Capturing Optimal Image Networks for Planar Camera Calibration

Brendan P. Byrne John Mallon Paul F. Whelan
Centre for Image Processing and Analysis
Dublin City University
brendan.byrne@eeng.dcu.ie

***Abstract*—This paper details a novel approach to specifying the optimal pose of planar targets in camera calibration that both reduces the number of images required, and improves the parameter estimates. This is accomplished within a semi-supervised strategy where virtual images of planar calibration targets are generated and displayed. These virtual targets are then replicated by the user to generate an image network with optimal geometry for the recovery of the camera parameters. Optimal planar pose is specified by enforcing maximum independence within the calibration constraints offered by each image within the network. This solution space is further refined to ensure that the generated target pose is suitable for easy acquisition and subsequent feature extraction processes. The results on simulated and real data demonstrate that proper consideration of image network geometry directly leads to more accurate camera parameter estimates.**

I. Introduction

The calibration of cameras is an area which has received much attention in the computer vision community over the last 25 years. The drive towards increased resolution with smaller sensor sizes has seen a continued growth in the digital camera market and contributed to the need for flexibility when calibrating a camera. In recent years, the mobile phone camera market has become the largest market for digital sensor suppliers with forecasts for sales to reach one billion units for the first time in 2011 [1]. The emergence of the smartphone has led to increased onboard processing power and high resolution displays, which, coupled with a high resolution camera, is allowing these devices to become more accessible to computer vision researchers.

Camera calibration is emerging as a key factor for computer vision researchers in this area for a number of reasons. The lenses used in camera modules manufactured for mobile phones are of a lesser quality (and cost) than a conventional point and shoot digital camera. As a result, lens distortion and aberrations are increased in the images. The increase in demand for mobile phone applications that make use of augmented reality and metric depth information have also contributed to the renewed interest in camera calibration as it is a fundamental step to solving these problems.

Planar camera calibration [2], [3] is the preferred technique of the community due to the reduction of input requirements compared to the traditional approaches [4]–[7]. Readily available implementations of planar approaches[1][2][3] have also contributed to its popularity. The input requirements of planar approaches simply involve imaging a planar target. Planar calibration targets such as a chessboard or circular dot patterns are commonly used to specify the planar feature points [8]. The stability of these planar calibration methods has been well studied in the literature. González et al. [9] compares the camera parameter estimates of different methods by comparing the distance between the real 3D position of points to the 3D position from their 2D projections. Similarly, Salvi et al. [10] assess the stability of the internal parameters by examining the relationship between the position of 2D image points and projection of the 3D object points on the image plane. Sun [11] provides an empirical evaluation of the algorithms proposed in [2], [5], [6]. The accuracy is characterised by the reprojection error. In order to increase accuracy of the calibration, the number of control points per image is increased. However, it is noted that as the control point data set increases considerably, it has little effect on the results. In all cases, the findings indicate that planar calibration achieves a good level of accuracy compared to the traditional methods, without the methodical setup.

The accuracy of planar calibration parameter estimates can deteriorate based on the orientation of the planar targets. Planar target poses which cause this deterioration are known as critical configurations. Sturm and Zhang [2], [3] have identified poses which consist of pure rotation and/or translation between targets as critical. Configurations which are close to critical also lead to inaccurate parameter estimates. These configurations are more difficult to identify. Wang et. al [12] suggest a set of orientations for the planar targets but without any justification, and also limits the flexibility of the input requirements. Byrne et. al [13] devise an image selection strategy in which the algorithm selects the most suitable images from a database of captured calibration images (forming an *image network*) to estimate the camera parameters. The selection process is based on geometric properties of the planar target configuration. Since the image network is formed from a large database of calibration images, the problem of excessive input requirements and the possibility that the image

[1] http://research.microsoft.com/~zhang/Calib
[2] http://www.vision.caltech.edu/bouguetj/calib
[3] http://www.opencv.willowgarage.com

database does not contain ideal geometry remains.

This paper aims to address the issue of image network geometry for planar camera calibration by providing foolproof, semi-automatic guidance for obtaining an optimal image network geometry. As outlined, at present the general practitioner has no clear indication of how the planar calibration grids should be orientated for the calibration images. We remove this ambiguity by providing synthetic planar poses via a display, which the user can replicate. We define this as the Generated Image Network (GIN). Our experiments clearly show that GINs directly lead to improved calibration estimations with reduced number of input images.

The main contribution of this work is the proposed method of generating optimal image networks. The approach requires one input image of the planar target captured by the user. Subsequently, this is used to generate synthetic optimal images of the calibration target. These synthetic images are then replicated by the user to form the optimal image network. Calibration with optimal image networks ensures that the precision of calibrations performed by general practitioners is increased with less input requirements.

