

A New Method for Automatically Labeling Aircrafts in Airport Video

Xiao Luo, Yong Tang*, Honggang Wu & Donglin He.
The Second Research Institute of CAAC, Chengdu, Sichuan, China

ABSTRACT: For the problem that the airport video monitoring could only provide the image information while the label information including the flight number is not provided, a new method for automatically labeling aircrafts in airport video through the fusion of video and ADS-B data has been proposed. First, the image coordinates of aircrafts will be obtained through the image tracking of video. Then, the homography matrix between two projection planes will be calculated with the four and above point and line correspondences selected from the airport map and video image, respectively to the map image coordinates into the map coordinates. Finally, the aircrafts in video can be automatically labeled through the fusion of image tracking data and ADS-B monitoring data. Because an image coordinate measurement error exists at the time of selecting points from the image, the resulting coordinate conversion error is derived and the impact of point correspondence geometric layout on mesh coordinate mapping error is analyzed. Experiments have been conducted based on the actual data of Chengdu Shuangliu International Airport. The results show that the method can automatically label aircrafts in video in an effective way.

Keywords: automatically labeling; video monitoring; ADS-B; data fusion; Homography Matrix; error analysis

1 INTRODUCTION

The movement of aircrafts is monitored through installing monitoring cameras in important areas such as runway, taxiway, gate position, and so on to help controllers to control the airport ground traffic in an effective way. Some large airports have a number of runways and the ground layout is complicated. There are often some dead angles invisible by visual inspection in the control tower. Then, the camera becomes an important means for controllers to intuitively grasp the ground movements. As a non-cooperative sensor, the monitoring video only provides the targets' image information and cannot provide the label information of targets. Controllers are also required to determine the label information of targets in video through viewing the flight plan, ATC monitoring screen, and so on. As a cooperative sensor, ADS-B (Automatic Dependent Surveillance-Broadcast) can provide target location, speed and other positioning information as well as the label information including the flight number.

In this paper, the label information of ADS-B is added into the video monitoring through data fusion to achieve automatic labeling of the video tracking targets.

2 SYSTEM SCHEME

The overall system scheme is as shown in Figure 1. Firstly, a frame of image is extracted from the video and then feature points are extracted from the image to find out the obvious point or line marks on the airport

ground. Meanwhile, the accurate map coordinates of these point or line correspondences will be found out on the airport map. The homography matrix between the video image and airport map can be calculated with over four pairs of correspondences. Background differencing [1,2] and KLT algorithm are applied to the airport video to achieve detection and tracking of image targets, and the center of image target tracking is taken as the position coordinates of the aircraft image. The homography matrix is utilized to transform the map coordinates of aircraft in ADS-B monitoring data into the image coordinates in video. Finally, the target position data in ADS-B and video tracking data are subject to the data fusion. The K-nearest neighbor algorithm and others are applied to achieve the data association [8], thus correlating the label information including the flight number in ADS-B to the video to achieve the automatic labeling.

3 CAMERA CORRECTION

Camera correction is for achieving inter-conversion of the image coordinates of video tracking targets and the actual airport map coordinates and is a critical step for achieving automatic labeling in video. If the airport map is a projective plane and the video image is another projective plane, the projection transformation relationship between both planes can be described with a 3×3 homography matrix H [9-11], as shown in Figure 2.

A pair of points $x \leftrightarrow x'$ on the airport map and video image has the following coordinate transformation relationship:

