

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 correspondence 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* based on the Euclidean invariants. The proposed descriptor computes unique Euclidean invariants from known 3D model and Euclidean invariants with two fold ambiguity from its image projections. The proposed matching approach follows three steps to obtain correspondences between the circular features against the ambiguity. In this paper, we have included a detailed account of factors affecting the computation of invariants from conic projections. We have conducted 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. We prepared 3D models with artificial circular features and obtained the ground truth 3D data with a Photogrammetric measurement system. The results of the correspondence matching algorithm are evaluated against the ground truth. 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. Pose estimation and object recognition applications require accurate knowledge of correspondence relation [Yuji: $(m_i \leftrightarrow M_i)$] $m_i \leftrightarrow M_i$ between 3D model features [Yuji: (M_i)] M_i and 2D image features [Yuji: (m_i)] m_i . In case of monocular systems the [Yuji: $m_i \leftrightarrow M_i$]2D-3D correspondence problem becomes more challenging as the depth information is lost. Various Computer Vision and Augmented Reality applications demand correspondence matching from a single view. Popular approach to achieve this is to compute projective invariant from features [Yuji: like]such as points, lines and conics [5]. Such features are [Yuji: selected]preferred because they are easier to detect from the images. Invariants are extensively studied topic in vision community, Forsyth et al. [6] and Gros [7] covered a detailed study on [Yuji: projective invariant descriptors]projective invariants and their stability under projective transformation.

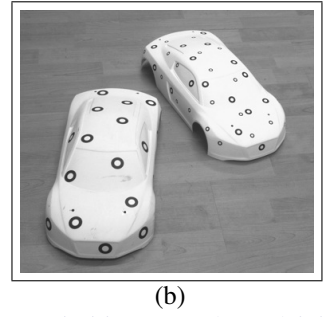
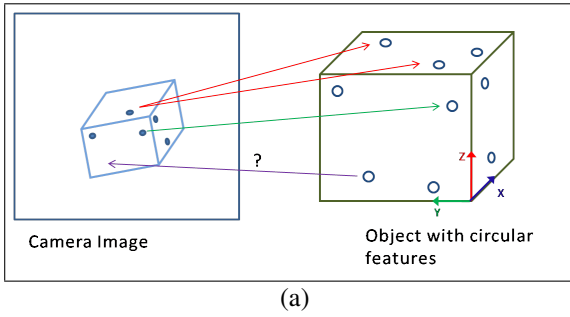


Figure 1: Matching problem for non-coplanar circles ; (a) A primitive example explaining the problem when image features have same properties and correspondence relation with model is ambiguous (b) A realistic model prepared with circular fiducials used for Photogrammetric measurements in Industry [Comments: (1) better to add caption to (a) and (b) so that the readers can easily get what they are. (2) better to save all figures as vector image files as .eps and .pdf so that the figures can keep its visual even when they are zoomed up.]

In this paper we will focus on a specific class of features, that is *circles*. Circles are one of the most primitive features [Yuji: .]. In case of industrial scenario circles are widely present on the model as natural features or circular fiducial are used for Photogrammetric measurements [1] ([Yuji:]see Fig. 1). In such cases it is favourable to build tracking approach depending on such features. A single view correspondence matching problem with coplanar circles and ellipses is widely studied [8][Yuji: .]. However matching problem with non-coplanar circles has not been addressed. Figure 1 shows a basic example of the 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.

3D objects retain Euclidean invariants rather than projective invariants [9]. In case of 3D objects Euclidean invariants are difficult to extract from images due to perspective mapping. Circles have a special property to retain depth information under projective transformation. A world circle always produces an elliptical curve on the image plane. If size of circle is known orientation of circle plane can be defined in 3D (camera coordinates) with a two fold ambiguity [9] [10]. Further, Forsyth et al. [9] proposed that up to three projective invariant can be computed from a non-coplanar pair of circles. They explain that angle between circle planes (angle between surface normals) and distance between centre of the circles are invariant quantities. The concept was proposed in early 90s, however these invariants have remained unexplored. We propose using these invariants to solve correspondence problem when multiple 3D circular features exist on a model. In this approach we bring problem from 2D to 3D by computing 3D invariants from image features, then solve 3D-3D matching problem with model. We introduce *ConicPairDescriptor*, which encapsulate invariants computed from elliptical image features to solve the conic ambiguity and provide accurate matching with 3D features. The proposed method is first attempt to use these invariants, therefore we also carried out simulations to show stability of invariants against change of perspective.

