e}}_F^H$ denotes the use of homogeneous coordinates. Also, we have added the subscript $()_F$ to distinguish these filter-induced distortions from the ones caused by chromatic aberrations, which are discussed next.

Even though we use an apochromatic corrected lens, there is a slight amount of chromatic aberration (CA) caused by the lens: CA causes a wavelength dependent pincushion or barrel distortion of the image [33], [34], introducing an additional circular displacement vector field. The description of this circular displacement vector field cannot be separated for x- and y-coordinates. As our results in section IV confirm, CA

can be captured well by an affine transformation matrix

$$\mathbf{T}_{\mathrm{C}} = \begin{pmatrix} g_{00,\mathrm{C}} & g_{01,\mathrm{C}} & t_{x,\mathrm{C}} \\ g_{10,\mathrm{C}} & g_{11,\mathrm{C}} & t_{y,\mathrm{C}} \\ 0 & 0 & 1 \end{pmatrix} . \tag{18}$$

This model remains valid when, without loss of generality, the effect of chromatic aberration is assumed to be located at the position $z = 0$, i.e., at the projection center. Taking into account CA, the distortion model (17) becomes

$$\begin{aligned} \Delta\hat{\mathbf{e}}^{(H)} &= \mathbf{T}_F \mathbf{T}_{\mathrm{C}} \begin{pmatrix} B_x \\ B_y \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} g_F \cdot g_{00,\mathrm{C}} & g_F \cdot g_{01,\mathrm{C}} & g_F \cdot t_{x,\mathrm{C}} + t_{x,F} \\ g_F \cdot g_{10,\mathrm{C}} & g_F \cdot g_{11,\mathrm{C}} & g_F \cdot t_{y,\mathrm{C}} + t_{y,F} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} B_x \\ B_y \\ 1 \end{pmatrix} , \end{aligned} \tag{19}$$

where $\mathbf{T}_{\mathrm{C}}$ mathematically precedes $\mathbf{T}_F$ because the lens (causing chromatic aberration) precedes the optical bandpass filter. Amongst others, $\mathbf{T}_{\mathrm{C}}$ adds elements on the secondary diagonal to the transformation matrix, which were equal to zero in (17). As shown in section IV and Fig. 18, we measured the effects caused by chromatic aberration alone to be below 0.2 pixels. The new matrix elements will therefore be quite small. Defining

$$\mathbf{T} = \begin{pmatrix} g_{00} & g_{01} \\ g_{10} & g_{11} \end{pmatrix} = \begin{pmatrix} g_F \cdot g_{00,\mathrm{C}} & g_F \cdot g_{01,\mathrm{C}} \\ g_F \cdot g_{10,\mathrm{C}} & g_F \cdot g_{11,\mathrm{C}} \end{pmatrix} \tag{20}$$

and

$$\begin{pmatrix} t_x \\ t_y \end{pmatrix} = \begin{pmatrix} g_F \cdot t_{x,\mathrm{C}} + t_{x,F} \\ g_F \cdot t_{y,\mathrm{C}} + t_{y,F} \end{pmatrix} \tag{21}$$

yields the final physical model for the distortions between two color components

$$\Delta\hat{\mathbf{e}} = \mathbf{T} \begin{pmatrix} B_x \\ B_y \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \end{pmatrix} . \tag{22}$$

III. Registration

The geometric distortions between two spectral channels recorded through different optical filters can be estimated and compensated with image registration. Additionally, we will see that the obtained displacement vector fields verify our physical model of the filter-induced misalignment and the CA model.

We develop two different approaches for registration: Our region-based approach divides the images into small nonoverlapping regions (size: 96 × 96 pixels) and computes a displacement vector for each region via matching. The resulting initial displacement vector field is then used to estimate the parameters of our model in Eq. (22). Alternatively, our global registration approach uses the entire image for registration rather than splitting it up into regions. The model parameters are computed by an iterative search algorithm which optimizes similarity of the relevant spectral channels by applying tentative geometric transformations according to our model with different parameter sets.

A. Similarity measurement

To align the spectral channels, one of them is selected as reference image to which all other channels are registered by optimizing a similarity criterion. Using cross correlation may not be adequate since spectral channels may exhibit a "contrast inversion" [29]. Also the camera may have a non-linear camera transfer function [35] relating the incident radiation to gray levels. Both issues may result in a non-linear relationship between the reference image and the image to be registered, as shown in Fig. 4b.

Registration algorithms for medical images therefore often use mutual information (MI) as similarity criterion [36], [37], which addresses the mentioned problems. Denoting G as reference region and F as examined region, the optimum geometric transformation

$$\begin{aligned} \tilde{\mathbf{T}} &= \operatorname*{argmax}_{\mathbf{T}} I(\mathbf{T}) \\ &= \operatorname*{argmax}_{\mathbf{T}} \sum_{f,g} p_{\mathbf{T}(F)G}(f,g) \log_2 \frac{p_{\mathbf{T}(F)G}(f,g)}{p_{\mathbf{T}(F)}(f) p_G(g)} \end{aligned} \tag{23}$$

is found by maximizing the mutual information $I(\mathbf{T})$. The term $p_{\mathbf{T}(F)G}(f,g)$ denotes the joint probability density function (PDF) of the transformed input region $\mathbf{T}(F)$ and the reference region G, with the corresponding pixel intensities f, g as independent variables. $p_{\mathbf{T}(F)}(f)$ and $p_G(g)$ denote the PDFs of the regions.

Figure 4 shows two examples for estimated PDFs: On the left side, an identity transformation $\mathbf{T} = \mathbf{I}$ without correction of the distortions has been applied, i.e., the regions are *unregistered*. The mutual information measure thus takes a low value. The lowest possible value occurs when the considered regions are independent. The joint PDF then is a product of the regions' PDFs according to

$$p_{\mathbf{T}(F)G}(f,g) = p_{\mathbf{T}(F)}(f) \cdot p_G(g) . \tag{24}$$

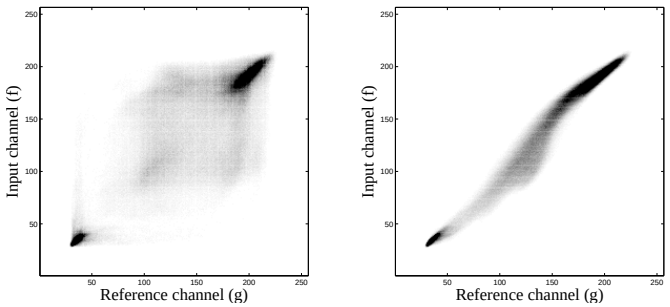
In this case, the numerator and denominator of Eq. (23) are identical and the logarithm is zero.

