Automatic Extraction of Interior Orientation Data in Aerial Photography Using Image Matching Method

Helmy Mukti Wijaya, Teguh Hariyanto, and Hapsari Handayani

Department of Geomatics Engineering, Institut Teknologi Sepuluh Nopember, Surabaya

e-mail: helmy.18033@mhs.its.ac.id

Abstrak—**The Interior Orientation is a set of parameters that have been determined to transform the coordinates of the camera photo. The pixel coordinates of Fiducial Mark in the base image (Search window) are obtained automatically. The template of the fiducial mark is designed on single-frame aerial photographs. The concept of photogrammetry with Image Matching techniques is applied in the programming works. The Normalized Cross-Correlation (NCC) method coupling with the Area Based Matching technique is precisely used in the automatic computation for measuring the coordinate of fiducial marks. Herein, three templates are provided for this calculation. The coordinates from the manual rectification are employed. The results reveal that the third template is more accurate than the others with the RMSE value of 0.0066. The accuracy regarding the results of manual rectification depends on the operator when they are identifying for pixel values.**

Kata Kunci—**Interior Orientation, Fiducial Mark, Iamge Matching, Least Square, Normalized Cross-Correlation.**

I. INTRODUCTION

IN THE era before the digital era, the process of combining aerial photographs was carried out manually, including in the process of searching for orientation parameters. This orientation parameter consists of calibrated and equivalent camera focal lengths, lens distortion, principal point, camera resolution, flatness of the focal plane and fiducial mark location. The location value of fiducial mark is needed to find the middle value of the photo. In this research, a program is created to automatically search for the value of the orientation parameter (Fiducial Mark) with the image matching process. To find the fiducial mark value automatically needed an engineering program to find the conjugation points in two or more photos or images that overlap automatically which are the basis in digital photogrammetry. This process is usually referred to as photo matching or better known as Image Matching [1].

In the process of matching images from patch templates and search window is one of the core areas of research in Computer Vision and Digital Photogrammetry there are 2 techniques, namely Area Based Matching and Feature Based Matching. Technological developments in image matching have progressed using the Area Based Matching Normalized Cross-Correlation technique to get more precise matching. The main purpose of matching the images is to restore the suitability values of the two or more images provided [2].

If the overlap in the image is considered sufficient, then the same photo with the value can be described with a transformation between objects in both photos, this transformation is called Affine Transformation where the transformation is used to transform the coordinate values of a two-dimensional coordinate system into another two-dimensional coordinate system. Determination of the parameter values of a transformation is determined based on the availability of coordinate data of allied points of each two-dimensional system and the calculation technique of determining the transformation parameters [5].

One of techniques of image matching is Normalized Cross-Correlation (NCC).. In this method, the basic unit used for matching is a regular pixel-sized environment. The position of a given pattern is determined by the pixel comparison of the image with the given pattern which contains the desired pattern. For the position (m, n) of the patch template to be shifted towards the x and y directions of the search window, comparisons are calculated over the template area for each position (m, n) against the position (x, y) of the search window [3].

II. METHODOLOGY

A. Image Processing

In general, based on a combination of colors in pixels, the image is divided into three types namely RGB image, grayscale image, and binary image. Each color channel has a pixel intensity value with a bit depth of 8 bits which means it has a color variation of 2^8 (0 to 255). In the red channel, the perfect red color is represented by the value 255 and perfect black with the value of 0. In the green channel, the perfect green color is represented by the value of 255 and the perfect black with the value of 0. Likewise on the blue channel, the perfect blue color is represented by the value 255 and perfect black with a value of 0 [4]. Conversion RGB to grayscale can see Figure 1.

