Advances in Close-Range Photogrammetry

Clive Fraser, Melbourne

ABSTRACT

This paper presents a general overview of notable developments in close-range photogrammetry (CRP) over the past four decades, during which time both automated processes were introduced and the full transition from film-based to digital imaging cameras occurred. With these developments and more recent innovations centred upon so-called structure-from-motion approaches to network orientation, along with dense image matching, CRP has over recent years attracted a broader user community who are applying the technology across a host of new application areas. The focus of the paper is upon the evolution from manual to automatic image orientation, and from manual feature point measurement through to automatic generation of dense 3D point clouds. Three main orientation and 3D point determination data processing options are discussed, along with the need to consider processing pipelines that integrate these three principal workflows, which are illustrated via practical examples. The paper highlights that while targetless multi-image orientation and dense matching have provided significant impetus to CRP, such fully automatic photogrammetric measurement does not necessarily provide a panacea for all measurement ills.

1. INTRODUCTION

As well chronicled through presentations at this and past *Photogrammetric Week* conferences, Photogrammetry has dramatically changed over the last four decades. This period from the mid-seventies to the present day corresponds to that in which our retiring *PhoWo* host, Prof. Dieter Fritsch, enjoyed his very successful career. During this period there has been a complete evolution from analog photogrammetry with manual image mensuration to fully automated digital photogrammetric measurement processes. Across the spectrum of application areas for photogrammetry, close-range photogrammetry (CRP) stands out as a conspicuous early-adopter of automation and digital techniques, although much of this development occurred on the sidelines of mainstream photogrammetry and it has thus gone under recognised to some extent.

The aim of this presentation is to highlight 40-odd years of advances in CRP through the prism of evolving photogrammetric orientation and 3D object measurement/reconstruction approaches. Ironically, whereas fully automated measurement processes have become routine in some application domains, such as vision-based industrial metrology, their use in others is still quite limited due to the nature and requirements of the tasks. This is exemplified by traffic accident reconstruction via photogrammetry.

The author has avoided citing specific published works in this paper, not because of any shortage of comprehensive accounts of developments in CRP, but more because the present account has been prepared as a general overview.

2. ORIENTATION AND 3D POINT DETERMINATION APPROACHES

Shown in Figure 1 is a workflow diagram that indicates the three presently adopted approaches for close-range photogrammetric network orientation, along with three categories of processing for 3D point determination. As can be seen from the figure, the distinction between the three orientation options is basically in the degree of automation, from manual image measurement (A), though automatic measurement of targets (B), to a fully automatic and targetless process (C). Irrespective of the image measurement and initial relative orientation approach, the assumption is that a final, refined

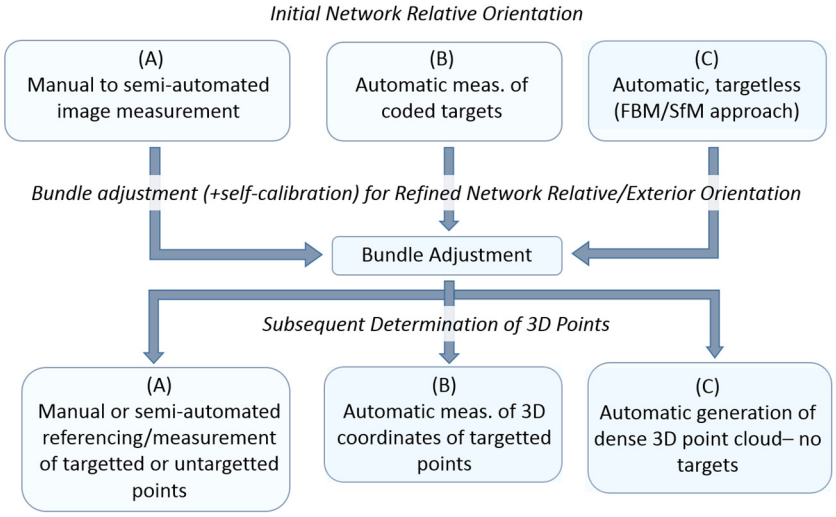


Figure 1: Data processing options for CRP orientation and 3D point determination.

exterior orientation will be generated through bundle adjustment, with or without self-calibration, and with or without constraints imposed through supplementary information, such as redundant scale constraints or positional constraints upon camera stations and/or object 'control' points.

With a refined exterior orientation in place, there are again three options for object point determination, these being manual or semi-automated measurement of targets, fully automatic target measurement, and automatic generation of point clouds. Although these three processes are presented in Figure 1 as post exterior orientation operations, they invariably form an integrated component of the bundle adjustment process, at least for cases of automatically measured targets and sparse point clouds arising from feature-based matching (FBM) or structure from motion (SfM) approaches. Dense point cloud generation via image matching would still be carried out as a post-orientation process.

