A NEW APPROACH TO AUTOMATIC JUNCTION OF OVERLAPPING AERIAL IMAGERY DATA

Yuri B. Blokhinov [a], Dmitry A. Gribov [b]

State Research Institute of Aviation Systems (FGUP GosNIIAS), Moscow, Russia
[a] blokhinov@gosniias.ru, [b] gda@gosniias.ru

Working Group III/2

KEY WORDS: Automation, matching, recognition, detection, feature, imagery, mosaic, aerial

ABSTRACT:

The original approach to image matching is proposed. The method itself can be classified as relational matching, bases on point features. For robust extraction and filtration of features the special procedure, based on dynamic resampling technique, was elaborated. Further the rotation invariant relations among the features are used to confirm or reject initial hypothesis. All calculation procedures are time effective and invariant to images rotation. Finally, the approach given is applied to two different tasks: automatic mosaic creation from video camera sequence frames and automatic relative orientation of photographic camera images.

1. INTRODUCTION

Image matching is the task, aroused in many different applications. Both input data and practical aims can differ, but the underlined principles are the same. So the task under consideration should be of interest for wide range of specialists.

Here we try to develop the sort of feature based relational matching as the most suitable for comparison of large images. Considerable efforts was done by the investigators in this branch, many interesting results was obtained (Heipke, C., 1996, Woozug, C., ., 1996), each optimal to use in it's specific domain. The method, described below, was elaborated for real technical applications and two properties was obligatory for it: to work in the near real time (minutes, not hours) and to give reliable results.

In short, after some kinds of special pre-processing procedures, image can be represented as a set of spatially distributed features. Each feature is unique and, in general, can be described by some digital parameters and hence can be distinguished among another features. Main features types are (Henricsson, O., 1996) points, lines and regions. For each type the specific methods are elaborated to extract it from image. When all substantial features in the image are extracted, their relative coordinates with respect to each other can be fixed. Now we can say, that image is described by the finite set of numbers, features' parameters and their relative coordinates, and to compare different images in the formal mathematical way. The main problem is that most of methods used at present for image recognition require considerable time to implement. This is due to the fact that complex feature extraction by known algorithms is very time-consuming procedure. This paper introduces one approach to relational image matching, suitable for performance in near real time.

2. VIDEOCAMERA SEQUENCE FRAMES JUNCTION

2.1 Task and data

The input data are video shooting obtained by swinging camera from airplane. Raw material can be cut into sequence frames, which are considered as a set of digital images, the overlapping is 40-80%. The total sequence sometimes includes thousands of frames and cover large area of the surface. Mosaic of these frames, built up on-the-fly, is of considerable interest in some practical applications. In the given case "to build on-the-fly" means to build automatically due to very large number of input images. Substantially that all algorithms should rotational invariant and non sensitive to variations in brightness level among different frames.

Figure 1. Sample of video frames

2.2 Features extraction

Proper choice of features is the key part of relational matching (Henricsson, O., 1996). A reasonable compromise should be found between the informativity and complexity of the features at hand. Lines and regions are informative and stable though, they requires much time for extraction and handling. So for "build on-the-fly" algorithm only point features were taken into consideration. As the index for interest points extraction the variance of image brightness $V(x, y)$ was taken. Variance for window of size N×N, centred at x_0, y_0 is defined as follows:

$$V_N(x_0, y_0, N) = \frac{1}{(2N+1)^2} \sum_{x=-N}^{N} \sum_{y=-N}^{N} f(x+x_0, y+y_0)^2 - \left(\frac{1}{(2N+1)^2} \sum_{x=-N}^{N} \sum_{y=-N}^{N} f(x+x_0, y+y_0)\right)^2 \quad (1)$$

where $f(x, y)$ - image brightness in point with coordinates x, y.