Grayscale image is an image whose pixel intensity values are based on gray degrees. In 8-bit grayscale images, degrees of black to white are divided into 256 degrees of gray where perfect white is represented by a value of 255 and perfect black with a value of 0. RGB images can be converted to grayscale images. The captured diseased leaves image is in RGB image. So it is necessary to convert from RGB to Grayscale for Grayscale Image pre-processing. This method matches the luminance of the grayscale image to the luminance of the color image. First get the values of three primary colors (Red, Green and Blue) and encodes this
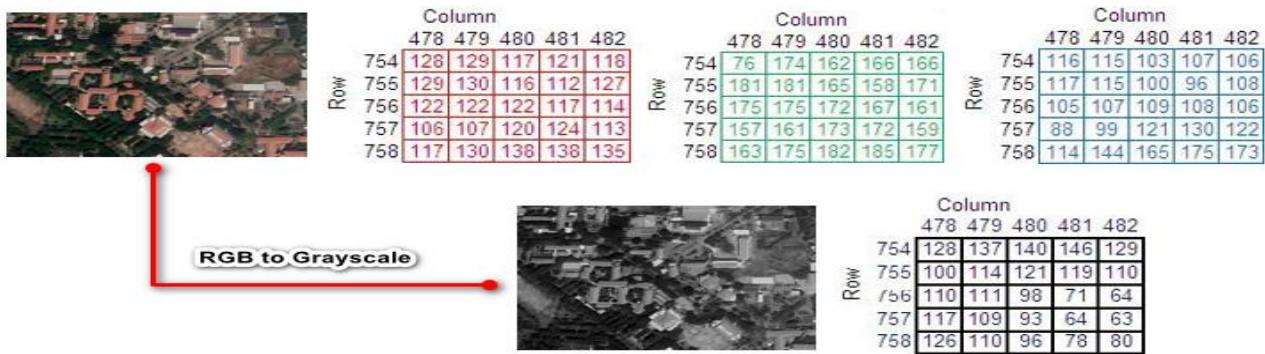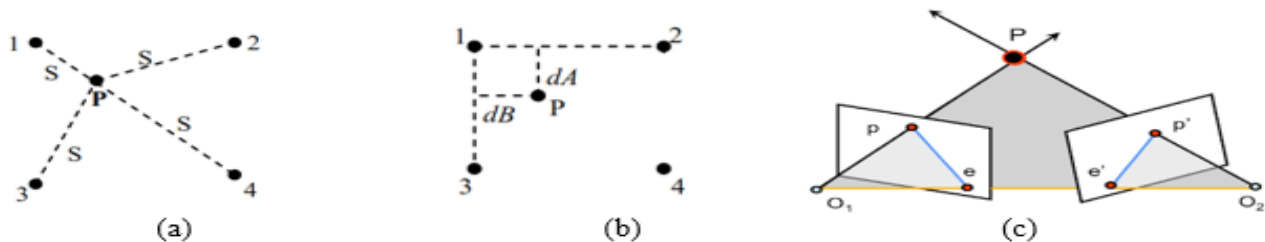
Figure 1. Conversion RGB to Grayscale.

Figure 2. (a) Nearest Neighbor; (b) Bilinear Interpolation; (c) Epipolar Plane.

linear intensity values using the gamma expansion [5]. The gamma expansion is:

$$C_{linear} = \begin{cases} \dfrac{C_{rgb}}{12.92} & \\ \dfrac{(C_{rgb}+0.065)}{1.065} & C_{rgb} \leq 0.04045 \\ & C_{rgb} > 0.04045 \end{cases} \qquad (1)$$

Here, C_{rgb} is RGB primaries which has the range from 0 to 1 and Clinear is the linear-intensity value which also has the range from 0 to 1 and C_{linear} the luminance of the output image is obtained using weighted sum of the three linear intensity values. the conversion is obtained using the function:

$$y = f(x) \qquad (2)$$

Here, x is the original input data and y is the converted output data. the function f(x) converts RGB values to grayscale values using weighted sum of the R, G, and B components:

$$f(x) = 0.2989 \, x \, R + 0.5870 \, x \, G + 0.1140 \, x \, B \qquad (3)$$

B. Image Matching

In general, Image Matching (also called Image Correlation) is based on checking and matching the gray level of a small part (Template Patch) of each stereopair photo, or matching the Image Patch of one photo with an Image Template. Matching may be on a pixel-by-pixel basis (Area Based Matching) or by checking and matching individual features of the Image Patch (Feature-Based Matching). Regardless of the methodology used, the most important Image Matching applications are: the determination of interior and exterior orientation parameter; automatic DTM manufacturing from stereo photography; and feature extraction in three dimensions such as roads, buildings, and

natural boundaries. Finding the conjugation points in two or more overlapping photos automatically is the basis in digital photogrammetry.