A very important aspect to note at this point is that in CRP measurement the project requirements often dictate the adoption of 'mixed' network orientation and 3D point determination approaches. For example, it is not uncommon to utilize coded targets or the FBM approach for automated exterior orientation, followed by the manual measurement of 3D feature points of interest. Moreover, an initial precise network orientation via coded targets may well be followed by the generation of a 3D point cloud via dense matching. Thus, it is very beneficial to have all the possibilities offered through Figure 1 within a single data processing system. The *iWitnessPRO-Agilis* software (www.photometrix.com.au) exemplifies such a system.

A primary aim of CRP measurement, either as an initial or final outcome, is typically the generation of 3D coordinates of specific feature points of interest, which will then form the basis of subsequent mapping, diagramming or dimensional analysis. It is noteworthy that this aim is directly realized through both the traditional manual measurement approach and the use of targets. However, it is not a direct outcome of the targetless automated orientation and point cloud generation process. Here, it could be said that another form of raster data is being generated, albeit one that is structured in a more convenient 3D form. In situations where the measurement requirement is a fully triangulated (meshed) model this is of little consequence, but if coordinates of specific feature points are required then a further feature extraction or measurement process must be undertaken.

The advances in CRP that have accompanied the evolution from analog to digital imaging, along with the introduction of measurement process automation, can be conveniently summarized through consideration of current practices in each of the workflows (A), (B) and (C), indicated in Figure 1.

3. MANUAL IMAGE MEASUREMENT

In this era of automation, with SfM techniques being viewed by some as image-based measurement panaceas, a reasonable question to ask is just why we are considering manual image measurement at all any more. The simple answer is that for a very large user base in CRP, there is no other option. Situations that compel the use of manual measurement to establish feature correspondences include:

- Objects or scenes where image texture does not support feature detector-based matching;
- Network geometry where perspective disparity between images is such that image matching is precluded (e.g. convergent and wide-baseline configurations); and
- Objects or scenes where targetting cannot be used.

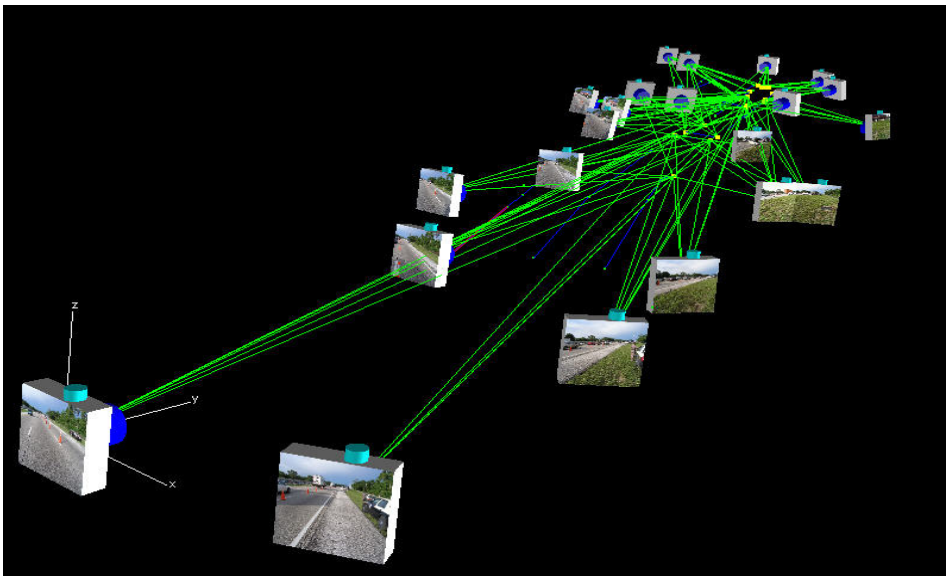
Applications that fall into this category are annoyingly frequent, one prominent example being traffic accident reconstruction. On the basis of software license sales, there are in excess of 2000 users of CRP amongst the traffic accident reconstruction and forensics community in the US alone, the vast majority being police and highway patrol departments. The adoption of CRP for accident reconstruction has primarily been driven by the need to reduce road closure periods at accident scenes. The fact that the imagery forms a permanent data record, which supports 3D measurement at any time after the event, along with the economy and ease-of-use of a measurement process needing only inexpensive consumer grade cameras and basic user skills are also important factors.

Returning to the issue of why manual image measurement is still needed, consider the example of the fatal accident shown in Figure 2. What is required in the court system of the US is invariably a 'diagram' as represented by Figure 2c. This mapping highlights features of the scene pertinent to both analysis of the event (e.g. calculation of speed and trajectories based on skid marks and vehicle 'crush') and any subsequent litigation. The diagramming is basic, and deliberately so, in order to enhance full comprehension of the event within a court of law. However, the imaging network required to generate the 3D data is anything but simple in a photogrammetric context. The network shown in Figure 2b is representative of most; the scene is effectively planar, with a height variation of generally a few metres at most over a 100+ m linear extent. Moreover, the camera stations are close to being all in the same plane (a potential nightmare for subsequent orientation, unless it is handled in a sequential network-building process). So, here there is a situation where two of the reasons for manual measurement listed above are illustrated.

