

Article

# MDPI

# A Novel Image Alignment Algorithm Based on Rotation-Discriminating Ring-Shifted Projection for Automatic Optical Inspection

# Chin-Sheng Chen \* and Chien-Liang Huang

Graduate Institute of Automation Technology, National Taipei University of Technology, 1, Sec. 3, Zhongxiao E. Rd, Taipei 10608, Taiwan; chien-liang@hotmail.com

\* Correspondence: saint@ntut.edu.tw; Tel.: +886-2-2771-2171 (ext. 4325); Fax: +886-2-8773-3217

Academic Editor: Chien-Hung Liu

Received: 24 December 2015; Accepted: 29 April 2016; Published: 9 May 2016

**Abstract:** This paper proposes a novel image alignment algorithm based on rotation-discriminating ring-shifted projection for automatic optical inspection. This new algorithm not only identifies the location of the template image within an inspection image but also provides precise rotation information during the template-matching process by using a novel rotation estimation scheme, the so-called ring-shifted technique. We use a two stage framework with an image pyramid searching technique for realizing the proposed image alignment algorithm; in the first stage, the similarity based on hybrid projection transformation with the image pyramid searching technique is employed for quick selection and location of the candidates in the inspection image. In the second stage, the rotation angle of the object is estimated by a novel ring-shifted technique. The estimation is performed only for the most likely candidate which is the one having the highest similarity in the first stage. The experimental results show that the proposed method provides accurate estimation for template matching with arbitrary rotations and is applicable in various environmental conditions.

Keywords: image alignment; rotation-discriminating; ring-shifted projection; image entropy

# 1. Introduction

Image alignment is a fundamental technique that has many applications for machine vision and image processing, including image retrieval, object recognition, pose estimation, industrial inspection, and target tracking. Because industrial inspections are usually conducted in well-controlled environments (e.g. the working distance between the lens and inspection objects is fixed), it is common practice to perform image alignment using constrained two-dimensional (2D) rigid transformation. This creates the possibility of establishing a fast image alignment algorithm that can be used for industrial inspection applications, and is the focus of this study. The research results are usually used in the quality control of manufacture automation. For example, two cases show the feasibility of this proposed research: (1) a robust vision system can inspect components fitted to several ranges of diesel engines at a critical stage and align in the engine assembly; (2) the quality assurance of surface mount devices (SMD) mounted on the printed circuit board (PCB).

Image alignment techniques are generally divided into two major categories: a feature- or geometry-based method and an area- or intensity-based method [1]. The feature-based methods, of Huttenlocher *et al.* [2,3] apply the directed Hausdorff distance to several algorithms to allow image alignment. Kown *et al.* [4] proposed a robust hierarchical Hausdorff distance to compare edge maps in a multi-level pyramidal structure. Chen *et al.* [5] also used the Hausdorff distance for image alignment in an inspection system for printed circuit boards. Peripherally, local descriptors are commonly used in several real-world applications, such as object recognition. Lowe [6] proposed a scale invariant

feature transform (SIFT), which combines a scale invariant region detector and descriptor and uses the gradient distribution in detected regions. SIFT emphasizes local features at a particular point of interest that are invariant to scale and rotation. Mikolajczyk and Schmid [7] compared the performance of descriptors that were computed for local interest regions. Even though SIFT is a useful method to describe the region of interest when the region has scale and rotates, SIFT-based matching fails when there are only a few feature points in the template image. The Zernike moments method [8] is also used to pattern features for a pattern image and determines invariance to rotation, translation, and scale. To ensure computational efficiency, Mondal [9] further proposed an accelerated version that uses a multi-resolution wavelet technique, projection, and Zernike moments. Ullah [10] proposed a rotation-invariant template method that uses gradient information in the form of orientation codes. However, this method is time-consuming because the histogram must be computated. In order to increase the speed of the template-matching method that uses orientation codes, the novel rotation discriminative descriptor and an object search strategy were used [11]. The aforementioned invariant descriptors of points of interest in the image are resistant to partial occlusion and relatively insensitive to changes of viewpoint. The algorithms referred to are generally efficient in general applications,

but they can result in unnecessary processing when the image alignment only involves rotation and

limited scaling compensation, as with automatic optical inspection (AOI) applications. The area-based method is sometimes called the correlation-like or template-matching method and has been popular for decades because of its basic concepts. The small template is firstly applied to a large inspection image by sliding the template window on a pixel-by-pixel basis. The normalized cross correlation (NCC) is computed between the template and the inspection images. The maximum values or peaks for the computed correlation values indicate the matches between a template and sub-images in the inspection image. The NCC metric is often used to manage the registration of images that differ only by a translation. If images are deformed by complex transformations, the template window cannot cover the same regions of interest in the template and the inspection images, so several studies have proposed modified NCC metrics to circumvent the image registration problem that is associated with complex transformations. To allow image matching that is invariant to rotation, a ring-projection transformation was proposed that transforms a 2D gray-level image into a rotation-invariant representation, as in a one-dimensional (1D) ring-projection space [12]. Tsai and Chiang [13] further used the ring-projection to represent a target template in wavelet-decomposed sub-images. Specifically, they used only pixels with high wavelet coefficients at low levels of resolution to compute the NCC between two separate templates. Choi and Kim [14] also used a two-stage image-alignment method that first locates candidates by comparing the vector sums for ring-projection and then matches these candidates using Zernike moments. Table 1 summarizes the features, approaches, and disadvantages of image-alignment methods for each category of method. The feature-based category uses edge maps, points of interest, shape contours, and invariant descriptors as alignment features. The embedded approaches are the Hausdorff distance, feature correspondence, the Zernike moment, and measurement of dissimilarity. The disadvantage of the feature-based method is the alignment error that occurs when there is inaccurate feature extraction. The area-based category uses image templates as matching features and cross correlation or ring projection to register the images. However, computation time can be excessive for applications that involve complex image transformation.