The underlying theory of the method builds on the alternative geometric approach to planar calibration originally proposed by Gurdjos et al. [14]. Gurdjos et al. make use of Poncelet's theorem (see Figure 1). A closer examination of Figure 1 reveals that a world plane intersects the image plane in a line. If this intersection line is taken as an axis, about which the world plane may rotate freely, there is an infinite number of world planes which remain in homographic correspondence with the generated image on the image plane. Thus a planar homography H represents the transformation. Each world plane has a unique centre of projection through which it remains in homographic correspondence with the image. The locus of these centres of projection forms the Centre Circle (CC). The orthogonal projection of the CC onto the image plane forms a line, the Centre Line (CL). Thus, for each projected centre of projection onto the image plane, there is an associated principal point on the CL. It is from this perception of the problem that we formulate our approach to generating optimal image networks. In the case of planar camera calibration, when two or more images are used, multiple CL's are formed in which their intersection is the principal point of the camera. This allows the camera calibration to be recovered.

Generation of optimal image networks is accomplished by manipulating properties of the CLs. Byrne et al. [13] based their image selection algorithm solely on the slope of the CLs. By ensuring equal angle between image CLs, independence in the planar calibration equations is increased. Our method uses the slope constraint and two additional constraints to generate optimal image networks. The first is an additional constraint on the CL in which its y-intercept is manipulated to ensure that the generated image has similar camera properties to that of the initial captured image. The second additional constraint is based on the compression and expansion of pixels in the generated image. This constraint ensures that a practical image

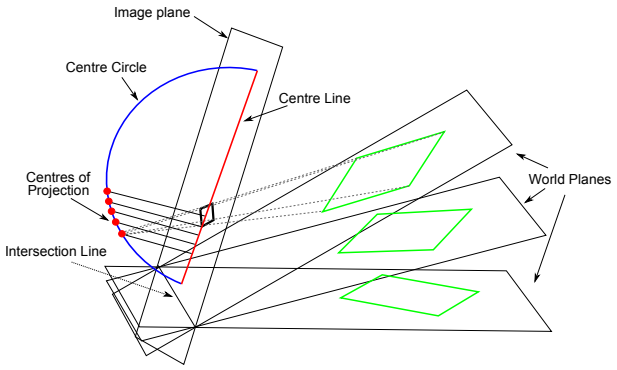


Fig. 1. Poncelet's Theorem [14]

is generated for the end user to capture. For an n image network, we need only generate $(n-1)$ CLs.

The remainder of the paper is organised as follows: Section II describes in detail the necessary steps in the formation of the optimal image networks. Section III presents the experiments conducted with real and synthetic data. The benefits of GINs in comparison to conventional image networks are highlighted. Section IV discusses applications of the developed theory and method. Section V concludes with a summary of the main contribution made in this paper.

The most significant findings were that by using GINs, better calibration results can be achieved with less images in the network. This has important implications for the practitioner in the field of planar calibration.

II. Optimal Calibration Networks

Optimal image networks are formulated based on a two stage approach. Firstly, planar target poses are generated from an initial image, with non fronto parallel orientation [4], supplied by the user. Then, images of the generated targets are captured (via replication) to complete the GIN.

The replication process is performed using one display. The optimal target pose is displayed to the user along with a rough estimate for angular orientation. An optimal image is captured when the target pose on the camera live-view and the optimal pose display, provided to the user, visually match. Section IV proposes an augmented reality approach using a single display for capturing and visualising optimal images.

Generation of valid optimal pose deems it necessary to approximate properties of the camera used for the initial image captured. Ideally this requires the construction of a planar homography matrix (H), which encodes the internal parameters (K) of the camera and the new pose of the target (R). Estimating (u_0, v_0) at the centre of the image allows an approximate calculation for the camera focal length (f). This gives $\hat{K}$ for the camera which reduces the problem to the construction of the pose $\hat{R}$.

The homography for the synthetic pose $\hat{H}$ is formed by $\hat{K}$ and $\hat{R}$ (detailed in Section II-A). $\hat{R}$ is decomposed into the pitch ($\hat{R}_\alpha$), yaw ($\hat{R}_\beta$) and roll ($\hat{R}_\gamma$) angles, which are used

[4] to avoid degeneracy [2]

to estimate the new pose of the planar target. Constraints are imposed reducing the solution space and ensuring a valid pose is determined (detailed in Section II-B).