*Corresponding author: tangyong1979@126.com

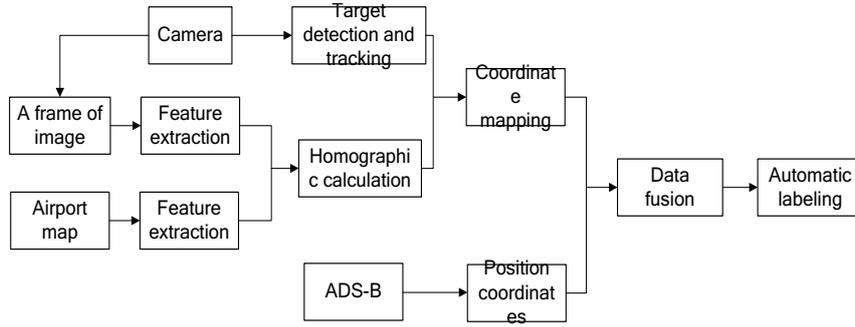


Figure 1. Overall System Scheme

$$X' = HX \quad (1)$$

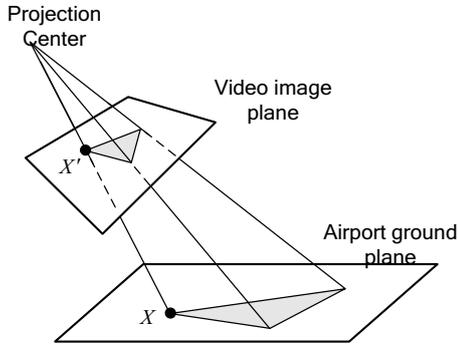


Figure 2. Schematic Diagram of Homograph

To put the map coordinates (x, y) and image coordinates (u, v) into the three-dimensional homogeneous column vectors $X = (x, y, 1)^T$ and $X' = (u, v, 1)^T$, the equation (1) can be written as follows:

$$\lambda \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = H \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (2)$$

What is significant to homography matrix H is the ratio of matrix elements. Among the nine elements of H , there are 8 independent ratios. Therefore, homography matrix H has eight degrees of freedom. In the equation, homography matrix H multiplied by a nonzero scale factor will not change the projection transformation relationship, that is, homography matrix can only be determined with the difference of only one nonzero scale factor λ .

Two linear equations in relation to the elements of H are obtained after removing the λ in equation (2) [12]:

$$-h_1x - h_2y - h_3 + (h_7x + h_8y + h_9)u = 0 \quad (3)$$

$$-h_4x - h_5y - h_6 + (h_7x + h_8y + h_9)v = 0 \quad (4)$$

The way to write equations (3) and (4) into a matrix form is as follows:

$$A_i h = 0 \quad (5)$$

Wherein: $A_i = \begin{pmatrix} -x & -y & -1 & 0 & 0 & 0 & ux & uy & u \\ 0 & 0 & 0 & -x & -y & -1 & vx & vy & v \end{pmatrix}$;

$$h = (h_1 \ h_2 \ h_3 \ h_4 \ h_5 \ h_6 \ h_7 \ h_8 \ h_9)^T.$$

It means that the correspondence of each pair of points can provide two independent linear equations. Then, four pairs of points should provide eight independent linear equations. Four 2×9 matrixes A_i can be superposed to constitute a 8×9 matrix A . The sole limit is that any three points out of four point correspondences are non-collinear. As the rank of A is eight and the equation set $Ah=0$, there is a one-dimensional null space, thus obtaining the determined solution h with the difference of only one scale factor λ . Once the homography matrix H is obtained, the inter-conversion between the image coordinates and map coordinates of any point can be achieved through equation (2).

When there are over four groups of point correspondence, $Ah=0$ is an over-determined equation set. If the measured values of all point correspondences are accurate, the rank of A is still eight and A has a one-dimensional null space, for which exact solution of h can be obtained. However, in fact, there is error in the coordinate measurements of point correspondence due to the impact of noise, resulting in no accurate solution of the over-determined equation $Ah=0$. Then, we can only define a cost function to obtain an approximate solution to let the cost function to have the minimum value. Generally, algebraic distance $\|Ah\|$ is taken as the cost function. In order to ensure that h is not the vector 0, a norm condition $\|h\|=1$ is added. The algorithm obtained is called direct linear transformation (DLT) algorithm.

Sometimes, it is difficult to select sufficient point correspondences on the airport map and video image. In addition to selection of point correspondence to calculate homography matrix, line correspondence can also be used to calculate homography matrix; the experiments in the literature by Wang[13] indicate that the homography matrix estimation method based on line correspondence has higher precision and better robustness than the point correspondence estimation. According to the duality principle of two-dimensional projective geometry, the theorem applicable to points is also applicable to lines, and points and lines are interchangeable. Therefore, in addition to point correspondence, the line correspondence can also be used to calculate homography matrix.