Our contribution is a new method to accurately identify image correspondences when multiple identical 3D circular features exist in the scene. We assume that calibration of cam-

era is known and 3D information of features is available. Often in Industry based model tracking applications 3D-CAD data is known. Our matching method is suitable for tracking any 3D objects having known circles on different planes. In close range photogrammetry multiple circular markers (Figure. 1) are placed on 3D models for surface measurements [10]. These measurements include computation of surface normal and 3D position of each marker. This process involves taking multiple images of the model with additional presence of encoded markers in the scene to solve correspondence. Once 3D measurements are done our method can be extremely useful to support tracking application without using coded patterns. Similarly various industrial parts having natural circles can be identified and tracked with this method. We prepared two car models with circular markers for evaluating our matching method. The proposed method can find corresponding circular marker from a single image, with high accuracy. We also show that our method is stable against false positives detected from the scene. Our method is fast enough to support real-time tracking applications.

2 Related Work

Object detection and pose estimation from conic features is widely studied in 3D vision literature [1][16][9][13]. Circular shape is also a popular choice for designing artificial fiducial. Detection of contour points from image and fitting ellipse is a well studied topic [4]. Quan [14] proposed a two view approach for finding correspondence and 3D reconstruction with conic section. Authors have proposed methods to compute invariants for coplanar conics [5][8]. A 3D problem is simplified to 2D when coplanar features are recovered and used for correspondence. Ying et al. [10] use a coplanar pair for camera calibration. Uchimaya et al. [18] developed invariant descriptors from multiple coplanar circles, and extended the work for deformable model [17]. Work of [9][11][12] propose a circular marker for 6D pose estimation, no invariants are computed as correspondence is solved by using unique coded pattern around the circle. [9][12] use circular shape to define circle plane in 3D, Additionally use a coded pattern is used encode 6D pose without ambiguity.(Fig). Luhmann [10] provides detailed account of methods using point circular fiducials in close range Photogrammetry. The current state of the art methods coded patterns are introduced to simplify correspondence problem. [GOM][AICON] are one of the industrial supplies for close range photogrammetry measurement equipments.

[Refer Thesis: Comment on catalogue based methods,]

Literature study suggests that, existing methods either provide a solution for coplanar circular features or non-coplanar coded circular features. The novelty of our method is that it addresses 2D-3D matching problem for non planar circles present in the scene.

3 New introduction

[Paragraph: what is correspondence matching

- correspondence matching is ...
- 2 solutions.

] Correspondence matching aims to find correspondences between images or ones between object model and its image projection, thus is one of the key problems in computer vision. Since perspective transformation differs appearance of a single object on images, correspondence matching needs feature descriptors, which can compute invariant quantities against some transformation, and establishes correspondence by matching the invariant quantities obtained by the descriptors.

[Paragraph: Projective invariants on planar objects or primitive features] [Yuji: ask Hemal whether this paragraph is right or not.] One of the popular approaches use invariant quantities on planar object against perspective transformation. This type of approaches are based on invariants theory, which has been well-studied in 90s. There exist several works that covered a detailed study on projective invariants and their stability under projective transformation [5, 6] The invariants theory revealed that primitive features such as points, lines, and conics have invariant quantities against perspective transformation [6] and the idea is to use these invariant quantities for matching. Especially, planar projective descriptors, which compute invariant quantities against perspective transformation from planar object, have been investigated. Quan used WHAT invariant quantity from conic section for finding correspondence between two images and 3D reconstruction [14]. Coplanar conics was used to compute WHAT invariant quantity [5, 6]. [Yuji: any more projective invariants on primitive features?] [Yuji: will mention why these invariants have used less in this decade.] [Yuji: is here the best place to mention triangulation based matching such as ART?]