The examples of variances for two consecutive frames are shown in the Figure 2. The operator of maximum variance was

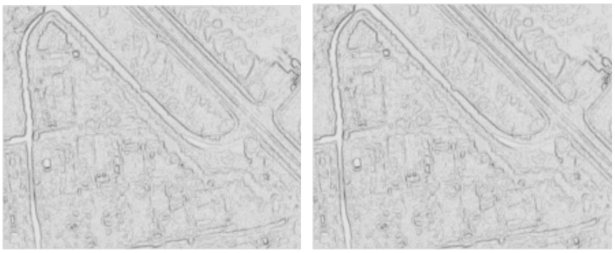

Figure 2. Variance fields (window size 3×3)

taken as the operator of interest for image due to simple structure and computational stability. At first step a list of candidate points is extracted by maximum operator with window size 5×5. The lists of candidate points for left and right images are shown in the Figure 3 and, in general, depends on the

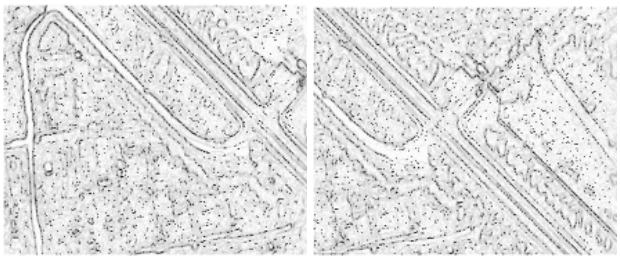

Figure 3. Candidate points lists for features

window size. At the next step the informativity size of each feature is defined and then used to select stable features. The informativity size is defined in the following way. Consider variance V in the given point as function of window size N, the typical form of the function is presented in the Figure 4.

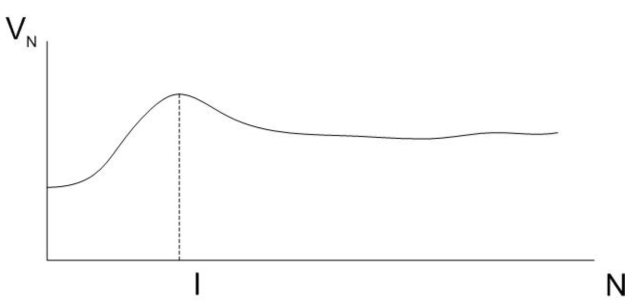

Figure 4. Informativity size

Abscissa of the maximum is considered to be the informativity size I of the given feature. This valuable parameter is used in filtering of the lists of candidate features for left and right images. Those features with $I < I_m$, where I_m is threshold, are filtered out. Small features should be filtered out because they could arise due noise maxima. The lists of candidate points for left and right images after filtering are shown in the Figure 5.

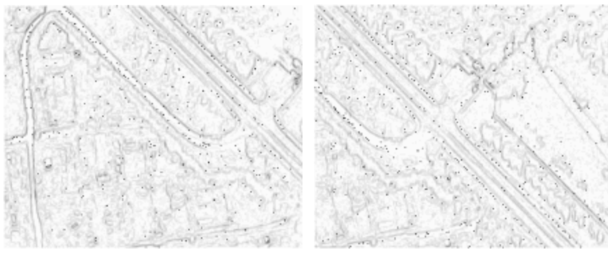

Figure 5. Lists of features after filtering, $I_m = 11$

2.3 Features matching

Now we should set up a correspondence between features at the left and right images. For that the set of K parameters describing features should be introduced. Two features, close in K-dimensional space of parameters, are considered as the conjugate pair. The obligatory condition the parameters should satisfy is the invariance to shift, rotation and scale. Suitable theoretical basis for tasks of this type was laid by Hu, M.K (Hu, M.K., 1962), who developed the algebraic theory of invariant moments for image recognition. He proposed to use seven invariants for this purpose:

