IMAGE MATCHING TOWARDS MATURITY

Dimitris P. SKARLATOS
Department of Surveying, National Technical University, GR-15780 Athens, Greece
dskarlat@survey.ntua.gr

Working Group III/2

KEY WORDS: Image Matching, Automation, Quality Control, Algorithms, Surface Reconstruction

ABSTRACT
This paper deals with the least squares matching algorithm, focused on stereo pairs, without the use of epipolar geometry. The least squares matching algorithm although widely used has still some shortcomings, which in certain cases keep it away from commercial production. The shortcomings are the low reliability of the "successful" matches and the automation/adaptation of the algorithm, which has numerous parameters. These problems will be addressed in this paper by introducing a new algorithm for dynamic adaptation of the template size and two new algorithms to enhance accuracy and reliability. Discussion and examples will be presented along with proposals for further work, based on the combination of the three algorithms and the use of epipolar geometry.

1 INTRODUCTION

It is well known that the main algorithm for least squares matching is an ingenious concept and very well documented, which nowadays is included in every digital photogrammetric software. This research is part of a Ph.D. thesis, concerning image matching in stereo pairs, since the bulk of the commercial production is still based on pairs. The least squares matching (LSM), which is the basic model under analysis, has two phases, namely the initial approximations and the least squares (LS) solution. The former stage is basically addressed with image pyramids or interest point operators, while the LS are responsible for the acceptance of the initial values and the accuracy of the final fix. The initial approximation problem seems to have been solved for aerial photography, while it is still challenging in close range photogrammetry where abrupt scale differences are causing large shifts, which cannot be dealt with the aforementioned methods. The well-established least square matching with 6 or 8 parameters was the basic model for investigation. Since the sub-pixel accuracy of the method is unquestionable, the problems that still pose are the quality control and the reliability over the matches (correct matches that are rejected or wrong matches, which are assumed correct) and the automated adaptation of the algorithm parameters. Epipolar geometry hasn't been included in the model yet, so that the algorithm is more general and doesn't relate only to central projection, nor to a priori knowledge.
Manual correction is still needed on automatically collected digital elevation models (DEMs) not only to fill gaps but to correct the wrong matches as well, which in some cases is even more painful and time-consuming than collecting manually the points from the very beginning. Therefore elimination of wrong matches would save a lot of time, not to mention that high confidence matching could give a new thrust to the initial approximation problem.
The automation over the LS parameters is something that one considers only when one tries to write the code. Not only they are too many, but often difficult to predict. All these parameters increase the complexity of the matching algorithm when used by unskilled users, while the experts cannot choose "average" parameters for the whole area. Therefore the parameters should change dynamically in the overlapping area.

2 ADAPTIVE TEMPLATE

Adaptive template addresses the problem of automation and better adaptation over the area. The template size should be "big enough for signal and small enough for minimum geometric distortions and good localization". The word "enough" cannot be measured in any terms and it's up to the user to interpret. It should be mentioned that some commercial packages offer automation on this aspect, claiming adaptation of the template size, without giving any information about how is this being done. Nevertheless the criteria used for the adaptation are not mentioned in published articles nor in manuals (none that the author is aware of) and therefore the research started from scratch.
It has been suggested that the template size could be increased in order to grab "signal that is close but was not originally included" (Gruen, 1985). The method has been applied and tested. Starting from a small template (e.g. 5x5) the solution converges rather easily, especially when the model incorporates the two radiometric parameters (against the application of Wallis prior to LS, Baltsavias 1991), which in small templates dominate and manage to force convergence even in wrong solutions. Although there should be an upper limit of the template size there is always the possibility for the template to grow until it contains too much signal so that the LS provides some solution, which if the

template is too big will not be correct. Therefore this method was abandoned because of the large percentage of wrong matches.

The best way to get information about the signal and content of a given template is statistical analysis. The most common statistical measures that can be used for such analysis are the mean, standard deviation, median, mode, range and mean deviation (Kennedy, Neville, 1976). The analysis is being done around the pixel of interest only in one image (i.e. the left one), because we expect the search window on the other image to have the same content as well.

In order to avoid the initial approximation problem, a pair of orthophotos (from left and right photograph) was created with 0.5 pixel size from 1:15000 scale photography over the suburbs of Volos in Greece. From the two orthophotos three pairs were selected so that they contain a variety of features, basically homogeneous areas, edges and houses. Of course the orthophotos created from left and right photographs are not exactly the same allowing the algorithm to work in sub-pixel matching. Some errors in DEM produced obvious distortions in the orthophotographs, where matching should fail. Some other features such as cars in roads should also be considered possible failures.

