A Survey of 2D and 3D Image Matching in Industrial Object Alignment

Chin-Sheng Chen, Chien-Liang Huang and Chun-Wei Yeh
Graduate Institute of Automation Technology, National Taipei University of Technology, Taipei, Taiwan, ROC

ABSTRACT
Image matching is a practical tool in the computer vision domain. 2D and 3D image matching can be based on (2D) features or on view and point (3D) which can be cloud based. Intensive practical research has been done on both 2D and 3D image matching and the methods are generally designed to meet the requirements of industrial alignment applications and include such aspects as full rotation, scaling and the handling of multiple objects. The aim of this study is the presentation of a review of 2D and 3D image matching and the provision of a comprehensive reference source for those involved in image matching.

1. Introduction
Recently, there has been a significant increase in the demand for both two-dimensional (2D) and three-dimensional (3D) image matching. Especially, the image matching is often used in industrial alignment, such as the Automated Optical Inspection (AOI) or industrial automation. This requires that real-world objects be automatically recognized and localized in a digital image by a computer. Additionally, the pose estimation of the object by 2D or 3D image matching can be used for various tasks, such as image retrieval, industrial inspection, target search, automated manufacturing, pick, and place operations or quality control. In the object alignment, the
The application illustrated in Figure 2 introduces further important aspects to be considered in image matching. Here, the metal parts on the conveyer shown in Figure 3 must be picked by the robot. In this example, image matching is used and the program should also be able to recognize several different instances of the object in an image simultaneously. In this example, the different metal parts may overlap and so the matching approach must be capable of handling a certain degree of occlusion. In addition, non-uniform illumination, in this case caused by an incorrectly mounted light source (see Figure 3), needs also to be taken into account. Uniform illumination that is of constant brightness is very important for most applications, but sometimes this is hard to achieve without an expensive setup. This means the image matching method used must also take variations in illumination into account. The image matching program should also be resistant to noise, but fortunately noise is a minor issue in most industrial applications. After the metal parts have been recognized, in the pick and place example, the world coordinates of the pick point are transmitted to the robot.

Another pick and place application is illustrated in Figure 4. This example shows an object with arbitrary xyz axis rotation. In this application, the metal parts will be picked by a 6-DOF robot arm. The matching method used in this case should be capable of recognizing the object image with any degree of rotation on the xyz axes.

The points below summarize the requirements of an image matching approach [1]:

1. The object must be matched in real-time.
2. The model representation of an object should be computed from a golden sample.
3. The image matching approach should be general with regard to the type of object.
4. Image noise and clutter should not interfere with image matching.
5. An object under rigid motion should be recognized.
6. The pose parameters must be of high accuracy.

Figure 1. Image matching in inspection applications: (a) input image and (b) inspected image.

Figure 2. Pick and place application by industrial robot.

classic or main key ideas of 2D and 3D of image matching were published within 20 years. In this paper, we do not contemplate into details of particular algorithms or describe results of comparative experiments. We summarize the papers which were proposed within 15 years and point out interesting parts of the 2D and 3D image matching methods, respectively. Several applications have been listed to illustrate the most important requirements that need to be fulfilled in real-world applications.

An inspection application is illustrated in Figure 1. Here, the task is to count the number of Light-Emitting Diode (LED) dies on the printed circuit board (PCB), and also to check for defects in the matched dies. Before defect detection can be done, the pose of the dies in the image must be determined using an image matching approach. In this case, a ‘gold’ sample die image (an image with no defects) is used to build a model for the matching process. In order to detect defective dies, the matching approach must detect all the dies mounted on the PCB. After the pose of the dies has been determined by the matching approach, defect detection can be done using a method such as image subtraction.
All instances of an object should be found in the image. The rest of the paper is organized as follows: Section 2 provides a review of 2D image matching based on area and features. Section 3 introduces a review of 3D image matching. The conclusions are set out in Section 4.

2. Approaches to 2D image matching

2D image matching techniques are particularly useful and widely used for many applications. They are simple and can handle many different types of objects. Existing 2D image matching approaches can be divided into two broad categories [1]: area based and feature based. In the following subsection, we adopt this classification and review a large number of existing 2D image matching approaches.

2.1. Area based

The area based method is sometimes called the correlation-like or template matching method and the basic concept has been very popular in the past. In this approach a small template is applied to a large inspection image by sliding the template window pixel-by-pixel and the correlation value is computed between the template and compared image blocks. The maximum values or peaks of the computed correlation values indicate the degree of match between the template and compared image blocks in the search image. Simple similarity measures such as the sum of absolute differences (SAD) or the sum of squared differences (SSD) are frequently used in template matching applications. However, these measures are not immune to brightness and contrast variations which occur in practical situations. The most popular and robust correlation-based
similarity measurement, normalized cross correlation (NCC) which is robust against illumination changes, is proposed. For template matching in AOI, the optimal matching results are indicated by the maximum NCC correlation, and the NCC coefficient is confined within the range of $-1$ and 1. The NCC coefficient is equal to 1 when there is perfect matching and the two compared images are the same. NCC is a computationally intense problem because to find an object in a scene the template must slide over the image pixel-by-pixel before the NCC coefficient can be computed. Template matching using full search correlation metric is exhaustive and time-consuming and this is especially so if rotation and multiple objects are to be detected. The efficient and rapid location of the best object in an image has always been an important issue in the matching process.