Studies that if overlapping is considered sufficient, then in the same photograph the value can be described by a transformation between objects in the two photographs. In this transformation there are 8 parameters and that can be approached using Affine Transformation (6 parameters). This approach is to equate accuracy to reach subpixels. Some photo matching techniques are image-based matching, feature-based matching, and symbolic matching [6-7]. The relationship between each method and its entities is shown in the following table 1.

C. Area Based Method

Gray value is an entity from an area-based method. The patch image is taken from the first photo which is then referred to as a template, and will be searched for in the second photo. Templates are usually mxn pixel in size, or m = n. The center of the template is at the center pixel of the template size, so the template is usually an odd size. The correlation value between the template and matching window is calculated to obtain the position of the object in the second photo (search window). To avoid mismatch, the position in the search window must be determined more thoroughly in this method.

Epipolar line is an intersection of epipolar plane and image plane. The epipolar plane ini (Figure 2 point (c))is obtained from the projections O_1, O_2 and point P objects. Therefore, the conjugation points P 'and P "are assumed to be the relationship between epipolar lines e' and e". So thatmatching along the epipolar line is easier, then photos can be transformed in advance or called photo normalization [8].

D. Convert Pixel Coordinates to Photos

In digital cameras the coordinate system used is the pixel

Table 1.
Image Matching Technique.

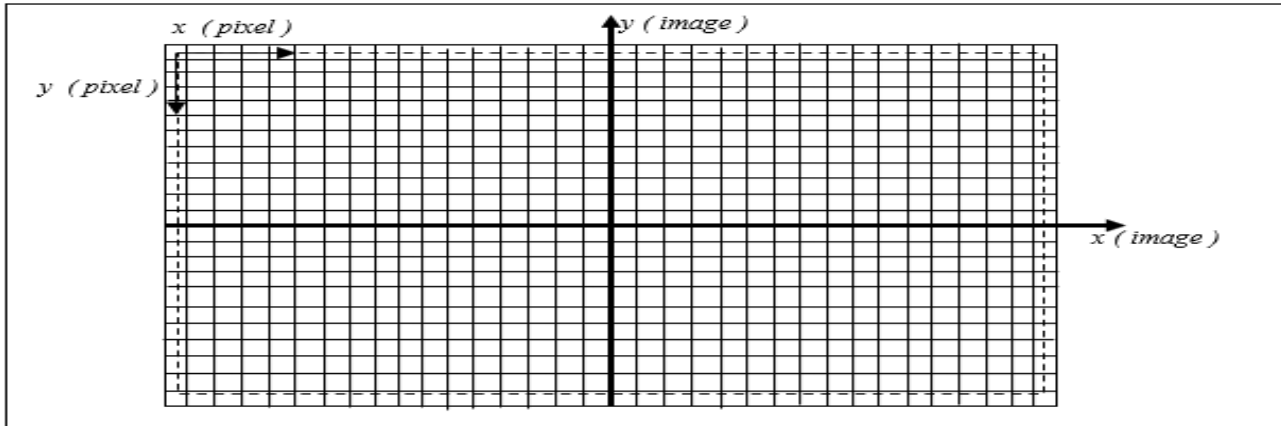| Image Matching Method	Photo Matching Calculation Techniques	Entity
Area-based	Normalized Cross-Correlation, Least Square Matching	Gray level
Feature-based	Cost function	Point, edge, area

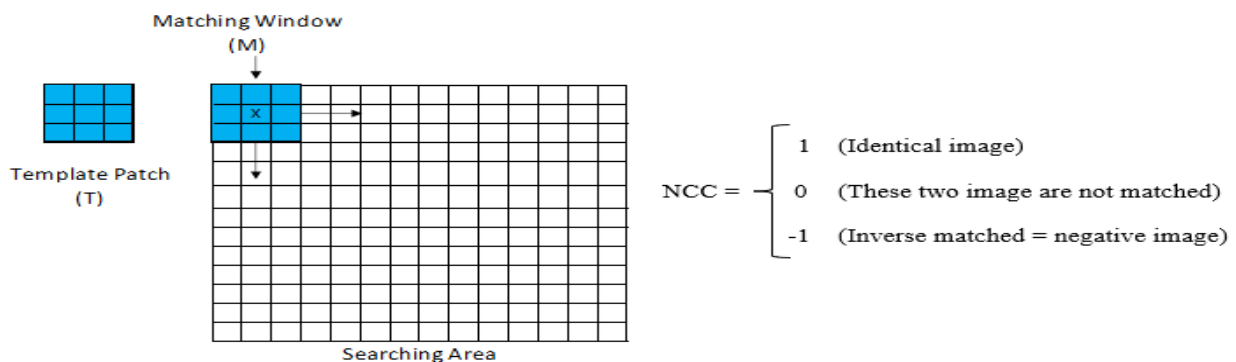Figure 3. Pixel coordinate system to photo coordinate system.

