Automatic recognition of coded targets based on a Hough transform and segment matching

Mark R. Shortis[a], James W. Seager[a], Stuart Robson[b] and Euan S. Harvey[c]

[a]Department of Geomatics, University of Melbourne, Australia
[b]Department of Geomatic Engineering, University College London, England
[c]Department of Botany, University of Western Australia, Australia

ABSTRACT

This paper will present details of a coded target system that employs a Hough transform and segment matching to automatically recognise and identify the targets in digital images. The code system is based on a square surrounding the central circular target and will be described at a level of detail that would allow the system to be readily duplicated. Pre-detection processes, developed to improve the success rate under unfavourable conditions, and the tests conducted to validate a correct target match will also be described. Finally, the paper will include some examples of the use of the coded targets, drawn from calibrations of digital still cameras and underwater stereo-video systems.

Keywords: target, code, automatic measurement, Hough transform, segment match, digital image

1. INTRODUCTION

In the last two decades there have been a number of advances in the efficiency and effectiveness of image-based metrology systems. The use of digital images and digital transmission of images, the expansions in image resolution provided by new CCD arrays, the rapid measurement of targets using centroid computations or least squares matching (LSM), and the increasing sophistication of software for data analysis and geometric fitting are all factors that have contributed to the gains in efficiency and effectiveness. However, arguably the most important factor has been the automatic detection, recognition, identification and measurement of artificial targets used to signalise points of interest. The points of interest may be control locations, in order to determine the position and orientation of the cameras, or measurement positions that are the objective of the metrology application. The artificial targets are likely to be circular retro-reflective targets, but can also be passive circular targets or self-illuminated point-source targets. Irrespective of the type of target, the aim is automatically detect, recognise, identify and measure the point of interest.

Without the ability to recognise and measure targets automatically, the efficiency and response times of industrial metrology systems are severely limited by the necessity of manual measurement of target images. With the ability to recognise, identify and measure targets automatically, industrial metrology systems can achieve real time response rates, successfully measure thousands of targets or enhance accuracy by employing many, many digital images.

Many different schemes[1,2] have been proposed in order to provide semi-automated or automated target measurement. Semi-automation may involve an operator indicating a search area or the use of algorithms that make some assumptions about the locations of particular targets or the overall pattern of the target array. Whilst these approaches do improve the efficiency of image-based metrology, typically any operator intervention rules out real time responses or the algorithms do not generalise well because of the implicit assumptions. In general, fully automated target detection, recognition, identification and measurement processes have relied on coded targets.

Typically, coded targets can take one of two forms. The less common form is that of a unique and identifiable pattern of a small number of standard circular or square targets[1]. The proximity and relative positions of the targets in the pattern are the key to the identification and recognition. Measurement of the position of the coded pattern uses one, usually central, target in the pattern. The more common form is an individual target with an additional pattern of material or retro-reflective sheeting[3,4,5]. So-called code bars or code rings form a surrounding pattern that allows the target to be detected and uniquely identifies the target. Measurement of the location is based on the centroid, or perhaps LSM, of the central circular target.

Coded patterns of targets have been used for orientation tools, to signalise control locations or for probes used in conjunction with optical coordinate measurement systems, also known as optical CMMs. The clear disadvantage of coded patterns of targets is that they must be significantly larger than a standard, single target and therefore tend to take up more space on the object to be measured. If the individual targets in the pattern have a smaller diameter than other targets, or the targets are in very close proximity in order to reduce the overall size of the code pattern, then there is the potential for poor measurement precisions or a poor detection rate for the targets, especially in conditions of high perspective distortion. Whilst these disadvantages can be overcome by careful selection of patterns, target sizes and placement of the code patterns, the limitations ensure that coded patterns are not a universal solution.