The largest possible value for $I(\mathbf{T})$ is generated by identical regions F and G (or when they are related by a non-zero scaling factor). The joint PDF then is a straight 45°-angled line and the mutual information is identical to the entropy of the region. Since the regions used in figure 4b have been registered, i.e., region F has been transformed with an optimum transformation matrix $\mathbf{T} = \tilde{\mathbf{T}}$ to match region G, the resulting PDF estimate is a narrow curve rather than spread over a broad area. The mutual information therefore is increased compared to Fig. 4a.

In our region-based approach, the transformation $\mathbf{T}$ is a translation, which is computed for each region separately to provide the initial displacement field from which the model parameters are then estimated. The global registration approach directly employs the complete physical model (22) for the entire image.

B. Mathematical model

We use the physical model (22) to characterize the geometric distortions between the different spectral channels. In our



(a) Unregistered image pair, $\mathbf{T} = \mathbf{I}$ (identity matrix, no correction).

(b) Registered image pair, $\mathbf{T} = \tilde{\mathbf{T}}$ (optimum transformation).

Fig. 4. Joint histogram $\hat{p}_{\mathbf{T}(F)G}(f, g)$ of image pairs.

implementation, the origin of the pixel [2] coordinates x, y is in the left upper corner of the image plane (Fig. 5). The derivation of the underlying physical model was based on centered image coordinates B_x, B_y, i.e., coordinates with origin in the image center $B_{x,c}$, $B_{y,c}$ given by the intersection between image plane and optical axis, as shown in Fig. 5. The relation between the image coordinates and physical coordinates depends on the pixel edge length s (which can be taken from the datasheet of the sensor but is not needed for the registration), and the unknown image center position $B_{x,c}$, $B_{y,c}$. By substituting the image coordinates $B_x = x \cdot s - B_{x,c}$, $B_y = y \cdot s - B_{y,c}$ and the displacements $\Delta\tilde{e}_x = \Delta x \cdot s$, $\Delta\tilde{e}_y = \Delta y \cdot s$ in (22) we derive

$$\begin{pmatrix} \Delta x \cdot s \\ \Delta y \cdot s \end{pmatrix} = \mathbf{T} \begin{pmatrix} x \cdot s - B_{x,c} \\ y \cdot s - B_{y,c} \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \end{pmatrix} . \qquad (25)$$

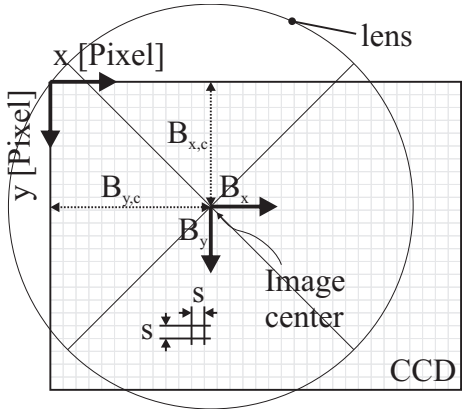


Fig. 5. Coordinate transform: The physical coordinates B_x, B_y, which originate in the image center ($B_{x,c}$, $B_{y,c}$), are transformed to pixel coordinates x, y by shifting and scaling with the pixel edge length s.

To convert (25) to a mathematical form similar to (22), we divide by the pixel edge length s and reorder the term elements

$$\begin{aligned} \mathbf{\Delta x} = \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} &= \mathbf{T} \begin{pmatrix} x - B_{x,c}/s \\ y - B_{y,c}/s \end{pmatrix} + \frac{1}{s} \begin{pmatrix} t_x \\ t_y \end{pmatrix} \\ &= \mathbf{T} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} x_c \\ y_c \end{pmatrix} \end{aligned} \qquad (26)$$

[2]We use subpixel accuracy to account for the small displacements.

with

$$\mathbf{x_c} = \begin{pmatrix} x_c \\ y_c \end{pmatrix} = \frac{1}{s} \begin{pmatrix} t_x \\ t_y \end{pmatrix} - \frac{1}{s} \mathbf{T} \begin{pmatrix} B_{x,c} \\ B_{y,c} \end{pmatrix} . \qquad (27)$$

Eq. (26) shows that the coordinate origin shift only induces an additive term, but does not influence the parametric form of the affine displacement model in (22). Using the notation $\mathbf{x}' = (x', y')^T$ for points in the input image to be transformed and $\mathbf{x} = (x, y)^T$ for reference points, we perform correction of our image with

$$\begin{aligned} \mathbf{x}' &= \mathbf{x} + \mathbf{\Delta x} \\ &= \begin{pmatrix} x \\ y \end{pmatrix} + \mathbf{T} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} x_c \\ y_c \end{pmatrix} . \end{aligned} \qquad (28)$$

Rewriting Eq. (28) with homogeneous coordinates by combining all terms into one matrix yields the compact notation

$$\begin{aligned} \begin{pmatrix} x' \\ y' \end{pmatrix} &= \begin{pmatrix} g_{00} + 1 & g_{01} & x_c \\ g_{10} & g_{11} + 1 & y_c \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \\ &= \mathbf{T_c} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} . \end{aligned} \qquad (29)$$

C. Region-based model fit

The block diagram of our region-based model registration algorithm including fitting of the model is shown in Fig. 6. The vector field is computed by dividing the image into nonoverlapping regions (region size: 96 × 96 pixel) and performing registration with the similarity measure described in section III-A. Since our affine distortion model indicates a smooth displacement vector field, we choose a relatively large region and estimate only translation between the corresponding regions. The result of this first step is a displacement vector field $\mathbf{V} = \{\mathbf{v}(x, y)\}$ with one vector $\mathbf{v}(x, y)$ for each region.

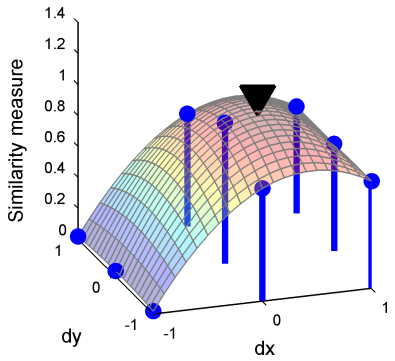


Fig. 7. Subpixel interpolation of similarity measure; stems denote pixel-precise measures, the triangle marks the analytic maximum.