Figure 4. Calculation of Normalized Cross-Correlation.

coordinate system, while in the analytical calculation process, the system used is the Cartesian coordinate system (metric). So in this case the coordinates of the pixel system must be transformed into a photo cartesian system [9]. Pixel coordinate system to photo coordinate system can see Figure 3.

$$x = (x' - xc') \; x \; x_{\text{PixelSize}} \tag{4}$$

$$y = (yc' - y') \; x \; y_{\text{PixelSize}} \tag{5}$$

Where

$$x_c' = \left(\frac{nx'}{2}\right) - 0,5$$

$$y_c' = \left(\frac{ny'}{2}\right) - 0,5$$

E. Normalized Cross-Correlation (NCC)

In statistics, the Normalized Cross-Correlation between two random variables is a measure of how closely the two variables differ simultaneously. Similarly, Normalized Cross-Correlation in Image Matching is a measurement of the degree of similarity between two images. This level of similarity is determined by the Normalized Cross-Correlation (NCC) which is defined as ρ in the equation below :

$$\rho = \frac{\frac{\sum_{i=1}^{n}\sum_{j=1}^{m}[(g_T(x_i,y_i)-\overline{g_T})(g_M(x_i,y_i)-\overline{g_M}]}{n.m-1}}{\sqrt{\frac{\sum_{i=1}^{n}\sum_{j=1}^{m}[(g_T(x_i,y_i)-\overline{g_T})^2]}{n.m-1}}\sqrt{\frac{\sum_{i=1}^{n}\sum_{j=1}^{m}[(g_M(x_i,y_i)-\overline{g_M})^2]}{n.m-1}}} \tag{6}$$

Normalized Cross-Correlation (NCC), ρ, is calculated by sliding the template up. The search area from left to right and from top to bottom, as in Figure 2 the resulting CCC is calculated based on equation (2) determined for the pixel in the middle of the window. This shows the degree of compatibility between the template and the point in the image. Because NCC is actually a statistical correlation coefficient, its value, which ranges between -1 and 1, some researches are set the threshold for NCC value of 0.7 [10]. Calculation of Normalized Cross-Correlation can see Figure 4.

Matching window it's moving (moving window) with increment 1 pixel along each row and column in the search area (search window). The correlation value (r) between template and matching was calculated window. Correlation values between two groups of gray value data are calculated based on mathematical formulas in the following equation [11].

