ESTIMATION OF HEIGHTS OF ROBBERS USING PHOTOGRAPHIC AND VIDEO SECURITY IMAGERY

J.H. Chandler[1] and D.M. Stirling[2]

1. J.H. Chandler BSc PhD ARICS, Contracts Manager.
2. D.M. Stirling BSc, Lecturer.
   The Engineering Surveying Research Centre
   Department of Civil Engineering
   The City University
   LONDON EC1V OHB, UK

RESUME:

This paper describes work carried out to provide expert evidence in criminal trials. Heights of bank robbers were determined using analytical bundle methods with various non-metric and video imagery from security cameras. Among the problems encountered were monoscopic coverage, poor control and low image resolutions. Data analysis techniques using a CAD system to estimate the highest point on the head from the photogrammetrically determined discrete points around it are detailed. Finally, problems associated with selection of the correct functional model and interpreting the results in a legal environment are highlighted.

KEY WORDS: Analytical, Photogrammetry, Bio-stereometric, Non-metric, Legal

1. INTRODUCTION

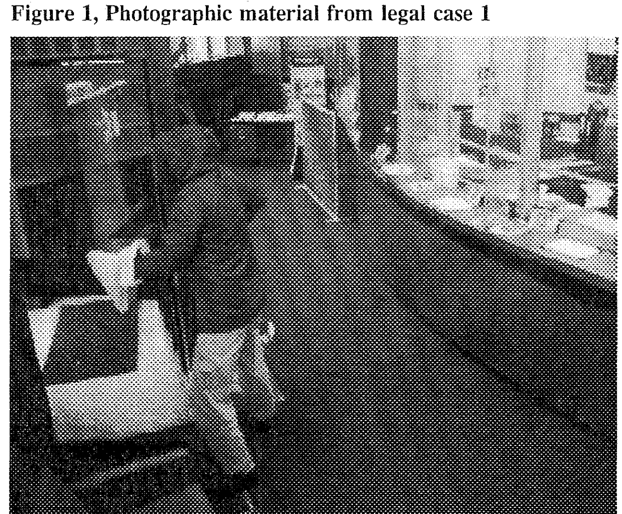
The aim of this paper is to indicate how the Engineering Photogrammetry Unit (EPU) made use of a variety of photographic and video security imagery to determine the height of persons robbing banks and a betting office in the UK. The work was carried out on behalf of both defence and prosecution counsel, for the purpose of providing legal evidence in the British Crown Court.

A major problem with this type of analysis is the design of the photogrammetry. Typically, geometric strength of the network is weak due to a limited number of cameras, frames and non-stereoscopic coverage. Such a weak network yields an (externally) unreliable estimation, compounds the problem of selecting the optimum functional model and inaccurate results can be obtained. This particular problem is illustrated in the analysis carried out by the Engineering Photogrammetry Unit and shows that not only is suitable software required but extreme care in measurement, processing and particulary analysis is essential. Despite these concerns, evaluation of the technique in a controlled environment prove that accurate results can be obtained, although selecting the correct functional model remains a critical problem.

2. THE LEGAL APPLICATIONS

During late 1990 and early 1991 the Engineering Photogrammetry Unit (EPU) was approached on three separate occasions to undertake photogrammetric analyses of material obtained by various security imaging systems during an armed robbery. The aim was to estimate the height of the robber using the visual evidence which could then be compared with the measured height of the suspect.

Figure 1, Photographic material from legal case 1



On two occasions photographic imagery was available (Figure 1, 2) and this type of material was found to be

Figure 2, Photographic material from legal case 2

infinitely more preferable than video (Figure 3) for this type of analysis. The resolution of the images was greater and allowed more detail to be identified and measured. The department had the necessary photogrammetric hardware for highly precise measurement and has built up robust and proven software necessary for restitution. The photographic material had been acquired with a 35mm camera with characteristics of inner orientation which are more predictable and stable than simple video security imaging systems. As photogrammetrists we were far more familiar with this type of medium and were fully aware of it's potential and likely pitfalls.

