Single View Correspondence Matching for Non-Coplanar Circles Using Euclidean Invariants

BMVC 2014 Submission # 123

Abstract

In this work we introduce a method to determine 2D-3D correspondences for non-coplanar circles using a single image, given that the 3D information is known. The core idea of our method is to compute 3D information from 2D features, thereby transforming a 2D-3D problem to a 3D-3D problem. Earlier researchers suggested that a pair of non-coplanar circles preserves Euclidean invariants under perspective projection. These invariants can be extracted from their image projections, but with a two fold ambiguity. In this paper, we propose *Conic pair descriptor*, which computes unique Euclidean invariants from known circular features on the 3D model and invariants with two fold ambiguity from its image projections. In this paper, we include a detailed account of factors affecting the computation of invariants from conic projections. We conduct experiments on real and synthetic models, in order to evaluate the proposed method. The experiment with synthetic images focuses on showing the impact of the size and plane orientation of the circles on the success of descriptor matching. 3D models with artificial circular markers are prepared and 3D data is measured by using a photogrammetric measurement system. The results of the correspondence matching algorithm are compared against the ground truth, which is also generated from the measurement system. We also show that our method is robust against false positives and capable of supporting real-time applications.

1 Introduction

Correspondence matching is one of the key problems in computer vision. Various vision problems such as pose estimation, object detection or model reconstruction rely on correspondence matching. Correspondence matching is identifying same set of features, either between two or more images (2D-2D matching) or between a model and its image (2D-3D matching). Perspective projection differs the appearance of an object in the images, therefore correspondence matching needs feature descriptors that are immune to the projection. Such feature descriptors consist of invariant quantities which can be computed from certain image features and matched directly with those obtained from the object features in case of the 2D-3D matching. This paper focuses on the 2D-3D matching problem.

There are mainly two types of feature descriptors. One computes invariant quantities defined by local texture in an image such as SIFT descriptor [13]. This type of feature descriptors is supposed to be used for textured scenes and textured objects so that we can use

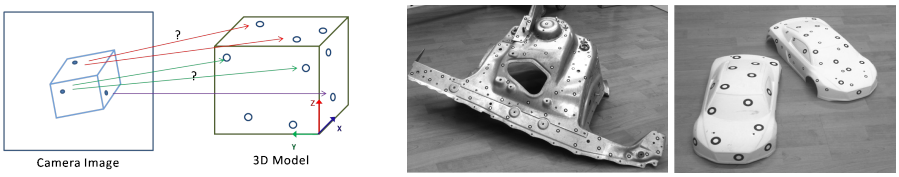


Figure 1: A primitive example explaining the correspondence problem when circular features exist on different planes of the model. Image on the left shows the problem with a simple cube. The other two images show complex 3D models with circular markers, Such markers are widely used in Industry for 3D surface measurements.

the descriptor for mobile applications, in which assumption on object shape is disfavored. Such methods focus on improving their invariance as well as efficient memory usage so that the matching can run even on powerless processors in mobile phones [3, 8]. One of the disadvantages is that their invariances are up to 2D transformation such as affine and rotation transformation around camera coordinate system. To handle view change caused by perspective transformation, we must explicitly learn how invariant quantities are affected by perspective transformation [9, 10].

The other type of feature descriptors assume primitive features such as points, lines or conics to be present in the scene or on the target object. Advantage of using such specific features is that they are easier to detect from images. Detailed studies on invariant quantities of planar and non-coplanar features as well as their stability under perspective projection have been conducted [6, 7]. The descriptors generated by plane projective groups have been proposed in various applications for correspondence matching [15, 19, 22]. In Augmented Reality [19, 20] and industrial tracking applications [1], planar patterns with lines or circles are specifically designed to support invariant computation for correspondence matching.