A. *Forming* $\hat{\mathrm{H}}$

The relationship between the image of the absolute conic (IAC) and a point homography is well known [2]. Coupled with the knowledge of the principal point and use of H from the initial image, $\hat{H}$ can be partially formed. For a complete formation, the target pose being generated ($\hat{R}$) must be considered. A general planar homography matrix H can be decomposed as,

$$H = K[R_{12}\ |t], \tag{1}$$

where R_{12} represents the first two columns of the rotation matrix. Image generation of the planar target is obtained by applying $\hat{H}$ to an image of the calibration target in its canonical position. A valid $\hat{H}$ is constructed similarly to (1). The first step of the formation takes advantage of the assumption of u_0 and v_0 (at the image centre) by eliminating these terms in expression (1). This is done by pre multiplying H by the matrix U where $\hat{H} = UH$ with $U = [I\ |\ -c]$ and $c = (u_0, v_0, 1)^T$. This leaves an expression containing the focal length (f) and the scaling factor (s) of the H matrix,

$$\hat{H} = \begin{bmatrix} \frac{fr_{11}}{s} & \frac{fr_{12}}{s} & .. \\ \frac{fr_{21}}{s} & \frac{fr_{22}}{s} & .. \\ \frac{r_{31}}{s} & \frac{r_{32}}{s} & .. \end{bmatrix} \tag{2}$$

The third column which contains the translation does not affect the forming of f. By manipulation of the calibration equations in [2], the parameter f can be estimated from $\hat{H}$ as,

$$\mathrm{f} = \sqrt{\frac{-\hat{\mathrm{h}}_{11}\hat{\mathrm{h}}_{12} - \hat{\mathrm{h}}_{21}\hat{\mathrm{h}}_{22}}{\hat{\mathrm{h}}_{31}\hat{\mathrm{h}}_{32}}} \tag{3}$$

which can now be used to find the scaling factor in expression (2). Calculating the scaling factor is straightforward, since it is well known that the columns of R form an orthornormal basis [15].

With the ability to decompose the planar homography matrix from the first image (H) into a calibration matrix $\hat{K}$ and scaling factor s, the planar homography matrix for the optimal generated image ($\hat{H}$) can be constructed via these parameters coupled with the optimisation parameters $\hat{R}_\alpha$, $\hat{R}_\beta$ and $\hat{R}_\gamma$ which form the estimated rotation matrix $\hat{R}$.

$$\hat{H} = s\hat{K}\hat{R} \tag{4}$$

B. *Constraints on* $\hat{\mathrm{H}}$

The criteria for generating a valid $\hat{H}$ have been identified. The assumption of the principal point at the image centre recovers an approximated $\hat{K}$ which, in turn, recovers the scale from the original image homography H. Therefore, the key to determining the GIN is the estimation of the pose ($\hat{R}$). Forming a $\hat{H}$ which yields a correct solution requires constraints to be imposed on $\hat{H}$ while $\hat{R}$ is estimated. These constraints are based on geometric properties of the image.

The CL of an image can be derived from its homography matrix. In [12] an expression for the CL is given as,

$$v_0 = \Gamma u_0 + \Lambda \tag{5}$$

where Γ is the slope of the CL given by,

$$\Gamma = \frac{h_{11}h_{32} - h_{12}h_{32}}{h_{22}h_{31} - h_{21}h_{32}} \tag{6}$$

and Λ is the y-intercept given by,

$$\Lambda = \frac{h_{21}h_{31} + h_{22}h_{32}}{h_{31}^2 + h_{32}^2} - \frac{h_{11}h_{31} + h_{12}h_{32}}{h_{31}^2 + h_{32}^2}\Gamma \tag{7}$$

The first constraint applied to $\hat{H}$ is on the slope of its CL. As previously discussed, equal angle between CLs is necessary to generate the optimal image networks. The slope (m_1) of the initial CL (provided by user via initial image) is calculated using equation (6). Depending on the number of images in the network to be generated (k), the optimal slope (Γ_i) is calculated as M_i such that $M_i|_{i=1:k} = \tan(\tan^{-1}(m_1) + (i+1)\alpha\pi/180)$ where α, the angle between the CLs, is calculated as $180/(k+1)$. To generate the synthetic image for the network, the slope being estimated ($\hat{\Gamma}_i$) must equal the optimal slope. As the slope is nonlinear in the parameters of $\hat{H}$, we develop a cost function to estimate the optimal slope value. The cost function $\mathfrak{C}_i(\mathtt{1}, \hat{R})$ for this constraint is expressed as,

$$\mathfrak{C}_i(\mathtt{1}, \hat{R}) = (\hat{\Gamma}_i - \Gamma_i)^2 \tag{8}$$

The implications of $\mathfrak{C}_i(\mathtt{1}, \hat{R})$ lead to multiple solutions of equal angle between CLs which are determined by the y-intercept ($\hat{\Lambda}_i$). Thus a second constraint is imposed on $\hat{H}$ with the aim of generating the correct CL. This is enforced by ensuring the generated CL passes through the assumed principal point. The optimal intercept value (Λ_i) is calculated as,

$$\Lambda_i = v_0 - \Gamma_i u_0 \tag{9}$$

The cost function $\mathfrak{C}_i(\mathtt{2}, \hat{R})$ is formed similarly to $\mathfrak{C}_i(\mathtt{1}, \hat{R})$.