If in the left photo (Figure 5) an object is determined as a reference point of search, the human eye (operator) will easily
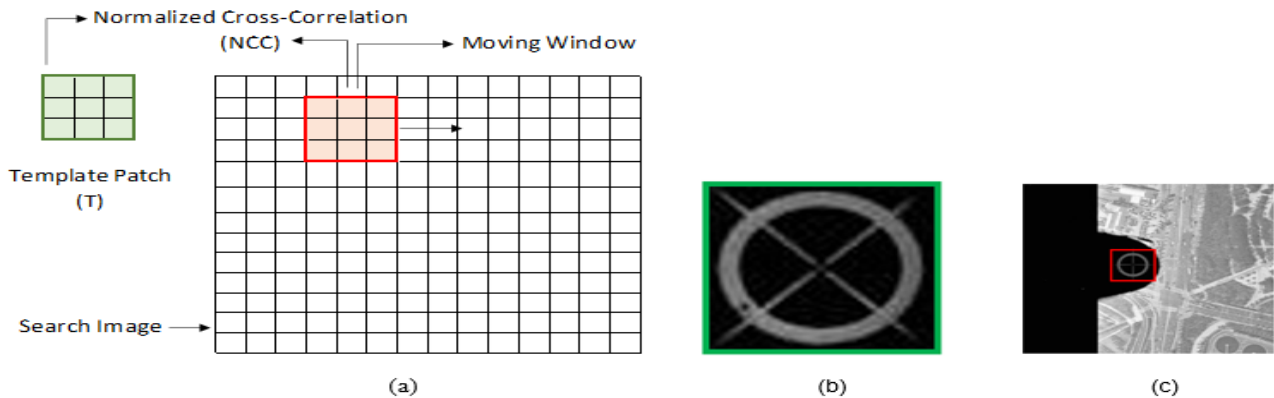
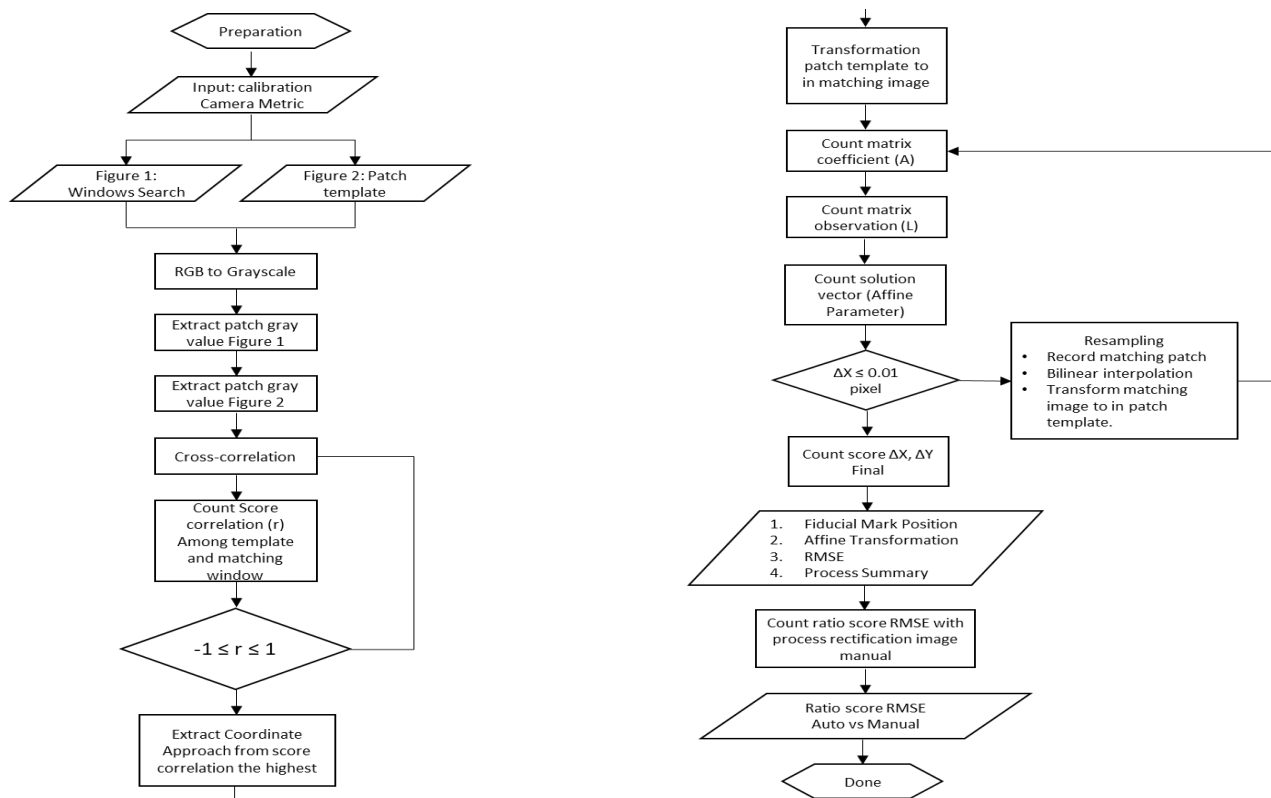Figure 5. (a) Matching scheme; (b) Template patch; (c) Search Window.

Figure 6. Flowchart Diagram.

recognize and find the object in the right photo. Not so in the digital correlation process. The computer must determine the object in the right photo by observing a set of gray values as illustrated in Figure 5. The variation of pixel values in the photo is influenced by several factors, among others, the quality and quantity of pixel values that make up the object.