### 2.2. Feature based

In contrast to the area-based method, the feature-based approach does not work directly with image intensity values. Instead, a search is made for pair wise correspondence between interest points, edge points, or other feature descriptors between the template and the search model. The features represent information on a higher level. This property makes the feature-based method suitable for a scenario where changes in illumination are expected. To use feature-based matching, basic features, regions, edges, or interest points need to be extracted. These regions are often represented by the center of gravity, which is invariant with respect to rotation, scaling, and skewing and is also stable under random noise and gray-level variation. However, the general criterion of closed boundary regions is prevalent. The Maximally Stable Extremal Regions (MSER) based on homogeneity of image intensities is proposed to obtain the regions feature. For line feature detection, the standard edge detection method by Canny [7] is employed. To detect interest points, the powerful SIFT [8] and SURF [9] methods are used. The planar object detection method was introduced [10] for matching the object, and the SIFT and SURF methods are used to obtain the key points; a nearest neighbor search being employed to find corresponding points in the two images. The use of edge features, where the sum of the normalized dot products of the gradient direction vectors of the model edge points is calculated, has been proposed by Steger [11]. Hinterstoisser et al. [12] suggested real-time template matching based on the dominant gradient orientations of edges. Furthermore, Huttenlocher et al. [13–15] applied the directed Hausdorff distance to several efficient algorithms for image alignment. The main problem with feature-based methods is that the objects should contain discriminative texture. The rotation-invariant descriptor method, in contrast features of regions, edges and interest points, transforms a 2D gray-level image into 1D ring-projection space. Tsai and Chiang [17] further used ring-projection to represent a target pattern in wavelet-decomposed subimages. Specifically, they used only pixels with high wavelet coefficients at low resolution to compute and compare the NCC between two patterns. Ullah and Kaneko [18] proposed a rotation-invariant template method based on gradient information in the form of orientation codes. However, the method involves histogram computation and is time-consuming. In addition, the Zernike moments method [19] is useful for making the pattern features of a template image and considers invariance to rotation, translation, and scale.

### 2.3. Improving real-world application

To meet the requirements, computation time, and full pose estimation in AOI applications, numerous approaches have recently been proposed to reduce the time taken for 2D image matching. The improved methods are listed in Table 1 which shows the numerous techniques aimed at speeding up the image alignment process. There are two main strategies, computation enhancement and skipping unnecessary computation to enhance the performance. The main goal of computation enhancement strategy is the rapid calculation of key values, such as image mean, image variance, edge gradients, etc. Integral image technique is very useful and only needs 3 addition and 3 subtraction similarity measurement, normalized cross correlation (NCC) which is robust against illumination changes, is proposed. For template matching in AOI, the optimal matching results are indicated by the maximum NCC correlation, and the NCC coefficient is confined within the range of $-1$ and 1. The NCC coefficient is equal to 1 when there is perfect matching and the two compared images are the same. NCC is a computationally intense problem because to find an object in a scene the template must slide over the image pixel-by-pixel before the NCC coefficient can be computed. Template matching using full search correlation metric is exhaustive and time-consuming and this is especially so if rotation and multiple objects are to be detected. The efficient and rapid location of the best object in an image has always been an important issue in the matching process.

### 2.2. Feature based

In contrast to the area-based method, the feature-based approach does not work directly with image intensity values. Instead, a search is made for pair wise correspondence between interest points, edge points, or other feature descriptors between the template and the search model. The features represent information on a higher level. This property makes the feature-based method suitable for a scenario where changes in illumination are expected. To use feature-based matching, basic features, regions, edges, or interest points need to be extracted. These regions are often represented by the center of gravity, which is invariant with respect to rotation, scaling, and skewing and is also stable under random noise and gray-level variation. However, the general criterion of closed boundary regions is prevalent. The Maximally Stable Extremal Regions (MSER) based on homogeneity of image intensities is proposed to obtain the regions feature. For line feature detection, the standard edge detection method by Canny [7] is employed. To detect interest points, the powerful SIFT [8] and SURF [9] methods are used. The planar object detection method was introduced [10] for matching the object, and the SIFT and SURF methods are used to obtain the key points; a nearest neighbor search being employed to find corresponding points in the two images. The use of edge features, where the sum of the normalized dot products of the gradient direction vectors of the model edge points is calculated, has been proposed by Steger [11]. Hinterstoisser et al. [12] suggested real-time template matching based on the dominant gradient orientations of edges. Furthermore, Huttenlocher et al. [13–15] applied the directed Hausdorff distance to several efficient algorithms for image alignment. The main problem with feature-based methods is that the objects should contain discriminative texture. The rotation-invariant descriptor method, in contrast features of regions, edges and interest points, transforms a 2D gray-level image into 1D ring-projection space. Tsai and Chiang [17] further used ring-projection to represent a target pattern in wavelet-decomposed subimages. Specifically, they used only pixels with high wavelet coefficients at low resolution to compute and compare the NCC between two patterns. Ullah and Kaneko [18] proposed a rotation-invariant template method based on gradient information in the form of orientation codes. However, the method involves histogram computation and is time-consuming. In addition, the Zernike moments method [19] is useful for making the pattern features of a template image and considers invariance to rotation, translation, and scale.