$$\mathfrak{C}_i(\mathtt{2}, \hat{R}) = (\hat{\Lambda}_i - \Lambda_i)^2 \tag{10}$$

The minimisation of $\mathfrak{C}_i(\mathtt{1}, \hat{R})$ and $\mathfrak{C}_i(\mathtt{2}, \hat{R})$ yields a manifold of possible solutions. In order to select poses that are realisable, a third constraint is applied on the solution space which examines the extent of compression/expansion in the generated synthetic images. This prevents the generation of GIN images that would be practically unattainable eg. a severely perspectively distorted image.

When a transformation ($\hat{\mathrm{H}}$) is applied to an image, it can result in the expansion and compression of pixels in the image. This compression and expansion can be measured locally in the image by examining the singular values of the Jacobian [16]. The Jacobian of a single point $p = (x_u, x_v)$ in the image is defined as:

$$\mathfrak{G}_p \Rightarrow \mathfrak{J}(\hat{H}, p) = \begin{bmatrix} \frac{\delta \hat{x}_u}{\delta x_u} & \frac{\delta \hat{x}_u}{\delta x_v} \\ \frac{\delta \hat{x}_v}{\delta x_u} & \frac{\delta \hat{x}_v}{\delta x_v} \end{bmatrix} \tag{11}$$

Each point (p) in the image has two corresponding singular values $\sigma_1(\mathfrak{G})$ and $\sigma_2(\mathfrak{G})$. For a transformation $\hat{H}$, if $\sigma > 1$ there is an expansion of pixels and if it is less, the overall effect is compression. It is desirable to ensure that the singular values of the Jacobian at each point in the image are as close as possible to 1. This limits perspective distortion and ensures realisable images. The Jacobian can be calculated at each point in the image, or alternatively over a grid of points in the image. This constraint is expressed as,

$$\mathfrak{C}_i(3, \hat{R}) = \sum_{i=1}^{l} \left[(\sigma_1(\mathfrak{G}_i) - 1)^2 + (\sigma_2(\mathfrak{G}_i) - 1)^2\right] \qquad (12)$$

where l is the number of points in the grid used. The total cost function ($\mathfrak{T}$) which incorporates all the constraints is,

$$\mathfrak{T}(\hat{R}) = \sum_{i=1}^{k} \left[\mathfrak{C}_i(1, \hat{R}) + \mathfrak{C}_i(2, \hat{R}) + \mathfrak{C}_i(3, \hat{R})\right] \qquad (13)$$

where k is the number of images to be generated in the image network. Initial estimates of $\hat{R_\alpha}$, $\hat{R_\beta}$ and $\hat{R_\gamma}$ for the optimisation are taken from the inverse orientation of the target in the original image. The minimisation of $\mathfrak{T}(\hat{R})$ will generate the synthetic target poses for the GIN through the formation of optimal $\hat{H}s$. The Levenberg-Marquardt algorithm is used for this purpose. The algorithm generally converges within 15 to 20 iterations.

C. Network orientation and sensitivity

Given that the rotational parameters are being estimated for the GIN optimisation, the sensitivity of $\hat{R}$ is examined.

The characteristics of the CL were investigated in [12] and it was noted that R_γ has no effect on the CL orientation. Therefore R_γ does not influence $\mathfrak{C}_i(1, \hat{R})$ and $\mathfrak{C}_i(2, \hat{R})$. However, it does influence the image grid constraint $\mathfrak{C}_i(3, \hat{R})$. If the roll angle is large in a perspectively distorted image, the feature extraction process is complicated. By including the R_γ in the optimisation, it counteracts this and yields better results than when it is not estimated.

The sensitivity of the $\hat{R}$ parameters to random perturbations is tested. These perturbations result from the practitioner being unable to reproduce the synthetic pose exactly, and are modelled as random variations ($\hat{R}_{err}$) of $\pm 15^\circ$ in $\hat{R_\alpha}$ and $\hat{R_\beta}$. Typically the user error is within $\pm 5^\circ$ of the optimum angle as shown in Section III-C.

The variations of $\hat{R}$ are then used to form the perturbed planar homography matrix H_{per} which gives a perturbed CL. H_{per} is formed similarly to $\hat{H}$ in Section II-A using equation (1).

$$\begin{aligned} \hat{R} &= (R_\alpha + \hat{R}_{err1}, R_\beta + \hat{R}_{err2}, R_\gamma) \\ H_{per} &= \hat{K}\hat{R} \end{aligned} \qquad (14)$$

Simulated testing was performed for 1000 instances of random perturbations in the $\hat{R}$ of a homography describing a CL with known angle relative to another CL. Figure 2 shows the relationship between the perturbed error on $\hat{R}$ and the corresponding induced angle error between the CLs.