Category	Features	Approaches	Disadvantages		
Feature-based	Edge maps Interest point Invariant descriptors Orientation code	Hausdorff distance Feature correspondence Zernike moment Dissimilarity measurement	Inaccurate feature extraction		
Area-based	Image templates	Cross correlation Ring-projection	Excessive computation time		

<b>Table 1.</b> A summary of the image alignment method
---

Although there have been many studies of area-based image alignment techniques, relatively few concern ring-projection alignment, which is a special type of area-based method. This method transforms a 2D gray-level image into a rotation-invariant representation in one-dimensional (1D) ring-projection space [12]. Tsai [15] used the ring-projection technique to represent templates in multi-channel images using a method that rapidly selects candidates by computing the NCC between color ring-projection information. The rotation is then estimated by rotating the template. To reduce the computational complexity, an elimination strategy that quickly detects a similar candidate is used [16]. In this approach, the rotation invariant ring-projection transformation is used to describe the image. Two stage-matching methods have also been proposed in which the candidates are selected using the vector sums of the ring projection at the first stage. Only a single candidate then undergoes Zernike-moment template matching using rotation-invariant properties [14,17]. Radial projection has also been shown to be applicable to template matching [18]. Radial projection transforms a 2D gray-level image into 1D space and the scale invariance is constructed in the form of radial lines. However, ring-projection descriptors are specific to rotation-invariant matching. The rotation angle between two corresponding shapes in different images can also not be estimated in order to perform image alignment as is necessary for industrial inspection.

In AOI applications, the geometric relationship between a template image *P* and an inspected image *S* is constricted as shown in Figure 1.



Figure 1. The geometric relationship between pattern image P and inspection image S.

An indicated inspection pattern (such as the diamond templates shown in Figure 1) is usually used for the images, in order to alleviate the computational burden that is associated with real-time image alignment. Because the working distance between the lens and the inspected target is fixed in AOI applications, the scaling variation between *P* and *S* is tiny and can be neglected. Therefore, the image alignment problem for AOI can be formulated to identify the optimal 2D rigid transformation between *P* and *S*, as constrained by translation vector  $\overrightarrow{T}$  and rotation angle  $\theta$ . The method presented here locates the template image within the inspection image and provides precise rotation information during the template-matching process by using a novel rotation estimation scheme. This is the so-called ring-shifted technique which uses the architecture of a two-stage template-matching method. In the first stage (the candidate-searching stage), the only likely candidates are determined using rotation-invariant features. Subsequently, the ring-shifted technique is applied to the most likely candidate that is obtained from the first stage in order to estimate the rotation in the second stage (the rotation-estimation stage). Because the rotation is directly calculated using hybrid-projection data, the

proposed method minimizes computational complexity. Even though the newly matching method is proposed, it is not invariant to scale changes.

The main contributions of this paper include the following: (1) The novel image alignment algorithm based on rotation-discriminating ring-shifted projection, which is efficiently applied to recognize and align the template and inspection images; (2) The proposed hybrid projection transformation based on image entropy can provide unique and robust features for rapidly selecting the most likely candidate based on the image pyramid searching technique; (3) A new rotation-estimation technique known as the ring-shifted technique is proposed to detect the rotation of only the best of the output of the candidates; (4) An image pyramid searching technique is applied to significantly enhance the computation performance.

#### 2. Architecture of the Proposed Method

Before giving an overview of the proposed method, two fundamental projections, ring- and variance-projection, should be mentioned. The ring- and variance-projection transform the 2D image into 1D signals as function of radius, which are defined as the mean and variance of intensities along the radius, respectively. The main concepts of ring- and variance-projection are shown in Figure 2. The ring-projection calculates the average intensity for the specified radius, here three different rings,  $r_1 = 60$ ,  $r_2 = 120$ , and  $r_3 = 180$  pixels are shown. Furthermore, the corresponding variance-projections can be seen in the bottom figure. Obviously, the value of 180 pixels ring is lower than others because the intensity is almost uniform in this ring.



Figure 2. Main concepts of ring- and variance-projection.

The architecture of the proposed algorithm is shown in Figure 3. It consists of two phases: template preprocessing and online alignment. In the template preprocessing phase, template image P is the input specified manually. A pyramid technique is then used to build a multi-resolution image from the template image. Hybrid-projection values are then determined using ring- and variance-projection values and the weighting coefficients,  $w_{m,l}(r)$  and  $w_{\sigma,l}(r)$ , in conjunction with corresponding multi-resolution images. The coefficients  $w_{m,l}(r)$  and  $w_{\sigma,l}(r)$  are those of the ring and variance projection, respectively, depending on the intensity distribution at the specific radius of the template image. In the online alignment phase, the inspection image S is the input and a sliding window, which is the same size as the region of interest (ROI) in the multi-resolution template image, is moved for multi-resolution image  $S_l$ . This detects the most likely candidates by evaluating

the similarity coefficient of the hybrid projection values. The locations of the candidates with scores that exceed the similarity threshold  $T_{\min,l}$  are registered. The translation is estimated using the vector

T between the centers of the target detected in  $P_L$  and  $S_L$ . These candidates and their corresponding ring intensity profiles for each specified radius undergo the rotation-estimation process. This proposed method is called the ring-shifted technique to estimate the rotation angle. The accuracy of the rotation angle is finally refined by using a parabola curve fitting. The details of the two phases of the algorithm are described in the following subsections.



Figure 3. Architecture of the proposed algorithm.

# 3. Candidate Selection and Localization

The candidates are first selected and located using a rotation-invariant descriptor and the second stage uses the rotation-estimation method. For computational efficiency, the ring-projection transformation and the image pyramid technique [19] are used during the candidate selection process. These techniques identify the best matching results by using low-cost features in each image pyramid level. In the candidate-selection and localization process, the selection of a template image for the proposed method is a crucial step. This template image is used in the template preprocessing phase to generate hybrid-projection values. The candidate-selection and localization uses these hybrid-projection values. Therefore, the features of the template image in the template preprocessing phase considerably influence the matching results. In the hybrid-projection transform, the hybrid-projection values are calculated using a circular region between  $R_{l,\min}$  and  $R_{l,\max}$  on image pyramid level *l*. For illustrating the hybrid projection calculation, a concise example is illustrated in Figure 4, the hybrid projection values on current image pyramid *l* are calculated from the  $R_{l,\min} = 20$  to  $R_{l,\max} = 200$ . And the details of hybrid projection calculation are described in the following subsections.



Figure 4. An example to illustrate the circular region of hybrid projection calculation.