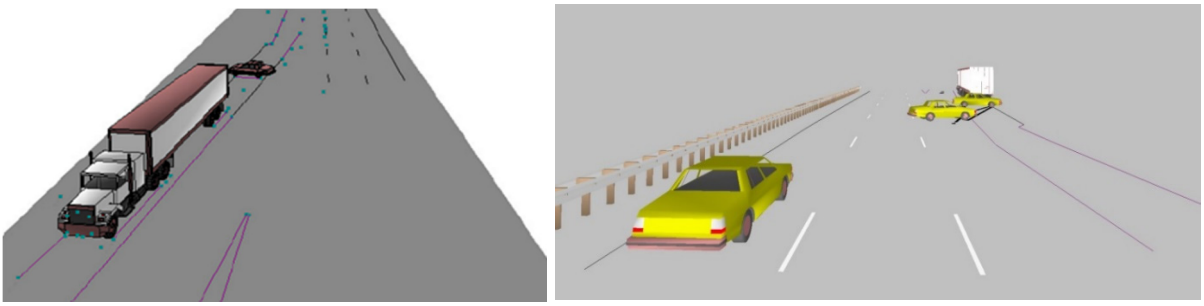
The third reason, the object/scene being non-conducive to targetting, need not apply in accident reconstruction. Seen in Figure 2a are 'evidence marker' targets. A degree of semi-automated measurement is afforded by such targets, which are used to highlight particular features of interest. An image scanning process can be used to find and measure the centroids of these markers, yet manual 'referencing' (conjugate point identification) is still required within the initial network building phase. Once relative orientation is established, tools such as 'resection-driveback' (automated scanning for targets based on predicted location) and other operator-assisted mechanisms can be employed, but the orientation remains a largely manual process. There are advantages here, especially in regard to error detection and quality assurance. Also, on-line quality control with errors being rectified as they are made, leads to robust and blunder-free solutions.

Many feature points of interest in accident reconstruction are not targetted, necessitating manual referencing in the post exterior orientation, 3D point determination phase. Once again, operator-assisted functions such as driveback can be employed to provide a degree of semi-automation, but the inability to utilize automated image matching typically remains. The network geometry and scene characteristics might strike photogrammetrists more familiar with short-baseline multi-view geometry as extreme, yet it must be recalled that this is the norm in arguably the largest user-community of low- to moderate-accuracy CRP measurement.

(a)



(b)



(c)

Figure 2: Traffic Accident Reconstruction: (a) Image from CRP network, (b) Network geometry and (c) Result of diagramming.

4. AUTOMATED MEASUREMENT USING TARGETS

Automated high-accuracy CRP gained its commercial start in the mid-1980s with the scanning of film from large-format cameras. Already at this time, researchers in CRP were having only modest success in exploiting video cameras, which proved deficient for moderate- to high-accuracy measurement. It was not until the early 1990s that the then termed high-resolution (1 – 1.5 megapixel) digital camera was incorporated into commercial vision metrology systems. These systems utilized targets, invariably retro-reflective targets, with coded targets being employed to provide an automated solution to the image point correspondence problem. In terms of the basic approach, vision metrology

systems, single or multi-camera, off-line or on-line, had reached maturity by the turn of the century. Whereas specially built cameras, and more lately modified higher-resolution DSLRs, have found application, and initial network orientation based on relative orientation has supplanted the resection-based exterior orientation (EO) device, the fundamental processing pipeline has remained as per process B in Figure 1.

Here, there is little crossover to the manual or FBM approaches, for reasons that are distinct for processes A and C. In the manual measurement case, untargetted feature points will have an accuracy that can be an order of magnitude less than targetted points, due to the disparity in image measurement accuracy, say 0.03 pixels for centroiding versus 0.3 pixels for manual feature measurement. This 10-fold accuracy difference is generally unacceptable, as are the time and operator skill levels required.

