SAR image matching area selection based on actual flight real-time images

WANG Jianmei[a], WANG Zhong[a], ZHANG Shaoming[a], FENG Tiantian [a]

[a] College of Surveying and Geo-informatics, Tongji University, Shanghai 200092, P. R. China
{jianmeiw, Zhangshaoming}@tongji.edu.cn

ABSTRACT

Matching suitability analysis is a key issue of INS/SAR integrated navigation mode. The existing suitability area selection methods are based on the simulated real-time images which are mostly satellite images. If the imaging mode of the aircraft sensor is the same as that of the satellite sensor, it does well, otherwise it doesn't work. In order to address this issue, a novel method is proposed in this paper. The sample dataset is built on the actual flight real-time images, then a hybrid feature selection method based on D-Score and SVM is used to select the suitability features and build the suitability area selection model simultaneously. Experimental results show that the consistency between the prediction results of the model and the expert labelling reaches 81.29%.

Keywords: INS/SAR integrated navigation, suitability area selection, suitability features, real-time image, reference image

1. INTRODUCTION

Inertial/Synthetic Aperture Radar (INS/SAR) integrated navigation mode is an ideal navigation mode because of its high autonomy and strong anti-jamming capability (Yin 2009). The principle of the INS/SAR navigation system is to match the real-time image taken by the aircraft with the pre-loaded satellite SAR reference image to get the real-time coordinate position, and to correct cumulative errors of the inertial navigation device after long-time working. In order to ensure the matching reliability, the suitability areas needed to be selected before the flight.

The suitability areas are the scene areas that meet the requirement of the matching probability. The matching probability refers to the probability that the real time image can be located correctly by the image matching algorithm when the real time image is in the specified area of the reference image, which size is determined by the parameters of the aircraft platform, the sensor, and the inertial navigation system. The problem of the suitability area selection is to estimate the matching probability of a scene area without the real time image. Some researchers, such as Jiang (2007), Bu (2009), Chen et al (2013), Zhang et al (2018), tried to construct a synthesized feature, which is homomorphic with the matching probability. Unfortunately, the synthesized feature that meets the need is still not found. Some researchers simplify the matching probability estimation problem into a binary classification problem (Yang & Chen 2009; Ye *et al* 2015). First, the matching probabilities are calculated based on the reference images and the simulated real-time images, then the scene areas are labeled as "suitability" or "unsuitabiliy" according to a specified threshold, finally the function between the suitability features and the scene area classes is established by supervised learning. Now the simulated real-time images are mostly satellite images. If the imaging mode of the satellite sensor is the same as the imaging mode of the aircraft sensor, the simulation method does well, otherwise it doesn't work.

The remainder of the paper is organized as follows. A novel SAR image suitability area selection method based on actual flight real-time is detailed in Section II. In Section III, taking the aerial RAR images and TerraSAR images as an example, the effectiveness of the proposed method is verified. Finally, some conclusions are reported in Section IV.

2. METHODOLOGY

2.1 Sample dataset

At the stage of test flight, the aircraft flied along the route designed by experts. The real-time image was taken in every specified area and located by image matching with the reference image. Each real-time image can be looked as a stochastic sampling. Statistically, small probability event is unlikely to happen in a single test. According to the small probability principle, the scene areas are labeled as "Suitability" as long as the real-time images match the reference images correctly, otherwise they are labeled as "unsuitability".

2.2 Suitability features

The core issue of the scene area matching suitability analysis is extraction and evaluation of the suitability features. Although the suitability features selected by different experts are various, they all follow the principles of informativeness, robustness, and uniqueness (Fu *et al*. 2003; Shen *et al*. 2010). Referring to the research results from other scholars and own practical experience, the image features in this paper include information entropy, image gray variance, Harris corner response value, edge density, surface density, autocorrelation coefficient, correlation plane statistic features. Edge density is defined as the percentage of edge pixels derived from image edge detection in whole image. Surface density is defined as the percentage of the foreground pixels derived from image segmentation in whole image. The Correlation plane is a two-dimensional surface that measure the similarity between two images in each position.

For SAR images, besides the image features, the suitability features should include the topographic features. In this paper, elevation variance and maximum height difference are used. The maximum height difference equals the maximum elevation value subtract the minimum elevation value.