Individual coded targets are more versatile, take up less space and, when the target diameter is identical to that of non-coded targets, realise commensurate target image measurement precision. For these reasons, coded targets are widely employed in commercially available industrial metrology systems and image-based metrology systems used for research. Commercial industrial metrology systems often have proprietary or patent protected coded target systems and claim the automated detection, recognition, identification and measurement processes as one of the major advantages of the systems. A number of examples of research systems have been reported in the literature, including some with detail of the algorithms used to carry out the recognition and identification[6].

2. CODED TARGET RECOGNITION ALGORTHIM

The coded target system described in this paper employs a Hough transform and segment matching to automatically recognise and identify the targets. The code system is based on a square surrounding the central target. The 124 codes are uniquely identified using different combinations of the sides and part-sides of the square. Part-sides must be one third of the length of the side of the square. The Hough transform is used to convert the Cartesian locations of the code bars into a Hough space of direction and magnitude. The peaks in the Hough space are then identified to locate the corners of the code pattern. Segment matching is then used to identify target code by comparing the segment profile for each side of the square against known profiles.

Figure 1 shows four different coded targets in the available sequence of 124 different types. Each coded target consists of a central dot, surrounded by a different combination of sides that may be any of the 6 segments shown in figure 2. The size of the central dot should match that of other, non-coded targets used to signal points of interest, which in turn is based on the image scale and the typical requirement for targets to span 5-10 pixels in the image. The size of the square should be sufficient to clearly separate the central dot from the sides of the square, without being over-large and occupying too much space on the object to be measured.

Template 1 Template 22 Template 50 Template 119

Figure 1: Some examples of coded targets

When constructing the sides that define the coded target, there are 3 conditions that must be satisfied:

- Each side must consist of one of the six segments (see figure 2)
- Each corner must have a line touching it.
- A coded target must be rotationally unique to all other coded targets.

In the absence of the above rules, a four-sided figure with 6 possible segments per side would result in 1296 possible codes. Inclusion of the 3 restrictions results in 124 possible codes. The choice of part-sides as one third of the side length is a compromise between the available number of codes and the reliability of the code detection. Half-sides would result in too few unique codes, whilst quarter-sides, or smaller, would reduce the success rate of code detection under conditions of poor contrast or perspective distortion. A further consideration is that quarter-sides makes manual construction of the codes more complex. Although the codes may be prepared using automated techniques such as screen printing, a strong rationale for the design is the ability to readily prepare coded targets by hand using cut strips of retro-reflective sheeting or other materials.

2.1 Hough Algorithm

The matching algorithm is based around a Hough transformation. The image segment is subjected to a threshold extraction, and then Hough transformed. The four largest peaks in the transform are found, and these correspond to the corner points of the coded target. Knowing the four corner points of the coded target, a profile along each of the four sides is generated and matched against the known segment types. The combination of the four profiles allows a unique code match to be determined. The algorithm exhibits good stability in conditions of template rotation, scaling, and projective distortion.

Figure 2: Valid line segments for the Hough algorithm

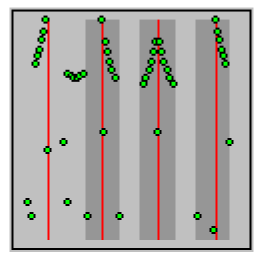

Figure 3: Hough transform of code 50

The following step-wise summary of the algorithm uses the example of code 50, shown in figure 1. This coded target image has had a threshold extracted, leaving a clean image where the target is distinct from the background. The centre of the image is adopted as the origin for the Hough transform.

In a Hough transform, points are transformed from a cartesian coordinate system into a Hough space that consists of a direction and magnitude. In figure 3, the vertical axis represents magnitude, and the horizontal axis represents direction. Since the transform origin is based at the centroid of the centre dot of the template, the points in Hough space with the largest magnitude represent the corner points of the square. Again, with reference to figure 3, the vertical (red) line that is not surrounded by grey shading is the largest magnitude in the Hough space. Once this point is found, making use of the geometry of the square, the other corner points will be increments of 90 degrees direction away from the point of largest magnitude. Due to projective distortion of the coded target images during real use, other corner points will not always be at 90 degree angle increments from the maximum point. To compensate for this a search area around the 90 degree increments is defined, and the maximum magnitude within these bounds is sought. In figure 3, the search areas are shown in grey shading, and the three remaining corner points are shown as vertical (red) lines.