[Paragraph: some invariants on fiducial markers] Another approach also using primitive features embeds binary patterns into a planar object, called fiducial markers. The embedded patterns are extracted from images by adaptive thresholding for matching the correspondence between the marker model and its image projection and then the extracted patterns are decoded for identifying the extracted marker. The original work by Kato and Billinghamurst [?]] uses the contour of a rectangle but does not use any invariant quantity. Matsunaga et al. made a special chessboard with different pattern sizes and use cross ratio, which is a projective invariant quantity, defined by intersection of the patterns [?]. Recent works compute perspective [?], affine [?], rotation invariants [?] on key-points distribution defined by a set of key-points' location. [Yuji: any disadvantage?]

[Paragraph: local texture descriptors] Recent trend computes some invariant quantities defined by local texture in image such as SIFT descriptor [?]. Contrast to the above approaches, this type of approaches does not assume any primitive features in target scene. Alternatively, the approaches assume textured scene so that we can use the approaches for mobile applications, in which primitive shapes cannot be assumed. The approaches focus on improving their invariance as well as efficient memory usage so that the matching can run even on powerless processors in mobile phones [? ?]. One of the disadvantages of these approaches is that their invariance are up to 2D transformation such as affine and rotation transformation. To handle view change caused by perspective transformation, we must explicitly learn how invariant quantities are affected by perspective transformation [? ?]

3.1 Conic invariants under perspective transformation

[Paragraph: conic invariants under perspective transformation] [Yuji: will add more stuff later.]

3D objects are known to retain Euclidean invariants rather than projective invariants [6]. In case of 3D objects, Euclidean invariants are difficult to extract from images due to perspective projection. Circles have a special property to retain depth information even under

projective transformation. A circle in the 3D world coordinate always appears as an elliptical curve on the image plane. When the size of circle is known, orientation of circle plane can be defined in 3D (camera coordinates) with a two fold ambiguity [5, 16]. Further, Forsyth et al. [5] proposed that up to three projective invariant can be computed from a pair of non-coplanar circles. It is explained that angle between circle planes (angle between surface normals) and distance between centre of the circles are perspective invariant quantities. Although those invariants were theoretically validated even in early 90s, its practical application have not been explored.

3.2 Our contribution

[Paragraph: Our contribution] [Yuji: will touch later]
 [Comments: h: our work based on conic invariants] We propose using these invariants to solve correspondence problem when multiple 3D circular features exist on a model. In this approach we bring problem from 2D to 3D by computing 3D invariants from image features, then solve 3D-3D matching problem with model. We introduce *ConicPairDescriptor*, which encapsulate invariants computed from elliptical image features to solve the conic ambiguity and provide accurate matching with 3D features. The proposed method is first attempt to use these invariants, therefore we also carried out simulations to show stability of invariants against change of perspective.

[Comments: i: our contribution] Our contribution is a new method to accurately identify image correspondences when multiple identical 3D circular features exist in the scene. We assume that calibration of camera is known and 3D information of features is available. Often in Industry based model tracking applications 3D-CAD data is known. Our matching method is suitable for tracking any 3D objects having known circles on different planes. In close range photogrammetry multiple circular markers (Figure. 1) are placed on 3D models for surface measurements [16]. These measurements include computation of surface normal and 3D position of each marker. This process involves taking multiple images of the model with additional presence of encoded markers in the scene to solve correspondence. Once 3D measurements are done our method can be extremely useful to support tracking application without using coded patterns. Similarly various industrial parts having natural circles can be identified and tracked with this method. We prepared two car models with circular markers for evaluating our matching method. The proposed method can find corresponding circular marker from a single image, with high accuracy. We also show that our method is stable against false positives detected from the scene. Our method is fast enough to support real-time tracking applications.

4 Method

We will assume that both the 2D and 3D data are already available and focus on the matching method in detail. The 3D data includes the surface normal N_i , centre position M_i and size R_i (diameter) of the circles present on the model. The 2D data includes centre points m_i and conic matrices C_i recovered from the undistorted image. The camera intrinsics (K) and distortion parameters are known. We have followed the approach presented by Naimark [16] for ellipse detection and Farin et al. [4] for computation of conic matrices from ellipse parameters.