2.3 Suitability feature selection and suitability area selection model

Generally, more features will provide more information, but too many features will increase the computing complexity greatly. Worse still, redundant information, irrelevant information, and even noise can mislead decisions and judgments. Feature selection from initial feature sets should be carried out firstly. There are two aspects should be considered in feature selection: one is evaluation criteria, and the other is search strategy.

Obviously, for suitability feature selection, the evaluation criterion should be maximizing matching probability. If the matching probability estimation problem is simplified into a binary classification problem, the evaluation criterion should be minimizing classification error probability. There are three types of search strategy: global optimal search, random search and heuristic search. These three search strategies have their own advantages and disadvantages. In practice, if the rules are well designed, the heuristic search strategy can achieve the similar results of the other two search strategies and compute faster (Mao *et al*. 2007).

There must be some errors in the sample dataset, so support vector machine with soft-margin is used to classification. The hybrid feature selection method based on D-Score and SVMs is used to spit out the suitability features from the candidate feature set (Weston et al. 2000; Xie et al. 2011; Dai et al. 2014). The flow chart is shown in Figure 1, （1） sort the candidate feature set A descending by D-score value, (2) initialize the suitability feature set B as an empty set, (3) move the first element in set A into set B, (4) train the SVM classifier based on the suitability features in set B， （5）if the classification accuracy is not better than before, delete the new element from set B, (6) repeat (3) – (5) until set A is empty. When the suitability feature selection is done, the suitability area selection model has been generated simultaneously.

The D-Score value uses the ratio of the inter- class variance to the intra-class variance to measure the intrinsic characteristics of the data, and could fix the difference of order of magnitude among the different features. The larger the value is, the better the classification performance of the corresponding features is. The calculation formula is as follows:

$$D_i = \frac{\frac{1}{l-1}\sum_{j=1}^{l}\frac{(\bar{x}_i^j - \bar{x}_i)^2}{\bar{x}_i}}{\sum_{j=1}^{l}\frac{1}{n_j\text{-}1}\sum_{k=1}^{n_j}\frac{(x_{k,i}^{(j)} - \bar{x}_i^{(j)})^2}{\bar{x}_i^{(j)}}} \tag{1}$$

In the formula, $\overline{x}_i$ stands for the mean value of the ith feature in the whole data set, and $\overline{x}_i^j$ stands for the mean value of the ith feature in the jth class. l is the number of classes, $x_{k,i}^{(j)}$ is the ith feature of the kth sample in the jth class.

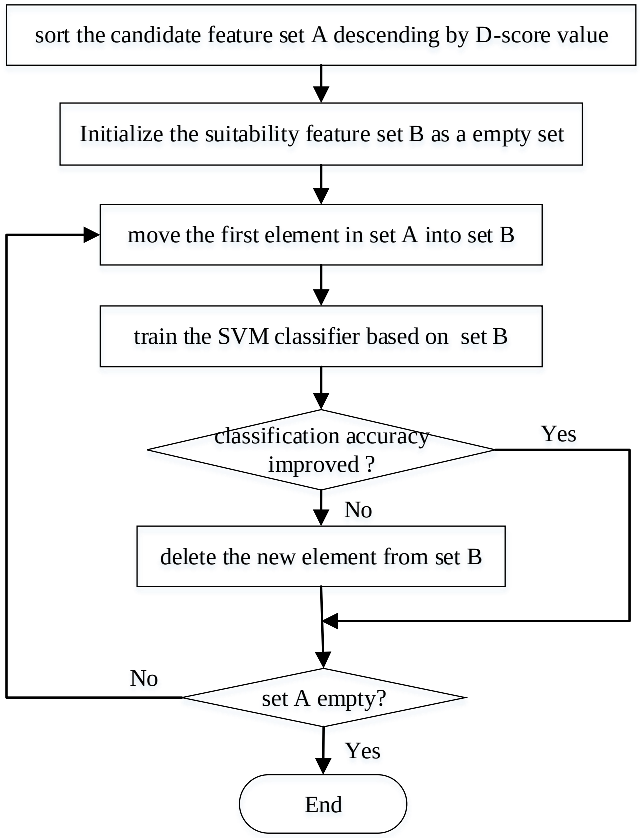
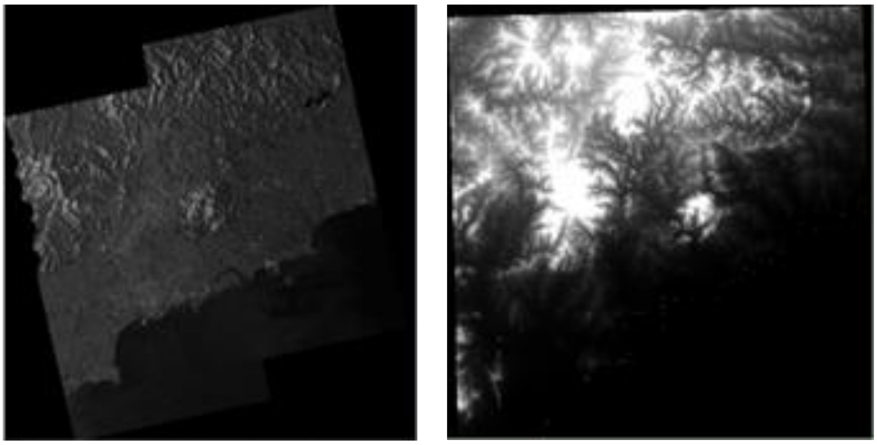