The linear equation on plane is $ax+by+c=0$, wherein a, b and c are parameters of straight line. Therefore, it is the same with the homogeneous vector representation of points that straight lines can also be represented as homogeneous vectors $(a,b,c)^T$. Assuming a pair of line correspondence l, l' on airport map plane and video image, then the point X on the straight line l and the point X' on the straight line l' can be respectively expressed as follows:

$$l^T X = 0 \tag{6}$$

$$l'^T X' = 0 \tag{7}$$

It can be seen from equation (1) that $X'=HX$. The following equation can be obtained after putting $X'=HX$ in equation (7):

$$l'^T HX = 0 \tag{8}$$

Through comparing equations (7) and (8), the following equation can be obtained:

$$l = H^T l' \tag{9}$$

Assuming $l=(x,y,1)^T$ and $l'=(u,v,1)^T$, then:

$$\lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = H^T \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{bmatrix}^T \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \tag{10}$$

The matrix as shown below can also be obtained after removing the λ in the equation (10):

$$A_j = \begin{pmatrix} -u & 0 & ux & -v & 0 & vx & -1 & 0 & x \\ 0 & -u & uy & 0 & -v & -vy & 0 & -1 & y \end{pmatrix},$$

Assuming that there are n point correspondences and m line correspondences, a $(2n+2m) \times 9$ matrix A can be obtained. The optimal solution h can be obtained from the over-determined equation set $Ah=0$.

4 COORDINATE CONVERSION ERROR

The key for realizing aircraft automatic labeling in the video is to make the image coordinates of target and ADS-B survey coordinates realize mutual conversion possibly. The accuracy of coordinate conversion will directly influence the effects of automatic labeling. In this section, we will discuss the conversion error between the image coordinates and map coordinates.

The coordinates corresponding to point have measurement error, which results in that calculated homography matrix generally will not map object point x_i on the airport map as the image point X_i in the image accurately, that is to say, error $X_i(x_i, \bar{x}_i)$ exists between the calculated image point X_i and real image point \bar{x}_i . Assuming that corresponding measurement on the airport map is accurate, this can be ensured by accuracy of airport map. We only discuss the coordinate conversion error due to corresponding coordinate measurement error of the points selected in the image.

For unreliability of homograph, it can be measured by covariance matrix Σ_h of homography matrix H . Map $X'=(x',y',w')^T=HX$, and write down

$$\tilde{X}' = (\tilde{x}', \tilde{y}')^T = (x' / w', y' / w')^T,$$

\tilde{X}' is derived as below:

$$\partial \tilde{X}' / \partial X = \frac{1}{w'} \begin{bmatrix} h^{1T} & -\tilde{x}' h^{3T} \\ h^{2T} & -\tilde{y}' h^{3T} \end{bmatrix},$$

H is derived as below:

$$\partial \tilde{X}' / \partial h = \frac{1}{w'} \begin{bmatrix} X'^T & 0 & -\tilde{x}' X'^T \\ 0 & X'^T & -\tilde{y}' X'^T \end{bmatrix}, \quad \text{where } h \text{ is}$$

nine-dimensional column vector composed by elements of H ,

Write down the j th row of H as h^{jT} , and it can

be obtained that $h = \begin{bmatrix} h^1 \\ h^2 \\ h^3 \end{bmatrix}$.

Coordinate x and coordinate y of the n th matching point $X'_i (i=1, \dots, n)$ in the image make up $2n$ -dimensional measurement vector X .

The process for calculating an estimative converted covariance matrix is as following [9, 10]:

(1) Estimate converted \hat{h} according to data of marching point and elements of \hat{h} constitutes nine-dimensional vector \hat{h} ;

(2) Calculate the value of Jacobian matrix $J = \partial X' / \partial h$ at \hat{h} , and the calculation formula is given in (b);

Write Jacobian matrix in the form of partitioning, namely $J = (J_1^T, J_2^T, \dots, J_i^T, \dots, J_n^T)$, and the following equation can be obtained:

$J_i = \partial \tilde{X}'_i / \partial h = \frac{1}{w'_i} \begin{bmatrix} X_i^T & 0 & -\tilde{x}'_i X_i^T \\ 0 & X_i^T & -\tilde{y}'_i X_i^T \end{bmatrix}$, where X_i^T indicates vector $(x_i, y_i, 1)$

(3) Estimate covariance matrix $\Sigma_h = (J^T \Sigma_X^{-1} J)^+$ of h . Assuming that measurement of X'_i is error-free and measurement of X_i is mutually independent, namely, X_i is an independent gaussian random variable. If there is a standard deviation of m pixels at each coordinate direction, that is, covariance matrix

of X'_i is $\Sigma_{X'_i} = \begin{bmatrix} m^2 & 0 \\ 0 & m^2 \end{bmatrix} = m^2 E$, accordingly,

$\Sigma_h = (J_h^T \Sigma_{X'}^{-1} J_h)^+ = m^2 (J_h^T J_h)^+$, thereinto,

$J_h = \partial X' / \partial h$.