4.1 Conic Invariants : Theory and Computation

In this part we will explain the theory and computational aspects behind generating Euclidean invariants from the image conics. Let the camera projection centre be assumed as the vertex of a cone having the world circle as its base. Now the image plane can be considered as a cutting plane π , which always creates an elliptical cross section of the cone. A rotation transformation R_c can be applied to the camera coordinate system such that the new image plane π' intersects the cone as a perfect circle, with the z-axis passing through the centre of both the circles. The normal of the plane (π') is same as the base of the cone, therefore a normal N_{c_i} to the plane π can be computed using R_c . Similarly, using R_i , distance between the plane π' and the base of the cone can be computed. Additionally, 3D position of circle centre M_{c_i} can be computed in the original camera coordinate frame. The normal N_{c_i} and the point M_{c_i} are sufficient to define the orientation of the circle in camera coordinate system. R_c is a combination of two rotations, one of the two has $\pm\phi$ rotation ambiguity, which in turn introduces a two fold ambiguity in the solution (See Eq.1). We will refer the ambiguity as *Conic ambiguity* and the method to obtain plane orientation as *Ellipse backprojection*. The mathematical model for *Ellipse backprojection* is extensively covered by both Forsyth [et al.](#) [\[9\]](#) and Diego [\[9\]](#).

$$\begin{aligned} \text{Ellipse Backprojection}(m_i, C_i) &\rightarrow \pi(N_{c_i}^1, M_{c_i}^1), \pi'(N_{c_i}^2, M_{c_i}^2) \\ \text{Ellipse Backprojection}(m_j, C_j) &\rightarrow \pi(N_{c_j}^1, M_{c_j}^1), \pi'(N_{c_j}^2, M_{c_j}^2) \end{aligned} \quad (1)$$

Forsyth [et al.](#) [\[9\]](#) explained that in case of three dimensional objects invariant descriptors consists of Euclidean invariants rather than projective invariants. Various invariants can be computed from a pair of non-coplanar circles based on Ellipse backprojection. We use following invariants for our method,

- **Angle between planes (θ)** : It is same as the angle between the surface normals (i.e. $\angle(N_{c_i}, N_{c_j})$). θ can be recovered from the image conics without knowledge of the circle size, however due to *Conic ambiguity* θ will have 4 solutions out of which only one is correct.
- **Distance between circle centres** : This is the distance between M_{c_i} and M_{c_j} . The length of the vector d_c is invariant (object scale should be known) and consistent despite of the ambiguity.

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) [\[9\]](#). We performed simulations to understand behaviour of Ellipse Backprojection with respect to both r and η . The parameter r is varied from 0.5 to 2 m and η from 0-70° in step wise manner, while recording 100 iterations at each step (Fig. 2). Realistic values have been assumed for camera intrinsics and size of circle ($R_i = 5, 8, 12$ mm). Figure. 2 shows results recorded with $R_i = 12$, and image noise as $\sigma = 0.3$. The results suggest, at low viewing angles 0-10° both normal and centre estimation errors are high, at any given distance. This can be explained by the fact that for lower values of η image projection of a circle is circular and not elliptical, therefore recovery of elliptical parameters may have errors. We also observe that the estimation error grows with camera distance, however the error in normal recovery appears less sensitive to camera distance than the error in centre recovery. The results obtained with $R_i = 5, 8$ mm show similar pattern, however the magnitude of error is higher as

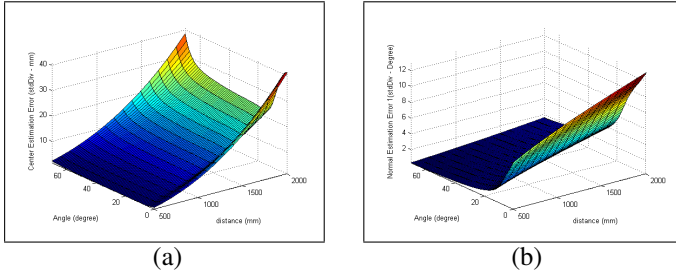


Figure 2: Ellipse backprojection results at different camera distance and viewing angles, Image noise = $\sigma = 0.3, \phi = 12mm$; (a) Error in Centre (M_{c_i}) (b) Error in Normal (N_{c_i})

the image projection become smaller in the image. We also computed and compared the results of estimated centres $M_{c_i}^1$ and $M_{c_i}^2$. The maximum distance recored between the two is ≤ 0.1 mm for $R_i = 12$ mm, this suggests that ambiguity can be neglected for the recovered centre position. The centre recovery is prone to higher errors with higher camera distances, however it produces a consistent solution for invariant d_c . After performing *Ellipse backprojection* we have essentially transformed our problem from image space to three dimensional space.