In order to examine the statistical measures and the best template size for matching that they suggested, a large number of sample points were collected from the test areas (Fig. 1)

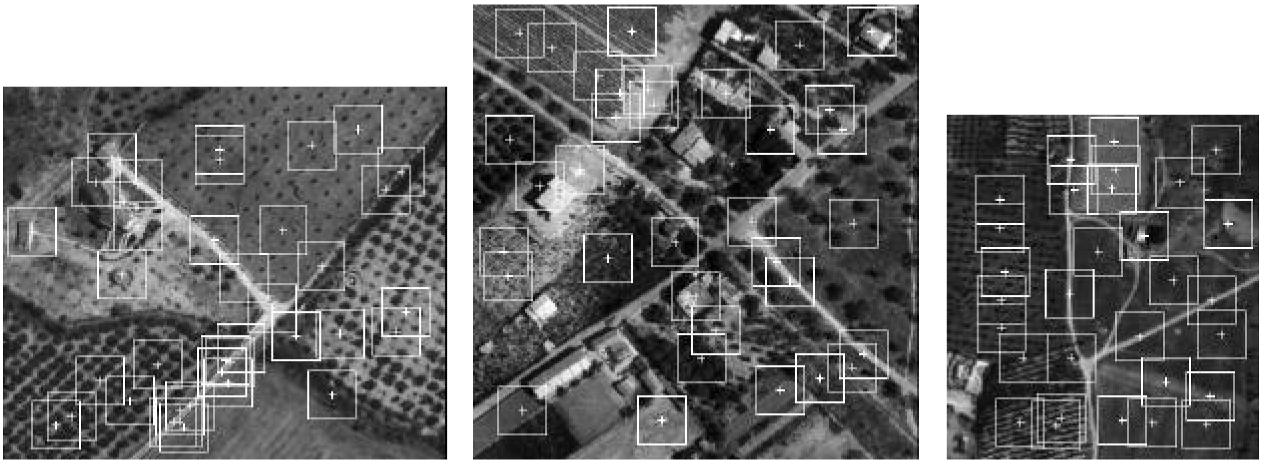

Figure 1. The three test areas with pixel size of 0.5 m. The characteristic points selected with the 41x41 template can be viewed. The original photographs of 1:15000 scale was scanned at 600 dpi.

From the three areas 99 representative points were selected, containing a variety of features such as homogeneous areas, line features (roads of variable widths and ploughed land), point features (trees and bushes in fields), and random formations. The sample is big enough to export confident results. The statistical measures from 5x5 to 41x41 template sizes were collected and diagrams were drawn in order to detect any obvious correlation between the measures and the "correct" template size. All 99 samples were visually checked and the best template was subjectively selected and marked. Correlation coefficient was also used so that a measure of the goodness of the template was available, apart from the subjective criteria. Three sample areas can be viewed in figure 2.

It is quite clear that finding criteria among the 6 lines, which could be applied in all 99 cases and expressing them with logical functions is extremely difficult. There are numerous thresholds that could be applied (i.e. when standard deviation is more that 30) and numerous combined functions (i.e. intersection between median and mode). Another puzzling factor that was whether there should be a penalty function for the size of the template so that to keep it to a minimum.

The algorithm that was finally adopted was quite simple: the intersection between range and mean. If this is not fulfilled for the 41x41 template then the standard deviation is checked against a given threshold and finally the range is checked against the augmented mean in the certain template size. If none of the above criteria is fulfilled then no attempt for matching is being done and the next point is selected.

The results of the aforementioned algorithm produced quite satisfying results, which could be seen in figure 3. It is clear that the templates are bigger in homogeneous areas, while in areas with enough signal the template is kept small. In cases where the initial point is in homogeneous areas but close to some characteristic the template is just big enough to enclose the additional signal.

The intention to find a penalty function, so that to keep the template to a minimum (for better localization) wasn't utilized because of the good results. All successful matches were correct, wrong matches were avoided without any further changes to the algorithm. This phenomenon could be explained by the fact that homogeneous areas are generally flat, and therefore the affine transformation can model the perspective geometry quite well. In sudden slope changes there is also change in illumination, hence the algorithm keeps the template small and the affine can still model adequately the geometry. The ratios of successful matches to attempts were 88.6%, 86.7% and 88.46% respectively. The ratio of correct to successful matches is 100%. There might have been correct matches that haven't passed the criteria of template size or have exceeded the iteration threshold without convergence.