Since we evaluate mutual information between reference image and transformed image only at integer shift positions, the displacement so far is not subpixel-precise. Subpixel accuracy is obtained by interpolation of the mutual information "grid" (see Fig. 7) around the pixel-precise maximum with a second order two-dimensional polynomial "surface"

$$I(dx, dy) = A_1 + A_2 dx + A_3 dy + A_4 dx dy + A_5 dx^2 + A_6 dy^2 \qquad (30)$$

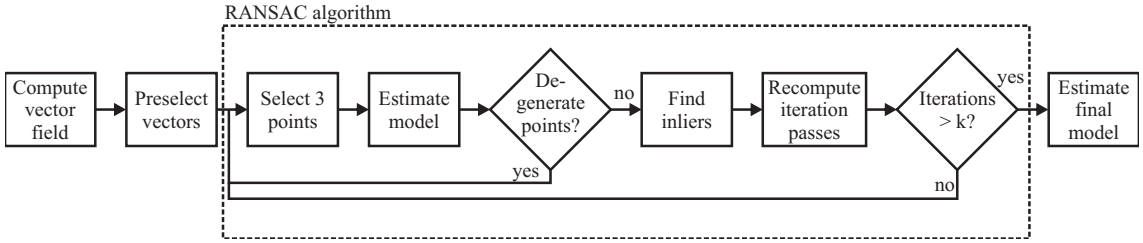


Fig. 6. Region-based algorithm. See Eq. (36) for estimation of iteration count k.

and analytic computation of the exact maximum. We compute the parameters $A_1 \dots A_6$ by fitting the polynomial to the MI surface values of the positions in the 3×3 neighborhood of the pixel-precise maximum. The same interpolation algorithm is used in [31] for refinement of 3D registration.

Depending on the image, large areas may not contain sufficient structure to enable reliable registration, such as homogeneous areas. To speed up computation and avoid registration errors in these critical regions, we determine them by evaluation of an entropy measure for each region Ω. All regions with an entropy value below the 40% percentile are skipped for the initial displacement estimation.

Now the iteration starts, being initialized with the estimated translation vector field. We use the "Random sample consensus" (RANSAC) algorithm [38] for removal of stochastic errors in the displacement vector field. First, three randomly selected point pairs $\mathbf{x}_0, \mathbf{x'}_0$, $\mathbf{x}_1, \mathbf{x'}_1$ and $\mathbf{x}_2, \mathbf{x'}_2$ are selected from the vector field $\mathbf{V}(x, y)$, since three pairs are required to compute the transformation matrix $\mathbf{T}$ in Eq. (29). With

$$\mathbf{x} = \begin{pmatrix} x & y & 1 \end{pmatrix}^T \qquad \mathbf{x'} = \begin{pmatrix} x' & y' \end{pmatrix}^T \tag{31}$$

and

$$\mathbf{X_S} = \begin{pmatrix} \mathbf{x}_0 & \mathbf{x}_1 & \mathbf{x}_2 \end{pmatrix} \qquad \mathbf{X'_S} = \begin{pmatrix} \mathbf{x'}_0 & \mathbf{x'}_1 & \mathbf{x'}_2 \end{pmatrix} \tag{32}$$

we compute the transformation matrix to

$$\mathbf{T} = \mathbf{X'_S} {\mathbf{X_S}}^{-1} \tag{33}$$

for cases where $\det(\mathbf{X_S}) \neq 0$. When $\mathbf{X_S}$ is not invertible, i.e., $\det(\mathbf{X_S}) = 0$, the set of point pairs is regarded as degenerated and another set is selected.

Now a consistency assessment of the estimated model parameters is carried out, which detects stochastic errors in the displacement vector field caused by registration failures in the corresponding regions. Towards this end, the current estimated transform $\mathbf{T}$ is applied to all points $\mathbf{x_i}$, yielding the transformed points $\mathbf{T x_i}$. Then the distances

$$l_i = |\mathbf{T x_i} - \mathbf{x'_i}| \tag{34}$$

to the corresponding points of the initially estimated displacement vector field are computed for all point pairs. The ones with distances below a certain threshold τ (e.g., $\tau = 0.5$) are regarded as inlier pairs.

Several iterations are carried out for different transforms $\mathbf{T}$ and the one with the most inliers is finally taken. The number of iterations k is computed [38] from the probability

$$P_0 = 1 - (1 - w^n)^k \tag{35}$$

of obtaining at least one point set which has $n = 3$ inlier pairs, because we need three pairs to compute the transformation matrix. The inlier probability w can be computed from the number of inlier vectors normalized by the number of all vectors. It is updated each time a new optimum of inlier vectors has been found. Then, w^n denotes the probability of selecting at least n point pairs passing the threshold τ, and $1 - w^n$ describes the probability of obtaining a corrupted point set, i.e., a set with less than n good points pairs. The term $(1 - w^n)^k$ is the probability of retrieving solely poor point sets after k iterations and P_0 denotes the opposite, viz. the probability of obtaining at least one good data set after k iterations. The number of iterations to be carried out thus is

$$k = \frac{\log 1 - P_0}{\log 1 - w^n} \ . \tag{36}$$

The algorithm so far reduces stochastic noise by identifying outlier vectors. However, since the transformation is computed from only three vectors pairs, noise might still compromise the accuracy of the transformation matrix. Therefore, the transformation matrix

$$\mathbf{T} = \mathbf{X'_P} {\mathbf{X_P}}^+ \tag{37}$$

is re-estimated by the pseudoinverse $()^+$ with all inlier vectors $\mathbf{X_P}$ and $\mathbf{X'_P}$, which are computed similarly as in (32).

D. Global model fit

Our global registration algorithm performs registration directly on the whole image rather than on regions, thus taking the affine model in section III-B into account from the beginning.

As Fig. 8 shows, we start with an initialization of the parameters by, e.g., the identity matrix (no distortion), which are used to transform the input image, if needed with subpixel accuracy using linear interpolation. Starting from the initialization, the model parameters are estimated by optimizing the mutual information criterion described in section III-A by the Nelder-Mead simplex method [39]. Actually, we use a variant of Matlab's® `fminsearch` function, which performs optimization in a restricted parameter search space. Since the geometric transformation has to be performed for evaluation of each parameter set, a gradient descent search algorithm would entail a high computational cost and is therefore not used here.