In industrial scenario, circles are widely present on a model, tracking target, as natural features or circular markers are attached to the model for photogrammetric measurements [14] as shown in Fig. 1. The state of the art measurement system involves taking multiple images of the model with encoded markers temporarily added in the scene to solve the correspondence problem. The encoded markers are removed after the measurements are done. In such cases it is favourable to build a tracking approach without the support of coded markers. A single view correspondence matching problem with coplanar circles and ellipses has widely been studied [5, 6, 11]. However matching problem with non-coplanar circles has not been addressed. Figure 1 shows an example of the matching problem, where multiple identical circular features exist on different planes of a 3D model. In this case, features existing on the model is an advantage, but coplanar invariants can not be used for correspondence matching. In our work we focus on solving this problem of single view matching of multiple non-coplanar conics.

In case of non-coplanar features on 3D objects, Euclidean invariants are preserved rather than projective invariants [6]. Euclidean invariants are difficult to extract from images due to perspective mapping, however circles are a special case. A world circle always produces an elliptical curve on the image plane. The ellipse in the image can be backprojected and the plane of the circle can be defined in camera frame with a two fold ambiguity [6, 13]. Further, Forsyth et al. [6] showed that for a given pair of non-coplanar circles in 3D, the angle between the circle planes and the distance between their centre positions are invariant quan-

tities. These invariants can be recovered from image projections of conics with a two fold ambiguity. The concept was proposed in early 90s, however these invariants have remained unexplored.

We propose using the aforementioned Euclidean invariants to solve the 2D-3D correspondence problem when multiple non-coplanar circular features exist on a model. In our approach we bring the problem from 2D to 3D by backprojection image conics, then compute Euclidean invariant quantities to solve the matching problem from a single image. In this paper, We introduce the *Conic pair descriptor*, which encapsulates the invariants computed from elliptical image features. Our contribution is a new method to accurately identify image correspondences when multiple identical non-coplanar circular features exist in the scene. The method assumes that the camera intrinsics are known and 3D information (i.e. size, normals, centre position) of features is available. In industry based model tracking application, the 3D-CAD data is often available. The evaluation consists of synthetic experiments to understand factors influencing invariant computation and matching. Additionally, real 3D models with circular markers as shown in Fig. 1 are used to evaluate the matching method. The proposed method can find corresponding circular markers from a single image, with high precision. We also show that our method is stable against false positives and that it is fast enough to support real-time tracking applications.

2 Conic Invariants : Theory and Computation

This section explains the theory of conic invariants derived by Forsyth et al. [6]. In this section we will describe Euclidean invariants preserved under perspective transformation defined by a pair of non-coplanar circles.

A rigid 3D model can be assumed as a set of rigidly coupled planes. Suppose there exists a set of circles on different planes of the model and the normal vector $N_i \in \mathbb{R}^3$ and the centre position $M_i \in \mathbb{R}^3$ is known in object frame. Forsyth et al. pointed out that between each pair of non-coplanar circles the following Euclidean invariant quantities are preserved under perspective transformation: angle between the conic planes θ and the distance between their centre positions d. The angle between the planes is equivalent to the angle between the surface normals, thus is obtained as

$$\theta_{i,j} = \angle(N_i, N_j), \quad i \neq j, \tag{1}$$

where i and j represent the index of the conics. The distance between the conic centres d is computed as

$$d_{i,j} = \text{dist}(M_i, M_j), \quad i \neq j. \tag{2}$$

We can recover both invariants, the angle θ and the distance d, from image projections if the size of the circles is known. This is possible because image projection of a circle is always elliptical. This property can be exploited further by backprojecting the ellipse in camera space ($\mathbb{R}^3$) to obtain plane orientation of the circle, using ellipse parameters C_i and centre position m_i of each image conic. A method to backproject ellipses is described in [6], we can compute the normal vector $Nc_i{}^1, Nc_i{}^2 \in \mathbb{R}^3$ and the centre position $Mc_i{}^1, Mc_i{}^2 \in \mathbb{R}^3$ of i-th ellipse up to a two fold ambiguity as

$$\{(Nc_i{}^a, Mc_i{}^a)\} = \text{EllipseBackprojection}(m_i, C_i), \quad a = \{1, 2\}, \tag{3}$$

where the superscripts a, b represent the ambiguous solutions. The two fold ambiguity is referred as *Conic ambiguity* in this paper. Once the normals and the centre positions are recovered for each conic, we can compute the angle $\theta_{i,j}$ and the distance $d_{i,j}$ between the two conics as Eqs. 1 and 2.