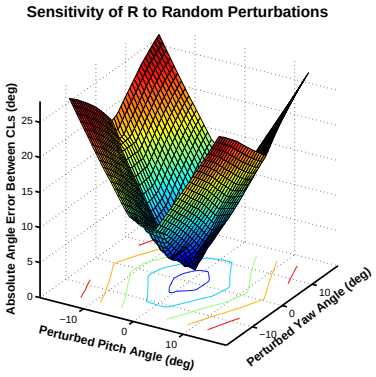


Fig. 2. Effects of perturbations when estimating the ideal image via the planar homography H and the rotational parameters R

In Figure 2, it can be seen that for small perturbations of $\hat{R}$, the induced angle error between CLs remains bounded. This is an important result as it ensures that small error in the $\hat{R}$ estimation will not dramatically affect the calibration results.

III. Testing

Experiments are performed with real and synthetic data to illustrate the advantages of GINs over the conventional Random Image Networks (RINs). The planar calibration target used for the real tests is a checkerboard pattern as shown in Figure 4. For simulated tests, a synthetic grid of control points is used.

Synthetic testing is conducted to highlight the main benefit of GINs, which is that less images are needed in the network to achieve better calibration accuracy (Section III-A). Real images are used in the remaining tests. Section III-B examines the effects of variation of the angle between CLs on the parameter estimation accuracy. Section III-C examines the implications of the angle between the CLs in terms of parameter estimation accuracy, and the general users capability of capturing the synthetic images. Section III-D aims to verify the simulated results in section III-A by investigating the calibration accuracy/performance of three and four image networks (GINs v RINs). The comparison is conducted by analysing the uncertainties of the parameter estimates. Note: The calibration method of Zhang [2] is used in all testing sections.

A. Synthetic testing

A synthetic camera was formed with internal parameters as shown in equation (15). Synthetic grids (of control points) were simulated in P^3 and subsequently projected onto the image plane in P^2. Each grid formed an image of 400×600 pixels. A two parameter lens distortion model was applied to the image and random noise with standard deviation 0.5 pixels was added, to the grid point locations, to simulate image noise.

$$K = \begin{bmatrix} 500 & 0 & 200 \\ 0 & 500 & 300 \\ 0 & 0 & 1 \end{bmatrix} \qquad (15)$$

The test set consists of 50 trials of each image network instance from 2 to 25 images. For the RINs, each network

TABLE I
RESULTS FOR CALIBRATION WITH IMAGE NETWORKS WITH NON-OPTIMUM ANGLE BETWEEN CLS. PARAMETER ESTIMATIONS AND SD GIVEN IN PIXELS

Angle (deg)	u_0	v_0	f_u	f_v
5	808.2	617.4	1796.5	1794.1
10	806.9	607.7	1789.4	1787.9
15	812.2	638.6	1808.5	1798.0
20	804.85	609.8	1791.4	1790.3
25	792.1	620.9	1801.2	1803.7
SD	7.6159	12.2738	7.7210	6.2809

TABLE II
RESULTS FOR CALIBRATION WITH IMAGE NETWORKS WITH OPTIMUM ANGLE AND CLOSE TO OPTIMUM ANGLE BETWEEN CLS. PARAMETER ESTIMATIONS AND SD GIVEN IN PIXELS

Angle (deg)	u_0	v_0	f_u	f_v
80	802.4	617.6	1800.2	1797.8
85	801.2	619.3	1798.8	1795
90	804.1	620.9	1803.5	1801.7
95	802.6	619.6	1805.5	1802.7
100	803.5	620.1	1803.5	1801.4
SD	1.1104	1.2227	2.7285	3.2244

was drawn randomly from the range $[-40^\circ \ \ 40^\circ]$ on R_α and R_β. GINs were formed by taking one image from each RIN and then generating the optimal image networks from these images. The metric used to compare image network configurations is the mean of the absolute error between the ground truth data and the RIN/GIN parameter estimates for each network trial.

The results presented in Figure 3 reveal that the GIN estimates are more accurate than the RIN estimates for each intrinsic parameter. The number of images needed in a GIN to achieve better calibration accuracy than RINs is significantly reduced. The results also indicate that three and four image GINs yield better estimates than the 25 image RINs. However it is noted that as the number of input images increases beyond 5 for the GINs and 7 for RINs, there is little improvement in the overall accuracy of the parameter estimates.

B. Angle between CLs in image network

Two image networks are used for this test, thus the optimal angle between the CLs is 90°. The initial images used for the calibration networks are represented by the 15° and 90° angle entries in Table 1 and Table 2 respectively.