#### 3.1. Ring-Projection Transformation

Using the rotation invariant properties, the ring-projection transformation is used. The center point,  $(x_c, y_c)$ , of the block image is calculated and used to transform the pixel coordinate (x, y) of the block image from a Cartesian coordinate into a polar coordinate by:

$$x = x_c + r\cos\theta, \ y = y_c + r\sin\theta \tag{1}$$

where  $\theta \in [0, 2\pi]$ ,  $r = \sqrt{(x - x_c)^2 + (y - y_c)^2}$ ,  $r \in [R_{0,\min}, R_{0,\max}]$ ,  $R_{0,\max} = \min(M/2, N/2)$ ,  $R_{0,\min}$  is defined by the user, and *M* and *N* are the width and height of an image, respectively.

 $C_{P,l}(r)$  represents the ring-projection values for the template image at a specific radius r on image pyramid level l, which is defined as:

$$C_{P,l}(r) = \frac{1}{Q_r} \sum_{k} P(x_c + r \cos \theta_k, y_c + r \sin \theta_k)$$
<sup>(2)</sup>

where  $Q_r$  is the number of pixels that are located on the circle of radius  $r = R_{l,\min}, R_{l,\min} + 1, \ldots, R_{l,\max}$ .

The ring-projection value,  $C_{P,l}(r)$ , is the mean of the values for pixel intensity at the specific radius r on image pyramid level l. Because the value is constructed along rings of increasing radius, the one-dimensional ring-projection values are invariant to any rotation of the two-dimension template image. An example of the ring projection of an image is shown in Figure 5.

Figure 5 shows that the profiles of these two images remain nearly identical after rotation changes have been completed. In the candidate-selection process, the ring-projection value in the inspection image that is located at (x, y) with specific radius r on image pyramid level l is defined as follows:

$$C_{S,l}(x,y,r) = \frac{1}{Q_r} \sum_k S(x + r\cos\theta_k, y + r\sin\theta_k)$$
(3)

where the "S" is inspection image shown in Figure 4. The lookup table and optimal step angle  $\phi_{opt,l}$  are used to reduce computational complexity. The lookup table is created using rounded integer values for the distance from the center of the template image to the corresponding locations at the specific radius. The optimal step angle  $\phi_{opt,l}$  is shown in Figure 6 and is defined as follows:

$$\Phi_{opt,l}(r) = \cos^{-1}(1 - \frac{d^2}{2r^2}) \tag{4}$$

where *d* is the minimum shift distance between the original and the corresponding rotated points at the specified radius *r*. Using the optimal step angle, the number of pixels  $Q_r$  at the specific radius *r* is calculated by  $2\pi/\phi_{opt,l}(r)$  and  $\theta_k = \theta_{k-1} + \phi_{opt,l}$ .



**Figure 5.** An image at two distinct angles of rotation. (**a**) The original image; (**b**) The rotated image ( $30^{\circ}$  counterclockwise, CCW); (**c**, **d**) Show the corresponding profiles of ring projection.



Figure 6. Definition of the optimal step angle for image pyramid level *l*.

#### 3.2. Robust Features

As previously mentioned, the ring-projection transformation is used to describe the features of an image and to reduce computational complexity. However, the features should have unique and robust properties to correctly recognize the candidates. In despite of the ring-projection transformation having a rotation invariant property, it cannot describe the unique features of an image when the image contains a homogeneous region. For example, texture-less searching images containing different types of homogeneous regions are shown Figure 7a–d. The corresponding features that are described by ring-projection transformation are shown in Figure 7e. As can been seen in Figure 7e, the features of searching images 3 and 4 are easily distinguished; it also means the searching images 3 and 4 can be correctly recognized by evaluating the similarity of the features in the template and searching image in the candidate selection and localization process.



Figure 7. (a-d) Textures-less searching images 1-4; (e) Corresponding ring projection features.

By contrast, the features of images 1 and 2 cannot be correctly recognized by evaluating the similarity of the features in the candidate selection and localization process. Consequently, the features must describe the homogeneous ring regions in a distinct rather than a uniform manner. To overcome this problem, the variance of the specific radius in the image is considered and the variance-projection transform is used [20]. The variance projection transformation in the template and searching images on image pyramid level *l* are defined as  $\sigma_{p,l}$  and  $\sigma_{s,l}$ :

$$\sigma_{P,l}(r) = \frac{1}{Q_r} \sum_k \left[ P(x_c + r\cos\theta_k, y_c + r\sin\theta_k) - C_{P,l}(r) \right]^2$$
(5)

$$\sigma_{S,l}(x,y,r) = \frac{1}{Q_r} \sum_k \left[ P(x + r\cos\theta_k, y + r\sin\theta_k) - C_{S,l}(x,y,r) \right]^2$$
(6)

In Figure 8, the features are described by the variance-projection transformation which actually reflects the variation for image 2 and other images (images 1, 3 and 4). However, this is not sufficient to describe the features of image 1, 3, and 4. Fortunately, the features of image 1, 3, and 4 can be distinguished by ring-projection transformation as shown in Figure 7e.

For the sake of unique and robust features description, a hybrid-projection transformation that combines the ring- and variance-projection transformations is used in this paper. The hybrid-projection transformation in the template and searching images on image pyramid level l are defined as  $H_{p,l}$  and  $H_{s,l}$ :

$$H_{P,l}(r) = w_{m,l}(r) \times C_{P,l}(r) + w_{\sigma,l}(r) \times \sigma_{P,l}(r)$$
(7)

$$H_{S,l}(x,y,r) = w_{m,l}(r) \times C_{S,l}(x,y,r) + w_{\sigma,l}(r) \times \sigma_{S,l}(x,y,r)$$
(8)

where  $w_{m,l}(r)$  and  $w_{\sigma,l}(r)$  are weighting coefficients for the ring and variance-projection transformation, respectively. In Equations (7) and (8), the weighting coefficients at the specific radius on image pyramid level *l* are calculated by the concept of image entropy.



Figure 8. Variance projection features of texture-less searching images.

The image entropy is a quantity which is used to describe the randomness of an image, the image entropy *H* with associated probabilities of occurrence pb(i) can be defined as:

$$H = -1 \times \sum_{i=0}^{g} pb(i) \log(pb(i))$$
(9)

where pb(i) is the probability of gray level *i*.