$$\begin{aligned} d_{i,j}^{a,b} &= \text{dist}(Mc_i{}^a, Mc_j{}^b), \\ \theta_{i,j}^{a,b} &= \angle(Nc_i{}^a, Nc_j{}^b), \end{aligned} \qquad i \neq j, \quad a,b = \{1,2\}. \tag{4}$$

It is evident that due to *Conic ambiguity* we have four solutions for each invariant. Forsyth et al. explained that the solutions for $d_{i,j}^{a,b}$ are consistent, whereas only one solution of $\theta_{i,j}^{a,b}$ is correct [6]. Following this, we regard that a pair of two non-coplanar conics in an image derives a unique $d_{i,j}$ and four $\theta_{i,j}^{a,b}$.

3 Method

This section first introduces the proposed descriptor, called *Conic pair descriptor*, in Sec. 3.1. Further, the matching method using *Conic pair descriptors* is explained in detail in Sec. 3.2. We assume that the 2D data from the image and 3D data from the model is already available. The 3D data includes the surface normal N_i, centre position M_i and size R_i (diameter) of each circle on the model. The 2D data includes circle centre m_i and conic matrix C_i. The surface normals $Nc_i{}^a$ and centre positions $Mc_i{}^a$ are recovered from the 2D data, where $a = 1,2$ denotes the index of ambiguous solution. The 2D data is extracted from an input image by the following procedure: ellipse detection given an input image [16]; conic parameters estimation from the detected ellipse [4]; and the normal vector and the centre position recovery from the conic parameters [12].

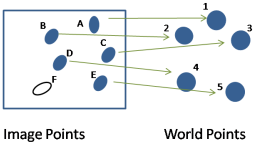
3.1 Descriptor Generation

This part mainly discusses the generation of *Conic pair descriptor* from conic invariants. The invariants for 3D model are computed from available 3D data (M_i, N_i) using Eq. 5. This descriptor does not contain any ambiguity. The same set of invariants can be computed from corresponding image features using Eq. 4, where the recovered d component is unique and θ component has four different solutions (Eq. 6). This *Conic ambiguity* can not be resolved without knowing the correspondence $m_i \leftrightarrow M_i$. In our approach, we pursue the idea that the existence of multiple features on the model can be used to find correspondences without solving the conic ambiguity. The conic invariants are used to generate *Conic pair descriptors*, V and v represent the descriptors computed from the model data and the image data respectively. The principle idea is to perform a descriptor matching between v and V to obtain $m_i \leftrightarrow M_i$ correspondence. The proposed *Conic pair descriptor* structure,

$$\textit{Conic pair descriptor}_{model} = V_p = \langle d_{i,j}, \theta_{i,j} \rangle \tag{5}$$

$$\textit{Conic pair descriptor}_{image} = v_q = \langle d_{i,j}, \theta_{i,j}^{1,1}, \theta_{i,j}^{1,2}, \theta_{i,j}^{2,1}, \theta_{i,j}^{2,2} \rangle \tag{6}$$

where index p represents world circles i, j and index q represents image conic pair i, j. The reader should note that given a unique set of points and their corresponding normals in 3D camera space (e.g. depth images), PFH descriptor [17] computes similar invariants. In our case, we use a single 2D image to extract ambiguous centre and normal information in



Step 1 : Pairwise Initial Match

ID	V <-> v	ID	V <-> v
H_1	V_{12} <-> v_{AB}	H_7	V_{25} <-> v_{BE}
H_2	V_{13} <-> v_{AC}	H_8	V_{35} <-> v_{CE}
H_3	V_{23} <-> v_{BC}	H_9	V_{12} <-> v_{DF}
H_4	V_{14} <-> v_{AD}	H_{10}	V_{14} <-> v_{CF}
H_5	V_{24} <-> v_{BD}	H_{11}	V_{24} <-> v_{CD}
H_6	V_{34} <-> v_{CD}		