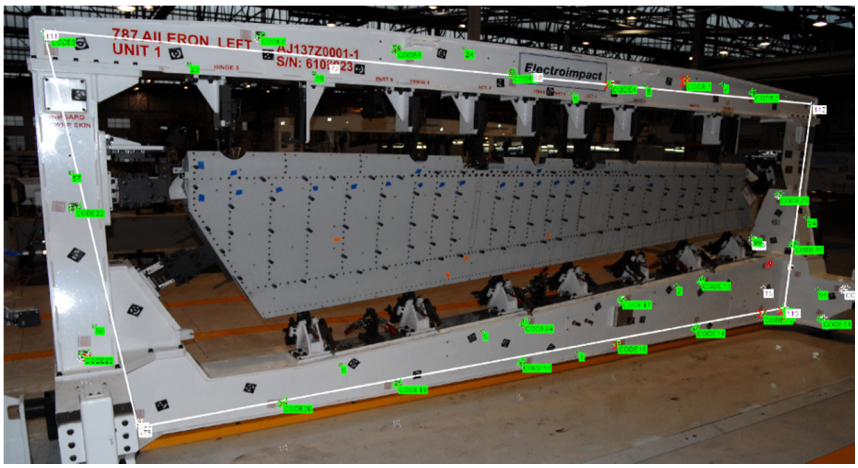
In the case of automated FBM or SfM-based point cloud generation, there are a number of problems associated with industrial measurement applications. Firstly, object texture (e.g. on machined parts and metallic surfaces) is not conducive to FBM. Secondly, imperatives of avoiding wide base-lines means that a weaker network geometry results. Thirdly, the previously mentioned 10-fold difference in accuracy is again present. Finally, and perhaps most importantly, the final 3D feature points generally do not represent the feature points of interest, which might be tooling points, corners, intersections, etc. Extraction of these features requires a further processing step.

In order to illustrate both the processing flow with automated target measurement, and shortcomings associated with the use of FBM-based orientation and point cloud generation in industrial photogrammetric measurement, consider the example of the mock 11-image measurement network for the 7 x 2.5m Boeing 787 Aileron tool shown in Figure 3a, which has been processed via the targetted and untargetted automatic orientation approaches, B and C (Figure 1), respectively. The results are summarised in Table 1.

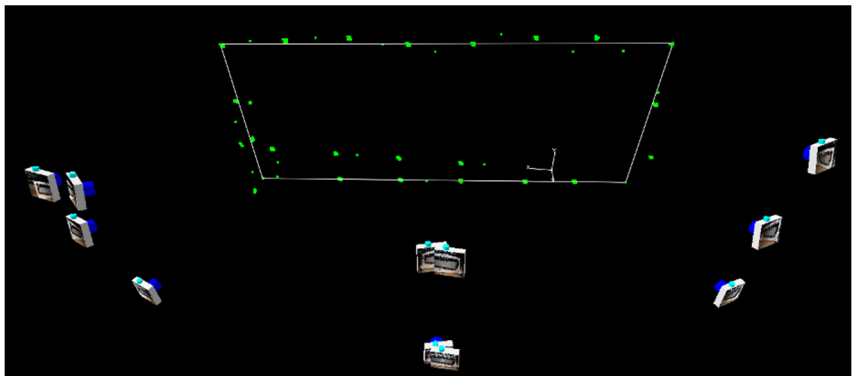
Table 1: Accuracy results from tool measurement via the targetted (Process B) and untargetted (Process C) approaches.

Bundle adjustment details	**Process B: targetted** (min. rays = 5)	**Process C: untargetted**, FBM (min rays = 3)	**Process C: untargetted**, FBM (min. rays = 5)
No. of 3D points	200	1980	490
RMS v_{xy}	0.07 pixels, 0.45μm	0.20 pixels, 1.2μm	0.23 pixels, 1.3μm
RMS σ_X (horiz. in plane of tool)	0.11 mm	1.0 mm	0.21 mm
RMS σ_Y (vert. in plane of tool)	0.08 mm	0.46 mm	0.42 mm
RMS σ_Z (horiz., depth)	0.14 mm	1.4 mm	0.20 mm
RMS σ_{XYZ}	0.12 mm (1: 63,000)	1.0 mm (1:7,000)	0.30 mm (1:19,000)

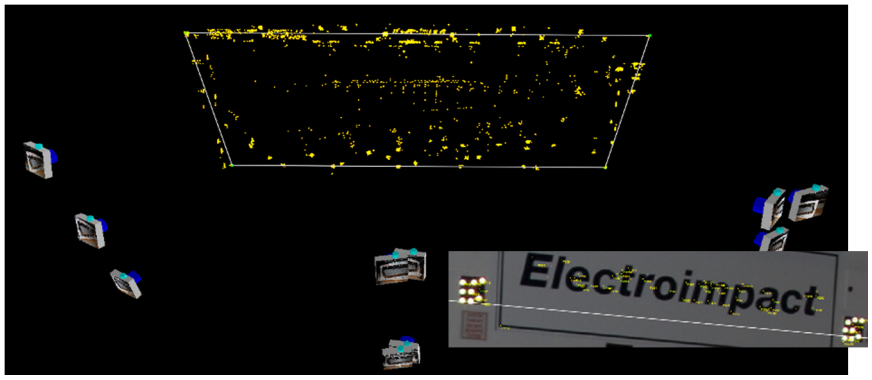
The camera used was an off-the-shelf Nikon D200 with a 17mm unifocal lens. In the targetted case (Figure 3b), which employed 22 coded targets for orientation and a further 24 'tooling point' targets (total of 200 points), 3D points with a minimum of 5 imaging rays were determined with a mean standard error (RMS σ_{XYZ}) of 1:63,000 of the principal dimension of the tool, with an image measurement accuracy (RMS v_{xy}) of 0.07 pixels. In the first of the targetless FBM orientations, where only 3D points with three or more rays were included (Figure 3c), the corresponding accuracy measures for the 1980-point network were 1:7,000 in object space and 0.20 pixels in image space. When only points with five or more rays are included (Figure 3d), the number of feature points drops to 490, but the accuracy improves to 1:19,000 in object space whereas image space precision falls a little to 0.23 pixels.



