Scene Matching Techniques Using Satellite Imagery Data

A. A. SHAHIN
Professor in National Authority for Remote Sensing & Space Sciences, NARSS, Cairo, EGYPT

Abstract: - The problem of scene matching is a challenging problem in the field of image processing and pattern recognition. Therefore, it is modeled and its influencing factors are analyzed. According to sources, influence factors can be catalogued into three types: 1- changes of scenes 2- changes of image conditions 3- changes of sensors. For each factor, its mechanism is discussed. Given a pictorial description of a region of a scene, it is desired to determine which region in another scene is similar. The most efficient algorithms for scene matching are discussed. Those are the sequential hierarchical scenes matching algorithms for grey-scale and binary images and the two-stage template-matching algorithm. Experimental results are presented for matching satellite images of AI-Minea (EGYPT) and Montana (USA) using those approaches. The results prove efficiency and success in reaching the best match location with minimum required computations. A comment on the results is presented as well as a comparison between the applied methods.

Keywords: scene matching, image registration, feature space, remote sensing

Received: October 11, 2022. Revised: September 9, 2023. Accepted: October 12, 2023. Published: November 2, 2023.

1 Introduction

The human vision and memory system can sense and store a scene, and when observe the same scene again, even with different illuminations and view points, human can still recognize it [1]. In computer vision this phenomena is called Image Matching/Registration [5]. In vision navigation of an aircraft or a satellite, it is named scene matching.

Scene matching is fundamental vision navigation missions, but it has many problems such as how to improve the correct rate, accuracy and efficiency of matching in different viewpoints, illuminations, and times. In addition, we can expand the adaptability of the algorithm to match images from different sensors, or to area with less salient objects.

The studies of scene matching algorithm consist of the followings:

1- Selection of scene matching area;
2- Feature space;
3- Similarity metric;
4- Search Space and Strategy.

Many surveys [16-19] have discussed the state-of-art of scene matching methods, but seldom addressed the model and influencing factors. To answer these questions, the scene matching problem is modelled, and then, its influencing factors are analyzed.

The target of this paper is to reach the best match location with minimum computations required and this leads us to deal with the most efficient scene matching algorithms. Those are the basic sequential hierarchical scene matching, the sequential hierarchical scene matching rising edge features, and the two-stage template matching for binary images. An overview of the techniques is provided as well as an analogy between the different algorithms applied.

2 Definition and Modeling

Scene matching algorithm geometrically aligns the sensed images and the reference images of the same scene, which were taken at different time, from different viewpoints and/or by different sensors, according to similarity measurements [20-25]. Since the reference images are calibrated, the coordinate of targets in the sensed images could be known.

An image is a 2D function of greyscale on the coordinates (x, y), the reference image could be denoted as $F_{reference}$.
I$F_{real\text{-}time}$ and $F_{reference}$ should be the same if no influences in viewpoints, times, sensors and illuminations. Therefore, the problem in scene matching is created due to these influences. These influences could be catalogued into: 1) changes of scenes; 2) changes of image conditions; 3) changes of sensors. These influences are functional since its independent and dependent variables are 2D functions.

$$F_{real\text{-}time} = M\,(F_{reference}) \qquad (1)$$

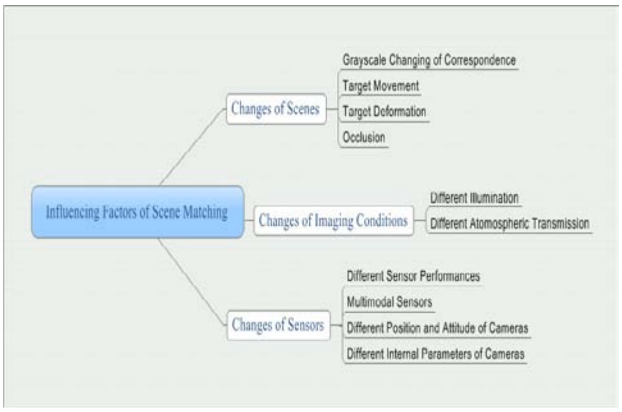

Figure: Influencing Factors of Scene Matching

2.1 Changes of Scenes

2.1.1 Greyscale Changing of Correspondence

Due to the different times at which the reference and sensed images are acquired, the albedo and radiance of the same point may change [26-30]. These kind of changes stochastic, but are decided by hidden mechanism. The green grassland in summer may change into yellow in autumn. The influence of greyscale changing can be modelled by:

$$F_{feature\ change} = M_{feature\ change}\,(F_{reference}, m). \qquad (2)$$

m denote the generalized materials but not limited to the materials of target points (x, y).