Step 2 : Pointwise Triplet Matching

V <-> v Triplet	M <-> m Triplet
H_1- H_2- H_3	A-1, B-2,C-3
H_1- H_4- H_5	A-1, B-2,D-4
H_2- H_4- H_6	A-1,C-3,D-4
H_3- H_7- H_8	B-2,C-3,E-5
H_9- H_{10}- H_{11}	F-1,D-2,C-4

Step 3 : Correspondence Hypothesis

	A	B	C	D	E	F
1	3	0	0	0	0	1
2	0	3	0	1	0	0
3	0	0	3	0	0	0
4	0	0	1	2	0	0
5	0	0	0	0	1	0

Figure 2: The image shows an example matching problem with real correspondence relations, Image has one false positive detection F. A pictorial overview of the three steps used in the matching process are shown with respect to the given problem. H_i represents index of matched descriptor pair.

3D camera space. The *Conic ambiguity* introduced by *Ellipse backprojection* forces us to include ambiguity in the descriptor structure (Eq. 6). Unlike popular methods, a *Conic pair descriptor* represents two features at the same time. In order to uniquely represent a single conic using Euclidean invariants, at least three conic features are required. The addition of each conic feature adds 3 wrong solutions of θ in the descriptor, moreover the matching must rely on detection of all the conics used for descriptor computation. Descriptors $v_{\{1..q\}}$ are generated for each pair of detected n image conics, where $q = \binom{n}{2}$. $V_{\{1...p\}}$ are generated off-line as the 3D model data is already available. In case of l circles on the model $p \leq \binom{l}{2}$, as the circle pair not likely to appear in the same image can be rejected. After computing the *Conic pair descriptors* the following 3 step matching approach is used to achieve $m_i \leftrightarrow M_i$ correspondences.

3.2 Descriptor matching

3.2.1 Step 1: Pairwise Initial Matching

In this stage we compare the *Conic pair descriptors* and find all possible $v \leftrightarrow V$ relationships. Each descriptor represents a pair of conics, therefore this step is called pairwise matching. The objective is to reduce the complexity of the problem by finding possible pair correspondences ($v \leftrightarrow V$). Figure 2 shows the possible outcome of step 1 with respect to the problem shown in the figure. First, the unique d component between the descriptors is compared with T_d as threshold. If a match is found, the unique θ component of V is compared with all four values of θ component in v with T_θ as threshold. The match is considered positive if any one of the four θ values match, this may result into false matching as only one of the θ is correct for a given pair. This explains that *Conic ambiguity* may contribute into false matching. The results may show *one to many* type of relation between the descriptors, observe H_4 and H_{10} in Fig. 2. False descriptor matching can also happen if false positives are present in the image. Similarly, symmetric orientation of features on the object may generate same invariants between multiple pairs, which can also contribute towards false descriptor matching.

3.2.2 Step 2: Pointwise Triplet Matching

In this stage we simplify the problem further and obtain hypothesis on pointwise matching ($m_i \leftrightarrow M_i$) by performing a verification on $v \leftrightarrow V$ matching results. The objective is to compare the results of step 1 to identify correspondences and reject the false descriptor matches. We seek three $v \leftrightarrow V$ results, such that they complement each other to form a unique three point $m_i \leftrightarrow M_i$ hypothesis, as shown in Fig. 2. A simple two stage approach is used to generate a triplet matching hypothesis.

Stage 1 Find any two results of Pairwise Initial Matching in which both the image and the model descriptors represent one and only one common conic. If such results exist then an initial triplet matching hypothesis can be proposed. As example we consider results H_1, H_2 from Fig. 2.

$$V_{12} \leftrightarrow v_{AB}, V_{13} \leftrightarrow v_{AC} \xrightarrow{\text{Triplet Hypothesis}} [1\ 2\ 3] \leftrightarrow [A\ B\ C]$$

In the example above we can see that conic 1 is common in V and image conic A is common in v among the two solutions. We can form an initial 3 point matching hypothesis with these results.