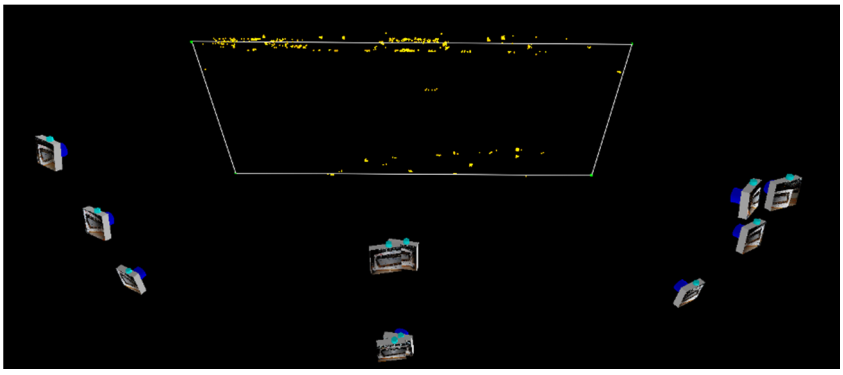
(a)



(b)



(c)



(d)

Figure 3: Automatic network orientation of a tool inspection measurement via both coded targets and the targetless FBM approach: (a) 7m x 2.5m tool, (b) targetted network of 22 codes and 24 tooling points, (c) targetless FBM orientation using 1980 feature points with 3 or more rays, and (d) FBM network with 490 points with 5 or more rays.

The results in Table 1 highlight the difference in accuracy between the targetted and untargetted approaches, this being attributable to the large difference in image measurement accuracy, though the three-fold difference in RMS v_{xy} values does not alone account for the nine-fold difference in object space accuracy between the targetted and three-ray FBM orientations. The differences in RMS σ_{XYZ} values are due to the impact of weaker geometric network strength: the targetted case can sustain much larger convergence angles and more imaging rays per point than the FBM network, which has a fundamentally weaker geometry because of the smaller maximum convergence angles and fewer imaging rays per point. Note also the relatively poor distribution of points in the FBM network. This arises from a concentration of feature point matches in areas of text on the homogeneously textured tool background surface (see the inset in Figure 3c). This example illustrates the practical difficulties associated with FBM-based targetless orientation in many industrial and engineering measurement applications, and it should be noted that the 3D points of interest, the tooling targets themselves, are not measured at the network orientation stage, whereas they are via Process B.

5. AUTOMATED TARGET-FREE MEASUREMENT

Beyond the field of industrial measurement, development of SfM-based multi-image network orientation over the last decade, along with its introduction into CRP over the past five years or so, has had a profound impact on the usability of image-based 3D measurement and modelling. Automatic, target-free 3D object reconstruction has been adopted across a very broad range of applications, from architecture and archaeology, through to cultural heritage recording, to 3D modelling for animation, and to human body measurement and modelling. There are many other applications, perhaps the most well-recognised to photogrammetrists being the generation of 3D models with associated DSMs and orthoimagery from networks of images recorded with UAVs.

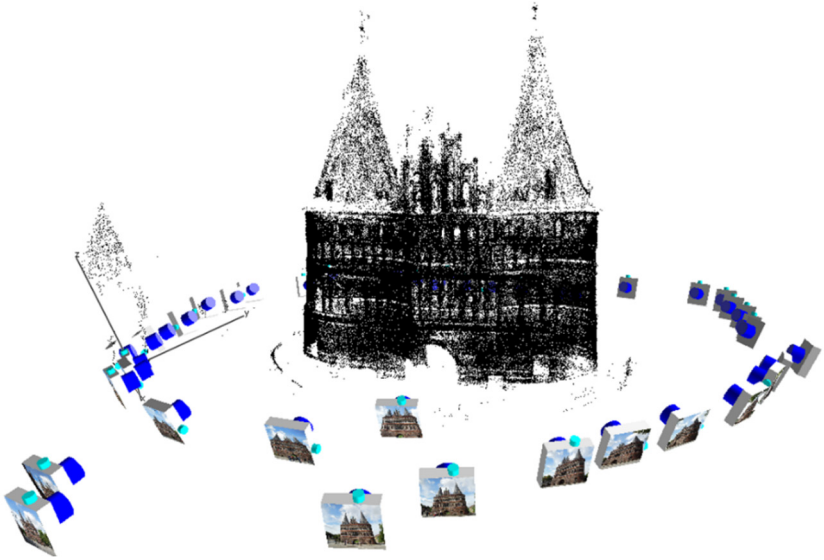
The term SfM is used rather loosely in the photogrammetric literature and it does not denote a specific workflow. To this author, at least, SfM is a means to solve the image point correspondence problem through pair-wise FBM, which implicitly includes associated filtering (e.g. via RANSAC and geometric consistency checks based on the Essential or Fundamental matrices). Once the image point correspondences have been established, network orientation and 3D point determination can follow traditional, rigorous photogrammetric processes, which are fully suited to automation and which possess all the internal quality control and accuracy measures familiar to photogrammetrists.