$$I_0 = \mu_{20} + \mu_{02}$$

$$I_1 = (\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2$$

$$I_2 = (\mu_{30} - 3\mu_{12})^2 + (3\mu_{21} - \mu_{03})^2$$

$$I_3 = (\mu_{30} + \mu_{12})^2 + (\mu_{21} + \mu_{03})^2 \quad (2)$$

$$I_4 = (\mu_{30} - 3\mu_{12})(\mu_{30} + \mu_{12})[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] + (3\mu_{21} - \mu_{03})(\mu_{21} + \mu_{03})[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2]$$

$$I_5 = (\mu_{20} - \mu_{02})[(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] + 4\mu_{11}(\mu_{30} + \mu_{12})(\mu_{21} + \mu_{03})$$

$$I_6 = (3\mu_{21} - \mu_{03})(\mu_{30} + \mu_{12})[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] + (3\mu_{12} - \mu_{30})(\mu_{21} + \mu_{03})[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2]$$

where $\mu_{pq} = \sum_{x \in \Omega} \sum_{y \in \Omega} (x - \bar{x})^p \cdot (y - \bar{y})^q \cdot f(x, y)$ = central moment of the order (p+q) for window centered in x,y

p,q = 0,1,2,…

$f(x, y)$ = normalized image brightness

Ω - image area in x,y coordinates

These invariants are taken to form the K-dimensional parameter space to compare the point features. Account must be taken of the fact that invariant propertied were established for continuous case. In discrete scheme some errors of discretization can arise, especially in rotation of images more then 45 degrees.

Correspondence between set of features at the left and right images is now established in the 7-dimensional parameter space. Features are compared by distance:

$$S_{ij} = \sum_{k=0}^{6} \frac{| I_k^L(x_i, y_i) - I_k^R(x_j, y_j) |}{\max(I_k^L(x_i, y_i), I_k^R(x_j, y_j))} \qquad (3)$$

where $I_k^L(x_i, y_i)$ = value of k-th invariant in x_i, y_i of the left image

$I_k^R(x_j, y_j)$ = value of k-th invariant in x_j, y_j of the right image

Let $i=1,N_L$ and $j=1,N_R$, where N_L, N_R are the numbers of candidates on the left and right images respectively. Point j is considered as conjugate for i, if

$$S_{ij} = \min(S_{i,k}),\ k \in \{1, N_R\} \qquad (4)$$

At given stage all conjugate points are tied in N pairs using criteria (4), $N=\min(N_L, N_R)$.

2.4 Features examination

In order to verify that conjugate pairs of point features was tied properly, the additional information about relative coordinates of points positions was used. In short, space distributions of features at left and right images should be similar, the distribution itself can be described as a set of distances. Consider the set of points A_1, A_2,…, A_i,…, A_N in the plane image, Figure 6.

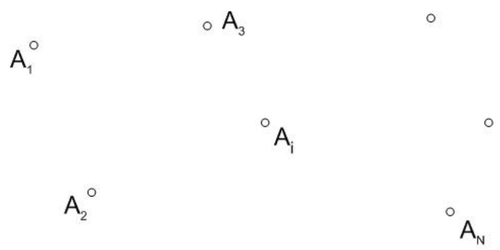

Figure 6. Point features distribution

Distances between points can be arranged in the form of N×N matrix $\|A_{ij}\|$ as follows:

$$\begin{array}{c|cccccc} & A_1 & A_2 & \dots & A_i & \dots & A_N \\ \hline A_1 & 0 & r_{12} & \dots & r_{1i} & \dots & r_{1N} \\ A_2 & & 0 & \dots & r_{2i} & \dots & r_{2N} \\ \dots & & & \dots & \dots & \dots & \dots \\ A_i & & & & 0 & \dots & r_{iN} \\ \dots & & & & & \dots & \dots \\ A_N & & & & & & 0 \end{array} \qquad (5)$$

where $r_{ik} = \sqrt{(x_i - x_k)^2 - (y_i - y_k)^2}$ = Euclidian distance between points A_i and A_k