One critical issue regarding the geometric interpolation with subpixel precision is the occurrence of artifacts in the similarity measure "surface" [40]. These are produced by averaging the image noise [41] and introduce false local maxima, which

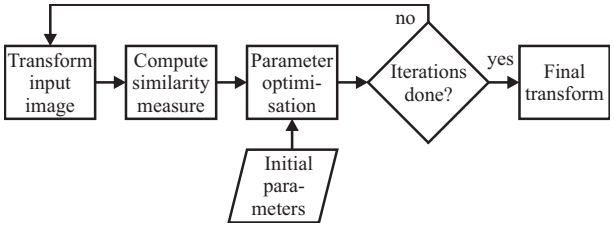
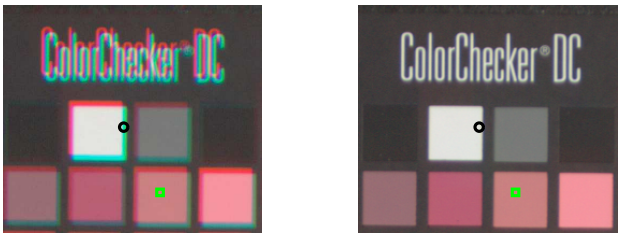


Fig. 8. Global registration algorithm.



(a) (b)

Fig. 9. Detail crop of RGB images, which have been reconstructed from unregistered (a) and registered (b) multispectral color components. The reconstructed spectra within the square and within the circle are shown in Fig. 10.

could lead to an erroneous registration. However, in most cases the global maximum is by far larger than the local ones. In our region-based approach we avoid this problem by using pixel-precise registration and polynomial interpolation.

IV. Results

A. Experimental setup

Fig. 1 shows our multispectral camera system which internally features a Sony XCD-SX900 CCD camera with a chip size of 6.4mm × 4.8mm and a resolution of 1280 × 960 pixels. We use a Nikkor AF-S DX 18-70mm lens on the external F-mount, while the internal camera is a C-mount camera. The thread of this camera has been shortened to place the optical filters as close as possible to the sensor. The filter wheel holds seven optical bandpass filters in the range from 400nm to 700nm in discrete steps of 50nm with a bandwidth of 40nm respectively. Without geometric distortion compensation, the resulting images exhibit clearly visible rainbow-like artifacts as depicted in Fig. 9a. The misalignment between two spectral channels may be up to ±10 pixels. Our compensation algorithm makes the color fringes vanish completely, as shown in Fig. 9b.

B. Impact on multispectral imaging

The typical workflow for multispectral image reconstruction transforms the acquired color components into an image in the sRGB color space (Fig. 9) or other color spaces [16]. Briefly, the processing chain includes a linearization of the image data, a spectral calibration, an estimation of the spectrum for each spatial position with a Wiener inverse and the transformation to the target color space.

The images in Fig. 9 of the ColorChecker DC test chart were acquired and processed with the described procedure. The spectra shown in Fig. 10 were reconstructed from the data within the square and circle regions in Fig. 9. As Fig. 10 illustrates, the spectra reconstructed from the homogeneous area within the square region are not affected by the color channel misalignments and result in the same spectra for unregistered and registered color components. However, when reconstructed from within regions with edges or other details, as highlighted by the circle in Fig. 9, unregistered color components generate a considerable spectral error.

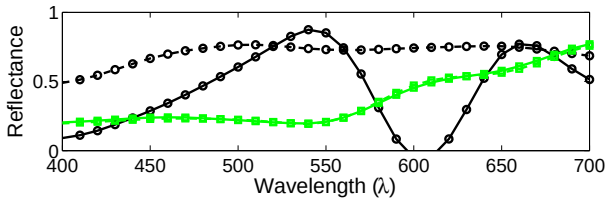


Fig. 10. Comparison of spectra derived from unregistered (solid lines) and registered (dashed lines) images. Symbols (squares and circles) correspond to Fig. 9.

C. Verification of our model

We verified our model with the multispectral acquisition and registration of graph paper shown in Fig. 11. Since the 550nm channel has the best signal to noise ratio and is the center channel of our system, we selected it as reference channel. For the 500nm spectral channel, we determined the transformation matrix in (29) to

$$\mathbf{T_c} = \begin{pmatrix} 1.0022 & -0.0007 & -0.2372 \\ -0.0006 & 1.0027 & -0.7797 \end{pmatrix} \tag{38}$$

and the matrix $\mathbf{T}$ in equation (28) to

$$\mathbf{T} = \begin{pmatrix} 0.0022 & -0.0007 & -0.2372 \\ -0.0006 & 0.0027 & -0.7797 \end{pmatrix} \tag{39}$$

by subtracting the identity matrix. In the matrix $\mathbf{T}$, the elements on the main diagonal represent mainly the distortions caused by the optical filters, whereas the influence of chromatic aberrations is comparably low. The elements on the secondary diagonal correspond to chromatic aberrations, and are of three to four times lower magnitude. The point of zero distortion can be computed by an eigensystem analysis of

$$\begin{pmatrix} & \mathbf{T_c} & \\ 0 & 0 & 1 \end{pmatrix} \mathbf{x} = \lambda \mathbf{x} \; . \tag{40}$$

and selecting the eigenvector corresponding to the eigenvalue $\lambda = 1$, thus indicating the identity transform. In the depicted case, this eigenvector is $(x, y, 1)^T = (214.9,\, 336.53, 1)^T$, as can be verified in Fig. 11. The first term in (40) represents the full homogeneous version of $\mathbf{T_c}$. Because of the homogeneous coordinates, the resulting eigenvector has to be normalized such that its last component is one.

To compare the resulting model with the measured vector data, we computed a synthetic displacement vector field by

$$\Delta\mathbf{x} = \begin{pmatrix} x' - x \\ y' - y \end{pmatrix} = \mathbf{T} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} - \begin{pmatrix} x \\ y \end{pmatrix} \tag{41}$$

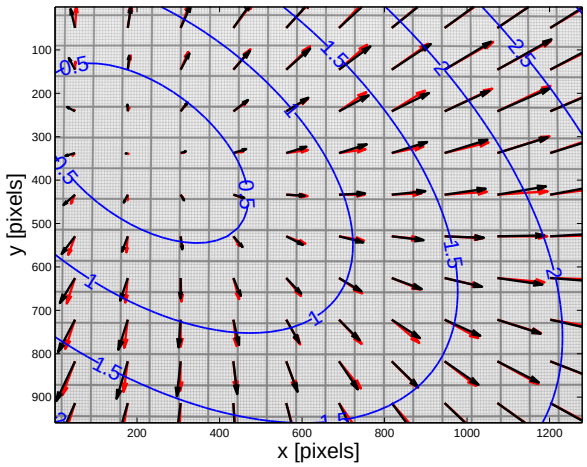


Fig. 11. Model verification: Scaled displacement vector field between channel 3 (λ_{reg} =500nm) and 4 (λ_{ref} =550nm), measurement data (⟶), model data (⟶), background: spectral channel at 500nm.

and inserted our estimated transformation matrix (38). The resulting error between the model vector data and the measurement vector data is only 0.13 pixel and will not be noticeable in a registered image.