In the case of an image, if all of the pixels have the same value, the image entropy of the image is zero. On the other hand, a high image entropy occurs when an image has chaotic gray levels. For the hybrid-projection transformation, the data along the specific radius having high variance should have high weights. Therefore, the weighting coefficients at the specific radius are defined as follows:

$$w_{\sigma,l}(r) = (1 - T_{min,l}) \frac{H}{\log(g)}$$
(10)

$$w_{m,l}(r) = 1 - w_{\sigma,l}(r)$$
(11)

where the range of variance-projection weighting coefficient is between 0 and 1,  $1 - T_{min,l}$  is designed for keeping the robustness against the noise in the image ( $T_{min,l}$  is the similarity threshold on image pyramid level *l*), the weighting coefficient of ring-projection transformation is based on Equation (10), and in this study it is designed to get a symmetrical weighting coefficient.

As mentioned in Equations (7) and (8), the weighting coefficients are separately determined at the specific radius within the circular region on each template image pyramid level. Using the hybrid-projection transformation in Figure 7a–d and normalizing the hybrid values to a range between 0 and 1, yield the results shown in Figure 9. All features of the texture-less searching images, *i.e.*, Figure 7a–d, are clearly identified by this new descriptor. It means the features of the template image and searching image can be correctly recognized by evaluating the similarity based on the described features in the candidate selection and localization process. For this new descriptor, the weighting coefficients  $w_{m,l}(r)$  and  $w_{\sigma,l}(r)$  represent the weighting for the ring- and variance-projection values at the specific radius in the image pyramid level. Using the proposed method, the hybrid-projection transformation gives unique and more robust rotation-invariant features.



Figure 9. Hybrid projection features of texture-less searching images.

#### 3.3. Similarity Measurement

When the robust features are determined, the NCC is used as a similarity measure for selection of the possible candidates. The correlation coefficient for the template image and a searching block at point (x, y) in the inspection image on image pyramid level *l* are then defined as:

$$\delta_{l}(x,y) = \frac{\sum_{r=R_{\min}}^{R_{\max}} (H_{P,l}(r) - \overline{H}_{P,l}(r)) \cdot (H_{S,l}(x,y,r) - \overline{H}_{S,l}(x,y,r))}{\sqrt{\sum_{r=R_{\min}}^{R_{\max}} (H_{P,l}(r) - \overline{H}_{P,l}(r))^{2} \cdot \sum_{r=R_{\min}}^{R_{\max}} (H_{S,l}(x,y,r) - \overline{H}_{S,l}(x,y,r))^{2}}$$
(12)

where  $H_{P,l}(r)$  and  $H_{S,l}(x, y, r)$  are the values of the hybrid-projection transformation for the template and inspection images, respectively. Note that *l* is the index of the image pyramid levels. The correlation value  $\delta_l(x, y)$  is between -1 and 1. It is equal to 1 when a perfect match occurs between the template image and searching block. The candidate selection process allows only the most likely candidate to be fed into the refinement process.

#### 3.4. Image Pyramid Search Technique

The image pyramid search framework allows this type of target search. The image pyramid hierarchically splits the candidates using information from different image pyramid levels to disjointed subsets. This process allows many candidates to be discarded early in the search process because their similarity coefficients are lower than the pre-defined threshold for similarity. One case involving three image pyramid levels of the search process is shown in Figure 10.

The search process starts at the highest pyramid level (level 2) of the candidate. The candidate on the highest pyramid level is identified by computing the similarity using hybrid-projection values,  $H_{P,2}(r)$  and  $H_{S,2}(x, y, r)$ , at each position (x, y). The candidate,  $obj_{2,1}$ , on level 2, which is indicated by a green circle in Figure 10, is also identified. The candidate with a similarity score that exceeds the similarity threshold  $T_{min,2}$  is stored in the candidate list. The value of  $T_{min,2}$  is also dependent on the application and must be set to the minimum expected object visibility. On the next lower pyramid levels (level 1), the search process uses the candidate that is located on the previous pyramid level. This includes candidates  $obj_{1,\{1,2,3,4\}}$ , which are denoted by a yellow circle and are inherited from their parent candidate  $obj_{2,1}$  using a 2 × 2 searching area on image pyramid level 1. Only candidate  $obj_{1,2}$ , which is denoted by a red square, has a similarity score that exceeds the similarity threshold  $T_{min,1}$ .

which, in this case is level 0. Finally, the best matching result with a similarity score that exceeds the similarity threshold  $T_{\min,0}$  is denoted by the red fill square in Figure 10.



Figure 10. The image pyramid search technique.

In the search process, the similarity threshold  $T_{\min,l}$  is set in a different image pyramid level. The most intuitive way for the user to determine the value is to choose it between 0 and 1. If the threshold is set to be a low value, more candidates will be selected in the searching process. It also means the computation burden will increase. Furthermore, the similarity score of the candidates should be decreased on the higher pyramid level because the details are missed. Therefore, the similarity threshold  $T_{\min,l}$  must be slightly reduced on higher pyramid levels in order to avoid a candidate that is missed. For example, the similarity threshold of each higher image pyramid level is decreased by 0.1 by rule of thumb.

#### 4. Estimation of the Rotation

The candidate-selection process, as described in the previous section, selects the most likely candidates. A new rotation-estimation technique known as the ring-shifted technique is proposed to detect the rotation of only the best of the output of the candidates. Although, the original ring-projection  $C_{P,l}(r)$  shown in Equation (2) is rotation invariant, the ring-shifted technique will only inherit the ring intensity profile  $P(x_c + r\cos\theta_k, y_c + r\sin\theta_k)$  for each specific radius r. So, it does not need any operation for preparing the ring intensity profile in the rotation estimation process, which considerably reduces the computation time. For the rotation invariance, the hybrid-projection value is constructed along circular rings with successively increasing radii. In the ring-shifted technique, the ring intensity profile at each radius is used to estimate the rotation. To illustrate the ring-shifted technique, Figure 11 shows images with different rotation angles and the corresponding ring intensity profiles  $p_r$  and  $p'_r$  at the specific radius of r = 17.