2.2 Segment Matching

When the four corner points are known, the locations can be transferred back to the original image (figure 4). Segment profiles along each of the (red, green, blue and magenta) lines at the sides of the square in figure 4 are generated. To allow for small errors in detection of the corner points, pixels to the left and right of these lines are considered when generating the profiles (this is highlighted along the left (blue) line, where pixels slightly to the left of the line must be considered to generate the correct profile). The resulting profiles are shown in figure 5.

The colours of the segment profiles in figure 5 correspond to the line colours in figure 4. The extracted segment profiles are matched against the valid line segments shown in figure 2, based on a pre-set number of samples along each profile within the image. A measure of the match quality is obtained by comparison of the extracted profiles with the known profiles, and the ordered combination of segment profiles leads to a unique template identification.

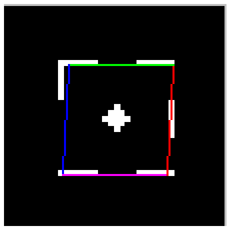

Figure 4: Code 50 with corner points

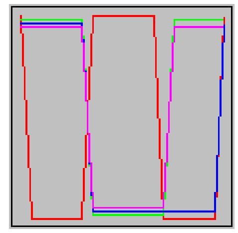

Figure 5: Segment profiles generated for code 50

2.3 Practical Considerations and Pre-detection

Within real images, candidates for coded targets must be identified and separated from non-coded targets and background clutter within the image. The initial steps in this process are as follows:

- A global threshold, based on an analysis of the histogram of image intensity values, is applied to the entire image or a manually identified region within the image.
- Areas of the image above the threshold and within pre-set limits of contiguous pixels are identified using an image scanning algorithm. These become candidate target "blobs" whilst all other areas are discarded as either noise if too small, or background objects if too large[7].
- A local image window is extracted around each blob and a local threshold computed and extracted[8]. The centroid is located and a number of geometric tests are applied to ensure that the blob is most probably an image of a circular target of appropriate size[7].

At this stage the target image centroid location can be passed to other routines that attempt to identify the target code from epipolar line searching or other mechanisms that attempt to determine the identity of the target image using only the location in the image space. The identification of a coded target image requires a further series of steps:

- A larger local image window is extracted around each target image. The size of the image window should be slightly larger than the ratio of the size of the square to the circular target. For the sake of consistency in the software and to obtain a clear separation, the recommended ratio between the sizes of the central circular target and the code square is 1:5. The corresponding ratio for the coded target image window is 1:7 to ensure the entire coded target is captured within the window.
- A local threshold is computed and extracted.
- Hough transform and segment matching are applied. The criterion for a successful identification of a coded target is based on achieving detection of all four peaks in the Hough transform and then a segment match greater than a pre-set threshold. A suitable criterion is that 80% of the samples in the segments match one of the six cases shown in figure 2.

This straightforward approach is prone to a high rate of failure for images that are very cluttered with dense target images or background detail. Intrusions into the coded target image window will cause the identification to fail because the Hough transform will not successfully locate the four corners of the square or interference in the segment matching will cause a mis-identification or matching failure. Strong perspective distortion and low contrast exacerbate these problems by increasing the probability of intrusions. A number of additional strategies can reduce the number of identification failures and mis-identifications:

- The size of the image window for the coded target is proportional to the size of the central circular target. In cases of perspective distortion this results in a rectangular window that reduces the probability of intrusions.
- The threshold can be applied locally to each target and varied through a range of increasing values, based on the image intensity histogram. The principle is that a low threshold value will remove the background noise, whilst higher threshold levels may eliminate intrusions of background clutter. Local thresholds can be set to remove the peak of the background pixels, select the mid point between the peaks of background and target pixels, and eliminate all but the peak of the target pixels (figure 6).