### 2.3. Improving real-world application

To meet the requirements, computation time, and full pose estimation in AOI applications, numerous approaches have recently been proposed to reduce the time taken for 2D image matching. The improved methods are listed in Table 1 which shows the numerous techniques aimed at speeding up the image alignment process. There are two main strategies, computation enhancement and skipping unnecessary computation to enhance the performance. The main goal of computation enhancement strategy is the rapid calculation of key values, such as image mean, image variance, edge gradients, etc. Integral image technique is very useful and only needs 3 addition and 3 subtraction

### Table 1. Summary of efficiency improvement methods.

| Category       | Strategy                     | Improvement method                                      |
|----------------|------------------------------|--------------------------------------------------------|
| Area-based     | Computation enhancement      | Integral image [2,23]                                  |
|                |                              | GPU [24,25]                                            |
|                |                              | SIMD [26–28]                                           |
|                | Skipping unnecessary computation | Coarse-to-fine [22,29]                                  |
|                |                              | Bounded conditions [30–36]                              |
|                |                              | Elimination strategy [37–41]                            |
|                |                              | SIMD [12,42–45]                                         |
| Feature-based  | Computation enhancement      | Dominant gradients orientation [12,42–44,46]            |
|                |                              | Branch and bound [28,48]                                |
|                |                              | Invariant descriptor [49–54]                            |
|                | Skipping unnecessary computation | Bounded conditions [51,55]                              |
operations to summarize the value of a specific area. Another powerful way to achieve parallel computing is using modern CPUs for single instruction multiple data operation (SIMD) [20,21] where one SIMD instruction is simultaneously applied to multiple data points. This can dramatically reduce the computational burden. Another enhancement strategy is the exclusion of unnecessary computation by leaving out impossible candidates as early as possible, or by discarding impossible searching space in the matching process. Image searches can be made much more efficient using the image pyramid technique [22] in a coarse-to-fine strategy. Here, a full exhaustive search is first applied to find the candidate at the top of the image pyramid; the search is repeated downwards, level-by-level, until a candidate is found in the lowest level. Search time is dramatically reduced because the search space is inherited from the candidate parent.

By skipping impossible candidates, methods based on the bounded condition, elimination strategy, and winner update estimate of the index of the partial region between the template and the compared image, can be substantially improved. The index is calculated using bounded or partial correlation in both the bounded condition method and elimination strategy. The index will show if the compared image is a possible candidate or not. If the compared image is not a possible candidate, the matching process will skip the location of the compared image. In the invariant descriptor method, the invariant feature can be obtained using ring-projection and Fourier transformation. Furthermore, the features can be used to reduce the computational burden of the matching process.

The NCC method with rotation was introduced to satisfy the full pose estimation requirements in an AOI application needing full rotation and scaling.[56,57] The pre-computed score set from rotated templates to the original template is used to recognize the target object by skipping unnecessary computation in the first step and then estimating the rotation angle of the target object in the refinement step.[56,57] Here, piecewise linear models are used to evaluate the rotation angle by straight lines. Furthermore, bi-cubic interpolation in the matching scores is used to obtain the subpixel-level location and rotation angle in both steps was proposed by Cui et al. [57]. An accelerating method for matching of an object with full rotation and scaling was proposed by Chen et al. [28]. In this method, the image pyramid search scheme is combined with SIMD parallel computation. This search scheme can be used to quickly find certain objects with rotation, translation, and scaling in both monochrome and color images.

In contrast to the correlation-based methods, a gradient-based NCC metric for the detection of texture-less objects has been presented by several authors.[12,43,44] Here, the efficient template matching method only considers image gradient to detect an object, and then makes further use of the binary representation of the gradients. SIMD units are then used to reduce the computational burden of gradient-based NCC metrics. The iterative closest point (ICP) is then employed to optimize the pose estimation.[42]

2.4. Summary

A survey of 2D image matching has been discussed in the sections above, several practical methods were proposed for use in real-world applications. Area-based methods are useful in applications where the window contains images with distinguishable features, but mismatching will occur if window images have smooth areas without any prominent detail. Feature-based methods are typically applied when the local information is more significant than the information carried by image intensity. Despite its limitations, area- and feature-based 2D image matching is still in common use because it is so easy to implement.