2.1 <u>Case Study- video imagery</u>

Although all three cases were distinct, the procedures adopted for analysis were broadly similar. Images suitable for measurement had to be identified and a control survey carried out at the premises of the bank. Image measurement and consequent processing produced three dimensional data representing pertinent features of the robber. Analysis of these data within a CAD environment produced the final result. A written legal report was then submitted and the case contested in the British Crown court.

The case involving video imagery reveals the importance of selecting the correct functional model during photogrammetric restitution and will be used in this paper to illustrate this key, problem. The video case study will also be used to portray some of the other difficulties common to all three contracts. The procedures were similar in all cases, the only important difference concerned initial point measurement. In the photographic cases this phase was carried out using an analytical plotter.

Before the Engineering Photogrammetry Unit agreed to carry out the analysis using the video imagery (Figure 3), several problems were immediately

Figure 3, Video material from legal case 3

apparent:

1. Although two cameras were installed in the betting office the recorded images were alternately sequenced so that no stereoscopic overlap was created at any instant of time. This created the problem that true three dimensional coordinates could not be obtained for any moving object such as a robber.
2. Although a number of static and well defined points which imaged on the video could serve as photogrammetric control, their distribution was poor and total number limited.
3. Two CCTV video cameras were used to record the images and such cameras are not designed for photogrammetric use. The image quality was also poor.
4. Only two weeks were available for the whole project during which all survey, measurement, processing, analysis and report writing had to take place.
5. The human body is not a fixed and well defined linear object and so a photogrammetric estimation of a person's height is dependant upon other factors. Typically these are associated with posture, stance and thickness of hair. In this particular case the problem was compounded by the robber wearing a woolly hat.

Despite these concerns it was felt that the Engineering Photogrammetry Unit could produce a result, qualified by a stochastical measure of quality which would reflect the uncertainties indicated above. It would then be up to the jury to decide upon the relevance of a photogrammetric estimation compared with the known measured height of the defendant. The decision was taken to proceed.

2.2 Control survey

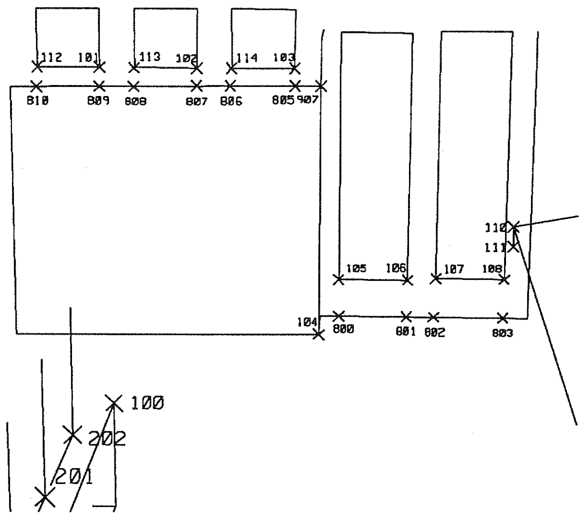


Figure 4, Control points

A survey was carried out at the premises of the betting office using a surveyor's tape and level. From these measurements the three dimensional coordinates of prominent features imaged on the video were determined, (Figure 4). These points would serve the function of photogrammetric control, points determined to a precision of +/-0.003m in plan and +/-0.002m in height. The survey also encompassed the positions of the perspective centres of the two video cameras, their precision determined to +/-.05m.

2.3 Frame grabbing

Critical frames were identified and attempts made to frame grab and store these data on disc for consequent measurement. Access to a Personal Computer (PC) with a frame grabbing board was obtained but the task proved difficult. One major requirement of a video surveillance system is to record many hours of business before the videotape needs substitution. This particular system achieved this by recording a frame once every three seconds and alternating between each of the two video cameras. The robber was only on the premises for a few minutes and so the number of frames where the robber was standing still, erect and with both feet and top of head imaged together was strictly limited. Moreover, it was found that if the video player was placed on 'pause' for display and grabbing of any particular image, the frame often broke up and distorted badly. Frame grabbing was only possible when the video player was in 'play' mode and it was then extremely difficult to grab any one specific image. After many hours spent attempting to grab specific frames only four acceptable frames were obtained, three recorded by one of the cameras. The frame grabber had a resolution of 512 x 512 pixels, each with 256 grey levels.