- Rather than accepting a full set of 124 codes, only those codes that are identified as having been included in the object space are acceptable. Once a code is identified, a repeat identification is an indicator of a mis-identification.

The final strategy that can be applied is a pre-detection step to substantially decrease the frequency of non-coded targets being subject to the identification algorithms for coded targets. This is effectively a filtering process that reduces the occurrence of mis-identifications caused by intrusions around non-coded targets. The pre-detection process is based on a straightforward 16 way test (figure 7). Commencing at the centroid of the target image, line searches are carried out at 22.5 degree intervals. The search directions are efficiently coded as steps in whole pixels. A flag is set every time each search reaches a boundary at which the intensity crosses over the local threshold. Three such changes in any line search indicates that a segment of the square has been encountered, and if more than three line searches encounter segments then the probability is very high that the target is coded. Furthermore, the identified locations of the segments of the square can be used to adjust the minimum size of the local window.

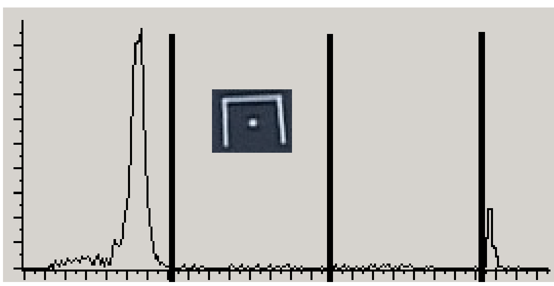

Figure 6: Intensity histogram for a coded target image (shown inset) – strong vertical lines show possible threshold locations.

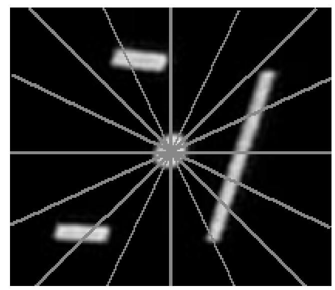

Figure 7: Pre-detection of coded targets using a 16 way search for line segments on the sides of the square.

3. EXPERIMENTAL TESTING

A first example of the use of the coded targets is taken from the requirement to periodically calibrate RS-170 video and digital still cameras systems in use for a variety of engineering measurement purposes at NASA Langley Research Center[9]. In this example 90 targets have been attached to the connecting rods of a 1m by 1m by 0.5m, rigid metal calibration fixture constructed from 25mm square section aluminium (see figure 9). The non-coded circular retro-targets have a diameter of 10mm in order to realise a span of 5-10 pixels on images obtained by RS-170 video cameras with a resolution of 752 by 480 pixels, or digital still cameras with resolutions of 1000 to 2000 pixels. The coded targets were constructed on 100mm by 100mm flat metal plates mounted at five positions on the fixture. Five coded targets provide some redundancy over the minimum of four targets required for an initial estimation of the exterior orientation of any image. The code bars were sized in accordance with the recommended 1:5 ratio to yield a coded target size of 50mm by 50mm. Figure 8 shows one example image from 20 images captured for a convergent, multi-photo calibration network taken with a Kodak DC265 digital still camera. Several additional targets seen in this image are associated with calibrated length bars to provide an accurate scale for the network. In this image the retro-target images exhibit peak intensities of approximately 250 grey levels whilst the image background is in the range of 5-15 grey levels.

All images and starting value coordinates for the 90 targets located on the calibration cube were loaded directly into the VMS application[10] for processing. This software includes the algorithms for the coded target identification, as well as algorithms for scanning images for targets based on a global threshold, computation of target image centroids with locally applied thresholds, automatic orientation of each image based on a modified Zeng-Wang computation[11], backdriving to non-coded targets based on object space locations and a self-calibrating free network adjustment with scale constraints. Of the 98 coded targets that were fully visible within the 20 images, 2 coded targets were not identified. The identification failures were carefully analysed and found to have been due to a combination of strong perspective distortion and a complex template shape. The algorithmic failure is generally caused by a segment match quality of 65% to 75%. Five repeats of this network, captured to test the stability of the camera and the calibration fixture, exhibited similar results. Subsequent calibration work with a range of cameras including Kodak DC4800, Kodak DCS460, Olympus C-1, Olympus E-20 and Canon EOS-D30 digital still cameras, and a variety of Pulnix, Hitachi and other CCD video cameras, has produced very similar success rates for the identification of the coded targets.