2.1.2 Target Movement, Deformation and Occlusion

The target movement and deformation could also cause differences between reference and sensed image. The greyscale of the same point didn't change but its position changed. This can be modelled by:

$$F_{feature\ movement} = M_{feature\ movement}\,(F_{reference}) = f_{reference}\{x\,(u, v), y\,(u, v)\} \qquad (3)$$

Where x (u, v), y (u, v) a re the movement correspondence points. If it could be recognized, the influences functional are modelled.

Occlusion is another influencing factor, which could be considered as greyscale changing.

$$F_{mask} = M_{mask}\,(F_{reference}, m) \qquad (4)$$

2.2 Changes of Imaging Conditions

In scene matching, the major light source is the sun.
2.2.1. The illumination in a sunny day could be modelled as parallel light (I_x, I_y, I_z) whose magnitude denotes the intensity and the direction represents the direction of light. The scene could be

also modelled by a 3x1 vector $S(S_x, S_y, S_z)$ whose magnitude denotes the albedo and the direction represents the surface normal [10].

$$F_{illumination} = M_{illumination}(F_{reference}, L, S) \quad (5)$$

2.2.2. Different Atmospheric Transmission

Atmosphere may dissipate and absorb the signal, causing changes in brightness and contrast of images.

$$F_{dissipation} = M_{dissipation}(F_{reference}, d)$$

$$= d \times F_{reference} \quad (6)$$

Where $d \in [0, 1]$ denotes the dissipation factor in transmission.

2.3 Changes of Sensors

2.3.1 Different Sensor Performances

Even in scene matching between the same modal of images, the different sensor performances might also influence the images.

$$F_{sensor\ performance} = M_{sensor\ performance}(F_{reference}, para) \quad (7)$$

The *para* describes the signal-to-noise ratio (SNR), the sensitivity and resolution and so on.

2.3.2 Multimodal Sensors

For the same scene, the optical, infrared and synthetic aperture radar (SAR) image sensed different features and images.

$$F_{sensor\ difference} = M_{sensor\ difference}(F_{reference}, m) \quad (8)$$

The mechanism of multimodal images is similar to greyscale changing of correspondence, but the differences are greater. It has a strong connection with the materials m.

2.3.3 Different Position and Attitude of Cameras

The image difference caused by camera position and attitude are modelled by

$$F_{position\ attitude} = M_{position\ attitude}(F_{reference}, R, T) =$$

$$f_{reference}\{x(u, v), (y(u, v)\} \quad (9)$$

The position changes of correspondence points x (u, v), y (u, v) follow the prospective transformation which is the dependent variables of the translation and rotation.

2.3.4 Different Internal Parameters of Cameras

Focal length difference, lens distortion, angle of view and other factors are the reason of image distortion. x(u, v), y(u, v) can also model the difference between. (10)

$$F_{inter\ parameter} = M_{inter\ parameter}(F_{reference})$$

$$= f_{reference}\{x(u, v), y(u, v)\}$$

To conclude Eq. (1~ 10), we get

$$F_{real\ time} = (M_{inner\ parameter} * M_{position\ attitude} * M_{sensor\ difference} * M_{sensor\ performance} * M_{dissipation} * M_{illumination} * M_{mask} * M_{feature\ deformation} * M_{feature\ movement} * M_{feature\ change}) * (F_{reference}) \quad (11)$$

Considering the following reasons, Eq. (11) could be simplified.