D. Reduction of stochastic errors

Figure 12 shows the effect of the entropy detection algorithm (see section III-C), which excludes homogeneous regions since they are likely to contain no valuable registration information. Thick arrows in the figure denote regions with a high entropy measure, thin ones regions with too low an entropy, which were determined by the lower 40%-percentile. This threshold has been derived experimentally from a set of test images, and was used for all our experiments. Most parts of the sky and the roof in the figure are excluded from estimation.

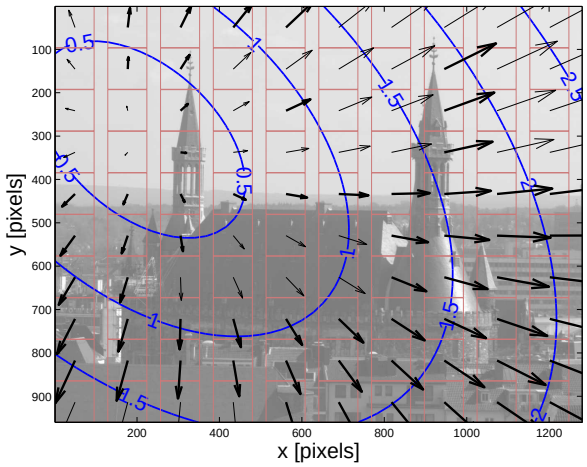


Fig. 12. Homogeneous regions (low entropy, thin arrows) are not considered for further computation; figure setup: see Fig. 11.

Even though homogeneous regions are excluded from the vector field used for final estimation of the transformation matrix, there may be still regions where registration fails. Here the RANSAC algorithm integrated into our algorithm (see Fig. 6) detects probably incorrect vectors and rejects them. We set the threshold $\tau = 0.5$, i.e., vectors with an Euclidean distance of less than half a pixel to the computed vector field are regarded as inlier vectors. Figure 13 demonstrates the effectiveness: Vectors which are classified as *inliers* are marked with *thick* lines, the *outliers* are marked with *thin* lines. All vectors being evidently false in the vector field are correctly classified – they are not included in final transformation matrix estimation. This – in combination with previous entropy detection – filters out stochastic noise robustly. Since the application of the transformation intrinsically computes the displacement vector field for each pixel coordinate based on the transformation matrix, the displacements in the excluded regions are interpolated consistently. To improve the accuracy of the estimation, statistical noise is reduced by taking *all* inlier vectors and computing a final transformation matrix using the pseudoinverse according to Eq. (37).

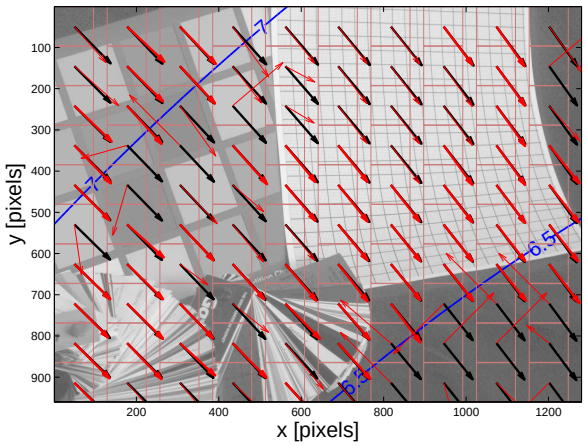


Fig. 13. Registration failures (thin arrows ⟶, 47 outlier vectors) are excluded from computation of the final vector field (⟶) by the RANSAC algorithm, λ_{reg}=600nm, λ_{ref}=550nm.

The performance of our global registration algorithm from section III-D is illustrated in figure 14: The initial transformation matrix of the algorithm is set to the identity matrix, i.e., it does not yield any displacement (not shown in figure). Further iteration results are depicted with a color fade of the vectors from red to blue. The latter color denotes the final transformation matrix result

$$\mathbf{T} = \begin{pmatrix} 1.0024 & -0.0006 & -0.3165 \\ -0.0006 & 1.0030 & -0.9210 \end{pmatrix}. \qquad (42)$$

As Fig. 15 shows, 145 iterations have been performed, while 241 similarity measurements, i.e., comparisons of the actual transformed image and the reference one, were needed.

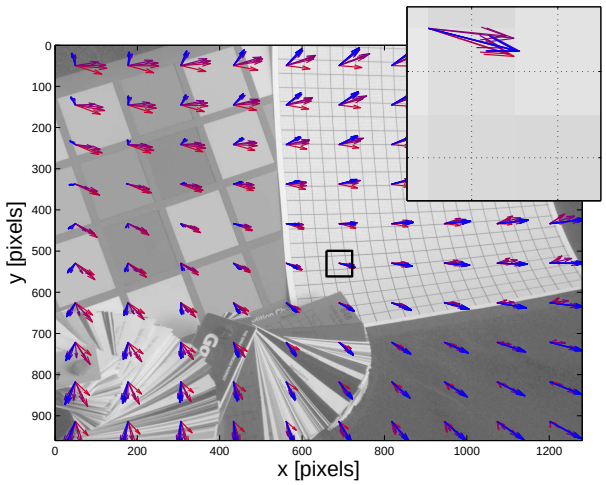


Fig. 14. Iterative operation of global algorithm: ⟶ denotes vectors from the iteration start, ⟶ the final ones; the cropped detail view shows unscaled vectors on pixel level; λ_{reg}=500nm, λ_{ref}=550nm.

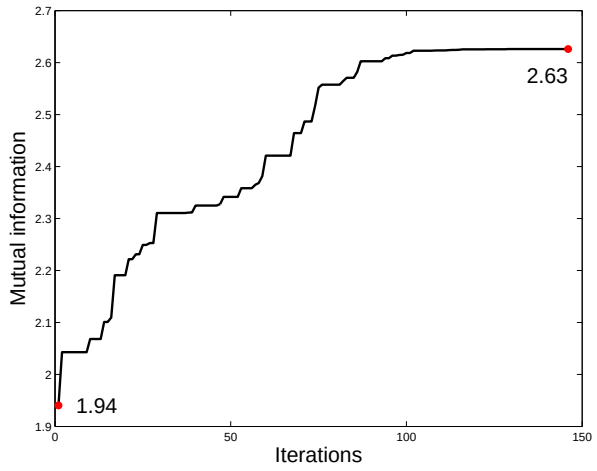


Fig. 15. Evolution of the mutual information function over iteration number; after 145 iterations the stopping criterion is fulfilled. Corresponds to Fig. 14.