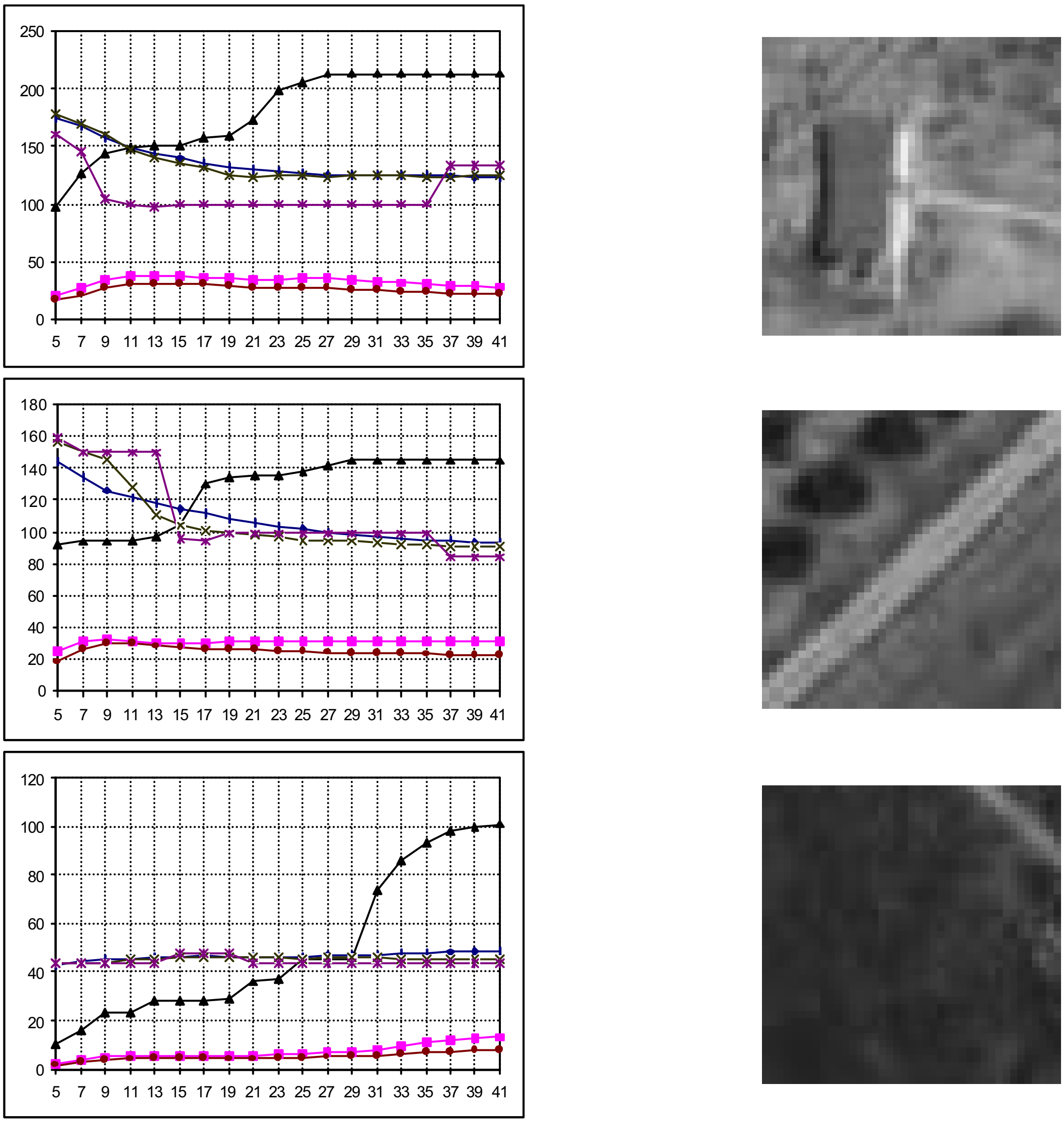

Figure 2. The image patches (41x41pixels) with the corresponding diagrams. The lines on the diagrams represent mean, standard deviation, median, mode, range and mean deviation