3. 3D image matching approaches

While 2D image matching has been extensively investigated and is a relatively mature research area, 3D image matching has several advantages. For example, (1) range images provide depth information; (2) Features extracted from range images are not affected by illumination; (3) An estimated 3D pose of the object is more accurate compared to an estimated pose from 2D images. However, 3D image matching is difficult and there are many problems such as those of scaling, viewpoint variation, illumination changes, partial occlusion, and background clutter. Modeling technique that uses the 3D point cloud is becoming increasingly popular. 3D image matching can be defined as the problem of finding all instances of the models in a database in an arbitrary scene and then determining the pose of the detected objects. Several reviews related to the 3D image matching problem can be found in the literature, these are typically registered together using the ICP algorithm because it is popular and simple. The Point Cloud Library (PCL) is a good example of state-of-the-art implementations in the point cloud technique.[58] The existing 3D image matching schemes that use point clouds are: local feature-based, global feature-based, and view-based methods.

3.1. Local feature-based methods

In local feature methods, local features are extracted from around specific key points. They generally handle occlusion and clutter better than the global feature methods.
By describing the important points in the point cloud, Point Feature Histograms (PFH) was introduced by Rusu et al. [61,62], in which a PFH of two-point descriptors is built for all neighboring points of the reference point. Four features, the angles between the points' normal and the distance vector between them, are measured and categorized using a 16-bin histogram. However, defining local invariant features depends very much on the local surface information directly related to the quality and resolution of the acquired and model data. A more efficient method, Fast Point Feature Histograms (FPFH), based on PFH was later proposed by Rusu et al. [63], it retains most of the discriminative power of the PFH method and full descriptions of both PFH and FPFH are given. To consider the efficiency, a fast voting scheme based on the features which describe the pair of points was proposed by Drost et al. [60]. The pair features are defined using surface points with their normal vector which is more discriminative than individual primitives. The pair features can also reduce the complexity of the matching task. In this work, the model is represented by a set of pair features in a hash table for fast retrieval off-line. In the on-line phase, all the points in the scene are paired with the reference point to create point pair features. These features are efficiently matched to those of the model using the hash table. An industrial part handling system was proposed by Skotheim et al. [64] that makes use of this efficient voting scheme. Fast voting based on the pair features is employed to recognize the object at a coarse level and the ICP algorithm is applied to obtain results at the fine level. Choi et al. [65] include three additional pair features, described by oriented surface and boundary points and boundary lines, to increase the robustness and efficiency of pose estimation. In addition, a comparative study of local feature descriptors has been carried out by Morell-Gimenez et al. [66].

### 3.2. Global feature-based methods

In contrast to local feature matching, the global feature-based methods process the object as a whole. They define a set of global features which effectively and concisely describe the entire 3D model. This means they are not suitable for the matching of a partially hidden object in a cluttered scene. Rusu et al. [67] describe extension descriptors for FPFH, the Viewpoint Feature Histogram (VFH). In this work, additional VFH statistics are computed between the viewpoint direction and the estimated normal at each point. The results show a high degree of discrimination. To keep the advantages of VFH, clustered VFH (CVFH) was proposed to enhance robustness against occlusions. [68] This work proposed that a partial view of an object is described as a cluster of CVFH. The results show better performance than those obtained with VFH. A novel global registration method, oriented bounding box (OBB) regional area-based descriptor, was proposed for accurately matching. [69] The descriptor consists of two parts, the OOB information and the surface area of the subdivided boxes, and it is used to estimate the similarity of two scanning object data by NCC method. The results show better accuracy than the local feature-based methods.