E. Comparison between both algorithms

To compare both registration algorithms, viz. the region-based model fit in section III-C and the global one in section III-D, we fed them with synthetically distorted images. Towards this end, we transformed, e.g., a spectral image at 550nm with the inverse transformation matrix of (38), simulating the distortions between bandpass channels 500nm and 550nm. We thus can compute a ground truth vector field from the transformation matrix and compare it to the results of both algorithms. A slight drawback is that the original geometric distortion might not be simulated perfectly.

The comparison is depicted in Fig. 16 for $\lambda_{\text{reg}} = 400$nm: The original vector field is estimated quite well by both algorithms. However, the region-based one slightly outperforms the global algorithm: The mean Euclidean vector error between original vector field and the one estimated with our region-based algorithm is 0.04, compared to 0.13 for the global one. The maximum Euclidean vector errors are 0.13 and 0.28, respectively (see Table I). We also logged the execution time of our Matlab® implementations on a Pentium 2.8 Ghz computer: The region-based algorithm needs about 60% of the global one's execution time. These results pertain to the registration and correction of *all* spectral channels; and for the worst case that no regions are excluded by the region-based algorithm. Additional speedup of the latter one is achieved when excluding the most homogeneous regions as described in section III-C. The execution time of the global algorithm can be reduced by initialization with the previous optimization results; when focus and zoom remain unchanged, the geometric transformation parameters should be estimated quite fast.

algorithm / measure	mean error	max error	execution time [s]
region-based	0.04	0.11	64.74
global	0.13	0.28	108.64

TABLE I
COMPARISON OF REGION-BASED AND GLOBAL REGISTRATION ALGORITHM.

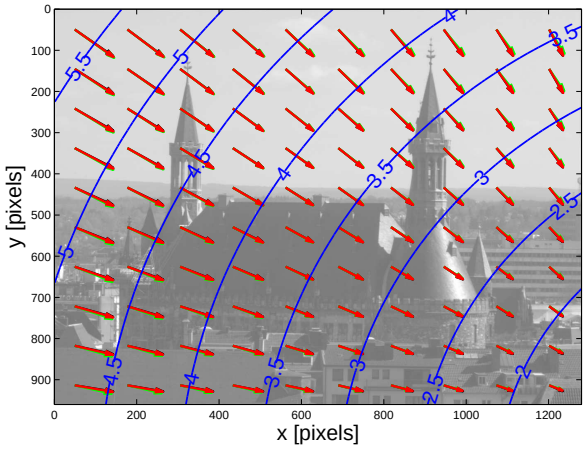


Fig. 16. Ground truth vector field (⟶), estimation by region-based approach (⟶) and global one (⟶), λ_{reg}=400nm, λ_{ref}=550nm.

F. Verification of the chromatic aberration model

The effects caused by chromatic aberration alone can be measured by acquiring two images of the same scene but with different wavelength ranges passing the lens. Towards this end we use the setup sketched in Fig. 17: A fixed bandpass filter of our filter wheel, e.g., with $\lambda_c = 550$nm, is selected to ensure that displacements are not caused by the properties of different filters. The light source is filtered by an additional spectral bandpass filter with wavelength $\lambda_f = 500$nm or $\lambda_f = 600$nm placed directly in front of it. Since the passbands of these filters with a bandwidth of $\Delta\lambda = 40$nm overlap with the one selected from the filter wheel, the effective spectral ranges reaching the CCD are the intersection areas shown shaded in Fig. 17. Their center wavelengths are around $\lambda_i = 525$nm

and $\lambda_i = 575$nm, thus differing by $\Delta\lambda = 50$nm. Since the optical filter in front of the CCD chip is the same for both acquisitions and the acquired scene is constant, any differences between the images are caused by the different wavelengths passing the lens. We measured a displacement vector field between both images by the region-based registration algorithm presented in section III.

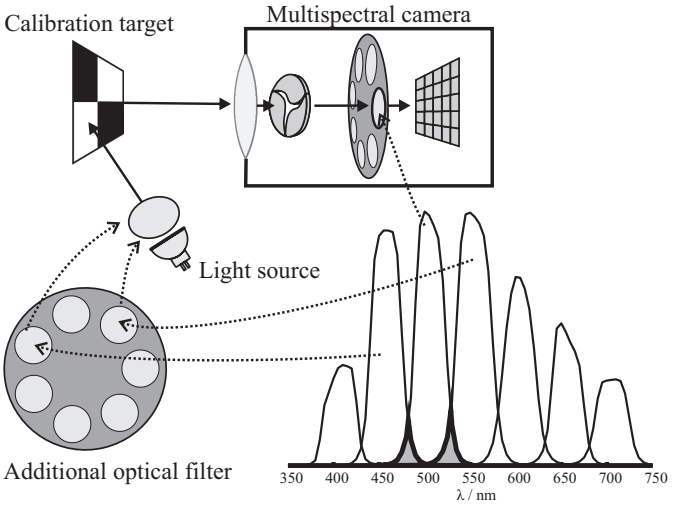


Fig. 17. Experimental setup of our chromatic aberration measurement: the wavelength of incoming light is changed using two different filters placed in front of the light source, while the filter wheel position in the camera remains unchanged.

The results are depicted in Fig. 18: Chromatic aberration causes a pincushion-like effect on the image. Even though the maximum vector length is only 0.2 pixel, we are still able to measure a systematic displacement vector field. Using an affine model for this vector field as suggested in section II, the remaining error between measurement data and the model fit drops to 0.076 pixel. Compared to the displacements caused by the optical filters of the filter wheel, the remaining errors almost vanish. Consequently, we can state that chromatic aberrations are modeled quite well for our application.

G. Verification of approximations

We verified the approximations in section II with our optical raytracer, whose results are shown in Fig. 19. The path of rays is simulated in 3D coordinate space and all refractions are computed with the exact equations (3) to (11) instead of the approximated one (15). The resulting displacement vector field between the unrefracted and refracted image points is compared with the one computed with the approximated equation (15). The resulting approximation error vector field is shown in Fig. 20 and exhibits a rotational symmetry. The reason is that the exact displacement vector field also obeys to a rotational symmetry, but the approximated one is separable. The maximum error is below one tenth of a pixel, thus confirming the validity of our approximations.