It can be seen that the intensity of point  $p_{17}(n^*)$  is equal to the point  $p'_{17}(n^* + k(17))$  when the consequent point shifts by k(17), as shown in Figure 11a. Actually, the shift value k(17) represents the ring that is rotated by  $\phi_{17}$  at a radius of 17 in the spatial domain. Furthermore, we can define the relationship between the rotation angle  $\phi_r$  and the shift value  $k_{opt}(r)$  at the specific radius, which is defined as follows:

$$\phi_r = k_{opt} \left( r \right) \times \phi_{opt,0}(r) \tag{13}$$

where  $k_{opt}(r)$  is the shift value at radius r and  $\phi_{opt,0}(r)$  is the optimal step angle described in Equation (4).



**Figure 11.** An example of the ring intensity profile at the specific radius (radius = 17). (a) Original image; (b) Rotated image ( $30^\circ$ , CCW); (c) Shift point *k*(17) in the profiles with rotation angles.

Before the rotation estimation, the shift value  $k_{opt}(r)$  at radius r should be calculated. The cross correlation measurement is employed to find the maximum correlation of the two ring intensity profiles by a shift value, the so-called ring-shifted correlation. Therefore, the shift value can be derived; with the measurement with maximum correlation, the shift value can be defined as:

$$k_{opt}(r) = \underset{k \in [0, N_r]}{\operatorname{argmax}} (\delta_r(k))$$
(14)

where  $N_r$  represents the number of pixels at radius r.

$$\delta_{r}(k) = \frac{\sum_{n=0}^{N_{r}} (p_{r}(n) - \overline{p}_{r}(n)) \cdot (p_{r}'(n+k) - \overline{p}_{r}'(n+k))}{\sqrt{\sum_{n=0}^{N_{r}} (p_{r}(n) - \overline{p}_{r}(n))^{2} \cdot \sum_{n=0}^{N_{r}} (p_{r}'(n+k) - \overline{p}_{r}'(n))^{2}}}$$
(15)

Based on the definitions given from Equations (13)–(15), the rotation angle of the candidate can be estimated at only a single radius. According to Equation (4), the accuracy of the rotation angle depends on the optimal step angle and the radius. Obviously, the optimal step angle is inversely proportional to the radius. Therefore, the rotation angle can be accurately estimated by integrating all ring weighting coefficients  $w_r$  in the range between  $R_{\min}$  and  $R_{\max}$ . Hence, the rotation angle of the most likely candidate is defined as follows:

$$\theta_{obj} = \sum_{r=R_{\min}}^{R_{\max}} w_r \cdot k_{opt}(r) \cdot \phi_{opt,0}(r)$$
(16)

where  $w_r = N_r/N_{sum}$ ,  $N_{sum}$  is the number of pixels for all radii and  $N_r$  is the number of pixels for the specific radius r, respectively

In addition, two neighbors of the shift value k(r) at radius r, k(r) - 1 and k(r) + 1, and the corresponding correlation coefficients,  $\delta_r(k(r) - 1)$  and  $\delta_r(k(r) + 1)$ , are input into the fitting model to refine the accuracy of the rotation angle. Computing the best fitting parabola equation involves solving a system equation, which is written as:

$$\delta(x) = ax^2 + bx + c \tag{17}$$

where x = k(r) - 1, k(r), k(r) + 1.

Three unknown parameters *a*, *b* and *c* are calculated using Cramer's rule. When the parabola parameters are determined, the optimal shift value is calculated by:

$$k_{opt}^{*}\left(r\right) = -\frac{b}{2a} \tag{18}$$

Therefore, the greatest accuracy for the rotation angle  $\theta_{opt}^*$  is refined as:

$$\theta_{opt}^{*} = \sum_{r=R_{\min}}^{R_{\max}} w_r \cdot k_{opt}^{*}(r) \cdot \phi_{opt,l}(r)$$
(19)

# 5. Implementation of the Proposed Method

According to the architecture mentioned in Section 2, the implementation details of the two phases are described in the following subsection.

# 5.1. Template Preprocessing Phase

A template preprocessing phase determines the multi-resolution hybrid-projection information for the object of interest in the template image *P*. The precise steps are presented as follows:

- Step 1: The template image *P* is input specified manually.
- Step 2: The multi-resolution template image  $P_l$  is constructed, where l = 0, 1, ..., L, l represents several image pyramid levels and L is the maximum image pyramid level.
- Step 3: Ring-projection transform is used to determine ring-projection values  $C_{P,l}(r)$  within different image pyramid levels, where  $r \in [R_{l,\min}, R_{l,\max}]$ .
- Step 4: After the ring-projection transform process, the variance-project values  $\sigma_{P,l}(r)$  for each image pyramid level are determined through variance-project transform, where  $r \in [R_{l,\min}, R_{l,\max}]$ .
- Step 5: The hybrid-projection values  $H_{P,l}(r)$  for each image pyramid level are determined using ring-projection values  $C_{P,l}(r)$ , variance-projection values ,  $\sigma_{P,l}(r)$  and the corresponding weighting coefficients for the ring projection  $w_{m,l}(r)$  and variance projection  $w_{\sigma,l}(r)$ . The weighting coefficients,  $w_{m,l}(r)$  and  $w_{\sigma,l}(r)$ , are calculated by image entropy.
- Step 6: An appropriate threshold of similarity,  $T_{\min,l}$ , for different inspection conditions is identified. The levels of the image pyramid, weighting coefficients, hybrid projection values, and threshold are used for the online-alignment phase of the image pyramid technique, and the similarity measure with the image pyramid search technique, respectively.