2.4 Image measurement

Using the PC and basic image processing software the pixel coordinates of all visible control points were recorded, each measured three times. The pixel coordinates of a series of points representing the positions where the robber's feet were in contact with the floor were also measured, (foot points). Similarly, coordinates representing points positioned around the robber's head were recorded, (head points). These points were taken where the woolly hat appeared to be pressing hard against the robber's head and where a curve, taken to represent the head, was visible on the image. These image coordinates were measured with an estimated precision of +/-0.5 pixel.

2.5 Data processing

The measured data were processed using analytical photogrammetry. Analogue methods would not have provided sufficient flexibility to obtain a solution.

A video camera is not necessarily designed to meet the condition of collinearity and the focal lengths of these particular cameras was unknown. A self-calibrating bundle adjustment is often used to overcome this type of problem (Brown, 1976; Granshaw, 1980; Chandler & Cooper, 1988) and four separate estimations were used in which the focal length of the camera was determined. Although the program could determine parameters to model lens distortion these additional parameters could not be estimated with statistical significance.

The lack of stereoscopic overlap due to the sequencing nature of the cameras prevented three dimensional coordinates from being derived by photogrammetric measurements alone. By making certain geometric assumptions it was possible to obtain the three dimensional coordinates of the measured head and foot points. Two geometric assumptions were made:

1. The robber was standing with his feet flat on the floor. The Z ordinate of the foot could be assigned that value derived by levelling the floor, so allowing the X and Y ordinates of the feet to be determined.
2. The plan position of the robber's head was mid-way between those of his feet. The orientation of the object coordinate system was selected so that the Y ordinate of the head points could be assigned the mean Y ordinate of the feet points. This allowed the X and most importantly the Z ordinates for the head to be computed.

These two geometric constraints could be enforced readily in the self-calibrating bundle adjustment by 'fixing' the appropriate ordinate. By processing these measured data for each frame it was possible to derive three dimensional coordinates for points representing the robber's feet and head.

2.6 Initial determination of height of robber

The coordinates of both foot and head points were read into a Microstation graphics design file for graphical display on the screen of an Intergraph CAD system. It was impossible to identify the top of the head directly because the robber was wearing a woolly hat and so the highest point on the robber's head was estimated by fitting a best fit arc to the measured head points. It was then possible to derive an estimate of the robber's height by measuring the vertical distance between the foot points and the highest point on the arc, (Figure 5). This procedure was repeated for all

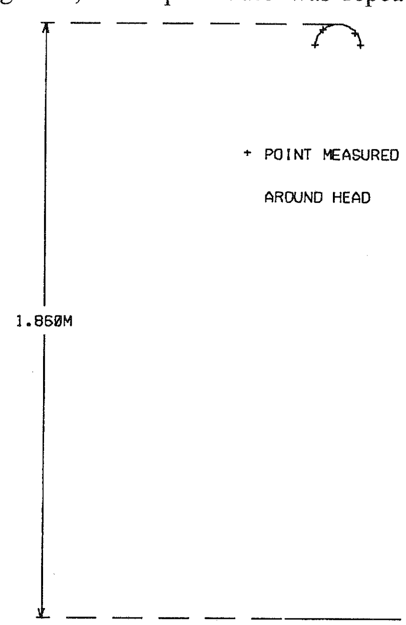


Figure 5, Estimation of height from CAD model

four frames and there was a variation of 0.084 (3.3") between these four estimates. This was thought to be attributable to differences in posture between frames, the poor distribution of control data and low resolution of the images.

Statistical analyses of these four estimations gave a mean of 1.888m and confidence, that the top of the robber's head was between 1.826m and 1.950m (ie between 5'11.9" and 6' 4.8") above the carpeted floor surface. No account could be taken of the thickness of the heel of shoe.

With the benefit of hindsight it is apparent that the *a posteriori* variance factors associated with the original bundle estimates were larger than would perhaps be expected. More exhaustive analysis was not carried out at this stage because the Metropolitan police confirmed that the height of their suspect was within 0.01m of the EPU mean estimate.