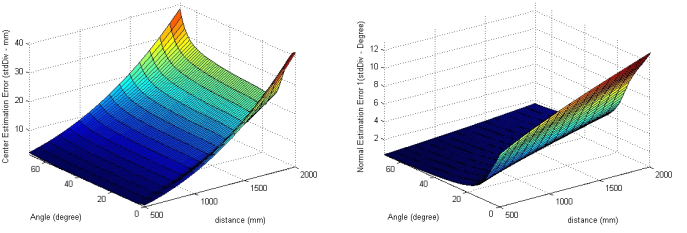
Stage 2 Find a new descriptor matching pair which can verify the triplet matching hypothesis formed in the previous stage. In context of the example given in stage 1, the result H_3 (Fig. 2) can verify hypothesis of stage 1. The verified triplets are saved and others are rejected. The results may also contain false triplet matches, $[F\ D\ C] \leftrightarrow [1\ 2\ 4]$ as depicted in Fig.2 is a wrong triplet.

3.2.3 Step 3: Correspondence Hypothesis

In this final step results of triplet matching are combined and a voting matrix is generated. For each $m_i \leftrightarrow M_i$ correspondence verified in step 2, one vote is added to the corresponding matrix entry. As a result, all unique $m_i \leftrightarrow M_i$ pair gaining maximum votes in the matrix are proposed as a correspondence hypothesis. In Fig. 2 the results are : A↔1, B↔2, C↔3 , D↔4, E↔5. A minimum of three correspondences are required compute a pose of the object [□], if the camera intrinsics are known. We propose computing the pose by selecting the three most voted $m_i \leftrightarrow M_i$ results from the matrix. Other correspondences having lower vote count can be verified by using the pose. In Fig. 2, A↔1, B↔2 and C↔3 can be selected to compute the pose P , and D↔4, E↔5 relation can be verified using the pose P. If only 3 out of n conics are detected in the image, the pose can not be verified.

4 Evaluation

In this section we will cover experiments carried out to comment on accuracy and robustness of the algorithm. The reader should note that the problem of achieving single image 2D-3D correspondences for non-coplanar circles has not been addressed earlier. Therefore, alternative methods for comparison are not available. We prepared two car models by attaching circular markers as shown in Fig. 1. Model 1 has 20 markers of size R_i = 12 mm and Model 2 has 26 markers of size R_i = 5mm. These markers are widely accepted and used in the industrial domain for photogrammetric measurements. 3D measurements of the markers are done



(a) Centre estimation error (b) Normal estimation error

Figure 3: Ellipse backprojection results, Image noise = $\sigma = 0.3$, $R_i = 12mm$

with state of the art metrology system, and the ground truth is established by giving each model point a unique ID in the database. The markers are attached randomly and coplanar placement is avoided. A high resolution (2560x1920 pix) camera is used for the experiments and MATLAB [2] is used for synthetic experiments.

4.1 Preliminary experiment

The quality of recovered plane from *Ellipse backprojection* depends on distance from the camera r and the viewing angle η (angle between the image plane and the circle plane)[21]. We performed simulations to understand the behaviour of *Ellipse backprojection* with respect to both r and η. The parameter r is varied from 500 to 2000 mm and η from 0-70° in step wise manner, 100 iterations are performed at each position. The results (Fig. 3) suggest that, at low viewing angles (η) 0-10° both normal and centre estimation errors are higher, at any given distance. When $\eta \leq 10$, the image projection of a circle is more circular than elliptical, therefore recovery of ellipse parameters may have errors. The estimation error grows at higher camera distance, however the error in normal recovery appears less sensitive to increase in camera distance than the error in centre recovery. The results obtained with $R_i = 5$ mm show a similar pattern, although the magnitude of the error is higher as smaller image projections reduce the accuracy of computation of ellipse parameters. We also compared the ambiguous results of estimated centres $Mc_i{}^1$ and $Mc_i{}^2$. The maximum distance recorded between the two is ≤ 0.1 mm for $R_i = 12$ mm, this suggests that ambiguity can be neglected for the recovered centre position. This behaviour also explains consistency of invariant d between two image conics.