1) The occurring probability of some influencing factors is low or can be eliminated through the calibration, say the target movement, deformation, occlusion and the lens distortion.
2) Some factors share similar models and could be merged. The greyscale changing of correspondence is a slight version of multimodal.
3) Some problems have been solved by current methods, such as the difference of brightness and contrast.

$$F_{real\ time} = (M_{sensor\ difference} * M_{projection} * M_{gray\ mapping} * M_{illumination})(F_{reference}) \quad \textbf{(12)}$$

Where $M_{sensor\ difference}$ is the functional of multimodal; $M_{illumination}$ denotes the functional illumination difference; $M_{projection}$ describes the functional of perspective projection;

3. The Sequential Hierarchical Scene Matching Algorithms

These approaches incorporate a hierarchies search for a possible match location starting at a low resolution level. During the search at each resolution level, sequential and detecting rules are applied to further minimize the amount of computations.

A- The Basic Sequential Hierarchical Scene Matching Algorithms:

The simple rule by which the level-K scene is reduced level K-1 scene is simple four-point averaging [1], i.e.

BY this rule, it is possible to create a set of images which are of lower resolution and smaller size. Hierarchical search analysis is created in [12].

$$f_{K-1}(i,j)=(1/4)[f_K(2i,2j)+f_K(2i,2j+1) +f_K(2i+1,2j)+f_K(2i+1,2j+1)] \quad \textbf{(13)}$$

Two sets of these images are created, one for the window and the other for the search region. For a search region of size N x N and a window of size M x M in the highest resolution level, the number of possible match locations is (N-M+1). This number reduces to

$$[(N/2^L)-(M/2^L)+1]^2$$

When dealing with lower resolution levels, L is the search level.

Sequential Decision Rules:

Dealing with the lowest resolution level, each window pair (the window W and the sub image of the search region of the same size S) are compared and the error measurement is calculated as:

$$E_n\,K(u,v)=\sum_{i=1}^{n} e^K_{u,v}(s_i,w_i) \quad \textbf{(14)}$$

$E\,K_{u,v}(s_i, w_i) = |\, s_i - w_I \,|$ is the error measure of the *ith* window pair of test location (u, v), n =M x M, and K is the resolution level.

This is done for all possible test locations in the lowest resolution level, and then this error measure is compared against a threshold T.

Threshold Sequence Categories:

The threshold T_n. is determined as the average of the cumulative error, i. e.,

$$T_n = E_n. \quad \textbf{(15)}$$

As the search resolution increases, the threshold sequence

$$T_n^k = (\sqrt{2})^{m-k}\, r_m\, (n + g_k\sqrt{n}) \quad \textbf{(16)}$$

previously introduced must be modified. [3,4,5] suggested a method of a determining the threshold for every resolution level K :

Where,

r_m= the amplitude of the average error measure at the matched location of the lowest resolution level m.

g_k= amount of deviation from the mean.

As the value of g_k increases, the threshold increases, and so is the probability of match. However, the computational efficiency decreases.

Let this method be denoted as method A1 for g_k=0.

The most reasonable method of determining the threshold is to consider the accumulated error measurement for the number of successful test locations. Then the new threshold will be the average error calculated. In general, the threshold at any resolution level will be [6],

$$E_n^k = (1/n) \sum_{j1}^{n} Ej \quad (17)$$

Where,

E_j= the total error at each successful test location

n = the number of successful test location which is determined by the previous level.

Let this method be denoted as method A2.

$$T_n^k = (1/n\sqrt{2}).\sum_{j=1}^{n} Ej$$

(18)

Also, one may think of these error measurements as if they were the spectrum of computations. So, the effective number of results may be considered as the 3dB bandwidth of the last measured one of equation (5), so

Let this method be denoted as method A3.

The sequential decision rules can be formulated as follows,

Let N_k be a set of test locations (u, v) at search level K such that:

$$G_{k-1}(2i-1,2j-1) = \begin{cases} 1, (i,j) \in N_k \\ 0, (i,j) \notin N_k \end{cases}$$

(19)

$N_k = \{(u, v) \mid E_n^k (u, v) < T_n^k, 1 \leq n \leq M^2\}$

Where

$< T_n^k$ is threshold computed at search level k. to deal with the resolution level K-1 a location

$$G_{k-1}(2i-1,2j-1) = \begin{cases} 1, (i,j) \in N_k \\ 0, (i,j) \notin N_k \end{cases} \quad (20)$$

matrix G_{k-1} is generated whose dimension depends on the way the resolution decreases.