2.7 The missing systematic error

The Engineering Photogrammetry Unit were not the only expert witnesses who attempted to derive an estimate of the robber's height. Dr. Linney, a medical physician with expertise in facial mapping, had produced an estimate which differed from the initial EPU solution.

The 'Linney' estimate was based upon analysis of one single frame (Frame 2728, Figure 3). A known distance on the object (Figure 4, Point 104 --> 907) and the same distance measured on a photographic print was used to derive a single scale factor for the whole image, (Linney & Coombes, 1990). This single scale factor was applied to the distance measured on the print between the feet and head of the robber and used to estimate the robber's height at 5' 9.9", (Linney & Coombes, 1990).

Photogrammetrists are aware that there are several systematic error sources ignored by such an approach. Most obvious is relief and tilt displacement but also the inner geometry of the video camera is disregarded. Distortions are introduced also by taking a photograph of a video screen and producing the necessary paper print enlargement. Originally it was thought that these omissions in the simple Linney functional model were sufficient to explain the difference between the two estimates.

The effect of each type of normal photogrammetric source of systematic error was considered individually but none yielded a sufficient explanation for the large difference in estimations. Background reading on the nature of video imaging systems was carried out and it was realised that one other source of systematic error had been ignored in the original EPU solution. In a video camera variations in beam scanning velocity and geometry can lead to deformations in the image (McGlore, 1989). Tests carried out by El-Hakim *et al* (1989) show that electronic distortions of video tube cameras are largely systematic and stable and that a large proportion of the electronic distortion is caused by scale differences between the horizontal and vertical axes of the video image. More significantly, large amounts of video distortion can be removed by a simple scale correction, (El-Hakim *et al*, 1989).

In the original EPU approach image scale is defined by the relationship between control coordinates in the object space and their associated image coordinates. By making use of a standard self-calibrating bundle adjustment one mean scale factor was derived for both x and y image ordinates, introducing an important scaling error. The Linney solution used one height difference to define scale in the y image direction. Scale in the x direction was never determined but was not required for the purpose of deriving the height of the robber. The error sources inherent with the simple Linney approach discussed above were indeed present but their effects were insignificant in comparison to the scaling error introduced by EPU. Dr Linney's solution was therefore originally more accurate than the EPU approach, despite the simplifying assumptions underlying it. By modifying the functional model used in the bundle estimation to account for differential scale between x and y image coordinates, a more accurate and precise solution could be obtained.

2.8 Revised data processing

Additional programming was required to allow the General Adjustment Program (GAP) to derive a differential scale factor between x and y image coordinates. This was incorporated into the software by estimating corrections to two focal lengths, one associated with the x image ordinates, the second with those in the y image direction. The differential scale factor was calculated from simply the ratio between these two estimated focal lengths.

When these data were processed using the new version of GAP the scale factor was found to be 0.84 which represents a large difference of scale between the x and y image coordinates. With the effective removal of this particular source of systematic error it was found possible to include data from all four images in one combined solution, with two sets of inner orientation parameters associated with the two original cameras. This increased the geometric strength of the estimation to the extent that a parameter used to model the lens distortion of one camera, could be derived with statistical significance.

2.9 Revised determination of height of robber

The revised three dimensional coordinates of both the 'foot' and 'head' points were again read into an Intergraph Microstation graphics design file and the height of robber estimated using the techniques discussed earlier, (Section 2.5). There was now a variation of 0.049 metres (1.9") between the four estimates, considerably lower than the range obtained before. Statistical analyses of the four estimations gave a mean of 1.781m and a standard deviation of the mean of 0.011m. Therefore it was concluded, with 95% confidence, that the top of the robber's head was between 1.745m and 1.817m (ie between 5' 8.7" and 5' 11.5") above the carpeted floor surface.

3. CONCERNS

The difference between the initial and revised solution was over 0.10m (1.888m - 1.781m) which was of extreme importance when put into the context of a person's height and liberty! What concerned the Engineering Photogrammetry Unit most was missing the critical systematic error during analysis of output from the original self-calibrating bundle adjustments. The functional model selected originally was inadequate and examination of the output had not indicated this error.