Once covariance matrix of h is determined, the error of image point X' mapped in the image by any object point X on the map can be calculated, and covariance matrix of X' is as below:

$$\Sigma_{X'} = J_h \Sigma_h J_h^T = \begin{bmatrix} \sigma_{x'x'} & \sigma_{x'y'} \\ \sigma_{x'y'} & \sigma_{y'y'} \end{bmatrix}$$

What we generally concern about is the distance from the image target calculated by homography matrix and the real image target, that is, root-mean-square error at position X' :

$$RMSE = \sqrt{\sigma_{x'x'} + \sigma_{y'y'}} = \sqrt{\text{trace}(\Sigma_{X'})}$$

It can be proved that [9,10], for identity mapping, if the points selected are uniformly distributed on a unit circle correspondingly, then

$$RMSE = \sqrt{1 + (x^2 + y^2)} = \sqrt{1 + r^4}$$

That is to say, distribution of X' depends on radial distance of X' . The farther distance from the target to the center of circle is, the larger the error is.

Below we will use a checkerboard photo for test and carry out simulation verification for the above error calculation equation. Figure 3. is a checkerboard photo, where the red points numbered as 1-6 in the figure are corresponding to the points selected for calculating homography matrix, and the blue points are the test points used for calculating coordinate conversion result. Assuming that corresponding world coordinates of the points on the checkerboard have no error, measurement error of image coordinate complies with the Gaussian distribution with mean value of 0-pixel and standard deviation of three pixels.

For point correspondence in the image, gaussian noise generated every time that complies with mean value of 0 and standard deviation of three is added to coordinates corresponding to point to calculate H ,

and then map world coordinate of blue point on the chessboard to the image. Repeat for 300 times and draw 3σ ellipse distribution situation of subpoint. It can be seen that from Figure 4, projection scattering points are mostly located within two-dimensional

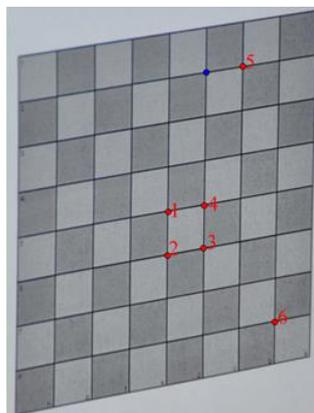


Figure 3. Chessboard Photo

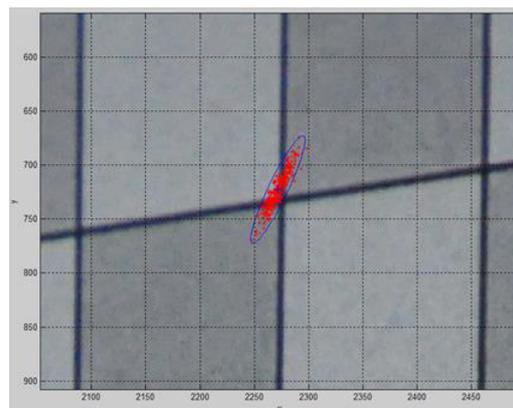


Figure 4. 3σ Ellipse and Scattering Points Perfectly Identical Through Monte Carlo Simulation Experiment for 300 Times Gaussian distribution 3σ ellipse, which indicates that the error calculation formula is reliable.

A projection of test point coordinate projection error finally obtained by respectively selecting two sets of points numbered as 1,2,3,4 and 1,2,3,5 to correspondingly calculate homography matrix is shown in Figure 5. It can be seen that, four points are used to correspondingly calculate homography matrix, but different geometric distribution causes different projection error, which indicates that corresponding geometric distribution of point has important influence on projection error, and homographic point shall try to surround the area needing coordinate conversion if we want to reduce coordinate conversion error. The error compari-

son between four and six point correspondences is shown in Figure 6, where the blue color indicates corresponding 3σ ellipse of four points, and the red color indicates corresponding 3σ ellipse of six points.