V. Conclusions

We derived a physical model for the geometric distortions caused by the optical bandpass filters of a multispectral camera. We also investigated chromatic aberration and showed that

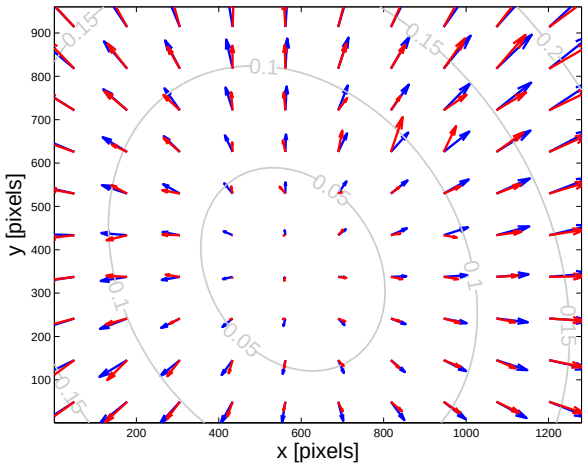


Fig. 18. Chromatic aberration between spectral channels with approx. 525nm and 575nm; measured vector field (⟶), affine model fit (⟶); remaining maximum error between both fields: 0.076 pixel.

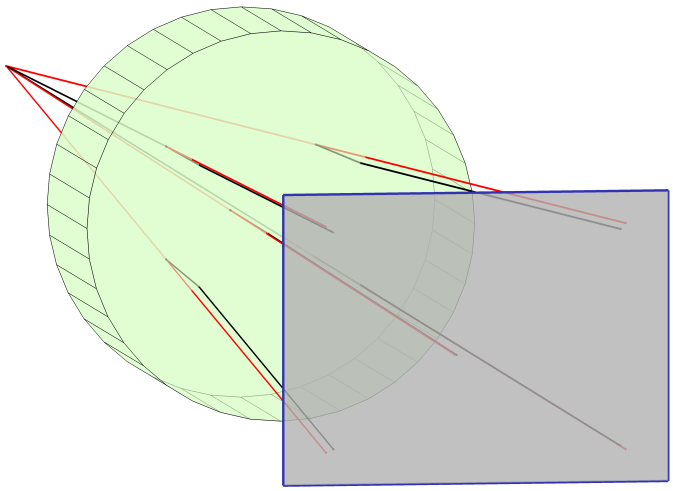


Fig. 19. Results of our optical raytracer: unrefracted rays are depicted with (⟶), refracted ones with (⟶). The sketch is true to scale concerning our optical setup: pixel size 4.65x4.65μm^2, image size 1280x960 pixel, refraction index $n_2 = 2.05$, filter thickness $d = 4$mm, image distance $z_b = 50$mm, $\gamma_x = 0.5°$, $\gamma_y = -0.5°$.

it can be modeled with an affine transformation matrix quite well. Based upon these models, we presented two correction algorithms and compared them in terms of accuracy and execution time: The region-based algorithm, which pre-registers small regions and estimates the final transformation matrix on vector field data slightly outperforms the one which estimates the transformation by tentatively applying transformations controlled by an parameter optimization algorithm. Detailed results are presented for both algorithms. After applying each one of our algorithms, the color fringes of the uncorrected image vanish completely.

Acknowledgment

The authors thank Stephan Helling and Bernhard Hill, Color Science and Color Image Processing Group of RWTH Aachen

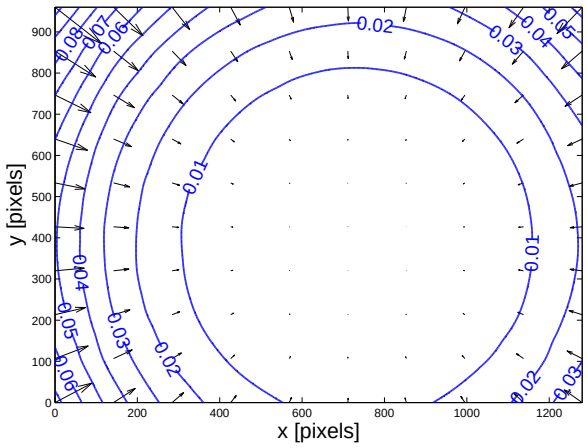


Fig. 20. Approximation error between exact equations (3) to (11) and approximation in (15). The maximum error is below one tenth of a pixel.

University, for valuable comments and discussions.

REFERENCES

[1] F. König and P. G. Herzog, "On the limitation of metameric imaging," *Proc. IS&Ts Image Processing, Image Quality, Image Capture, Systems Conference (PICS)*, vol. 2, pp. 163–168, 1999.

[2] R. Ramanath, W. Snyder, Y. Yoo, and M. Drew, "Color image processing pipeline," *IEEE Signal Processing Magazine*, vol. 22, no. 1, pp. 34–43, Jan 2005.

[3] P. M. Hubel, J. Holm, G. D. Finlayson, and M. S. Drew, "Matrix calculations for digital photography," in *IS&T/SID Fifth Color Imaging Conference: Color Science, Systems and Applications*, Scottsdale, Arizona, Nov 1997, pp. 105–111.

[4] R. Luther, "Aus dem Gebiet der Farbreizmetrik," *Zeitschrift für technische Physik*, vol. 8, pp. 540–558, 1927.

[5] B. Gunturk, J. Glotzbach, Y. Altunbasak, R. Schafer, and R. Mersereau, "Demosaicking: color filter array interpolation," *IEEE Signal Processing Magazine*, vol. 22, no. 1, pp. 44–54, January 2005.

[6] R. Lukac, K. Martin, and K. Plataniotis, "Demosaicked image postprocessing using local color ratios," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 6, pp. 914–920, June 2004.

[7] L. Zhang and X. Wu, "Color demosaicking via directional linear minimum mean square-error estimation," *IEEE Transactions on Image Processing*, vol. 14, no. 12, pp. 2167–2178, 2005.

[8] J. E. Greivenkamp, "Color dependent optical prefilter for the suppression of aliasing artifacts," *Applied Optics*, vol. 29, no. 5, pp. 676–684, February 1990.

[9] T. Aach, "Comparative analysis of shift variance and cyclostationarity in multirate filterbanks," *IEEE Transactions on Circuits and Systems–I: Regular Papers*, vol. 54, no. 5, pp. 1077–1087, 2007.

[10] A. Ribés, F. Schmitt, R. Pillay, and C. Lahanier, "Calibration and spectral reconstruction for crisatel: An art painting multispectral acquisition system," *Journal of Imaging Science and Technology*, vol. 49, no. 6, pp. 563–573, Nov/Dec 2005.

[11] A. Rabinovich, S. Agarwal, C. A. Laris, J. Price, and S. Belongie, "Unsupervised color decomposition of histologically stained tissue samples," in *Advances in Neural Information Processing Systems*, Whistler, British Columbia, Canada, Dec 2003, pp. 667–674. [Online]. Available: citeseer.ist.psu.edu/rabinovich03unsupervised.html

[12] R. M. Levenson, D. T. Lynch, H. Kobayashi, J. M. Backer, and M. V. Backer, "Multiplexing with multispectral imaging: From mice to microscopy," *ILAR journal*, vol. 49, no. 1, pp. 78–88, 2008.