This location search continues until one of two cases is encountered:

a) G_{K-1} (u, v) =1 for only one value of (U, V)

b) At K=0, there exist several locations (u, v) such that G_0,(u, v)=1. Select the location with the smallest accumulated error as the most likely match location.

(b) Sequential Hierarchical Scene Matching Using Edge Features:

For this method, it is the similarity between the two that is important, it is more appropriate to use edge and introduce a measure of similarity.

(1) Edge Extractions:

Edge images created for scene matching **(21)** must be capable of meeting some basic requirements [17]. Letting the grey-scale image to be S (x, y), we can generate a binary

image S (x,y) such that:

(22)

$$S(x,y) = \begin{cases} \mathrm{Im}\,ax \; if S(x,y) \geq T \\ \mathrm{Im}\,in \;\; f S(x,y) < T \end{cases}$$

(2) Pairing Functions:

In a matching process two image arrays are produced.

$$\begin{Bmatrix} N_{00} & N_{01} \\ N_{10} & N_{11} \end{Bmatrix}$$

Using the two quantization level (0 and 1), there will be four types of pairs: 0-0, 0-1, 1-0, and 1-1. The pairing functions matrix would be given by [8],

Where Nij= the number of I in W that pair with j in S.

$$R(u,v) = \prod_{i=0}^{n-1} [Nii(u,v) / \sum_{j=0}^{n-1} Nij(u,v)]$$

A similarity correlation R (u, v) can be constructed as:

(23)

Where n= the number of quantization levels.

R (u, v) is the product of the rations of the number of matched window pairs to the number of possible matches of each type. For binary case

(24)

$$R(u,v) = \left(\frac{N_{00}}{N_{oo} + N_{01}}\right)\left(\frac{N_{11}}{N_{1o} + N_{11}}\right)$$

For each resolution level, given a probability of match, one can determine the threshold required as in [7]:

$$R_T^K = \max\left[Rk(u,v)\right] - YR_b^k \quad (25)$$

Where:

R_b^k= the background level at resolution level K.

Y= A parameter determines the probability of match.

The test location which has a similarity measure less than this threshold is eliminated from discussion in higher resolution levels.

III- The binary Two Stage Template Matching:

This approach towards increasing the efficiency of template matching is to divide the matching process into two stages. The first stage applies the optimum sub template of the given template (window) at each location of the picture (the search region). The second stage applies the entire template, but only at locations where there is a sufficient match between the sub template and the picture. This match is determined for every test location by applying a mismatch measure which counts the number of mismatch points for each kind (0 and 1) normalized to the total number of points in the template, [19], i.e.

$$(1/n)\ [N_U\ (0) + N_Z\ (1)] \quad (26)$$

Where,

U ≡ the set of template points which are 1.
Z ≡ the set of template points which are 0.
N_U (0) ≡ the number of picture points in U which are 0.
N_Z (1) ≡ the number of picture points in U which are 1.

Each mismatch measure is examined against a threshold. The successful match locations are those which have mismatch measures less than this threshold which is determined for each Optimum sub template size. This optimum size is determined for minimum computational cost which is determined as follows [19],

$$E\ (p, q, t, m, n) = m + \phi\ (c\ m^{1/2})\ [n - m] \quad (27)$$

Where,

m is a sub-template size.
n is a template size.
ϕ (c $m^{1/2}$) is False alarm probability.
q is a fraction of template points which are 1.
P is Probability of occurrence of 1 in the background.

4. Experimental Work:

The Previous algorithms have been applied on satellite imagery of parts of AI-Minea (Egypt) and Montana (USA) of size 64x64 and different window sizes. The window is selected from the search region to be at the top left corner (position (1, 1).This work is done using the Remote Image Processing System (RIPS). Figures 1, 2 and 3 show the search regions under test. Samples of the results are listed in tables I, II, and Ill for a search region of Al-Minea and a window size of 24x24 at the highest resolution level.