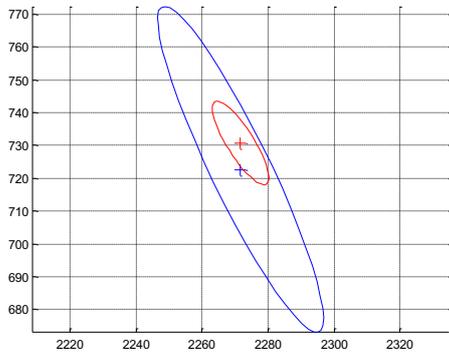


Figure 5. 3σ Error Ellipse of Four Homographic Points

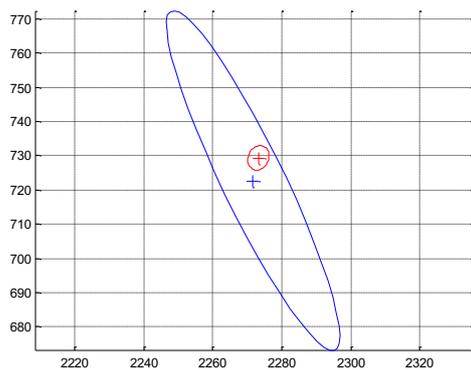


Figure 6. 3σ Error Ellipse of Four and Six Homographic Points

It shows that, the more point correspondences selected is, the smaller the error is.

An error distribution curve of homography matrix correspondingly calculated out by the points numbered as 1,2,3,4 is shown in Figure 7. And Figure 8 is an error distribution curve of homography matrix correspondingly calculated out by the points numbered as 1,2,3,5. It can be seen from Figure 7 and Figure 8 that,

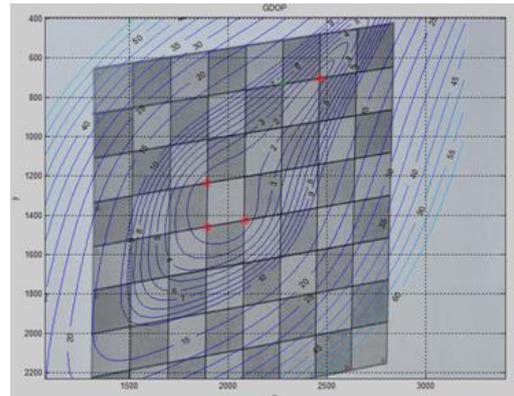


Figure 8. Error Distribution Curves Corresponding to Points 1,2,3,5.

error distribution curve changes with corresponding geometric distribution of point, and the closer distance to point correspondence is, the smaller the error is.

5 EXPERIMENTAL RESULTS

A Prosilica GC1350C industrial camera with the resolution of 1360×1024 produced by German AVT Company was used for the site video recording in the control tower of Chengdu Shuangliu International Airport on June 15, 2011. The ADS-B monitoring data

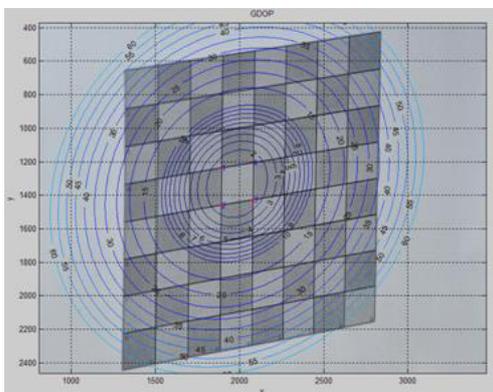


Figure 7. Error Distribution Curves Corresponding to Points 1,2,3,4

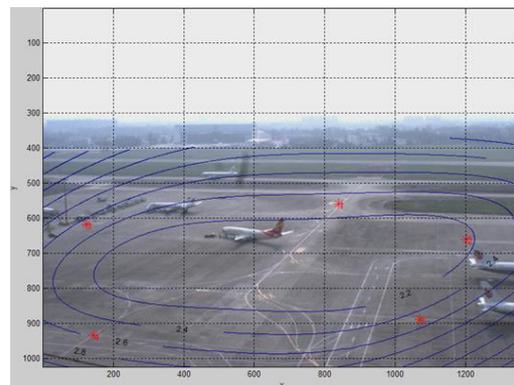


Figure 9. Error Distribution Curve of Airport Coordinate Conversion

that is synchronous with the video is from the ADS-B ground station developed by The Second Research Institute of CAAC installed in the airport.