F. *Least Square Adjustment*

Least Square Adjustment is a statistical technique used to estimate unknown parameters combined with a solution wherein the technique can also minimize the error value of the solution itself. In the photogrammetric technique the Least Square Adjustment method is used for the process including:

1. Estimating the value of Object Space points (X, Y, and Z) and their accuracy values.
2. Estimating and leveling the value of the Orientation parameter.

3. Minimize and distribute data errors through the observation network.

Approach Least Squareneeded for the iteration process until a solution is obtained. A solution is obtained when the residual or error value contained in a data is minimized. For a group of observations with the same weight, the main requirements must be imposed for adjustment Least Square that the number of residual squares is minimized. Furthermore, in the form of an equation the Least Square Adjustment's main requirements are stated as [12]

G. *Affine 2D Transformation*

Affine 2D Transformation is a transformation that is often used to transform coordinate values from a two-dimensional coordinate system to another two-dimensional coordinate system. Determination of the parameter values of a transformation is determined based on the availability of coordinate data of allied points of each two-dimensional
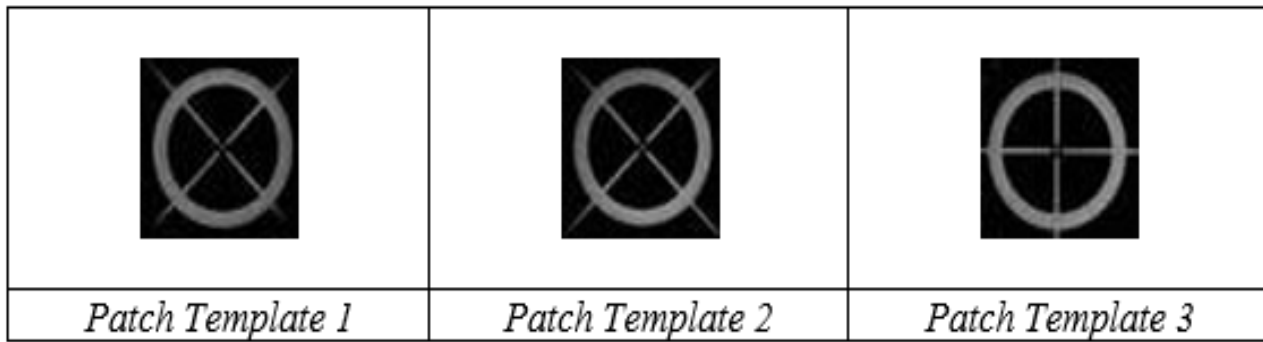
Figure 7. (a) variable template 1; (b) variable template 2; (c) variable template 3.

No	Parameter	Affine Transformation Count		
		Template 1	Template 2	Template 3
1	a11	0.041992	0.041890	0.041992
2	a12	0,000146	0,000143	0,000143
3	a21	0,000149	0,000146	0,000149
4	a22	-0.041986	-0.041984	-0.041986
5	Xt	-120.179611	-120.17421	-120.239861
6	Yt	116.904465	116.909943	116.862267

Figure 8. Affine Parameter Computing Results.

system and the calculation technique of determining the transformation parameters [13].

1) Rotation 2D

Rotational transformations involve rotation axis and rotation angle. To rotate points (x, y) with angle θ, the new position (x ′, y ′) is calculated by the following equation.

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{7}$$

2) Scaling 2D

Scaling is to change the size of the object based on scaling factors sx and sy, respectively in the x and y directions. For the point scale (x, y) with sx scale factor and sy angle, the scale point (x ′, y ′) can be obtained as follows:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} s_x & 0 \\ 0 & s_x \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{8}$$

3) Translating 2D

To translate or move objects from one position to a new position, it involves the translation distance in the x and y directions, tx and ty, respectively. To translate points (x, y) with translation distances tx and ty, the new position (x′, y′) is calculated by the following equation.

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \tag{9}$$

To get Affine Transformation, the functions above are arranged as Linear Transformation as presented in the equation below.

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \tag{10}$$

The general form of the Affine Transformation matrix is written as an equation.

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \tag{11}$$

One of the results of this work is to determine the Affine Transform parameters: a_{11}, a_{12}, a_{21}, a_{22}, t_x and t_y given photo coordinates and pixel coordinates of the Fiducial Mark.