### 3.3. View-based methods

An alternative way to determine the 3D pose of an object is to estimate the projection of the object location in 3D space onto a 2D camera image. [70] The existing methods manage to get by with just a single camera for the estimation of the so-called view-based 3D to 2D mapping transformation. [71–73] The objective of these methods is the determination of the position and orientation of a camera given its intrinsic parameters and a set of correspondences between 3D points and their 2D projections. In traditional Perspective-n-Point (PnP) methods, the pose is extracted from the homography matrix which exists between two projections of points on a 3D plane. Several algorithms have been proposed for the estimation of the homographic relationship between two images of a planar scene. [74,75] Consequently, the corresponding pose can be estimated by decomposing a homograph matrix. [76] In this investigation, different methods based on Singular Value Decomposition (SVD) [77] are thoroughly demonstrated. Recently, many authors have pointed out that non-iterative methods are time-consuming for large data point sets because they are computationally complex. [71] Therefore, an efficient non-iterative algorithm, PnP, has been proposed. It has linear complexity in n and expresses the solution as a weighted sum of eigenvectors. In addition, an optimal PnP method that applies Gauss–Newton optimization [72] has been established and a robust solution to a subsequent PnP problem has also been introduced. [73] In this work, an order 4 polynomial is used to describe the cost function, and a linearization technique is then applied to minimize the cost function. The results show the RPNP is suitable for applications that need to handle both small and large sets of points, such as those used in feature point-based image matching applications. A considerable number of studies have been undertaken on the template matching schemes related to the pose estimation techniques discussed above. Normalized edge-based templates rely on distance transformation which is used for matching and pose refinement. This, including real-time planar object detection and 3D pose estimation algorithms, is presented in some detail by Holzer et al. [78]. The matching result is useful for initial rough pose estimation and the Lucas-Kanade algorithm.
Table 2. Summary of 3D pose estimation techniques.

| Methods            | Data                  | Strength                          | Weakness                        |
|--------------------|-----------------------|-----------------------------------|---------------------------------|
| View-based         | CAD model or multiple images | Efficiency                       | Limited accuracy on viewpoints |
|                    |                       | Robust to occlusions, clutter, and contrast changes | Low-cost                        |
|                    |                       | High accuracy                     |                                 |
|                    | Point clouds          |                                   |                                 |

Point cloud-based  | Point clouds          | High accuracy                     | 3D sensor required              |

3.4. Summary

It is clear that the main difficulties encountered in 3D image matching are occlusion, clutter, and noise. To overcome these challenges, several investigations have been proposed and discussed. Table 2 gives a summary of 3D image matching. The point cloud-based methods are the most popular for 3D image matching, but computation is still highly complex and very efficient algorithms are needed to facilitate computation. Nowadays, view-based methods are the most frequently used in real-world application because of the cost and efficiency, despite the limited viewpoint accuracy.

4. Conclusions

This paper has presented a survey of the state-of-the-art 2D and 3D image matching methods. The intensive research that has been done to improve performance in real-world applications has been discussed and analyzed. Although, there are many improvement in the approaches to image matching, it still remains an open problem, such as how to improve the efficiency and the accuracy in 3D image matching when the object with viewpoint variation, partial occlusion, or cluttered background.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

[1] Zitová B, Flusser J. Image registration methods: a survey. Image Vision Comput. 2003;21:977–1000.
[2] Tsai D-M, Lin C-T. Fast normalized cross correlation for defect detection. Pattern Recognit. Lett. 2003;24:2625–2631.
[3] Lewis J. Fast normalized cross-correlation. Vision Interface. 1995;10:120–123.
[4] Briechle K, Hanebeck UD. Template matching using fast normalized cross correlation. In: Casasent DP, Chao T-H, editors. Aerospace/defense sensing, simulation, and controls. Orlando (FL): SPIE; 2001. p. 95–102.
[5] Ouyang W, Tombari F, Mattoccia S, et al. Performance evaluation of full search equivalent pattern matching algorithms. IEEE Trans. Pattern Anal. Mach. Intell. 2012;34:127–143.
[6] Matas J, Chum O, Urban M, et al. Robust wide-baseline stereo from maximally stable extremal regions. Image Vision Comput. 2004;22:761–767.
[7] Canny J. A computational approach to edge detection. IEEE Trans. Pattern Anal. Mach. Intell. 1986;8:679–698.
[8] Lowe DG. Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vision. 2004;60:91–110.
[9] Bay H, Ess A, Tuytelaars T, et al. Speeded-up robust features (SURF). Comput. Vision Image Understanding. 2008;110:346–359.
[10] Kottman M. Planar object detection using local feature descriptors. Assoc. Comput. Mach. Bull. 2011;3:59–63.
[11] Steger C. Similarity measures for occlusion, clutter, and illumination invariant object recognition, Pattern Recognit. 2001;2191:148–154.
[12] Hinterstoisser S, Lepetit V, Ilic S, et al. Dominant orientation templates for real-time detection of texture-less objects. IEEE Conf. Comput. Vision Pattern Recognit. 2010;10:2257–2264.
[13] Huttenlocher DP, Klanderman GA, Rucklidge WJ. Comparing images using the Hausdorff distance. IEEE Trans. Pattern Anal. Mach. Intell. 1993;15:850–863.
[14] Huttenlocher DP, Lilien RH, Olson CF. View-based recognition using an eigenspace approximation to the Hausdorff measure. IEEE Trans. Pattern Anal. Mach. Intell. 1999;21:951–955.

[15] Rucklidge WJ. Efficiently locating objects using the Hausdorff distance. Int. J. Comput. Vision. 1997;24:251–270.