x_i, y_i = image coordinates of point A_i

x_k, y_k = image coordinates of point A_k

To verify N pairs of conjugate points, matrix $\|A_{ij}^L\|$ for the left image and $\|A_{ij}^R\|$ for the right image should be compared. For qualitative estimation of erroneous tying variable δ_{ij} is introduced

$$\delta_{ij} = r_{ij}^R - r_{ij}^L \qquad (6)$$

Analysis of the histogram of variable δ_{ij} enables to estimate the threshold Δ to reject features according to the criteria stated below. Note that point with number (i) has N-1 connections with others, appropriate distances in matrix $\|A_{ij}\|$ are: r_{1i}, r_{2i},…, r_{ii}, $r_{i,i+1}$,…, $r_{i,N}$. Accordingly, the set of differences associated with conjugate pair with number (i) is

$$\delta_i = \{\delta_{1i}, \delta_{2i}, \dots, \delta_{ii}, \delta_{i,i+1}, \dots, \delta_{i,N}\} \qquad (7)$$

with $\|\delta_i\| = \min\{\delta_{1i}, \delta_{2i}, \dots, \delta_{ii}, \delta_{i,i+1}, \dots, \delta_{i,N}\}$ = norm of vector δ_i

Pair of conjugate points is accepted if $\|\delta_i\| < \Delta$, otherwise it is rejected. Verification procedure is performed for every i from 1 to N. Essential, that verification criterion based on analysis of matrix (5) is invariant to rotation of images.

To make the algorithm more effective, the pyramids of images were used. Initial approximation for sets of points is found at the highest pyramid level and then defined more exactly at next levels using cross-correlation. The example of performance of the algorithms above for video frames is presented in Figure 7

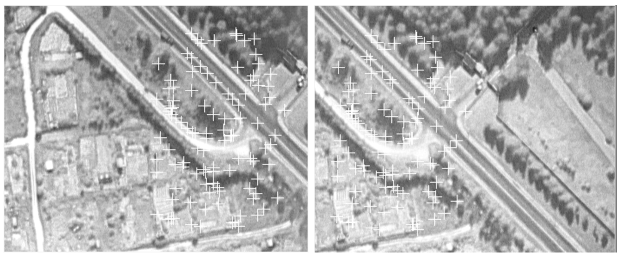

Figure 7. Accepted conjugate points, video camera frames

2.5 Invariance to rotation

At current stage it's worth trying to optimize the dimension and composition of parameter space, keeping in mind that features should be invariant to images rotation. Really the use of 7 invariants is more reliable, but requires considerable computational time. The results presented in Table 8 enable to conclude that optimal number of invariants, taken into consideration is 4. The worst percentage of success is taken place near $\alpha=45^\circ$, as it could be expected because the discrete resampling error is maximum at this angle.

Number of invariants K	2	3	4	5	6	7
Total pairs number, α=0°	333	293	277	271	248	238
Valid pairs number, α=0°	34	75	105	106	113	118
Total pairs number, α=30°	343	321	288	269	250	232
Valid pairs number, α=30°	19	28	66	69	92	99
Total pairs number, α=45°	339	320	276	270	245	246
Valid pairs number, α=45°	11	25	57	66	79	87

Table 8. Number of properly tied features depending on the number of invariants K and rotation angle α

From the above, verification distance criterion based on analysis of matrix (5) is also invariant to rotation of images. Hence the total image matching procedure under consideration is rotation invariant.

2.6 Mosaic creation

To perform the total montage of frames, the following successive tasks should be solved:

1. search of tie points for pairs of successive images for each row,
2. montage of separate raws using tie points,
3. search of tie points between successive rows of total frame montage,
4. montage of total mosaic using tie points.

In steps 1, 2 the coordinate table of successive frames is prepared and row montage is performed (Figure 9).

Figure 9. Example of row montage

In steps 3, 4 the coordinate table of successive rows is prepared and montage of total mosaic is performed (Figure 10). Total mosaic consists of 1584 frames, each of size 400 kb. Elapsed computing time for processor Pentium-IV, 2.0MHz, amounts to 40 min.