The results presented in Table 1 and Table 2 highlight the importance of the CLs in obtaining stable calibration results. The networks with non-optimum angle between their CLs have a Standard Deviation (SD) significantly larger than the results obtained with the optimal angle networks. This verifies that manipulating CLs of image networks is beneficial in achieving good calibration results. For the networks with an angle between their CLs within $\pm 10^\circ$ of the optimal angle, the SD of the parameters estimated is small indicating stable results. This also confirms that the sensitivity of $\hat{R}$ to small perturbations is insignificant on the results (discussed in Section II-C).

These results demonstrate the importance of the angle between CLs of an image network when calculating the camera

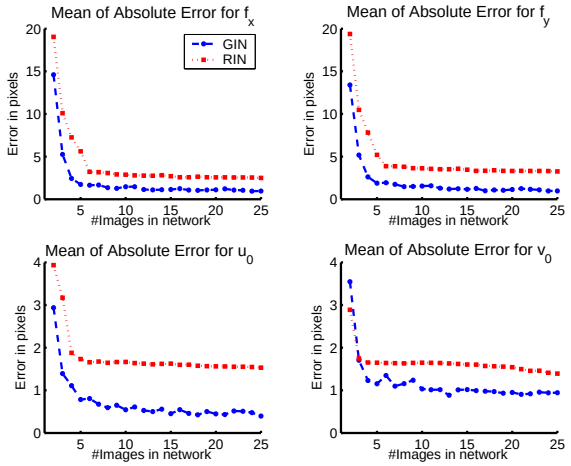


Fig. 3. Intrinsic parameter estimation results for RINs and GINs for 2 to 25 image newtorks

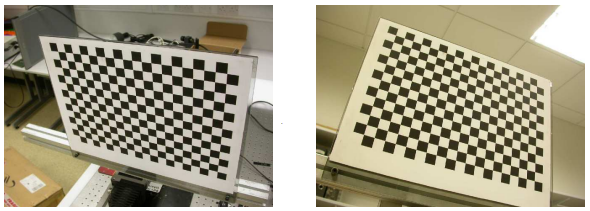


Fig. 4. Image network with an angle of 5^o between the CL's

parameters. It was found that as long as the generated image can be captured having an angle between CLs within the accuracy shown in Table II ($\pm 10^\circ$), the calibration results will be good. Figure 4 presents a situation where a general practitioner may be under the impression that the network captured is in a non critical configuration. However, the angle between the CLs for this network is 5° thus the image network will produce unreliable results. This highlights another benefit of GINs in that the practitioner will avoid critical configurations in the formed image network.

C. Image reproducibility

The accuracy needed to reproduce images from the GINs successfully (discussed in Section III-B) was within $\pm 10^\circ$ of the optimal CL angle. This level of accuracy ensures good calibration results.

An indicative test was performed were four non-expert practitioners were given 25 optimal images to reproduce. The orientation of the calibration target in the generated image was given as a guideline. Each practitioner captured only one image per generated image, thus the results presented are worst case scenario. The experimental setup is as discussed in Section II.

The results in Table III indicate that the non-expert practitioners were capable of reproducing the generated images with sufficient accuracy and thus capable of achieving better calibration results for the GINs rather than the RINs. This also validates the assumptions made of f, u_o and v_o in forming $\hat{H}$

TABLE III
NON-EXPERT PRACTITIONER RESULTS FOR OPTIMAL IMAGE REPRODUCIBILITY (NOTE: OPTIMAL ANGLE IS 90°)

Practitioner	Average Angle (SD)	Average Angle Err (SD)
1	88.45 (11.41)	8.61 (7.45)
2	89.22 (3.18)	2.53 (2.01)
3	90.23 (6.35)	4.74 (4.12)
4	88.50 (8.67)	6.73 (5.50)

for optimal image networks (see Section II-A) as the replicated images are producing the expected angles between CLs.

D. Benefit of generated vs random positions

The aim of this section is to examine the precision of the results obtained with three and four image networks. With real images there is no ground truth data available, therefore the measurement used to quantify the results is the estimated parameter uncertainties. This can be calculated from an estimate of the Fisher information matrix ($\mathfrak{F}$) which is formed upon convergence of the non-linear estimation process [17].