One does not need to look too far in the literature to find examples of multi-image network orientation via the FBM/SfM approach. With reference to Figure 1, it can be appreciated that fully automatic network orientation is invariably of more practical utility to the end user than the sparse 3D point cloud which is generated as an integral part of the process. Once the exterior orientation is established, both the manual extraction of feature points and the option of 3D point cloud generation through dense image matching (potentially to 1-pixel resolution) are available. Indeed, both can be employed such that the dense point cloud is supplemented with specific points of interest. Also, both point clouds, sparse and dense, can be employed to support monoplotting, whereby 3D feature points of interest can be digitised from single oriented images.

Figure 4a shows an example of an automatic targetless network orientation using the FBM approach. The sparse point cloud comprises > 200,000 points, though the automatic exterior orientation can be carried out with a lot less. All FBM points are imaged by three of more rays. Figure 4b shows a view of the final textured 3D model resulting from the post-orientation dense matching, with the dense point cloud comprising 50 million points.

As an illustration of both the potential and current limitations of FBM-based orientation followed by dense matching, consider its application to traffic accident reconstruction performed with imagery recorded from a UAV, as shown in Figure 5. An overview image of the accident scene is provided in

Figure 5a; the image geometry, automatic targetless orientation and accompanying sparse point cloud are indicated in Figure 5b; and a perspective view of the textured dense point cloud is shown in Figure 5d. Shown in Figure 5c is the manual measurement of feature points of interest made from the automatically oriented network. The 'diagramming' process, essentially simple 2D line diagrams currently used to document the scene, supports accident analysis and provides evidence for the courts. Of course the majority of the information recorded in the diagramming process could also be directly digitized from the 3D model of Figure 5d, and some would argue that the interpretability of the textured model is superior to that of a 2D 'map'. There are good prospects that such photo-realistic 3D representations will eventually find their way into routine accident scene reconstruction (at least where UAVs are employed), but this will not remove the need for the measurement and highlighting of specific features of interest, as required by the courts. So, for the foreseeable future this process will rely largely upon manual measurement, either from the 3D model directly, or from the oriented imagery via multi-view triangulation or monoplotting, or from both.

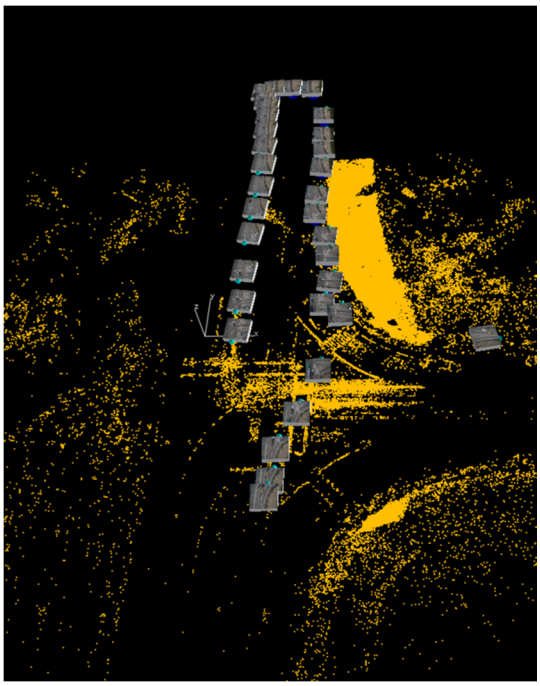


(a)