3. AUTOMATIC RELATIVE ORIENTATION OF AERIAL IMAGES

The relational matching approach developed above for camera frames junction, successfully works in automation of relative orientation for various types of imagery. All steps of processing, stated above for video frames, can be repeated here and supplemented with the items specific for epipolar geometry:

- create and employ image pyramids for both images in order to take the advantage of concept of hierarchy and make the algorithms more effective,
- extract sufficient number of features in both images on the highest level,
- match features to find candidates for conjugate pairs,
- verify that conjugate pairs of point features are tied properly, using relations among them,
- define features more exactly at next pyramid levels using cross-correlation,
- estimate orientation parameters,

Figure 10. Total mosaic, 1584 frames, 81 rows

- proceed with features verification using epipolar geometry restrictions,
- estimate orientation parameters using the coarse-to-fine approach to increase the accuracy of result.

The approach under consideration was implemented and tested in digital photogrammetric system Z_Space (Blokhinov, Y., Sibiryakov, A., Skryabin, S., 2000). Two examples of the tests are shown in Figure 11 and 12. Accuracies for standard deviation of the image coordinates are 0.36 and 0.27 pixels respectively.

Figure 11. Accepted conjugate points, automatic relative orientation of aeroborn images

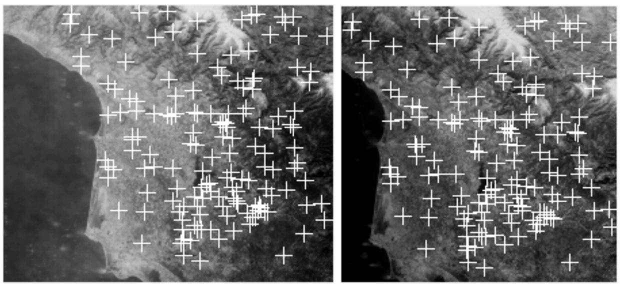

Figure 12. Accepted conjugate points, automatic relative orientation of spaceborn images TK-350

Now it's up to user to decide, whether the approach concerned meets the requirements of practice.

4. CONCLUSION

In the presented paper the universal approach to image matching for various types of imagery is proposed. At first, point features are extracted and filtered, using the original concept of "informativity size". The conjugate pairs of features are tied according to their vicinity in parameter space and filtered using the relational distance graph, both measures being invariant to image rotation. In what follows the conjugate points found can be used in different ways according to final goal of specific work.

In this study the approach was applied to the tasks of automatic mosaic creation from video camera sequence frames and automatic relative orientation of photographic camera images. In all cases the approach concerned shows high reliability and computational efficiency.

References

Blokhinov, Y., Sibiryakov, A., Skryabin, S., 2000. Z_Space - digital photogrammetric system for Russian TK-350 images. *The International Archives of the Photogrammetry and Remote Sensing*, Amsterdam, Vol. XXXIII , Part B2, Com. II, 81-90.

Heipke, C., 1996. Automation of interior, relative and absolute orientation. *The International Archives of the Photogrammetry and Remote Sensing*, Vienna, Vol. XXXI, Part B3, pp. 297-311.

Henricsson, O., 1996. Analysis of Image Structure using Color Attributes and Similarity Relations. Dr. sc. techn., Institute of Geodesy and Photogrammetry Swiss Federal Institute of Technology, Zurich.

Hu, M.K., 1962. Visual pattern recognition by moment invariants. IRE Trans. Information Theory, Vol IT-8, pp. 179 – 187.

Tang, L., Heipke, C., 1993. An approach for automatic relative orientation. Optical 3D Measurement Techniques II: Applications and Inspection Quality Control, and Robotics, pp 347-354.

Woozug, C., 1996. Relational Matching for Automatic Orientation. *The International Archives of the Photogrammetry and Remote Sensing*, Vienna, Vol. XXXI, Part B3, pp. 111-119.