$$\mathfrak{F}(\hat{R}) = \frac{1}{\sigma_i^2} \sum_{i=1}^{n} \frac{\delta e^T(c_i, \hat{R}_k)}{\delta \hat{R}} \frac{\delta e(c_i, \hat{R}_k)}{\delta \hat{R}^T} \tag{16}$$

where σ^2, the unknown noise variance can be approximated as

$$\sigma^2 = \frac{1}{n_t - n_p} \sum_{i=1}^{n_t} e_i^2 \tag{17}$$

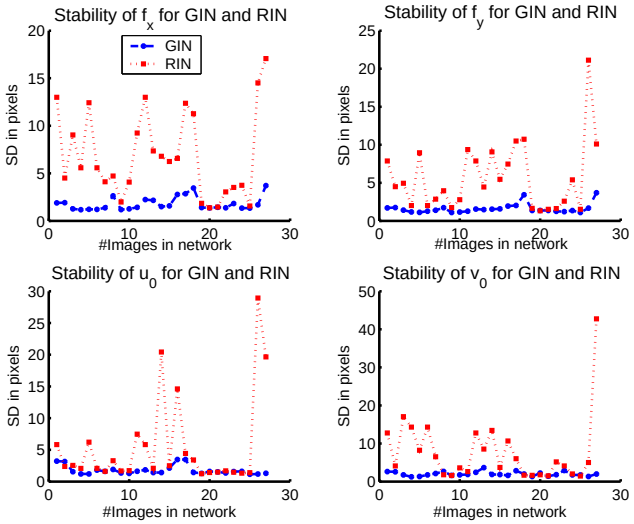
with n_t being the number of measurement points and n_p the number of parameters being estimated. e_i^2 is the norm of the residuals from the estimation algorithm. With the approximation of $\mathfrak{F}$, we use $\mathfrak{F}^{-1}(\hat{R})$ to find the uncertainty in the parameters by taking an estimate of the SD as the square root of each diagonal element in $\mathfrak{F}^{-1}(\hat{R})$,

$$SD(\hat{R}_i) = \sqrt{diag_i(\mathfrak{F}^{-1}(\hat{R}))} \tag{18}$$

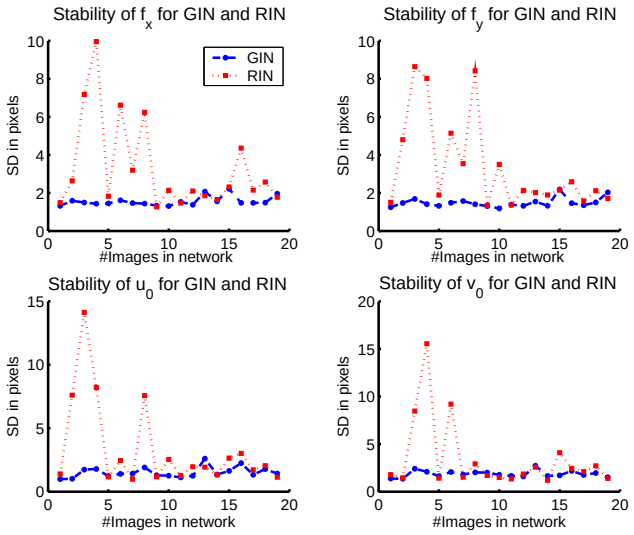
1) Three and four image networks: This section compares three and four image networks. Three image networks have an optimal angle between CLs of 60°, whilst for four image networks, the angle is calculated as 45°. 27 and 19 RINs and GINs were captured for three and four image network testing respectively. The accuracy of the calibration is given by the SD associated with the uncertainty of the parameters estimated. Figure 5(a) illustrates the results for three image networks. The uncertainty of the GIN parameters is significantly less than the RIN results. In some cases, the RINs have produced parameter estimates with SD as high as 20 pixels whereas the GINs are consistently below 5. Table IV presents the average of the SD across all image network instances. Four image network results are illustrated in Figure 5(b). Although the RIN uncertainty has decreased in comparison to the 3 image networks, the estimates remain unreliable with SD in some instances of 15 pixels. In comparison, the precision of the GINs estimates are stable with little uncertainty. The GINs are seen to significantly outperform RINs in terms of the accuracy of the parameter estimates both for three and four image networks.

TABLE IV
PARAMETER UNCERTAINTY FOR THREE AND FOUR IMAGE NETWORKS. NOTE: RESULTS GIVEN IN TERMS OF SD (IN PIXELS)

	f_u	f_v	u_0	v_0	φ	k_1	k_2
3 RIN	6.88	5.67	5.50	7.74	2.02	0.008	0.050
3 GIN	1.79	1.59	1.76	1.99	0.84	0.005	0.040
4 RIN	3.09	3.31	3.31	3.39	0.85	0.007	0.037
4 GIN	1.55	1.48	1.51	1.85	0.74	0.005	0.036



(a) Three image networks



(b) Four image networks

Fig. 5. Stability of intrinsic parameters for three & four image RINs and GINs for 2 to 25 image networks

IV. APPLICATIONS

The primary focus of this paper has been to enable non-expert practitioners determine precise camera parameters when calibrating a camera. An indicative experiment is performed (in Section III) in which the non-expert's ability to reproduce the generated images was tested. It was found that non-

experts were capable of reproducing the generated images with sufficient accuracy to achieve accurate calibration results. These results were achieved by using an auxiliary monitor to display the generated image. One proposed application of the method developed in this paper is to aid smartphone users to calibrate the camera on their phone. An Augmented Reality (AR) application in which the optimal generated image is superimposed in a semi-transparent manner onto the current camera view would be suitable for this task. Thus the image capturing process is reduced to simply visually aligning the semi-transparent synthetic image to the current image of the planar grid. This ensures that users will achieve accurate calibration results with little input to the calibration process. With accurate calibration parameters, image quality can be improved by removing lens distortions in addition to enabling better quality AR applications for mobile phones.