4.2 Correspondence Matching vs Threshold Settings

The aim of this experiment is to understand the role of the threshold values T_d and T_θ in $m_i \leftrightarrow M_i$ matching results. In order to perform this experiment we took 75 images of Model 1 from different camera positions (Distance Range 500-2000 mm). As suggested in Sec. 3.2.3, each image has at least four detected conics. The results are considered *Not Converged* (NC) in case of less than three $m_i \leftrightarrow M_i$ results.

The results (Table. 1) show that relaxed values of T_d impacts both precision and recall values in negative manner. On the other hand, very stringent thresholds lead to lower recall values. Therefore, a right balance of threshold can be selected to achieve higher precision and recall rates. Our preferred settings for experiments is $T_d = 10$ and $T_\theta = 5$. The selection may require change based on density of the features.

Table 1: Correspondence matching with varying threshold settings

T_θ	T_d	NC	Positive	FP	Precision	Recall
5	5	13	62	0	100	82.6
5	10	8	67	0	100	89.33
5	15	4	67	4	94.36	89.33
3	5	28	47	0	100	62.66
3	10	21	54	0	100	72
3	15	19	53	3	94.64	70.66

Table 2: Descriptor Matching Analysis

R_i (mm)	θ	d (mm)	Min-Max Success(%)
5	10°-40°	10-150	58-82
	40°-80°	10-150	35-65
12	10°-40°	10-150	64-86
	40°-80°	10-150	40-69
20	10°-40°	10-150	74-92
	40°-80°	10-150	48-77

4.3 Descriptor Matching vs Marker Orientation

This experiment is carried out synthetically to observe the effect of orientation of the circles on the individual descriptor matching. Two circle planes are placed in different orientations and images are rendered from 1000 random camera positions for each orientation. Gaussian noise is added to the images to simulate the camera behaviour. The control parameters, the distance between circle centres d is varied from 10 - 150 (mm) and the angle between the planes θ is varied from 10 - 80°. The objective is to recover the descriptor components d and θ from the images, compare them with the ground truth and measure the success rate. Realistic values are used for camera intrinsics, $T_d = 10$ and $T_\theta = 5$ are kept constant. Camera positions are chosen at random, x-y-z rotation range is $\pm 70°$, x-y translation range is ± 500 mm, z-translation (Camera distance) range is 500 to 2000 mm.

The table 2 provides summary of key observations made during the experiment. The matching success shows inversely proportional relation with θ, independent of d. We learn that descriptor matching is influenced more by the angle between the planes than the distance between the circle centres. It is also seen that success of matching can be improved by increasing the size of the circles. This experiments suggest that when feature placement is possible, it is advised to choose larger circles or surfaces with lower plane angles for improved matching results.

4.4 Robustness against false positives

This experiment aims to show the robustness of the matching method against presence of false positives in the scene. In order to introduce false positives in the scene, Model 1 and Model 2 are placed in the same scene and images are captured from different positions. The matching method is provided 3D information of one model at a time, which in turn make the markers present on the other model act as false positives in the image. The same set of images are used for the two experiments and the results of the experiment are shown in Figure 4. We learn that the false positives are completely rejected when matching is focused on Model

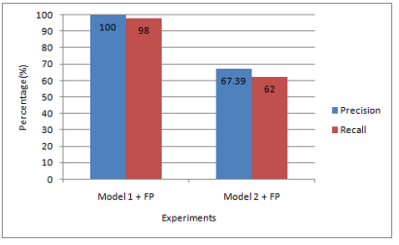


Figure 4: Robustness against False Positives

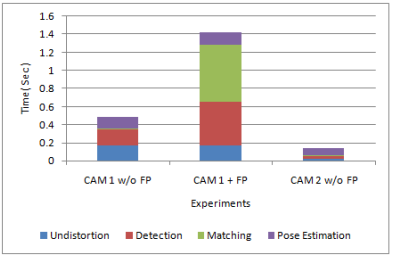


Figure 5: Time Analysis

1, On the contrary in case of Model 2, precision and recall values suffer due to presence of false positives from Model 1. Markers on Model 1 have bigger size and therefore invariant recovery is strong. This can explain higher robustness of Model 1 against false positives.