[13] B. Hill and F. W. Vorhagen, "Multispectral image pick-up system," Patent, 1991, U.S.Pat. 5,319,472, German Patent P 41 19 489.6,.

[14] F. Knig and W. Praefcke, "The practice of multispectral image acquisition," *Proc. SPIE*, vol. 3409, pp. 34–41, 1998.

[15] P. Colantoni, R. Pillay, C. Lahanier, and D. Pitzalis, "Analysis of multispectral images of paintings," in *Proc. European Signal Processing Conference (EUSIPCO)*, Florence, Italy, Sep 2006.

[16] S. Helling, E. Seidel, and W. Biehlig, "Algorithms for spectral color stimulus reconstruction with a seven-channel multispectral camera," in *IS&Ts Proc. 2nd European Conference on Color in Graphics, Imaging and Vision CGIV 2004*, Aachen, Germany, Apr 2004, pp. 254–258.

[17] R. Berns, L. Taplin, M. Nezamabadi, M. Mohammadi, and Y. Zhao, "Spectral imaging using a commercial color-filter array digital camera," in *Proc. of The 14th Triennal ICOM-CC meeting*, The Hague, The Netherlands, Sep 2005, pp. 743–750.

[18] K. Ohsawa, T. Ajito, Y. Komiya, H. Fukuda, H. Haneishi, M. Yamaguchi, and N. Ohyama, "Six band HDTV camera system for spectrum-based color reproduction," *Journal of Imaging Science and Technology*, vol. 48, no. 2, pp. 85–92, March/April 2004.

[19] H. Haneishi, S. Miyahara, and A. Yoshida, "Image acquisition technique for high dynamic range scenes using a multiband camera," *Wiley's Color Research & Application*, vol. 31, no. 4, pp. 294–302, 2006. [Online]. Available: http://dx.doi.org/10.1002/col.20231

[20] M. Yamaguchi, H. Haneishi, and N. Ohyama, "Beyond RedGreenBlue (RGB): Spectrum-based color imaging technology," *Journal of Imaging Science and Technology*, vol. 52, no. 1, pp. 010 201–1–010 201–15, Jan 2008.

[21] T. Boosmann, "Combined matrix based determination of control values for a 6-primary display considering different observers," in *IS&Ts Proc. 3nd European Conference on Color in Graphics, Imaging and Vision CGIV*, Leeds, UK, Jun 2006, pp. 463–468.

[22] T. Ajito, T. Obi, M. Yamaguchi, and N. Ohyama, "Expanded color gamut reproduced by six-primary projection display," *Proc. SPIE*, vol. 3954, pp. 130–137, 2000.

[23] "Cambridge Research & Instrumentation, Inc." [Online]. Available: http://www.cri-inc.com/

[24] T. S. Hyvarinen, E. Herrala, and A. Dall'Ava, "Direct sight imaging spectrograph: a unique add-on component brings spectral imaging to industrial applications," *Proc. SPIE*, vol. 3302, pp. 165–175, May 1998.

[25] A. Mansouri, F. S. Marzani, J. Y. Hardeberg, and P. Gouton, "Optical calibration of a multispectral imaging system based on interference filters," *SPIE Optical Engineering*, vol. 44, no. 2, pp. 027 004.1–027 004.12, Feb 2005. [Online]. Available: http://link.aip.org/link/?JOE/44/027004/1

[26] V. Cappellini, A. Del Mastio, A. De Rosa, A. Piva, A. Pelagotti, and H. El Yamani, "An automatic registration algorithm for cultural heritage images," in *IEEE International Conference on Image Processing*, vol. 2, Genova, Italy, Sep 2005, pp. II–566–9.

[27] H. Foroosh, J. Zerubia, and M. Berthod, "Extension of phase correlation to subpixel registration," *IEEE Transactions on Image Processing*, vol. 11, no. 3, pp. 188–200, March 2002.

[28] L. Lucchese, S. Leorin, and G. M. Cortelazzo, "Estimation of two-dimensional affine transformations through polar curve matching and its application to image mosaicking and remote-sensing data registration," *IEEE Transactions on Image Processing*, vol. 15, no. 10, pp. 3008–3019, Oct 2006.

[29] J. Kern, "Reliable band-to-band registration of multispectral thermal imager data using multivariate mutual information and cyclic consistency," *Proceedings of SPIE*, vol. 5558, pp. 57–68, Nov 2004.

[30] G. Caner, A. M. Tekalp, G. Sharma, and W. Heinzelman, "Local image registration by adaptive filtering," *IEEE Transactions on Image Processing*, vol. 15, no. 10, pp. 3053–3065, Oct 2006.

[31] L. Ding, A. Goshtasby, and M. Satter, "Volume image registration by template matching," *Elsevier's Image and Vision Computing*, vol. 19, no. 12, pp. 821–832, 2001.

[32] J. Brauers, N. Schulte, and T. Aach, "Modeling and compensation of geometric distortions of multispectral cameras with optical bandpass filter wheels," in *15th European Signal Processing Conference*, Poznań, Poland, Sep 2007, pp. 1902–1906.

[33] M. Rebiai, S. Mansouri, F. Pinson, and B. Tichit, "Image distortion from zoom lenses: modeling and digital correction," in *International Broadcasting Convention (IBC)*, Amsterdam, Netherlands, Jul 1992, pp. 438–441.

[34] S. B. Kang, "Automatic removal of chromatic aberration from a single image," in *IEEE Conference on Computer Vision and Pattern Recognition*, Minneapolis, Minnesota, USA, Jun 2007, pp. 1–8.

[35] A. A. Bell, J. N. Kaftan, D. Meyer-Ebrecht, and T. Aach, "An evaluation framework for the accuracy of camera transfer functions estimated from differently exposed images," in *IEEE Southwest Symposium on Image Analysis and Interpretation*, Denver, Colorado, Mar 2006, pp. 168–172.

[36] 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.

[37] J. P. W. Pluim, J. B. 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.
[38] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.
[39] J. Nelder and R. Mead, "A simplex method for function minimization," *Computer Journal*, vol. 7, pp. 308–313, 1965.
[40] J. Tsao, "Interpolation artifacts in multimodality image registration based on maximization of mutual information." *IEEE Transactions on Medical Imaging*, vol. 22, no. 7, pp. 854–864, Jul 2003.
[41] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "Interpolation artefacts in mutual information-based image registration," *Elsevier's Computer Vision and Image Understanding*, vol. 77, no. 2, pp. 211–232, Feb 2000.