V. Conclusion

This paper presents a new, semi-automatic approach to aid planar camera calibration practitioners obtain more accurate results. This is achieved by using optimal image networks, which are generated based on geometric properties of the image, that are represented by the Centre Line (CL). The proposed procedure manipulates characteristics of an initial image supplied by the user to generate optimal image networks which are subsequently captured by the practitioner.

Experimental results have shown that the relationship between CLs of images can be manipulated to achieve improved calibration results. Synthetic results illustrate that less images are needed in GINs than in RINs to achieve accurate calibration results. In addition, indicative tests with real data show that GINs are easily reproduced by the practitioner. Tests with real images were performed that demonstrate the improved accuracy of GINs over RINs for the cases of three and four image networks. The main benefits of using GINs are the significant reduction of images needed in an image network and the improved calibration precision. Our results show that an accurate calibration can be achieved with five image GINs in comparison to 25 image RINs. This is of major benefit to any planar camera calibration practitioner.

Acknowledgment

This research was conducted in the Centre for Image Processing and Analysis at Dublin City University, Ireland. It was funded by the Irish Research Council for Science, Engineering and Technology: funded by the National Development Plan.

References

[1] [Online]. Available: http://www.strategyanalytics.com/default.aspx?mod=reportabstractviewer&a0=6216

[2] Z. Zhang, "Flexible camera calibration by viewing a plane from unknown orientations," in *The Proceedings of the Seventh IEEE International Conference on Computer Vision, Vol.1, Iss., 1999*, 1999, pp. 666–673 vol.1.

[3] P. Sturm and S. Maybank, "On plane-based camera calibration: A general algorithm, singularities, applications," in *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on.*, vol. 1, 1999, pp. –437 Vol. 1.

[4] O. Faugeras, *Three Dimensional Computer Vision: A geometric Viewpoint*. MIT Press, 1993.

[5] J. Heikkila, "Geometric camera calibration using circular control points," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 10, pp. 1066–1077, October 2000.

[6] R. Tsai, "A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses," *Robotics and Automation, IEEE Journal of*, vol. 3, no. 4, pp. 323 –344, august 1987.

[7] G.-Q. Wei and S. D. Ma, "Implicit and explicit camera calibration: theory and experiments," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 16, no. 5, pp. 469 –480, may 1994.

[8] J. Mallon and P. F. Whelan, "Which pattern? biasing aspects of planar calibration patterns and detection methods," *Pattern Recognition Letters*, vol. 28, no. 8, pp. 921–930, June 2007.

[9] J. I. Gonzalez, J. C. Gmez, C. G. Artal, and A. N. Cabrera, "Comparative analysis of calibration methods for a static camera," *VI Workshop de Agentes Fsicos (WAF'2005)*, 2005.

[10] J. Salvi, X. Armangue, and J. Batlle, "A comparative review of camera calibrating methods with accuracy evaluation," *Pattern Recognition*, vol. 35, no. 7, pp. 1617–1635, July 2002.

[11] W. Sun and R. Cooperstock, "An empirical evaluation of factors influencing camera calibration accuracy using three publicly available techniques," *Machine Vision Applications*, vol. 17, pp. 51–67, March 2006.

[12] J. Wang and Y. Liu, "Characteristic line of planar homography matrix and its applications in camera calibration," in *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, vol. 1, 0-0 2006, pp. 147 –150.

[13] B. P. Byrne, J. Mallon, and P. F. Whelan, "Efficient planar camera calibration via automatic image selection," in *VISAPP (1)*, 2009, pp. 90–94.

[14] P. Gurdjos, A. Crouzil, and R. Payrissat, "Another way of looking at plane-based calibration: The centre circle constraint," in *European Conference on Computer Vision*, 2002, p. IV: 252 ff.

[15] R. Hartley and A. Zisserman, *Multiple view geometry in computer vision*, 2nd ed. Cambridge, UK: Cambridge University Press, 2003.

[16] J. Mallon and P. F. Whelan, "Projective rectification from the fundamental matrix," *Image and Vision Computing*, vol. 23, no. 7, pp. 643–650, 2005.

[17] E. Walter and L. Pronzato, *Identification of Parametric Models*. Springer, (1997).