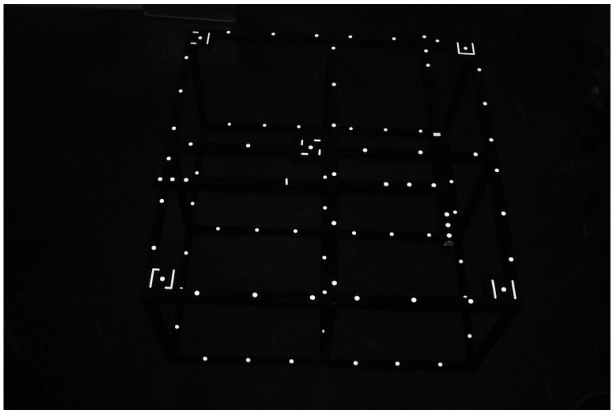

Figure 8: Example Kodak DC265 image of a calibration fixture and calibrated length bars.

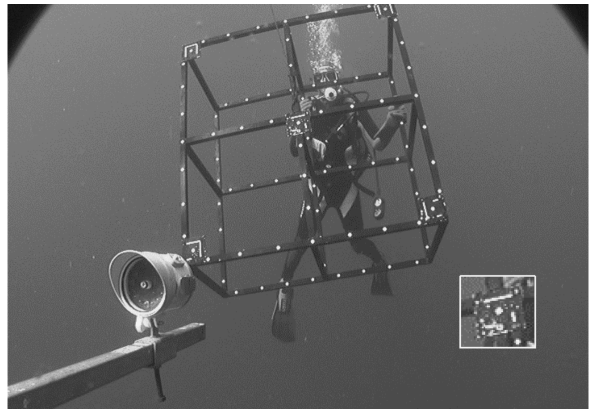

Figure 9: Example digital camcorder underwater image of a calibration fixture – centre coded target enlarged inset.

A similar process to that described above is also used to routinely calibrate pairs of digital video cameras mounted in underwater housings. This underwater stereo-video system is used for applications such as monitoring of fish and shellfish populations in marine park areas[12]. The cameras in this case are Sony TRV series camcorders recording onto miniDV tape with a resolution of 720 by 576 pixels. Calibration fixtures are of similar construction to those used in air, however passive targets are used. Retro-targets are not feasible because of the backscatter that would result from light sources near the camera lenses.

Significant challenges exist for the target code extraction algorithm in this circumstance. The optical properties of the marine environment, particularly light penetration and the influence of the housings, limits the image quality produced by the digital video cameras. Coded and non-coded targets can deteriorate quickly from the constant immersion in salt water, as well as accidental damage from constant transport. Figure 9 demonstrates a typical image from routine calibrations and demonstrates clearly that, due to a high level of background intensity, image scanning based on a global threshold will not be as successful for coded target identification. In this image the passive target images exhibit peak intensities of approximately 220 grey levels whilst the image background is in the range of 75-120 grey levels. Deterioration of the coded targets further reduces the success rate of identification (see inset in figure 9).

An average of 180 coded targets appear in the 40 images that are typically used to calibrate the underwater stereo-video systems. The success rate for the identification of coded targets in these images, on average, is quite low at 20-30%, based on many repeat calibrations during routine operations. Failures are most often due to intrusions of background or nearby non-coded targets within the match window. The algorithmic failure is most often from an inability to locate the four corners of the square. The identification success rate could be improved markedly by using larger corner plates to exclude the background from the image window, and re-positioning adjacent non-coded targets to avoid intrusions in the image window.