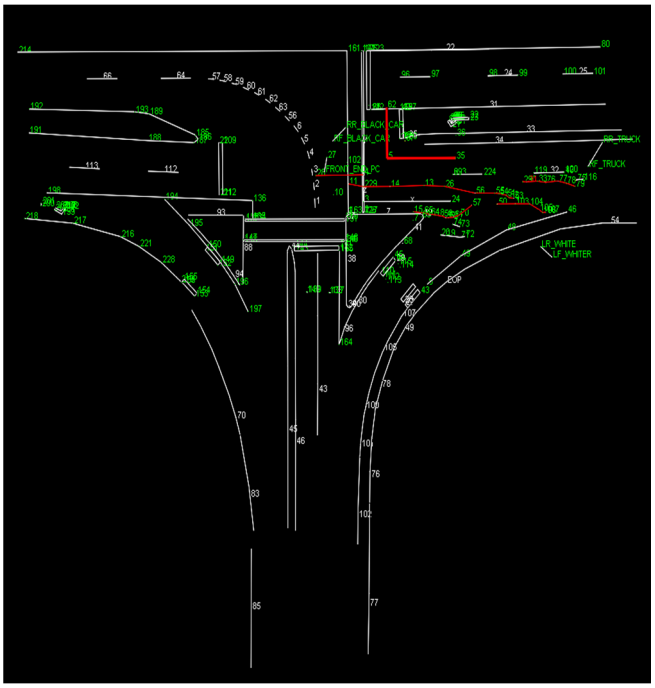
(b)

Figure 4: FBM-based automatic network orientation: (a) Network geometry of 43-station, 240,000-point network, (b) 3D model comprising 50 million points from SGM-based dense matching.

(a)



(b)



(c)

(d)

Figure 5: FBM-based automatic network orientation and dense matching in traffic accident reconstruction: (a) accident scene, (b) network of UAV-recorded images, (c) manual feature measurement process, and (d) 3D model.

6. AUTOMATIC CAMERA CALIBRATION

A currently under-recognised, yet very useful application of automated FBM-based orientation in CRP is camera calibration. Automatic calibration of digital cameras in CRP has, until recently, required the use of coded targets, and while targets continue to be used extensively, especially in high-accuracy industrial and engineering measurement, their use is nevertheless inconvenient in many applications, eg with UAVs. Nowadays, a practical, stand-alone targetless camera calibration is achievable via a process that combines SfM methods with rigorous photogrammetric orientation and self-calibration. The development of such a process has needed to account for perceived drawbacks of the FBM orientation approach, especially in regard to accommodating wide-baseline, convergent imagery as well as the lower image measurement accuracy of descriptor-based matching. Offsetting these potential limitations to high accuracy self-calibration is the huge observational redundancy and associated internal reliability that is afforded by having object point arrays comprising thousands of untargetted feature points, as opposed to only tens or hundreds of target points.

The CRP network shown in Figure 6 illustrates the effectiveness and practicability of automatic, target-free camera calibration, the camera in this case being an off-the-shelf 10.2 megapixel Nikon D200 DSLR with a 17mm unifocal lens. The only concession made in regard to rendering this camera metric was to tape the lens barrel so as to ensure a fixed focus (of nominally 5m). A convergent multi-station camera station configuration incorporating a diversity of camera roll angles and a non-planar object field was adopted, with the texture on the light sandstone wall (the object) being not 'rich', but rather moderate and favourable for feature point detection and matching. There was also some vegetation across the base of the 6m section of wall. Some 24,500 feature points were recorded within the 27 images (each point was seen in a minimum of 4 images). An array of 25 coded targets (200 points) was used to provide a baseline calibration against which to assess the targetless approach.

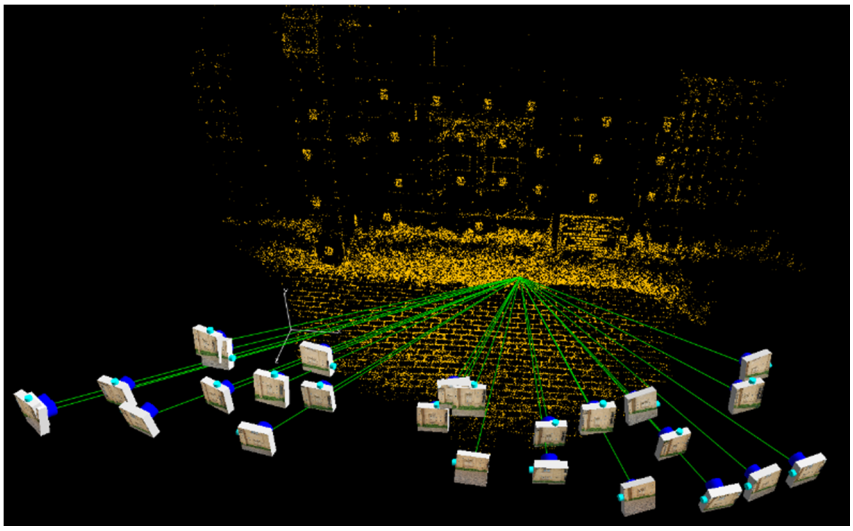


Figure 6: FBM-based automatic camera calibration via a network of 27 images and 24,500 points.