III. FLOWCHART DIAGRAM

A flow chart can see Figure 6 is shown of work flow process, which displays diagram represents an illustration or depiction of problem solving this case.

IV. RESUT AND DISCUSSION

A. Computing Image Matching Method

In this study is to find the value of computing Fiducial Mark automatically using the Area Based Matching method with computational techniques using Normalized Cross-Correlation (NCC) to get the approach value. The comparison value below is calculated by automatically method using Normalized Cross-Correlation, which is compared with the results of Least Square Adjustment which by manually calculation. This count uses variables from 3 different templates can see Figure 7. All three templates are randomly cropped images from the edges of the aerial photograph to get template patch.

B. Affine Parameter Computing Result

From the computation results of the three Variable Templates below, produce an affine transformation parameter value that is almost uniform in value, because the rotational, translational, and dilated values of each template have the same pixel value arrangement, only differing in the angle of rotation of the image, but not too influential to the results significant and can see Figure 8..

Where a_{11}, a_{12}, a_{21}, a_{22}, t_x and t_y (6 parameters) are transformational parameters, $a_{11} \neq a_{21}$, $a_{22} \neq a_{12}$. With this transformation formula, it does not produce conform shape. So, there will be changes in angle and distance. To be able to solve the transformation parameters of as many as 6

The 6th International Seminar on Science and Technology (ISST) 2020

July 25th 2020, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

Manual Computing		Automatic Computing								
X (Pixel)	Y (Pixel)	X (Pixel)	Y (Pixel)	NCC	X (Pixel)	Y (Pixel)	NCC	X (Pixel)	Y (Pixel)	NCC
337.344	260.496	337.000	261.000	0.891	337.000	261.000	0.922	338.000	260.000	0.771
2862.286	126.886	2862.000	127.000	0.593	2862.000	127.000	0.596	2863.000	126.000	0.809
5385.817	278.664	5385.000	279.000	0.999	5385.000	279.000	0.899	5387.000	278.000	0.791
185.678	2784.535	185.000	2785.000	0.770	185.000	2785.000	0.721	187.000	2784.000	0.903
5519.822	2803.719	5519.000	2804.000	0.782	5520.000	2804.000	0.787	5521.000	2803.000	0.999
320.248	5310.831	319.000	5310.000	0.955	319.000	5311.000	0.872	321.000	5309.000	0.766
2844.004	5461.368	2843.000	5462.000	0.743	2844.000	5462.000	0.762	2845.000	5461.000	0.914
5368.439	5327.419	5368.000	5328.000	0.899	5368.000	5328.000	0.999	5369.000	5327.000	0.789

Figure 9. NCC Computing Results.

FD	USGS DATA		MANUAL		AUTOMATIC IMAGE MATCHING					
	Calibration Data		Least Square		Template 1		Template 2		Template 3	
	X (mm)	Y (mm)	X' (mm)	Y' (mm)	X_1 (mm)	Y_1 (mm)	X_2 (mm)	Y_2 (mm)	X_3 (mm)	Y_3 (mm)
1	-105.994	105.998	-105.998	106.010	-105.99	105.996	-105.986	106.002	-106.009	105.996
2	0.001	111.995	0.012	111.989	0.021	111.999	0.018	111.997	0.002	111.999
3	106.001	105.994	106.004	105.985	105.989	105.993	105.979	105.985	106.012	105.993
4	-111.996	0.003	-112.007	0.017	-112.005	0.001	-112.008	0.013	-111.989	0.001
5	112.000	-0.002	111.991	-0.009	111.984	-0.002	112.008	-0.004	112.000	-0.002
6	-106.002	-105.998	-105.997	-106.030	-106.01	-105.995	-106.020	-106.018	-106.001	-105.995
7	-0.005	-111.998	0.004	-111.981	0.001	-112	0.025	-111.988	0.009	-112.000
8	105.997	-105.998	105.993	-105.988	106.012	-105.998	105.987	-105.993	105.978	-105.998
RMSE	0.000	0.000	0.008	0.015	0.012	0.002	0.017	0.010	0.011	0.002
RMSE	0.0000		0.0116		0.0072		0.0134		0.0066	

Figure 10. Comparison of Pixel Coordinate Values.

parameters, it takes at least 3 conjugation points from each point to give 2 equations.