#### 5.2. Online Alignment Phase

The online-alignment phase identifies the best candidates in the inspection image which have an arbitrary rotation angle. Also, the rotation angle is further determined using the ring-shift projection. The precise steps conducted at multi-resolution inspection image are presented as follows:

- Step 1: A multi-resolution inspection image  $S_l$  is constructed, where l = 0, 1, ..., L, and L corresponds to the maximum image pyramid level in the template preprocessing phase.
- Step 2: The search block initializes at location (x = 0, y = 0) in the highest inspection image pyramid level. This image is the same size as the template image on the highest image pyramid level.
- Step 3: Ring- and variance-projection transform are used to determine the ring-projection values  $C_{S,l}(x, y, r)$  and variance-projection values  $\sigma_{S,l}(x, y, r)$  from the search block at location (x, y) on image pyramid level *l*.
- Step 4: The hybrid-projection values  $H_{S,l}(x, y, r)$  on image pyramid level l are obtained using  $C_{S,l}(x, y, r)$ ,  $\sigma_{S,l}(x, y, r)$  and the weighting coefficients,  $w_{m,l}(r)$  and  $w_{\sigma,l}(r)$ , which are calculated in the template preprocessing phase.
- Step 5: When estimating the similarity coefficient between the template image and search block on image pyramid level l, if the similarity coefficient exceeds the pre-defined threshold for similarity, the location (x, y) is stored in the candidate list.
- Step 6: The search block in the inspection image is moved and Steps 3–6 are repeated until the similarity coefficient in the inspection image is computed for all locations.

When the candidate list for the highest image pyramid level is determined, the image pyramid search process is used to increase the speed of the candidate selection and localization processes. On the next lower image pyramid levels, the search block location inherits the location from the candidate list that has been stored on the previous pyramid level. The search block must only be searched within a small range of displacement. Steps 3–5 in the online-alignment phase are used to determine the most likely candidate on the lower image pyramid levels. The hybrid-projection transform for multi-resolution using the image pyramid search technique is described in Section 3.3. Finally, the best candidate is identified on the lowest image pyramid and then the rotation estimation process calculates the rotation angle using the ring-shifted projection for the best candidate.

### 6. Experimental Results

This section describes a series of experimental results that show the performance of the proposed method. Various images were used in different experiments to verify the proposed algorithm. To verify the proposed method, the results are compared with [15,21]. The similarity threshold  $T_{\min,0}$  is 0.8 for both this study and the compared methods. All experiments were performed on a personal computer with an Intel Core i7 3.4 GHz CPU and 8 GB of memory using Visual Studio 2008.

### 6.1. Rotation Estimation

Figure 12 shows the rotation test images, which are used to evaluate the rotation accuracy of the proposed rotation-estimation technique. To estimate the rotation accuracy, the simulated rotated images are generated by rotating the original image at rotation angles from  $0^{\circ}$  to  $359^{\circ}$  in increments of  $5^{\circ}$ . For obtaining the rotation angle, the least second moment method is utilized to estimate the angle between the template and inspection image in [15]. Also, the pre-constructed rotated templates score and piecewise linear model are applied to estimate the rotation angle in [21]. Detailed comparison results are shown in Table 2. The error  $E_r$  is calculated using the following equation:

$$E_r = |\theta_a - \theta_e| \tag{20}$$

where  $\theta_a$  and  $\theta_e$  are the actual and estimated rotation angles, respectively.

To provide an overall accuracy evaluation, three performance indices of error  $E_r$  are used to quantitatively show the performance: the mean of error  $E_{r_m}$ , the standard deviation of error  $E_{r_std}$ , and the maximum error  $E_{r_max}$ .



Figure 12. Images for the rotation test. (a) IC image 1; (b) Metal image; (c) IC image 2.

Case	Performance Index	Proposed Method	[15]	[ <b>21</b> ]
	$E_{r\_m}$ (°)	0.023	0.565	62.352
Figure <mark>12</mark> a	$E_{r \ std}$ (°)	0.020	0.376	48.950
-	$E_{r\_max}^{-}$ (°)	0.086	1.390	174.422
	$E_{r_m}$ (°)	0.010	1.004	90.306
Figure <mark>12</mark> b	$E_{r \ std}$ (°)	0.007	0.447	52.024
	$E_{r\_max}$ (°)	0.026	1.914	174.422
	$E_{r\_m}$ (°)	0.036	0.488	94.930
Figure <mark>12</mark> c	$E_{r\_std}$ (°)	0.027	0.393	47.098
	$E_{r\_max}^{-}$ (°)	0.102	1.607	174.422

Table 2. Errors for the rotation accuracy for both the proposed and compared methods.

As seen in Table 2, the errors in the mean, standard deviation, and maximum in all test cases, which are derived by using the proposed ring-shifted technique, are less than  $0.036^{\circ}$ ,  $0.027^{\circ}$ , and  $0.102^{\circ}$ , respectively. In [15], the least second moment method cannot provide precise results because of the simple integer data representation. Besides, the pre-constructed method for calculating the rotation angle in [21], the accuracy of angle depends on the number of pre-constructed rotation template images and the location which are estimated by the normalized cross-correlation (NCC) method. If the NCC method cannot give a correct location of the template within the inspection image, the rotation estimation will have failed. The number of pre-constructed rotation template images is nine in this test. The proposed method is seen to be superior to the other methods. For the ring-shifted technique, the rotation-estimation technique uses the optimal shift value  $k_{opt}^*(r)$  by increasing the radius from  $R_{min}$  to  $R_{max}$ . In this experiment, the proposed ring-shifted technique provides an accurate estimation when the template is rotated within the inspection image as a correct estimation image using arbitrary rotation angles.

#### 6.2. Performance on Images with Added Noise

This section considers the effect of noise on the inspection images. The test images and the corresponding template images are shown in Figure 13. The test images contain simple and complex backgrounds. The signal-to-noise ratio (SNR) for the Gaussian noise applied to the test images varies from 10 to 35 dB in increments of 5 dB. Figure 14 shows the errors in rotation at different noise levels. Table 3 lists the translation errors. The translation errors are calculated using the Euclidean distance between the real and matching positions. The rotation errors are calculated using Equation (20).

In this experiment, the maximum rotation errors for the proposed method are  $7.157^{\circ}$  and  $0.405^{\circ}$  when the SNR ratio is 10 and 15 dB, respectively. The maximum translation errors for the proposed method are 0 when the SNR is either 10 or 15 dB. In [21], the original template image is used to identify the location of the template image in the test images by using the NCC measurement. Because the

test images are not rotated in this experiment, the translation error for the test images with different Gaussian noise levels is not generated. By using the hybrid projection transformation, the features are obtained from the mean and variance values of the specific radius; therefore, it can significantly resist the noise in the inspection images. It is seen in this experiment that the overall performance of the proposed method is considerably better than that of the other methods when the SNR is greater than 10 dB.