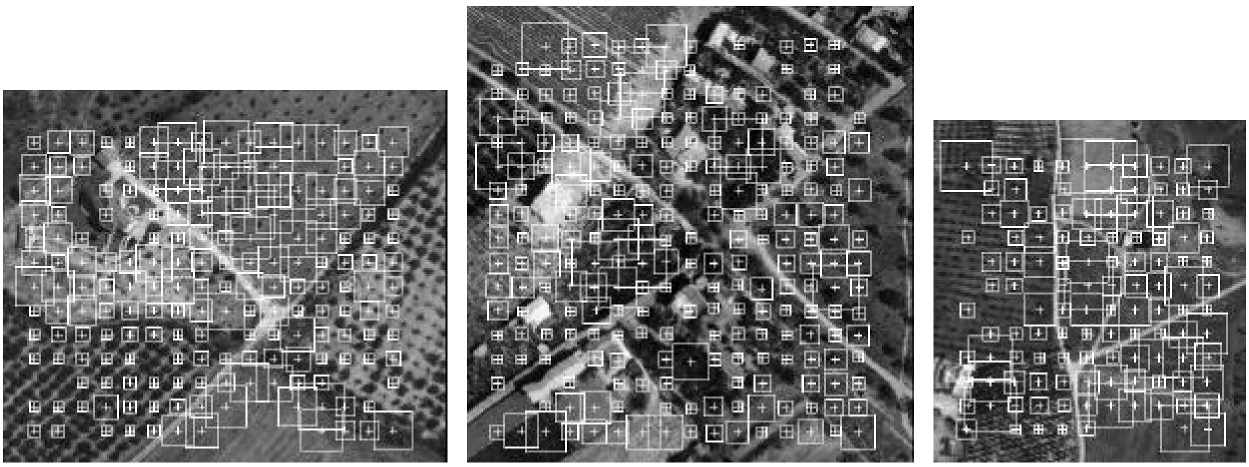

Figure 3. The automatically adapted templates in the test areas.

3 ROBUST ESTIMATORS.

Robust estimators address the problem of accurate matches, especially in cases of occlusions, shades or sudden slope changes. It was found that if more than one gross error exists in the data then the reliability of the LS solution decreases significantly (Pilgrim, 1996). In cases such as shades or occlusions a large number of pixels in the template might differ a lot and therefore the solution, if reached, would probably be wrong, In such cases robust estimators would be more accurate if the center of the template fails within the largest continuous region (Calitz et al, 1996). This enhances the matches near breaklines and can be used for their detection and localization, improving the surface reconstruction.
In order to avoid changing the Least Squares main algorithm, it is possible to simulate robust estimators by recalculating the weights in each iteration, based on the residuals. The proposed re-weighting scheme by Calitz and Rûther (1996) is based on the function

$$\mathbf{W_i} = \frac{1}{\left|v_i\right|^{2-p} + e} \tag{1}$$

where p=1,
e a small constant to avoid division by zero and
v_i the residual of the observation from the current iteration.
This approach is favorable since other weighting schemes can be applied and tested easily. There will be no further analysis over the particular weighting scheme (more in Calitz & Rûther, 1996).
The re-weighting of the observation in every iteration is quiet an old concept and there are many proposals (Baltsavias 1991, Cross 1990, Pilgrim 1996). In most of them the new weights are a function of the residuals themselves or their standard deviation, which causes huge computational cost, even for current computers. Depending heavily on the template size, in a Pentium 200 MMX with 11x11 template the matching speed was double of an operator, without any code or compiling optimization over speed. With 21x21 template the speed is by average 3 minutes per point. In addition all the weighting schemes did not score well in success rates, which were below 70% in all cases.
The results were surprisingly poor, but these and other weighting schemes would be tested in conjunction with the 6 parameter LS model in order to avoid overparameterisation.

4 DOUBLE IMAGE CHECKING.

The basic concept of the method is the fact that the matching function should be one to one. Therefore a successful match from left to right image should be repeated from right to left, ending on the initial pixel from left image. If the match is correct and the solution stable the sub-pixel on the left image should lie roughly within

$$\pm 3\sqrt{s^2_{xright} + s^2_{xleft}}, \pm 3\sqrt{s^2_{yright} + s^2_{yleft}}, \text{ (for 99\% probability level)}, \tag{2}$$

where ó_{xright}, ó_{yright} the standard deviations of the shifts from the first (left to right) match and
ó_{xleft}, ó_{yleft} the standard deviations from the second (right to left) match,
of the initial left pixel. The results were very encouraging since many spurious matches were removed. In figure 4 an example of the algorithm, with fixed template size of 7x7 pixels, can be viewed. In comparison with the normal case of single image matching, there has been a reduction of 3% in matches. These matches did not pass the double image check and do not appear. They are probably spurious matches.