[16] Tang YY, Cheng HD, Suen CY. Transformation-ring-projection (TRP) algorithm and its VLSI implementation. Int. J. Pattern Recognit. Artif. Intell. 1991;5:25–56.

[17] Tsai D-M, Chiang C-H. Rotation-invariant pattern matching using wavelet decomposition. Pattern Recognit. Lett. 2002;23:191–201.

[18] Ullah F, Kaneko S. Using orientation codes for rotation-invariant template matching. Pattern Recognit. 2004;37:201–209.

[19] Khotanzad A, Hong YH. Invariant image recognition by Zernike moments. IEEE Trans. Pattern Anal. Mach. Intell. 1990;12:489–497.

[20] Blake G, Dreslinski RG, Mudge T. A survey of multicore parallelism for image correlation. In: El Baz D, Bourgeois J, Spies F, editors. Computer Architecture for Machine Perception, CAMP Proceedings on Seventh International Workshop on Computer Architecture for Machine Perception, CAMP 2005. Los Alamitos (CA): IEEE; 2005. p. 193–197.

[21] Di Stefano L, Mattoccia S, Tombari F. ZNCC-based template matching using fast normalized cross correlation. In: El Baz D, Bourgeois J, Spies F, editors. Proceedings on Seventh International Workshop on Computer Architecture for Machine Perception, CAMP 2005. Los Alamitos (CA): IEEE; 2005. p. 193–197.

[22] Kadidlo J, Strey A. Exploiting data-and thread-level parallelism for image correlation. In: El Baz D, Bourgeois J, Spies F, editors. 16th Euromicro Conference on Parallel, Distributed and Network-Based Processing, PDP. Los Alamitos (CA): IEEE; 2008. p. 407–413.

[23] Chen C-S, Huang C-L, Yeh C-W, et al. An accelerating CPU-based correlation-based image alignment for real-time automatic optical inspection. Comput. Electr. Eng. 2015;49:207–220.

[24] MacLean J, Tsotsos J. Fast pattern recognition using gradient-descent search in an image pyramid. In: Sanfeliu A, Villanueva JJ, Vanrell M, et al., editors. Proceedings on 15th International Conference on Pattern Recognition. Los Alamitos (CA): IEEE; 2000. p. 873–877.

[25] Stefano LD, Mattoccia S, Mola M. An efficient algorithm for exhaustive template matching based on normalized cross correlation. In: Werner B, editor. Proceedings on 12th International Conference on Image Analysis and Processing. Los Alamitos (CA): IEEE; 2003. p. 322–327.

[26] Di Stefano L, Mattoccia S. Fast template matching using bounded partial correlation. Mach. Vision and Appl. 2003;13:213–221.

[27] Mattoccia S, Tombari F, Stefano LD. Fast full-search equivalent template matching by enhanced bounded correlation. IEEE Trans. Image Process. 2008;17:528–538.

[28] Mattoccia S, Tombari F, Di Stefano L. Reliable rejection of mismatching candidates for efficient ZNCC template matching. In: Chellappa R, Girod B, Manjunath BS, et al., editors. ICIP 2008. 15th IEEE International Conference on Image Processing. Los Alamitos (CA): IEEE; 2008. p. 849–852.

[29] Mattoccia S, Tombari F, Di Stefano L. Efficient template matching for multi-channel images. Pattern Recognit. Lett. 2011;32:694–700.

[30] Wei S-D, Pan W-H, Lai S-H. Efficient NCC-based image matching based on novel hierarchical bounds. In: Muneesawang P, Wu F, Kumazawa I, et al. editors. Advances in Multimedia Information Processing-PCM 2009. Berlin: Springer; 2009. p. 807–815.

[31] Mahmood A, Khan S. Early termination algorithms for correlation coefficient based block matching. In: Rodriguez J, editor. IEEE International Conference on Image Processing. ICIP. Los Alamitos (CA): IEEE; 2007. p. 469–472.

[32] Mahmood A, Khan S. Exploiting transitivity of correlation for fast template matching. IEEE Trans. Image Process. 2010;19:2190–2200.

[33] Mahmood A, Khan S. Correlation-coefficient-based fast template matching through partial elimination. IEEE Trans. Image Process. 2012;21:2099–2108.

[34] Sahani SK, Adhikari G, Das B. Fast template matching based on multilevel successive elimination algorithm. In: Sharma V, Tong L, Chockalingam A, editors. International Conference on Signal Processing and Communications (SPCOM). Los Alamitos (CA): IEEE; 2012. p. 1–5.

[35] Konishi Y, Kotake Y, Iijiri Y, et al. Fast and precise template matching based on oriented gradients. In: Fusiello A, Murino V, Cucchiara R, editors. Computer Vision—ECCV Workshops and Demonstrations. Berlin: Springer; 2012. p. 607–610.