**Figure 13.** Test images. The sizes of test images are (a)  $414 \times 525$ ; (b)  $800 \times 600$ ; and the corresponding template images are (a)  $170 \times 170$ ; (b)  $120 \times 120$ .



**Figure 14.** Rotation errors for the matching results with Gaussian noise. (**a**) The case of Figure 13(a); (**b**) The case of Figure 13(b).

SNR Ratio (dB)		Translation Error (pixel)						
		10	15	20	25	30	35	
	Proposed method	0	0	0	0	0	0	
Figure 13a	[15]	0	1	0	0	0	0	
	[21]	0	0	0	0	0	0	
	Proposed method	0	0	0	0	0	0	
Figure 13b	[15]	1.414	1	0	0	0	0	
	[21]	0	0	0	0	0	0	

Table 3. Translation error for the matching results with Gaussian noise.

# 6.3. Weighting Influence in Candidate Selection and Localization

The hybrid-projection transformation combined the weighted average of  $w_{m,l}(r)$  and  $w_{\sigma,l}(r)$  to select and localize the candidate. Here, the performance with the different weighted strategies of these two parameters are studied, the proposed image entropy method is compared with the user defined method. Three sets of parameters  $(w_{m,l}(r), w_{\sigma,l}(r))$  are selected as (0.3, 0.7), (0.5, 0.5), and (0.7, 0.3) corresponding to every specified radius for all image pyramid levels, respectively. The test images and corresponding template images are shown in Figure 15. The Gaussian noises from 10 to 35 dB in increments of 5 dB are applied in the test images for evaluation. Here, two performance indices of errors are used to quantitatively show the performance: the mean of error  $E_{t_m}$  and  $E_{r_m}$  for translation and rotation; the standard deviation of error  $E_{t_std}$  and  $E_{r_std}$  for translation and rotation, respectively. The results of translation and rotation errors are listed in Table 4, where the mean translation and rotation errors of our proposed method are 0 pixels,  $0.004^\circ$ , 0.17 pixels, and  $0.92^\circ$ , respectively. It is obvious that the results of our proposed methods are superior to the user defined method. According to the architecture in Figure 3, the accurate rotation estimated results are obtained based on correct locations of the matching candidates. As seen in the results of Figure 15a, wrong locations are obtained by the user defined method, thus, poor rotation estimated results are obtained in this case.



**Figure 15.** Test images. Sizes of all test images are  $800 \times 600$ , and the corresponding template images are (a)  $150 \times 150$ ; (b)  $120 \times 120$ .

Errors	Figure 15a				Figure 15b			
Methods	Translation (pixel)		Rotation (°)		Translation (pixel)		Rotation (°)	
	$E_{t_m}$	$E_{t\_std}$	$E_{r_m}$	$E_{r\_std}$	$E_{t_m}$	$E_{t_std}$	$E_{r_m}$	$E_{r\_std}$
Proposed method	0	0	0.004	0.001	0.17	0.37	0.92	2.02
$w_{m,l}(r) = 0.3, w_{\sigma,l}(r) = 0.7$	166.88	80.69	205.23	15.22	0.33	0.47	6.27	9.03
$w_{m,l}(r) = 0.5, w_{\sigma,l}(r) = 0.5$	160.91	11.02	241.88	9.92	0.33	0.47	6.27	9.03
$w_{m,l}(r) = 0.7, w_{\sigma,l}(r) = 0.3$	184.49	83.11	149.39	22.44	0.33	0.47	6.27	9.03

Table 4. Matching errors with different weighting coefficients.

# 6.4. Computational Performance in Real Captured PCB Images

Additional experiments were performed using real captured PCB images. The images were captured when an object was arbitrarily moving and rotating. All test and corresponding template images are shown in Figures 16–18. The sizes of the test images are  $800 \times 600$ . This experiment assessed performance of computation and correctness according to the matching results. The matching results for the proposed method and that of references [15] and [21] are denoted by red, blue and green boxes, respectively. To allow a fair comparison, the comparison methods are also optimized using the image pyramid search framework. The statistical results for efficiency and the effect of pyramid levels are summarized in Table 5. Here it is shown that the computational burden is significantly reduced when the image pyramid search framework is used.



**Figure 16.** Printed circuit board (PCB) case1 images. (a-c) real and (d) template image (200 × 200 pixels).



Figure 17. PCB case2 images. (a-c) real images and (d) template image  $(200 \times 200 \text{ pixels})$ .



Figure 18. PCB case3 images. (a–c) real images and (d) template image ( $200 \times 200$  pixels).

Time	Average Execution Time (s)							
Methods	Figure 16	Figure 17	Figure <mark>18</mark>	Figure 19				
Pyramid levels = 2								
Proposed method	0.316	0.316 0.124 0.12		0.322				
[15]	0.302	0.143	0.147	0.337				
[21]	0.057	0.057 0.03 0.031		0.062				
Pyramid levels = 1								
Proposed method	1.721	0.766	0.765	1.731				
[15]	3.701	1.393	1.401	3.842				
[21]	0.755	0.364	0.364 0.366					
Pyramid levels = 0 (without image pyramid search technique)								
Proposed method	19.5	8.318	8.329	21.183				
[15]	30.458	14.85	14.465	32.172				
[21]	11.526	5.694	5.737	11.526				

 Table 5. The computational burden for the proposed and compared methods.



Figure 19. Cont.



Figure 19. PCB case4 images. (a-c) real images and (d) template image ( $206 \times 201$  pixels).

In these test cases, the mismatch for the method of [21] is clearly seen in the PCB test cases, such as Figure 16b,c, Figure 17b,c, and Figure 18b,c (shown as green boxes). This method is sensitive to test images when the template is located in a complex background. In terms of the method of [15], the matching results provide a more accurate localization of the template image in the test images (shown as blue boxes). However, the results for the rotation angle are not sufficiently precise, such as in Figures 16a,c and 19a–c. By contrast, the matching results for the proposed method are seen clearly in Figures 16–18. They are superior to the results for the other methods. On the other hand, the image pyramid searching technique is used to enhance the performance in terms of efficiency. It can been seen the computation advantage of the proposed method, the [15,21] are around 66 times, 95 times, and 186 times.