Table I: Performance of the basic sequential hierarchical technique

Resol-ution level, k	SR Size	Window Size	Threshold			No. of successful test locations		
			A1	A2	A3	A1	A2	A3
2	16x16	6x6	1453			122		
1	32x32	12x12	8219	3010	2129	60		
1	64x64	24x24				60	39	15
						Minimum error at position (1, 1)		

Table II: Performance of the sequential hierarchical scene matching using edge features

Resolution level, k	R_b	Max [R]	Y	R_T	P_K	No. of successful test locations
2	0.241	1.0	3	0.276	0.998	121
1						49
0	0.240	1.0	2	0.519	0.977	31
			1	0.760	0.841	8
			0.5	0.860	0.691	1
						At position (1, 1)

Table III: Performance of the binary two stage template matching

DOI: 10.37394/232026.2023.5.18

It is seen that on the three matching methods, sequential hierarchical scene matching with edge features appeared to be the best candidate for scene matching. The second candidate is the basic sequential scene matching algorithm then finally the binary two-stage template matching.

The first candidate has an advantage of having a high probability of match. Excellent performance was obtained in the matching of images of regions that have a variety of contents to have a variability of gray values. The results show that the final decision (the true match location) is reached at greatly reduced computations than the other two methods.

Scene matching with the basic sequential hierarchical method provides good performance in matching of scenes that contain relatively man-made objects of varying background. A final direct decision of the true match is rarely reached from the first step and this needs investigations of the error measure of the previous resolution level. It is noted that as the window size decreases the number of successful test location increases. As a m atter of fact, the efficiency of this method will increase as the resolution levels increase.

The third candidate shows great efficiency in dealing with the image of Montana as seen in table IV. On the other hand, this method is not effective at all in dealing with the image of AI-Minea may be because the nature of this image is that the different details are rare besides that the selected window has the same nature of the search region as a whole.

Threshold	Stage	Window Size	No. of successful test locations
0.4	1	3x3	333
0.4	2	32x32	3 minimum mismatch at (1, 1)
0.25	1	4x4	69
0.25	2	32x32	2 minimum mismatch at (1, 1)

Table IV: Performance of the binary two-stage template matching for the Image of Montana with a window size 32x32

Threshold	Stage	Window Size	No. of successful test locations
0.3	1	3x3	809
0.3	2	24x24	569

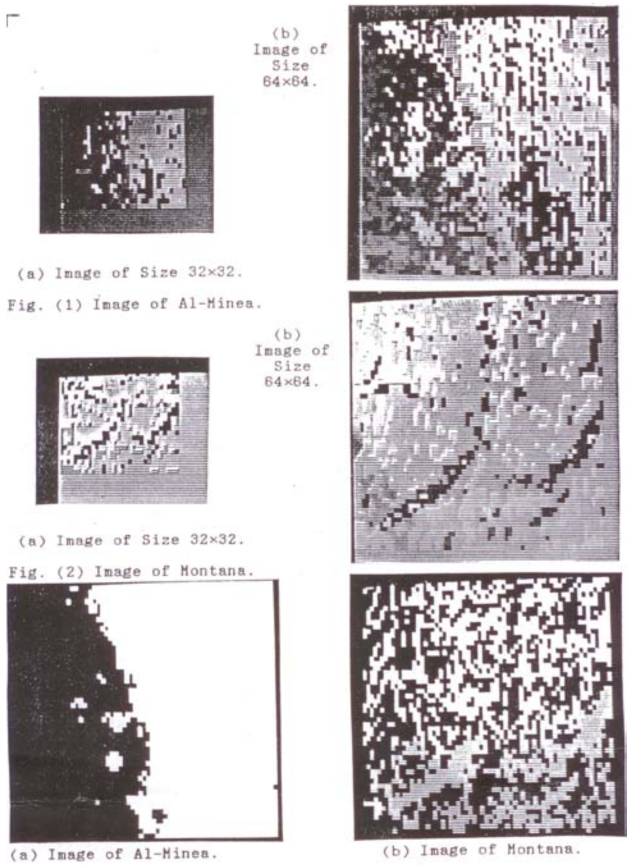

Fig. 4 Binary images of AlMinea and Montana

5. Conclusion

The disadvantages of the methods dealing with edge features, general, lie in the problem of edge extractions, which may result in a loss of the desired object with respect to background.