4.5 Time Analysis

In this experiment we measure the time consumed by the matching method when introduced into a tracking application. Two cameras CAM 1 (2560 x 1920 pix) and CAM 2 (640x480) are used for tracking Model 1, the results presented are averaged over 100 frames. Tracking process is divided into 4 stages, the graph in Fig. 5 depicts the time consumed by each stage. The results show that the *Matching* stage takes 3.09 ms with CAM 1 and 1.88 ms with CAM 2, which is $\leq 2\%$ of the total tracking time. In terms of performance, we achieve tracking speed of 3 FPS with CAM 1 and 8 FPS with CAM 2. The computer used has Intel Core i5 2.8 GHz processor with 8 GB RAM. Additionally, an experiment is performed with CAM 1 by adding 90-140 false positives in the scene. In this case, the *Matching* stage is exhausted and consumes 6.277 seconds out of the total tracking time of 6.7599 seconds. CAM 2 has limited tracking range (i.e. < 1000 *mm*) and low resolution, therefore accommodating 90-140 false positives in one scene within the tracking range is not possible.

5 Conclusion & Future work

In this paper we have demonstrated a successful approach for solving 2D-3D correspondence matching problem for non-coplanar circular features from a single image. We propose a new *Conic pair descriptor* which represents Euclidean invariants generated by a pair of non-coplanar circles. Our method can successfully define correspondences when more than three circular features are present in the scene. The proposed method is the first to address the correspondence matching using these invariants since its introduction in the 90s. Our contribution also includes providing detailed understanding of behaviour of invariants with respect to orientation and size of the circles. The major factors affecting matching are also discussed to optimize the method for best possible matching results based on the application (threshold settings, circle size, camera distance). The results of the experiments support our claim, that the method is fast, reliable and robust against false positives. Our method can be used for object tracking or object identification in the industrial environment, where natural or artificial circular features exist predominantly on the models. However, the method is generic and can be used for any application dealing with non-coplanar circles. The 3D information of the features on the model and camera calibration are the only prerequisites for

matching. The reader should also note that algorithm may not perform well with symmetric or coplanar arrangement of circular features.

In context of future work, we would like to improve the method to be able to handle features of different sizes simultaneously. Also a faster matching strategy is required to handle large number of feature points and false positives. We would like to use the same invariants to compute 2D-2D correspondence matching between two images in order to generate the 3D data which is a prerequisite now. We also consider using such a matching algorithm to support creating 3D markers for monocular Augmented Reality applications. This can be a cheap alternative to conventionally used 3D spherical markers.

References

[1] Aicon 3d systems. URL http://www.aicon3d.de/start.html.

[2] Mathworks. URL http://www.mathworks.de.

[3] Michael Calonder, Vincent Lepetit, Christoph Strecha, and Pascal Fua. Brief: binary robust independent elementary features. In Proceedings of the 11th European conference on Computer vision: Part IV, ECCV'10, pages 778–792, Berlin, Heidelberg, 2010. Springer-Verlag. ISBN 3-642-15560-X, 978-3-642-15560-4. URL http://dl.acm.org/citation.cfm?id=1888089.1888148.

[4] Gerald E Farin and Dianne Hansford. The geometry toolbox for graphics and modeling. A.K. Peters, Natick, Mass., 1998. ISBN 1568810741 9781568810744.

[5] M. Ferri, F. Mangili, and G. Viano. Projective pose estimation of linear and quadratic primitives in monocular computer vision. CVGIP: Image Underst., 58(1):66–84, July 1993. ISSN 1049-9660.