From the experimental results, the proposed method not only provides a correct location, but also estimates a correct and precise rotation angle for the template image in real captured PCB images. Based on the convincing results of this experiment, the proposed method can be used with real-world images. However, the feature-based matching method that use the SIFT descriptor is not appropriate for the complex images, such as Figures 16–19 because of the heavy computational burden. There are more than 2300 feature points in these real captured images. The computation times for the real captured cases are 642.22, 620, 448.28, and 489.1 msec, respectively. The efficiency of the feature-based method is slower than the proposed method.

# 7. Conclusions

A novel image alignment algorithm that uses rotation-discriminating ring-shifted projection for AOI applications is presented. It combines hybrid projection transformation and the ring shift technique. The hybrid projection transformation with the image pyramid searching technique can significantly reduce the computation burden and own the unique and robust features in the alignment process. The ring shift technique provides the rotation estimation between the template image and the searching image. The results show that the rotation estimation of the proposed method is superior to other comparative methods. Furthermore, the novel image alignment algorithm can obtain accurate and robust results in the scene image with rotation, translation, and noise. A series of experiments verified the efficiency, accuracy, and robustness of the proposed algorithm. Furthermore, the proposed method not only provides high accuracy with rotation, but also works well under noise influence and translation. The various experiment results indicate that this approach is suitable for accurate image alignment in AOI industrial inspections.

Acknowledgments: The authors would like to acknowledge the financial support of the National Science Council of Taiwan, through its grant MOST 104-2221-E-027-029-MY2.

**Author Contributions:** Contribute the ideas of the research and research supervision: Chin-Sheng Chen; Performing of the research and writing of the manuscript: Chien-Liang Huang.

Conflicts of Interest: The authors declare no conflict of interest.

# References

- 1. Zitová, B.; Flusser, J. Image registration methods: A survey. Image Vis. Comput. 2003, 21, 977–1000. [CrossRef]
- 2. Huttenlocher, D.P.; Klanderman, G.A.; Rucklidge, W.J. Comparing images using the hausdorff distance. *IEEE Trans. Pattern Anal. Mach. Intell.* **1993**, *15*, 850–863. [CrossRef]
- 3. Huttenlocher, D.P.; Lilien, R.H.; Olson, C.F. View-based recognition using an eigenspace approximation to the hausdorff measure. *IEEE Trans. Pattern Anal. Mach. Intell.* **1999**, *21*, 951–955. [CrossRef]
- 4. Kwon, O.-K.; Sim, D.-G.; Park, R.-H. Robust hausdorff distance matching algorithms using pyramidal structures. *Pattern Recogn.* **2001**, *34*, 2005–2013. [CrossRef]
- Chen, C.-J.; Lai, S.-H.; Liu, S.-W.; Ku, T.; Yeh, S.Y. Optical PCB inspection system based on hausdorff distance. In Proceedings of SPIE 5679, Machine Vision Applications in Industrial Inspection XIII, 53, San Jose, CA, USA, 17 January 2005; pp. 53–61.
- 6. Lowe, D.G. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* **2004**, *60*, 91–110. [CrossRef]
- Mikolajczyk, K.; Schmid, C. A performance evaluation of local descriptors. *IEEE Trans. Pattern Anal. Mach. Intell.* 2005, 27, 1615–1630. [CrossRef] [PubMed]
- 8. Khotanzad, A.; Hong, Y.H. Invariant image recognition by zernike moments. *IEEE Trans. Pattern Anal. Mach. Intell.* **1990**, *12*, 489–497. [CrossRef]
- 9. Mondal, T.; Mourya, G.K. An accelerated approach of template matching for rotation, scale and illumination invariance. In *Control, Computation and Information Systems*; Springer: Berlin Heidelberg, Germany, 2011; pp. 121–128.
- Ullah, F.; Kaneko, S.I. Using orientation codes for rotation-invariant template matching. *Pattern Recogn.* 2004, *37*, 201–209. [CrossRef]
- 11. Marimon, D.; Ebrahimi, T. Efficient rotation-discriminative template matching. In *Progress in Pattern Recognition, Image Analysis and Applications;* Springer: New York, NY, USA, 2007; pp. 221–230.
- 12. Tang, Y.Y.; Cheng, H.D.; Suen, C.Y. Transformation-ring-projection (TRP) algorithm and its VLSI implementation. *Int. J. Pattern Recogn. Artif. Intell.* **1991**, *5*, 25–56. [CrossRef]
- 13. Tsai, D.-M.; Chiang, C.-H. Rotation-invariant pattern matching using wavelet decomposition. *Pattern Recogn. Lett.* 2002, 23, 191–201. [CrossRef]
- 14. Choi, M.-S.; Kim, W.-Y. A novel two stage template matching method for rotation and illumination invariance. *Pattern Recogn.* **2002**, *35*, 119–129. [CrossRef]
- Tsai, D.-M.; Tsai, Y.-H. Rotation-invariant pattern matching with color ring-projection. *Pattern Recogn.* 2002, 35, 131–141. [CrossRef]
- 16. Lee, W.-C.; Chen, C.-H. A fast template matching method with rotation invariance by combining the circular projection transform process and bounded partial correlation. *IEEE Signal Process. Lett.* **2012**, *19*, 737–740.
- 17. Lee, W.-C.; Chen, C.-H. A fast template matching method for rotation invariance using two-stage process. In Proceedings of the Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP 2009), Kyoto, Japan, 12–14 September 2009; pp. 9–12.
- Kim, H.; Araújo, S. Grayscale template-matching invariant to rotation, scale, translation, brightness and contrast. In *Advances in Image and Video Technology*; Mery, D., Rueda, L., Eds.; Springer: Berlin Heidelberg, Germany, 2007; Volume 4872, pp. 100–113.
- 19. Tanimoto, S.L. Template matching in pyramids. Comput. Graph. Image Process. 1981, 16, 356–369. [CrossRef]
- 20. Feng, G.C.; Yuen, P.C. Variance projection function and its application to eye detection for human face recognition. *Pattern Recogn. Lett.* **1998**, *19*, 899–906. [CrossRef]
- Sassanapitak, S.; Kaewtrakulpong, P. An efficient translation-rotation template matching using pre-computed scores of rotated templates. In Proceedings of the 6th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON 2009), Pattaya, Chonburi, Thailand, 6–9 May 2009; pp. 1040–1043.



© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).