Undoubtedly a major factor in this omission was the poor design of the photogrammetric network, which was common to all three cases. If EPU had been able specify the number, position and type of the cameras, their recording rate and the number and type of control points the geometric strength of the network would have been significantly stronger. Not only would this have produced a precise solution but the resulting large and perhaps systematic pattern of residuals would have indicated the presence of undetected systematic errors. Once the correct functional model had been identified accurate results would then be obtained. A network with high internal reliability (Hottier, 1976) can yield both a precise, and perhaps more importantly, an accurate estimation.

In these types of analyses the photogrammetry cannot be designed and is generally, if not always, poor. Although it is normally possible to obtain a solution this will be only accurate if the selected functional model accurately simulates those original ray bundles. In these circumstances only experience, patient examination and thorough testing will help the photogrammetrist from making potentially disastrous mistakes. Unfortunately the pressure of commercial contracts leave little time for such analysis!

4. A CONTROLLED SIMULATION

The application of photogrammetric techniques to solve this general type of problem was of interest to the Engineering Photogrammetry Unit and the concerns indicated in Section 3.0 justified further examination into the various factors affecting a solution.

4.1 The rigid robber

It was decided to replicate the geometric relationships between robber, security video camera and photo-control in a controlled environment. The robber was replaced by a survey staff located in front of an array of control targets normally used for camera calibration. The relationship between the staff and the control field was determined using a combination of taped distances and height differences. Black circular targets were placed over the control points and on the staff at heights 0.000m, 1.600m and 1.800m. These targets were of sufficient size to be imaged on the video camera at a range of image scales. A single Panasonic NV-MC 10B CCD home video camera was used to obtain a series of images of the staff and control field from various positions, these images stored on TDK EHG EC30 video cassette.

4.2 Image measurement

Access to a PC based video frame grabber was provided and six images grabbed at a resolution of 748 x 548 pixels in red/green/blue bands. Only the red images were used for further analysis. Image positions of the control and staff targets were measured using similar procedures to those discussed in Section 2.3. The geometric configuration between the six measured frames and the targets was convergent (Figure 6) and represented a network with stronger internal reliability than was possible in the betting office.

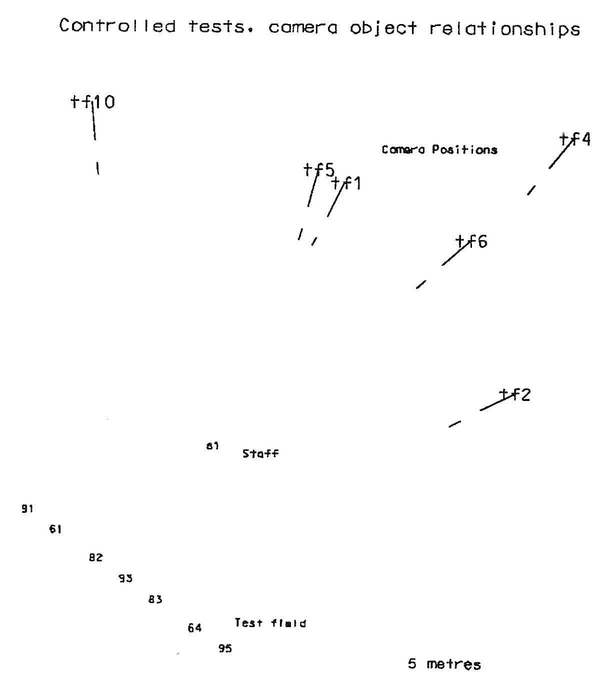


Figure 6

4.3 Data processing

These measured data were processed using the self-calibrating bundle adjustment with and without the differential scale factor carried as an additional parameter. Staff object points were treated in an identical manner as foot and head points previously, with the Z ordinate held fixed for the staff reading of zero and the resulting XY coordinates held fixed for upper staff points. During the running of the self-calibrating bundle adjustments a problem of slow convergence was encountered, although only when differential scale factor was included. This problem was found to be associated with correlation between the exterior and inner orientation parameters and was particularly acute where the control field was planar.

Four separate estimations were carried out in order to look at the role of the differential scale factor and the number of frames measured, either one or all six. An additional estimation was carried out using all measured frames, the optimum functional model and those taped distances and difference in height between the staff and the control points. This could be regarded as the best estimate for the staff position and staff height using all measured data.