5 CONCLUSSIONS – FURTHER RESEARCH

It should be noted that all proposed modifications were done upon the basic 8 parameter LS model. Until now no combinations among the methods has been done, although that by the time this paper is printed results of further modifications and combinations would be available.
Another way to check the similarity of the two templates before the match is the comparison of the statistical measures. In the case they are completely different the match shouldn't be attempted. On the other hand the template size which shows the best similarity on statistical measures between the left and right image should be used as the correct template.

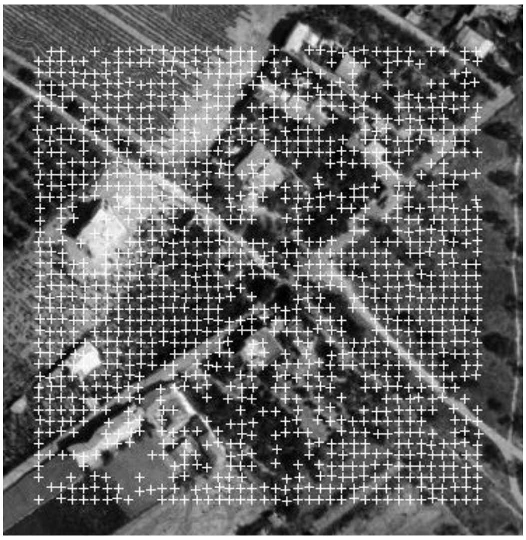

Figure 4. The application of the double image checking algorithm, with 7x7 fixed template size.

Since this is a first report, other types of photography are to be used, with different scales and resolutions. Close range images will be used when the initial approximation problem would have been addressed. Combinations over the proposed methods in order to find the best way to match images without any a priori knowledge (relative orientation), with the highest reliability factor are the next step. Theoretically the re weighting in conjunction with adaptive template to be a good solution and therefore research will be focused towards that direction. Computational effort is no object since computer power raises geometrically.

Check could also facilitate the affine parameters (Baltsavias, 1991) in terms of rotation, scale and x,y distortions. In such case application of thresholds on these parameters should be avoided and an adaptive algorithm should be used.

Other LS parameters which should be adapted automatically are:

- the window search area,
- the use of less than 8 parameters in each point,
- dynamically reducing the parameters by fixing those which do not change a lot (research is already being done on this subject, but results haven't been analyzed yet),
- the maximum number of iterations,
- the minimum corrections in affine parameters to end the iterations

It is expected that the use of the epipolar geometry would strengthen the matching.

In author's opinion the LS method is still alive; all it needs is a little refinement (Skarlatos, 1999).

ACKNOWLEDGEMENTS

The author would like to acknowledge the financial support from State Scholarship Fundation.

REFERENCES

Kennedy, J. B., Neville, A. M., 1964. Basic Statistical Methods for Engineers & Scientists. Dun-Donnelley, New York, p. 490. Chapters 3 and 4.

Calitz, M., F., Rûther, H., 1996. Least absolute deviation (LAD) image matching. ISPRS Journal of Photogrammetry & Remote Sensing 51, pp. 223-229

Cross, P. A.,1990. Advanced Least Squares Applied to Position-Fixing. Polytechnic of East London, London, p. 205.

Baltsavias, E., P., 1991. Multiphoto Geometrically Constrained Matching. Ph.D. thesis, Institute for Geodesy and Photogrammetry, Zurich.

Pilgrim, L., 1996. Robust estimation applied to surface matching. ISPRS Journal of Photogrammetry & Remote Sensing 51, pp. 243-257.

Shao, J., Fraser, C., 1998. Multi-Image matching for automatic object reconstruction, University of Melbourne, Australia. jsha@sli.unimelb.edu.au

Zhang, B., Miller, S. 1997. Adaptive automatic terrain extraction. Helave Associates Incorporated, California, USA. Helava@gdesystems.com

Gruen, A., 1985. Adaptive least squares correlation: A powerful image matching technique. South African Journal of Photogrammetry, Remote Sensing and Cartography. (14)3, pp.175-187.

Atkinson, K., B., 1996. Close Range Photogrammetry and Machine Vision. Whittles, Caithness, UK, p. 371.

Skarlatos, D., 1999. Orthophotograph production in urban areas. Photogrammetric Record, 16(94), pp643-650.