Figure 9 is a curve to calculate homography matrix and analyze error distribution by respectively selecting five relatively obvious mark points from the image and the airport map according to the intercepted one frame of video image. For point correspondence on the image and the map, manual selection is used and it is assumed that point correspondence coordinates on the map are accurate and point correspondence image error on the image complies with Gaussian distribution with five-pixel standard deviation. It can be seen from the figure that, the airport apron area surrounded by five points has relatively small error, the error of runway and taxi-way at the top left of the image is obviously increased because they are relatively far away from the point correspondence distribution area. Homograph is a projection conversion relationship between two planes, therefore, coordinate conversion between surface area in the image and the map can be realized only, and the air area in the image hasn't been in the airport plane, so such homograph relationship is not established and the error curve isn't given out.

Background differencing method is used to detect the image moving target. Gaussian Mixture Model [14, 15] is used to establish a background model to extract background image and make difference with the original image to obtain a difference imaging, so as to detect the image moving target. KLT algorithm is used to track the image moving target [3-7]. KLT is a tracking algorithm based on feature points, which conducts multiple iterations through building a Gaussian pyramid and predicts the withdrawn feature points, so as to achieve the purpose of tracking.

Assuming that measurement of image coordinate complies with Gaussian distribution for the convenience of error analysis, generally, such assumption is not verified [9,10]. Meanwhile, measurement standard difference of homographic point on the image is also from experience and assumption. Thus more meaningful point relative to specific value of coordinate conversion error is the change trend of the error distribution curve.

The detection of moving objects in image utilizes background differencing. The background picture is extracted through establishing a background model using Gaussian Mixture Model [14, 15], and is compared with the original image to obtain a difference image, thus detecting the moving objects in image. The tracking of moving objects in image applies KLT algorithm [3-7]. KLT is a tracing algorithm based on feature points. Prediction of the feature points extracted is made through constructing a Gaussian pyramid and multiple iterations, achieving the purpose of tracking.

For the measurement of image coordinates, it is assumed to comply with the Gaussian distribution, so as to facilitate error analysis. This assumption has gener-

ally not been validated [9, 10]. Meanwhile, the standard deviation of measurement of homographic points

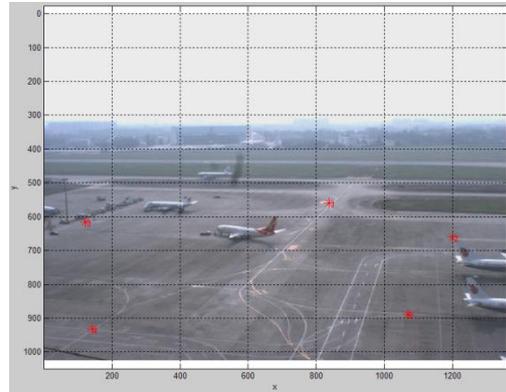


Figure 10. Selection of Homographic Points from the Airport Video Image

on image is also from experience and assumption. Therefore, compared to the specific value of coordinate conversion error, what is more significant is the variation trend of the error distribution curve.

Figure 10 shows the automatic labeling effects achieved by video data and ADS-B data fusion. The track formed by the red circles in the picture indicates the positions of aircraft's ADS-B monitoring data mapped into the image; the blue rectangle frame indicates the image target detected and tracked. The blue point in the blue rectangle frame indicates the center



Figure 11. Automatic Labeling in Airport video monitoring

of image target. The flight number of aircraft is marked on the rectangle frame. Figure 11 shows that automatic labeling has been achieved for both flights CA4462 and HU7856 in the video.

Due to the limited video recording time and limited visual angle of camera in the airport, and that not all aircrafts are installed with ADS-B airborne equipment, only two flights CA4462 and HU7856 are simultaneously recorded by video and ADS-B in the experiment. In order to further validate the reliability of coordinate transformation algorithm, a time frame not recorded

by video is selected and the geographic coordinates of CA4116's track as recorded by ADS-B data in the time frame are converted into the image coordinates,



Figure 12. CA4116 Flight

as shown in Figure 12. It can be seen from Figure 12 that the real measured coordinates of target as provided by ADS-B and converted into image coordinates are basically identical to the moving route of aircraft in image.