4.2 Descriptor Generation

This part mainly discusses generation of *Conic pair descriptor* from Euclidean invariants. The invariants for 3-D model are computed from available 3D data (M_i, N_i) without any ambiguity (See Eq. 2). The same set of invariants can be computed from corresponding image features, where the recovered d_c component is unique and θ component has 4 solutions (See Eq. 3). In our approach we pursue the idea that the existence of multiple features on the model can be used to overcome the Conic ambiguity problem. The principle idea is to generate descriptors from conic invariants to perform a descriptor matching to obtain $m_i \leftrightarrow M_i$ correspondence. The proposed *Conic pair descriptor* structure,

$$\text{Conic pair descriptor}_{\text{model}} = V_p \langle d_c, \theta \rangle_{i,j} \quad (2)$$

$$\text{Conic pair descriptor}_{\text{image}} = v_q \langle d_c, \theta_{11}, \theta_{12}, \theta_{21}, \theta_{22} \rangle_{i,j} \quad (3)$$

where V_p represents world circles i, j and v_q represents image conic pair i, j . The reader should note that given a set of points and their corresponding normals in 3D camera space, PFH descriptor [15] also computes similar invariants. In this case, the concept can not be applied to the results of *Ellipse backprojection* as the *Conic ambiguity* restricts us from computing a unique set of invariants. Unlike popular methods, a *Conic pair descriptor* represents two features at same time. In order to uniquely represent a single conic using Euclidean invariants, at least more than two conic features are required. Addition of each conic feature adds 3 wrong solutions of θ in the descriptor, additionally the matching must rely on detection of all the conics used for descriptor computation. We propose matching $v \leftrightarrow V$ first, thereby finding a corresponding conic pair, further we solve individual correspondence $m_i \leftrightarrow M_i$ problem. Descriptor $v_{\{1..q\}}$ are computed among each pair of detected n image conics, where $q = \binom{n}{2}$. $V_{\{1..p\}}$ are computed off-line as the 3D data is already available. In this case for l world circles $p \leq \binom{l}{2}$, as pairs not likely to appear in same image can be rejected.

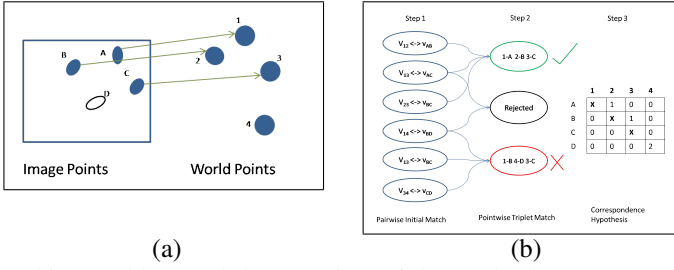


Figure 3: Matching problem and the overview of the method to generate correspondence hypothesis ; (a) A problem showing four image points, and their correspondence with world points (1-A,2-B,C-3) and one false positive (D) (b) Image shows the simplified matching process for the problem suggested in (a), partial results of all 3 steps are shown.

After computing the *Conic pair descriptors* the following 3 step matching approach is used to achieve $m_i \leftrightarrow M_i$ correspondences.

4.3 Step 1: Pairwise Initial Matching

In the first stage of the matching process we compare the *Conic pair descriptors*. The objective is to reduce complexity of the problem by finding the possible pair correspondences ($v \leftrightarrow V$). First the unique component (d_c) between the descriptors is compared, if matched then (θ) component of V is compared with all 4 values of θ component in v (Ref. Algorithm 1). T_{d_c} and T_θ are the threshold values used to compare the respective components. The stage may result in *one to many* type of relation between descriptors. This can be either due to similar feature orientation on object or due to presence of the *Conic ambiguity*. The reader should note that a descriptor represents a pair of conics, therefore the stage is called pairwise matching. The example given in Figure 3 shows possible outcome of step 1 with respect to the given example problem.