[36] Hinterstoisser S, Cagniart C, Ilic S, et al. Gradient response maps for real-time detection of textureless objects. IEEE Trans. Pattern Anal. Mach. Intell. 2012;34:876–888.

[37] Hinterstoisser S, Holzer S, Cagniart C, et al. Multimodal templates for real-time detection of texture-less objects in heavily cluttered scenes. In: Lin S, Schiele B, Soatto S, et al., editors. IEEE International Conference on Computer Vision (ICCV). Los Alamitos (CA): IEEE; 2011. p. 858–865.

[38] Taylor S, Rosten E, Drummond T. Robust feature matching in 2.3 μs. In: Huttenlocher D, Medioni G, Rehg J, editors. IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops CVPR Workshops. Los Alamitos (CA): IEEE; 2009. p. 15–22.
[46] Yoo J, Hwang SS, Kim SD, et al. Scale-invariant template matching using histogram of dominant gradients. Pattern Recognit. 2014;47:3006–3018.

[47] Marimon D, Ebrahimi T. Efficient rotation-discriminative template matching. Progress in Pattern Recognition, Image Analysis and Applications; 2007; Valparaiso, Chile.

[48] Korman S, Reichman D, Tsur G, et al. Fast-match: Fast affine template matching. In: Medioni G, Zabih R, editors. IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Los Alamitos (CA): IEEE; 2013. p. 2331–2338.

[49] Choi M-S, Kim W-Y. A novel two stage template matching method for rotation and illumination invariance. Pattern Recognit. 2002;35:119–129.

[50] Lee W-C, Chen C-H. A fast template matching method for rotation invariance using two-stage process. In: Chen Y-W, Pan J-S, editors. Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing IIH-MSP ’09. Los Alamitos (CA): IEEE; 2009. p. 9–12.

[51] 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.

[52] Rublee E, Rabaud V, Konolige K, et al. ORB: an efficient alternative to SIFT or SURF. In: Metaxas D, Quan L, Sanfeliz A, et al., editors. IEEE International Conference on Computer Vision (ICCV). Los Alamitos (CA): IEEE; 2011. p. 2564–2571.

[53] Azad P, Asfour T, Dillmann R. Combining Harris interest points and the SIFT descriptor for fast scale-invariant object recognition. In: Xi N, Hamel WR, editors. IEEE/RSJ International Conference on Intelligent Robots and Systems IROS. Los Alamitos (CA): IEEE; 2008. p. 4275–4280.

[54] Muja M, Rusu RB, Bradski G, et al. REIN-A fast, robust, scalable Recognition Infrastructure. In: Li Z, Zheng YF, editors. IEEE International Conference on Robotics and Automation (ICRA). Los Alamitos (CA): IEEE; 2011. p. 2939–2946.

[55] Lampert CH, Blaschko MB, Hofmann T. Beyond sliding windows: Object localisation by efficient sub window search. In: Huang T, Ahuja N, Shapiro L, editors. IEEE Conference on Computer Vision and Pattern Recognition CVPR. Los Alamitos (CA): IEEE; 2008. p. 1–8.

[56] Sassanapitak S, Kaewtrakulpong P. An efficient translation-rotation template matching using pre-computed scores of rotated templates. In: Krahkirkish M, editor. 6th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology ECTI-CON. Los Alamitos (CA): IEEE; 2009. p. 1040–1043.

[57] Cui X, Kim H, Park E, et al. Robust and accurate pattern matching in fuzzy space for fiducial mark alignment. In: Sukthankar R, editor. Machine vision and applications. Berlin: Springer; 2013. p. 447–459.

[58] Rusu RB, Cousins S. 3d is here: point cloud library (PCL). In: Li Z, editor. IEEE International Conference on Robotics and Automation (ICRA). Los Alamitos (CA): IEEE; 2011. p. 1–4.

[59] Bayramoglu N. Shape index SIFT: range image recognition using local features. In: Erçil A, editor. 20th International Conference on Pattern Recognition (ICPR). Los Alamitos (CA): IEEE; 2010. p. 352–355.

[60] Drost B, Ulrich M, Navab N, et al. Model globally, match locally: efficient and robust 3D object recognition. In: Davis L, Malik J, editors. IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Los Alamitos (CA): IEEE; 2010. p. 998–1005.

[61] Rusu RB, Blodow N, Marton ZC, et al. Aligning point cloud views using persistent feature histograms. In: Chatila R, Merlet J-P, editors. IEEE/RSJ International Conference on Intelligent Robots and Systems IROS. Los Alamitos (CA): IEEE; 2008. p. 3384–3391.

[62] Rusu RB, Marton ZC, Blodow N, et al. Learning informative point classes for the acquisition of object model maps. In: Soh YC, Nguyen KS, editors. 10th International Conference on Control, Automation, Robotics and Vision ICARCV. Los Alamitos (CA): IEEE; 2008. p. 643–650.