C. *Normalized Cross-Correlation Result*

Normalized Cross-Correlation (NCC), used for the 2D version, is routinely encountered in the Image Matching algorithm, as in this study the correlation results are obtained between the Patch Template and Search window to find the Fiducial Mark coordinate values. The following are the results of the experiments of the three different Variable Templates so that a suitable search scenario is obtained by generating the Normalized Cross-Correlation values as follows. The Figure 8 shows the count in the search for fiducial mark coordinates of the patch template that is automatically generated resulting in the Normalized Cross-Correlation (NCC) value of the matching window.

D. *Coordinate Transformation Values and RMSE*

The results below are a comparison of the photo coordinate values from the Fiducial Mark and the error tolerance value (RMSE) which is calculated based on data that is considered to be correct, namely the metric camera calibration report data issued by the United States Department of the Interior USGS. The Figure 9 results show the coordinate values that have been transformed from pixel coordinates to photo coordinates by calculating the RMSE value.

V. CONCLUSION

Automatic measurement in digital photographs is a basic procedure in photogrammetry. The Normalized Cross-Correlation (NCC) method with the Area Based Matching technique is precisely used in the computational method to automatically determine the Fiducial Mark value. Based on the automatic calculation of the random variable values of the three templates, the smallest RMSE result is 0.0066, while the result of the manual calculation is the RMSE value of 0.0116.

From the conclusions above show that the automatic method is more accurate than the manual method. It is possible that the manual rectification method can get better accuracy if a bundle iteration process is performed

REFERENCES

[1] Schenk, T. 1999. Digital Photogrammetry. USA: Terra Science, Ohio. Sachs, J. 2001. Image Resampling. Digital Light & Color. Thévenaz, P., Blu, T., and Unser, M. 1999. ImGE Interpolation and Resampling

[2] Zitova, B. and Flusser, J. 2003. Image Registration Methods a Survey. Image and Vision Computing 21 Vol.7 No.8 21, 977-1000.

[3] Fusiello A, Roberto, V., and Trucco, E. 2000. Synmetric stereo with multiple windowing. International Journal of Pattern Recognition and Artificial Intellegence, 14(8):1053-1066.

[4] Ratri Dwi Atmaja, et al. 2016. An Image Processing Method to Convert RGB Image into Binary. Indonesian Journal of Electrical Engineering and Computer Science Vol. 3, No. 2, August 2016, pp. 377-382. Telkom University

[5] Padmavathi1, Thangadurai. 2016. Implementation of RGB and Grayscale Images in Plant Leaves Disease Detection. Research and Development Centre, Bharathiar University, Coimbatore. Indian Journal of Science and Technology

[6] Potuckova, M. 2004. Image Matching and Its Application in Photogrammetry. Denmark: Department of Development Planning Aalborg University, Aalborg

[7] Leica Geosystems. 2006. Stereo Analyst User's Guide. United States of America. Luhmann, T., Robson, S., Kyle, S. and Harley, I. 2006. Close Range Photogrammetry: Principles, Techniques and Applications. UK: Whittles Publishing, Scotland. 510 pp.

[8] Potuckova, M. 2004. Image Matching and Its Application in Photogrammetry. Denmark: Department of Development Planning Aalborg University, Aalborg

[9] Mallat, S. 1999. A Wavalet Tour a Signal Processing. 2nd edition. Academic Press. Mikhail, E., Bethel, J., and McGlone, J. 2001. Introduction to Modern Photogrmmetry. New York: Jhon Wiley & Sons Inc

[10] Handayani, H.H. 2010. Digital Image Matching Method Using Normalized Cross-Correlation (Ncc). Geoid Vol. 06, No. 01, Agustus 2010 (001-005)

[11] Mitchell, H. and Pilgrim, L.J. 1987. Selection of an Image Matching Algorithm. University of Newcastle: Department of Civil Engineering and Surveying 334-341

[12] Wolf, P.R. and Dewitt, B.A. 2000. Element of Photogrammetry with Applications in GIS 3rd. McGraw-Hill. Higher Education. pp.

[13]   Makarovic, B. 1984. Consideration On Image Matching_An
        Engineering Perspective. Netherlands: International Institute for
        Aerospace Survey and Earth Sciences (ITC).