The problem of deciding which matching method is more effective for a certain image than the others depends on many factors such as: the kind of objects to be investigated, the dynamic range of the scene and the relation between the objects and the background. As a matter of fact the choice is restricted to the sequential hierarchical scene matching algorithms only for their superior performance which could be guaranteed for almost all images. Edge feature methods is the best for scenes which have objects that have grey scale values that vary considerably with respect to the background.

References:

[1] Li. Xiang, "Scene Matching Techniques"; Modeling and Analysis. Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 184–193, 2013.

[2] Deshmukh, M., Bhosle, U., "A Survey of Image Registration"; International Journal of Image Processing 5(3), 245–269 (2011).

[3] Wyawahare, M.V., et al., "Image Registration Techniques: An overview"; International Journal of Signal Processing 2(3), 11–28 (2009).

[4] Dawn, S., Saxena, V., Sharma, B., "Remote Sensing Image Registration Techniques: A Survey"; The International Conference on Image and Signal Processing", pp. 103–112 (2010).

[5] Guo, Q., "Modeling and simulation of the scene matching and orientation system"; Infrared and Laser Engineering 35(z1), 295–300 (2006).

[6] Wang, H., Zhang, K., Li, Y., "Research Progress on Image Matching"; Computer Engineering And Applications 40(19), 42–44 (2004).

[7] Zitova, B., Flusser, J., "Image registration methods: a survey"; Image and Vision Computing 21, 977–1000 (2003).

[8] Brown Gottesfeld, L., "A survey of image Registration Techniques"; ACM Computing Surveys 24, 325–376 (1992).

[9] E.M. Abdel-Raheem, "Microcomputer Applications in Scene Matching, "M. Sc. Thesis, Faculty of Eng., Ain Shams Univ, Cairo, December (1988).

[10] A. A. Shahin et al., "Analogy of Different Matching Methods using Satellite Imagery", Aeronautical Science & Aviation Technology ASAT; Cairo, 1985.

[11] Tanimoto, S. and Pavlidis, T., "A Hierarchical Data Structure for Picture Processing." Comp. Graphics Image Processing 4, 104-119, (1975).

[12] Hall E.L., "Computer Image Processing Recognition," Academic Press, New York, (1979).

[13] R.Y. Wong and E.L. Hall, "Sequential Hierarchical Matching," IEEE Trans. Comp. C-27, 359-366, (1978).

[14] Hall E.L., Wong, R.Y., and Rouge, J., "Sequential Scene Matching with Hierarchical Search," Proc. IEEE Southeast Conf., Williamsburg, Va., April, pp.402-405, (1977).

[15] Wong, R.Y., Hall, E.I. and Rouge, J., "Hierarchical Search for Image Matching," Proc. IEEE Conf. Decision and Control, Clearwater, Va., December (1976).

[16] Wong, R.Y., "Sequential Scene Matching Using Edge Features, "IEEE Aerosp. Electr. Syst., Vol. Aes-14, Jan. (1978).

[17] Green, G.S., Reagh, F.L. and H ibbs, E.B., "Detection Threshold Estimation for Digital Area Correlation, "Trans. Syst. Man Cybern. SMC-6, No.1, 65-70, (1976).

[18] G.J. VanderBurg and A. Rosenfeld, "Two-Stage Template Matching," IEEE Trans. Comput., Vol. C-26, No. 4, April (1977).

[19] R.Y. Wong and E.L. Hall, "Sequential Hierarchical Matching," IEEE Trans. Comp. C-27, 359-366, (1978).

[20] Hall E.L., Wong, R.Y., and Rouge, J., "Sequential Scene Matching with Hierarchical Search," Proc. IEEE Southeast Conf., Williamsburg, Va., April, pp.402-405, (1977).

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

The author contributed in the present research, at all stages from the formulation of the problem to the final findings and solution.

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself

No funding was received for conducting this study.

Conflict of Interest

The author has no conflict of interest to declare that is relevant to the content of this article.