[63] Rusu RB, Blodow N, Beetz M. In fast point feature histograms (FPFH) for 3D registration. In: Kosuge K, editor. IEEE International Conference on Robotics and Automation ICRA’09. Los Alamitos (CA): IEEE; 2009. p. 3212–3217.

[64] Skotheim Ø, Lind M, Ystgaard P, et al. A flexible 3D object localization system for industrial part handling. In: de Almeida AT, Nunes U, editors. IEEE/RSJ International Conference on Intelligent Robots and Systems IROS. Los Alamitos (CA): IEEE; 2012. p. 3326–3333.

[65] Choi C, Taguchi Y, Tuzel O, et al. Voting-based pose estimation for robotic assembly using a 3D sensor. In: Papanikolopoulos N, Christensen H, Sugano S, et al., editors. IEEE International Conference on Robotics and Automation (ICRA). Los Alamitos (CA): IEEE; 2012. p. 1724–1731.

[66] Morell-Gimenez V, Saval-Calvo M, Azorin-Lopez J, et al. A comparative study of registration methods for RGB-D video of static scenes. In: Passaro VMN, Reindl LM, Melesse AM, et al., editors. Sensors. Basel: MDPI; 2014. p. 8547–8576.

[67] Rusu RB, Bradski G, Thibaux R, et al. Fast 3d recognition and pose using the viewpoint feature histogram. In: Luo RC, Asama H, editors. IEEE/RSJ International Conference on Intelligent Robots and Systems IROS. Los Alamitos (CA): IEEE; 2010. p. 2155–2162.

[68] Aldoma A, Vincze M, Blodow N, et al., CAD-model recognition and 6DOF pose estimation using 3D cues. In: Metaxas D, Quan L, Sanfeliz A, et al., editors. IEEE International Conference on Computer Vision Workshops (ICCV Workshops). Los Alamitos (CA): IEEE; 2011. p. 585–592.

[69] Lysenkov I, Erushinov V, Bradski G. Recognition and pose estimation of rigid transparent objects with a kinect sensor. In: Roy N, Newman P, Srinivasa S, editors. Robotics: science and systems VIII. Cambridge: MIT Press; 2013. p. 273–280.

[70] Treiber MA. An introduction to object recognition: selected algorithms for a wide variety of applications. London: Springer Science & Business Media; 2010.

[71] Moreno-Noguer F, Lepetit V, Fua P. Accurate non-iterative o (n) solution to the pnp problem. In: Davis L, Bouthemy P, Ikeuchi K, editors. IEEE 11th International Conference on Computer Vision ICCV. Los Alamitos (CA): IEEE; 2007. p. 1–8.
[72] Lepetit V, Moreno-Noguer F, Fua P. EPnP: An Accurate O(n) Solution to the PnP Problem. Int. J. Comput. Vision. 2009;81:155–166.

[73] Li S, Xu C, Xie M. A robust O (n) solution to the perspective-n-point problem. IEEE Trans. Pattern Anal. Mach. Intell.. 2012;34:1444–1450.

[74] Hartley R, Zisserman A. Multiple view geometry in computer vision. Cambridge: Cambridge University Press; 2003.

[75] Agarwal A, Jawahar C, Narayanan P. A survey of planar homography estimation techniques. Hyderabad: Deemed University, Centre for Visual Information Technology; 2005.

[76] Malis E, Vargas M. Deeper understanding of the homography decomposition for vision-based control. INRIA Research Report #6303; 2007.

[77] Press WH, Teukolsky SA, Vetterling WT, et al. Numerical recipes 3rd edition: the art of scientific computing. Cambridge: Cambridge University Press; 2007.

[78] Holzer S, Hinterstoisser S, Ilic S, et al. Distance transform templates for object detection and pose estimation. In: Huttenlocher D, Medioni G, Rehg J, editors. IEEE Conference on Computer Vision and Pattern Recognition CVPR. Los Alamitos (CA): IEEE; 2009. p. 1177–1184.

[79] Steger C. Occlusion, clutter, and illumination invariant object recognition. Int. Arch. Photogramm. Remote Sen. Spatial Inf. Sci. 2002;34:345–350.

[80] Moré JJ. The Levenberg-Marquardt algorithm: implementation and theory. Heidelberg: Springer; 1978.

[81] Hinterstoisser S, Lepetit V, Ilic S, et al. Model based training, detection and pose estimation of texture-less 3D objects in heavily cluttered scenes. Presented at: Computer Vision – ACCV 2012: 11th Asian Conference on Computer Vision; 2013 Nov 5–9; Daejeon, Korea.

[82] Ulrich M, Wiedemann C, Steger C. Combining scale-space and similarity-based aspect graphs for fast 3D object recognition. IEEE Trans. Pattern Anal. Mach. Intell. 2012;34:1902–1914.