Goal : Find all possible V_p similar to v_q ;

Initialisation : $T_{d_c} = 10$, $T_\theta = 5$;

```

forall the 3D Feature Descriptors ( $V$ ),  $p \leftarrow 0$  to  $n$  do
  forall the 2D Feature Descriptors ( $v$ ),  $q \leftarrow 0$  to  $l$  do
    if  $\text{compare}_{d_c}(V_p, v_q) < T_{d_c}$  then // compares  $d_c$  component
      if  $\text{compare}_\theta(V_p, v_q) < T_\theta$  then // compares  $\theta$  component
        // All 4 solutions of  $\theta$  in  $v_q$  are checked
        ;
        SavePairResult( $p, q$ ) // Save matching descriptor pair
      ;
    end
  end
end

```

Algorithm 1: Pairwise Initial Matching algorithm

4.4 Step 2: Pointwise Triplet Matching

In this stage we simplify the problem further and obtain hypothesis on point wise 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 and reject false descriptor matches. We seek three $v \leftrightarrow V$ results, such that they complement each other to form a unique three points $m_i \leftrightarrow M_i$ hypothesis. A simple two stage approach is proposed to generate a triplet matching hypothesis,

- 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.

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

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

- 2 Find a new descriptor matching pair which can verify the triplet matching hypothesis formed in the previous stage (e.g. $V_{23} \leftrightarrow v_{BC}$).

The verified triplets are saved and others are rejected. The results may also contain false triplet matches (Fig. 3). If x number of results are generated in step 1, the number of pairs compared in this stage is $\binom{x}{3}$.

4.5 Step 3: Correspondence Hypothesis

In this final step results of triplet matching are combined and a voting matrix is generated (Fig. 3). A $m_i \leftrightarrow M_i$ pair having maximum votes in the matrix is proposed as a correspondence hypothesis. In case of conflicting votes the respective $m_i \leftrightarrow M_i$ relation is not considered. A minimum of 3 correspondence are required to compute the pose of the object [8], if camera intrinsics are known. A pose of the object can be computed by selecting top 3 correspondence results and verify other correspondences obtained from the matrix. If only 3 out of n conics are detected in the image, verification with pose is not possible and results may not be reliable.

5 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 test models by attaching circular markers of size $\phi = 12$ mm ($M_i = 20$) and 5mm ($M_i = 26$). The markers used are widely accepted and used in the Industrial domain for Photogrammetric measurements. 3D measurements for the markers are done 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) camera is used for the experiments and MATLAB is used for synthetic experiments.

Table 1: Descriptor Matching Analysis

R_i	θ	d_c	Camera Distance (mm)	Min-Max Success(%)
5	10-40	10-150	500-1200	58-82
	40-80	10-150	500-1200	35-65
12	10-40	10-150	500-2000	64-86
	40-80	10-150	500-2000	40-69
20	10-40	10-150	500-2000	74-92
	40-80	10-150	500-2000	48-77

Table 2: Correspondence matching with varying threshold settings

T_θ	T_{d_c}	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

5.1 Experiment 1 : Descriptor Matching vs Marker Orientation

This experiment is carried out synthetically to observe the effect of orientation of the circles on the descriptor matching results. Two circle planes are placed in different orientations and images are captured (noise $\sigma = 0.3$) from 1000 different camera positions for each orientation. The control parameters, the distance between circle centres d_c 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_c and θ from the images, compare them with ground truth and measure the success rate. Realistic values are used for camera intrinsics, $T_{d_c} = 10$ and $T_\theta = 5$ are kept constant. Camera positions are chosen at random from a defined range, range for x-y-z rotation is -70 to 70, x-y translation range is -500 to 500 mm, z-translation (Camera distance) range is 500-2000 mm.

The table 1 provides summary of key observations made during the experiment. The matching success shows inversely proportional relation with θ , independent of d_c . We learn that descriptor matching is influenced more by angle between planes than 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 features placement is possible, it is advised to choose larger circles or surfaces with lower plane angles for improved matching results.

5.2 Experiment 2 : Correspondence Matching vs Threshold Settings

In case of model with natural features choice over size or placement is not available. The aim of this experiment is to understand the role of threshold values T_{d_c} and T_θ in $m_i \leftrightarrow M_i$ matching results. In order to perform this experiment we took 75 images of car model with 12 mm markers (Distance Range 500-2000 mm). As suggested in Sec. 4.5, each image has at least 4 detected conics. The results are considered *Not Converged* (NC) in case of less than 3 $m_i \leftrightarrow M_i$ results.