6 CONCLUSION

(1) The four and above points are selected from the video image and the airport map to correspondingly calculate the homography matrix. The homography matrix is used to achieve conversion between image coordinates and geographic coordinates. The ADS-B data and video monitoring data integration is used to realize automatic labeling of aircraft in the monitoring video.

(2) The coordinate conversion error due to existing coordinate measurement error at the time of selecting points from the image is calculated, and the relationship between point-corresponding geometric distribution and error curve is analyzed. The error in the central area surrounded by point correspondences is relatively small and the error in the area far away from the point correspondences is relatively large. Theoretical derivation is also identical to Monte-Carlo simulation experiment.

(3) The ADS-B data and video monitoring data actually measured at Chengdu Shuangliu International Airport were used for experiments. The results show that the method applied in the paper can automatically label aircrafts in the airport video monitoring in an effective way.

ACKNOWLEDGEMENTS

This paper is supported by the National Natural Science Foundation of China (Grant No. U1333202, U1233103).

REFERENCES

- [1] Wang Yong, Tan Yihua. & Tian Jinwen. 2008. Video segmentation algorithm with Gaussian mixture model and shadow removal. *Opto-Electronic Engineering*, 35(3): 21-25.
- [2] Zhang Qiuyu, Wang Heng. & Zhang Moyi, et al. 2012. A real-time hand gesture tracking method based on mixture Gaussian model and mean shift algorithm. *Journal of Information and Computational Science*, 9(4):889-898.
- [3] Darby J, Hodson-Tole E F. & Costen N, et al. 2012. Automated regional analysis of B-mode ultrasound images of skeletal muscle movement. *Journal of Applied Physiology*, 112(1):313-327.
- [4] Ohmer J F. & Redding N J. 2008. GPU-accelerated KLT tracking with monte-carlo-based feature reselection. // *Proceedings of Digital Image Computing Techniques and Applications*. Canberra:[s.n.]: 234-241.
- [5] Mainali P, Qiong Yang. & Lafruit G, et al. 2010. Robust low complexity feature tracking. // *Proceedings of the 17th IEEE International Conference on Image Processing*. Hong Kong; *IEEE*: 829-832.
- [6] Kim J S, Hwangbo M. & Kanade T. 2009. Realtime affine-photometric KLT feature tracker on GPU in CUDA framework. // *Proceedings of the 12th IEEE International Conference on Computer Vision Workshops*. Kyoto; *IEEE Computer Society*: 886-893.
- [7] Nijad A N, Sara T. & E Edirisinghe, et al. 2012. An Automated Real-Time People Tracking System Based on KLT Features Detection. *The International Arab Journal of Information Technology*, 9(1):100-107.
- [8] He You, Wang Guohong. & Lu Dajin, et al. 2007. Information fusion and application of the multi-sensor. The 2nd Version. Beijing: Electronic Industry Press, 2007
- [9] Hartley R. & Zisserman A, et al. 2002. *Multiple View Geometry in Computer Vision*. Hefei: Anhui University Press.
- [10] Fadi Dornaika. & Fadi Chakik. 2012. Efficient object detection and tracking in video sequences. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision*, 29(6):928-935.
- [11] Elan Dubrofsky. 2009. *Homography Estimation*. Vancouver: The University of British Columbia.
- [12] Dimitropoulos K, Grammalidis N. & Simitopoulos D, et al. 2005. Aircraft detection and tracking using intelligent cameras. // *Proceedings of the IEEE International Conference on Image Processing*. 2005. Genova; *IEEE Computer Society*: 594-597.
- [13] Wang Guanghui, Hu Zhanyi. & Wu Fuchao, et al. 2005. Single view metrology from scene constraints. *Image and Vision Computing*, 23(9): 831-840.

- [14] Ali Sefidpour. & Nizar Bouguila. 2012. Spatial color image segmentation based on finite non-Gaussian mixture models. *Expert Systems with Application*, 39(10):8993-9001.
- [15] Greggio N, Bernardino A. & Laschi C, et al. 2012. Fast estimation of Gaussian mixture models for image segmentation. *Machine Vision and Applications*, 23(4):773-789.