4.4 Results

Table 1 indicates results obtained from these five different options.

Table 1

No. of Frames	Diff. Scale Factor	Est. Staff height (1.600m)	Plan position of staff (X)	Plan position of staff (Y)	Var. factor *a post.*	Opt
1	none	1.598 +/-15mm	102.445	112.065	2.56	A
6	none	1.568 +/-2mm	102.499	112.183	2.00	B
1	0.978	1.600 +/15mm	102.495	112.169	0.74	C
6	0.978	1.598 +/-3mm	102.504	112.183	0.90	D
6 + survey obs.	0.978	1.591 +/-3mm	102.506	112.195	0.87	E

The difference in scale between x and y image coordinates is clearly not as large as in the case cited earlier. This suggests that the value varies significantly between various PC based frame grabbing systems and so a means of estimating differential scale factor is clearly important. The estimated value is close too unity and the difference in estimated staff height with and without this additional parameter is not large, particularly with single frame estimations, (A/C). Indeed the estimated staff height is extraordinary close to the known value of 1.600m, although reference should be made to the low precision (+/-15mm) of these particular estimates. The most accurate and reliable estimate is represented by Option D in which all six frames are used and differential scale is included. The estimated plan position of the staff also agrees closely with known position represented by Option E and the *a posteriori* variance factor is also below unity suggesting that a suitable functional model has been selected.

5. OTHER CONSIDERATIONS

The relevance of a photogrammetric estimate of a persons height can perhaps be questioned, particularly for legal purposes. A comparison between an estimated height from an incidental photograph and a physical measurement of a person standing erect, is not necessarily valid. Posture, stance and footwear all effect the height of an individual. The linear effects of these sources of systematic error surely transcend the boundaries of photogrammetric expertise into the medical world. The additional 'expert witnesses' who would become involved with producing a result will inevitably confuse a jury already baffled by the science of photogrammetry. If the experts have not achieved this confusion with their own explanations then opposing counsel will gratefully assist and reduce the significance of the evidence.

It should be stated also that the human frame is not a fixed and linear object and a persons height is not a definitive parameter which can be used for identification. This type of evidence is perhaps best suited for defence purposes in which it may help to exclude an individual. The technique can not be used to prove that a particular person committed an alleged crime, even if the individual was exceptionally short or tall.

6. CONCLUSION

It has been demonstrated that with care photogrammetrists can determine geometric data from single camera security imagery. Security systems based upon photographic systems would appear more preferable than their video counterparts.

The value of strong network geometry should not be underestimated because such strength leads to an estimate which is both internally and externally reliable. Reliability can help identify a suitable functional model and will ultimately lead to both a precise and, perhaps more importantly, an accurate result. When network geometry is weak, extreme care is necessary at all stages of processing and analysis.

REFERENCES

BROWN, D.C., (1976). The bundle adjustment-progress and prospects. International Archives of Photogrammetry, 21(3) Paper 303:33pp.

CHANDLER, J.H. & COOPER, M.A.R., (1988). Monitoring the development of landslides using archival photography and analytical photogrammetry. Land and Minerals Surveying, 6:576-584.

EL-HAKIM, S.F., BURNER, A.W. & REAL, R.R., (1989). Video technology and real-time photogrammetry. In: Non-topographic photogrammetry, ed. by H.M. Karara, (Falls Church:American Society of Photogrammetry and Remote Sensing), pp. 279-304.

GRANSHAW, S.I., (1980). Bundle adjustment methods in engineering photogrammetry. Photogrammetric Record, 10(56):181-207.

HOTTIER, Ph. (1976). Accuracy of close range analytical solutions. Photogrammetric Engineering and Remote Sensing, 42(3):345-375.

LINNEY, A. & COOMBES, A.M., (1990). Report on comparison of photographs and video images. RE: Regina v McVicar. 4pp.

MCGLORE, J.C., (1989). Analytical data reduction schemes in non-topographic photogrammetry. In: Non-topographic photogrammetry, ed. by H.M. Karara, (Falls Church:American Society of Photogrammetry and Remote Sensing), pp. 37-57.