In Sec. 4.1, we learned that d_c recovery is weak and therefore a flexible threshold may

Table 3: Time Analysis

Algorithm Stage	CAM 1		CAM 2
	Model	Model + FP	Model
Image Undistortion	39.53%	12	11.79
Marker Detection	38.51	34	30.75
Correspondence Matching	0.35	44.2	1.34
Pose Estimation	21.61	9.8	56.12

be appropriate for matching. The results (Table. 2) show that flexible T_{dc} with stringent T_θ may allow more Pairwise matching results, but due to *conic ambiguity* of θ probability of wrong matching goes higher. On the other hand, very stringent thresholds lead to higher non convergence. Therefore, a right balance of threshold can be selected to achieve balanced results. Our preferred settings for experiments is $T_{dc} = 10$ and $T_\theta = 5$, as 90% of images show maximum 100% success with (0%) wrong matching. The selection may require change based on density of the features.

5.3 Robustness against false positives

5.4 Time Analysis

In this experiment we focus on analysing time consumed by the matching method when introduced into a tracking application. Two cameras CAM 1 (2560 x 1920) and CAM 2 (640x480) are used for tracking, the results presented are averaged over 100 frames. The results show that our method takes $\leq 1\%$ time from in the tracking pipeline. In terms of frame rates we achieve 2-3 FPS with CAM 1 and 7-8 FPS with CAM 2. Additionally, an exhaustive experiment with 90-140 false positives in the scene shows that matching consumes maximum time in the pipeline (0.7 FPS). Limited tracking range (< 500 mm) of CAM 2 does not allow experiment with such large number of false positives.

6 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 descriptor* which represents euclidean invariants generated by a pair of non-coplanar circles. Our method can successfully define correspondences for more than 3 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 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 Industrial scenario, where natural or artificial circular features exist 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 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] Michel Dhome, Jean-Thierry Lapresté, Gérard Rives, and Marc Richetin. Spatial localization of modelled objects of revolution in monocular perspective vision. In Computer Vision—ECCV 90, pages 475–485. Springer, 1990. URL <http://link.springer.com/chapter/10.1007/BFb0014897>.
- [2] Gerald E Farin and Dianne Hansford. The geometry toolbox for graphics and modeling. A.K. Peters, Natick, Mass., 1998. ISBN 1568810741 9781568810744.
- [3] 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.
- [4] Andrew Fitzgibbon, Maurizio Pilu, and Robert B. Fisher. Direct least square fitting of ellipses. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 21(5): 476–480, 1999.
- [5] 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.
- [6] 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.
- [7] Richard Hartley and Andrew Zisserman. Multiple view geometry in computer vision. Cambridge University Press, Cambridge, UK; New York, 2003.
- [8] Vincent Lepetit and Pascal Fua. Monocular model-based 3d tracking of rigid objects: A survey. Now Publishers Inc, 2005.
- [9] Diego Lo’pez de Ipin a, Paulo RS Mendonça, 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.
- [10] 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.

- [11] 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.
- [12] Alain Pagani, Johannes Köhler, and Didier Stricker. Circular markers for camera pose estimation. In 12th International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS 2011), 2011.
- [13] Long Quan. Invariant of a pair of non-coplanar conics in space: Definition, geometric interpretation and computation. In Computer Vision, 1995. Proceedings., Fifth International Conference on, pages 926–931, 1995. 00018.
- [14] Long Quan. Conic reconstruction and correspondence from two views. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 18(2):151–160, 1996.
- [15] Radu Bogdan Rusu. Semantic 3D Object Maps for Everyday Manipulation in Human Living Environments. PhD thesis, Computer Science department, Technische Universität München, Germany, October 2009.
- [16] Reza Saffae-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.
- [17] Hideaki Uchiyama and Eric Marchand. Deformable random dot markers. In Mixed and Augmented Reality (ISMAR), 2011 10th IEEE International Symposium on, pages 237–238, 2011.
- [18] Hideaki Uchiyama and Hideo Saito. Random dot markers. In Virtual Reality Conference (VR), IEEE, pages 35–38, 2011.
- [19] 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.
- [20] Xianghua Ying and Hongbin Zha. Camera calibration using principal-axes aligned conics. In Asian Convergence on Computer Vision, pages 138–148, 2007.