[6] D. Forsyth, J.L. Mundy, A. Zisserman, C. Coelho, A. Heller, and C. Rothwell. Invariant descriptors for 3d object recognition and pose. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 13(10):971–991, 1991.

[7] Patrick Gros and Long Quan. Projective invariants for vision. Technical Report RT 90 IMAG - 15 LIFIA, MOVI - IMAG-INRIA Rhône-Alpes / GRAVIR , Laboratoire d'Informatique Fondamentale et d'Intelligence Artificielle - LIFIA, 1992.

[8] Yan Ke and Rahul Sukthankar. Pca-sift: a more distinctive representation for local image descriptors. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), volume 2, pages II–506–II–513 Vol.2, 2004. doi: 10.1109/CVPR.2004.1315206.

[9] Daniel Kurz, Thomas Olszamowski, and Selim Benhimane. Representative feature descriptor sets for robust handheld camera localization. In IEEE International Symposium on Mixed and Augmented Reality (ISMAR), pages 65–70, 2012. doi: 10.1109/ISMAR.2012.6402540.

[10] V. Lepetit and P. Fua. Keypoint recognition using randomized trees. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 28(9):1465–1479, 2006. ISSN 0162-8828. doi: 10.1109/TPAMI.2006.188.

[11] Vincent Lepetit and Pascal Fua. Monocular model-based 3d tracking of rigid objects: A survey. Now Publishers Inc, 2005.

[12] Diego Lo\' pez de Ipin a, Paulo RS Mendon\cca, Andy Hopper, and Andy Hopper. TRIP: a low-cost vision-based location system for ubiquitous computing. Personal and Ubiquitous Computing, 6(3):206–219, 2002.

[13] D.G. Lowe. Object recognition from local scale-invariant features. In Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on, volume 2, pages 1150–1157 vol.2, 1999. doi: 10.1109/ICCV.1999.790410.

[14] Thomas Luhmann, Stuart Robson, and Stephen Kyle: Ian Harley. Close Range Photogrammetry: Principles, Techniques and Applications: Principles, Methods and Applications. Whittles Publishing, revised edition edition, October 2006. ISBN 1870325508.

[15] Chikara Matsunaga, Yasushi Kanazawa, and Kenichi Kanatani. Optimal grid pattern for automated camera calibration using cross ratio. IEICE Transactions on Fundamentals, E83–A:1921–1928, 2000.

[16] Leonid Naimark and Eric Foxlin. Circular data matrix fiducial system and robust image processing for a wearable vision-inertial self-tracker. In Proceedings of the 1st International Symposium on Mixed and Augmented Reality, page 27, 2002.

[17] Radu Bogdan Rusu. Semantic 3D Object Maps for Everyday Manipulation in Human Living Environments. PhD thesis, Computer Science department, Technische Universitaet Muenchen, Germany, October 2009.

[18] Reza Safaee-Rad, Ivo Tchoukanov, Kenneth Carless Smith, and Bensiyon Benhabib. Three-dimensional location estimation of circular features for machine vision. Robotics and Automation, IEEE Transactions on, 8(5):624–640, 1992. URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=163786.

[19] Hideaki Uchiyama and Hideo Saito. Random dot markers. In Virtual Reality Conference (VR), IEEE, pages 35–38, 2011.

[20] Arjen van Rhijn and Jurriaan D. Mulder. Optical tracking using line pencil fiducials. In Proceedings of the Tenth Eurographics conference on Virtual Environments, page 35–44, 2004. Cited by 0008.

[21] Naoufel Werghi, Christophe Doignon, and Gabriel Abba. Pose estimation of objects based on circular patterns in monocular computer vision. In SPIE's 1996 International Symposium on Optical Science, Engineering, and Instrumentation, pages 254–257. International Society for Optics and Photonics, 1996. 00002.

[22] Xianghua Ying and Hongbin Zha. Camera calibration using principal-axes aligned conics. In Asian Converence on Computer Vision, pages 138–148, 2007.