For these images, manual intervention is necessary to obtain at least four identified targets on the calibration fixture in order to estimate the exterior orientation, to allow the subsequent resection and back drive processes to measure all targets. The VMS software has the ability to click and drag in an image to identify coded targets. Selecting a small area around the coded target within the image has a much higher success rate as a remedial function carried out by an operator, however alterations to the calibration fixture would certainly improve the automatic coded target identification rate.

4. CONCLUSIONS

The straight sections of retro-reflective or passive material used to uniquely identify the codes are very straightforward to manufacture and assemble. Coded targets can be manufactured separately or, in some circumstances, added to pre-existing targets to enable automatic identification. The coded targets exhibit a high reliability of recognition and identification under favourable conditions. In a similar fashion to many other coded target systems, the success rate falls off under unfavourable conditions such as poor image quality, high perspective distortion and background clutter.

Whilst coded target identification failures do occur in the given examples, a higher success rate could be achieved by employing geometric constraints derived from the *a priori* knowledge of the location of each target within the object space, in order to predict the imaged location of any unidentified coded target images. Such information would be logically combined within the initial estimation of the exterior orientation of each image, which is based on a modified form of the Zeng-Wang algorithm[11]. The modification incorporates a majority voting and clustering algorithm that identifies consistent solutions for four or more targets. This algorithm could be extended to conduct image space tests of targets in predicted locations, including the dual cases possible when only three targets are correctly identified. This is an area of ongoing research.

REFERENCES

1. Fraser, C. S., 1997. Innovations in automation for vision metrology systems. *Photogrammetric Record*, 15(90): 901-911.
2. Wong, K. W., Lew, M. and Wiley, A. G., 1988. 3D metric vision for engineering construction. *International Archives of Photogrammetry and Remote Sensing*, 27(B5): 647-656.
3. Niederöst, M. and Maas, H.-G., 1997, Automatic deformation measurement with a digital still video camera. *Optical 3-D Measurement Techniques IV*, Wichmann Verlag, Karlsruhe, pp 266-271.
4. Schneider, C.-T. and Sinnreich, K., 1992. Optical 3-D measurement systems for quality control in industry. *International Archives of Photogrammetry and Remote Sensing* 29(B5): 56-59.
5. van den Heuvel, F. A., Kroon, R. J. G. A. and Le Poole, R. S., 1992. Digital close-range photogrammetry using artificial targets. *International Archives of Photogrammetry and Remote Sensing*, 29(B5): 222-229.
6. Knyaz, V. A. and Sibiryakov A. V., 1998. The development of new coded targets for automated point identification and non-contact 3D surface measurements. *International Archives of Photogrammetry and Remote Sensing* 32(5): 80-85.
7. Shortis, M. R., Clarke, T. A. and Short, T., 1994. A comparison of some techniques for the subpixel location of discrete target images. *Videometrics III*, SPIE Vol. 2350, pp 239-250.
8. Shortis, M. R., Clarke, T.A. and Robson, S., 1995. Practical testing of the precision and accuracy of target image centring algorithms. *Videometrics IV*, SPIE Vol. 2598, pp 65-76.
9. Shortis M. R., Robson, S., Pappa, R. S., Jones, T. W. and Goad, W. K., 2002. Characterisation and tracking of membrane surfaces at NASA Langley Research Center. Proceedings on CD-ROM, *ISPRS Commission V Symposium on Close-Range Imaging, Long-Range Vision*, Corfu, Greece.
10. http://geolier.eng.unimelb.edu.au/vms [accessed October 28, 2002]
11. Zeng, Z. and Wang, X., 1992. A general solution of a closed-form space resection. *Photogrammetric Engineering and Remote Sensing*, 58(3): 327-338.
12. Shortis, M. R. and Harvey, E. S., 1998. Design and calibration of an underwater stereo-video system for the monitoring of marine fauna populations. *International Archives of Photogrammetry and Remote Sensing* 32(5): 792-799.