Table 2 provides a summary of the self-calibration results; calibrated values are listed for the focal length c, and principal point offsets x_p, y_p, along with their estimated standard errors. Also listed are radial distortion correction values Δr at three selected radial distances, and two decentring distortion profile values *P(r)* for two radial distances. The reason for reporting lens distortion in this manner is that it provides a more easily interpretable indicator of the repeatability of the computed distortion profiles than would be the case if only polynomial coefficients were listed. The *RMS* v_{xy} value and the number of object points for each calibration are also shown in the table. The repeatability between

the targetted and untargetted cases can be seen to be very high for the lens distortion parameters, and high for the interior orientation elements, being to within 1 μm (0.16 pixel) for all values listed, other than the principal distance.

Table 2: Results of self-calibrations of the Nikon D200 camera for targetted and untargetted cases.

Case	Focal length, c (σ_c) mm	x_p (σ_{xp}) mm	y_p (σ_{xp}) mm	Δr @ r=8mm μm	Δr @ r=10mm μm	Δr @ r=12mm μm	P(r) @ r=10mm μm	P(r) @ r=12mm μm	RMS v_{xy} No of pts
Coded targets	17.632 (0.0011)	-0.040 (0.0008)	-0.193 (0.0008)	121.7	217.3	333.2	5.7	8.1	0.10 pixel 200
Untargetted (FBM)	17.620 (0.0007)	-0.040 (0.0004)	-0.192 (0.0004)	121.6	217.5	334.8	5.2	7.5	0.34 pixel 24,500

The 'true' calibration values of the D200 were not known (they rarely are!), and the quality and fidelity of the self-calibrations could only be assessed via internal means, with two measures being relied upon. These were the precision and repeatability of recovery of the calibration parameters, and the resulting discrepancies in object space coordinates when these parameters are subsequently applied. In regard to internal accuracy indicators there was, as anticipated, a 2 – 3 times discrepancy between the accuracy of image coordinate measurement in the targetted and untargetted cases, with the RMS v_{xy} values being 0.1 pixel for the targetted case and 0.35 pixels for the FBM case. It is interesting to note that there were no common points between the two cases. Although the same images were used, none of the coded target 'dots' were matched. On the basis of the difference in triangulation closure alone, it could be anticipated that the precision of recovery of calibration parameters would be better for the targetted case. However, the standard errors of calibration parameters were in fact superior for the untargetted network adjustments, simply because there were so many more matched feature points than coded targets, in this case more than 60 untargetted points for every artificial target. This illustrates that the FBM approach, coupled with very dense point fields of thousands of points, can yield camera calibration parameters to higher precision than is achievable from targetted arrays comprising a few hundred points. The same phenomenon was experienced with the development of image matching-based relative orientation on photogrammetric workstations back in the 1990s.

In regard to convergent imaging, there was one noteworthy surprise. Whereas it could be anticipated that accurate centroiding on high-contrast circular targets would be possible to incidence angles of 30 degrees to the target plane, it was thought less likely that descriptor-based matching of feature points would be able to accommodate moderately high convergence angles and hence the imaging geometry of the targetted array would be stronger due to points having a higher number of imaging rays over a wider diversity of viewing angles. This example illustrated that FBM is able to accommodate relatively wide baseline configurations in favourable situations, resulting in many points having effective convergence angles between imaging rays of 60 – 90 degrees. Care should be exercised, however, in situations where baselines are reduced to enhance FBM, for the resulting reduced convergence angles can increase the likelihood of excessive correlation between the parameters of interior and exterior orientation, thus potentially reducing the degree of scene-independence of the self-calibration. The success with the multi-ray matching of features in the reported case, also resulted in object points with better distribution in three dimensions than with the targetted points.

This example of target-free camera calibration demonstrates that camera parameters of greater precision and equal accuracy to those recovered using coded targets and the 'standard' automatic self-calibration approach can be achieved. The poorer image point measurement accuracy of descriptor-

based feature point matching is more than offset by the provision of potentially 100-fold more object points within the photogrammetric network. Within CRP, both approaches fit well into automatic data processing pipelines (Processes B and C). From a practical standpoint, if the scene or object being imaged is texture rich and conducive to the target-free approach, then this is arguably the more flexible automated camera calibration option and it represents a significant milestone in modern-day CRP.

7. CONCLUDING REMARKS

Few would argue with the observation that CRP has never been more widely employed across a broader range of 3D measurement applications than it is today; the technology is thriving at present! From the outset of the era of automation in the mid-1980s, though to the present, there have been notable development steps, each of which has been accompanied by an increase in the adoption of multi-image CRP. These steps have included image scanning for automatic image coordinate mensuration; coded target-based automatic network orientation; the adoption of off-the-shelf digital cameras (modified or not); SfM-based targetless network orientation; and dense image matching for 3D model generation.

Moreover, active development continues, for example in the area of utilization of non-traditional imaging systems such as time-of-flight 3D cameras; hemispherical imagery; mirror systems; and all-reflective optical systems. Regardless of new innovations, there is every chance that data processing options will remain for the foreseeable future as characterised by Figure 1, with the approaches described in this paper, from manual image measurement through to fully automatic 3D object reconstruction, being embraced.