Fig.1 Hybrid feature selection method based on D-Score and SVM

3. EXPERIMENT

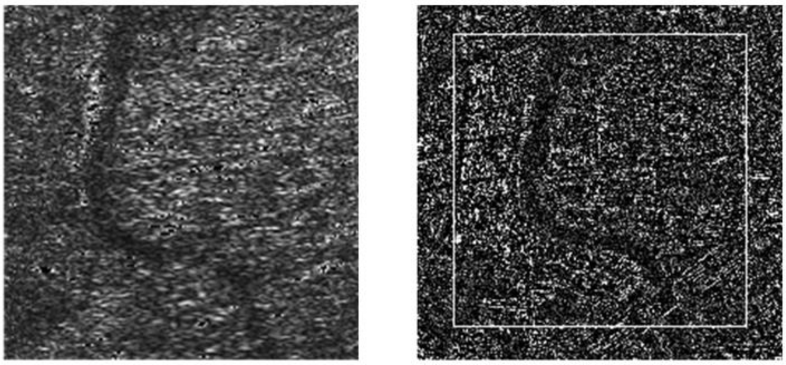
3.1 Study areas and data

The reference images used in the experiment are the X band images of TerraSAR after ortho-rectification and image mosaic. There are various types of land cover in the study areas, such as city, mountain, river, desert and so on. Figure 2 shows the SAR image (Figure 2(a)) and the corresponding DEM (Figure 2(b)) in Shanhaiguan, the resolution of both is 5m. The size of the matching area is 1500m×1500m. The real-time image is real aperture radar image (RAR), shown as Figure 3 (a). The real-time image taken by aircraft along the planned route is located by image matching with the reference image based on normalized cross-correlation algorithm. In figure 3(b), white rectangle denotes where the real-time image locates in the reference image. Both images in Figure 3 are processed by image enhancement for visualization. The matching results are checked by experts. If right, the areas are labeled as “suitability”, otherwise as “unsuitability”.



(a) TerraSAR image (b) DEM

Fig.2 Reference image and DEM image of the study area



(a) RAR image (b) TerraSAR image

Fig.3 Real-time image and reference image, the white rectangle in (b) displays the position of real-time image (a)

3.2 Results and discussions

The D-Score value of each feature calculated by formula (1) in descending order is shown in Table 1. It is obvious that topographic features have a great influence on SAR image matching. The suitability features are selected from the initial feature set by using sequential forward selection method, then the SVM classifiers based on different suitability features are trained by supervised learning, their overall classification accuracy are shown in Table 2. The optimized suitability feature set contains 2 topographic features and 4 image features, which are elevation variance and the maximum height difference, image gray variance, autocorrelation coefficient, edge density extracted by Sobel operator and surface density extracted by OTSU. The highest classification accuracy is 81.92%. The kernel function used in SVM is radial basis function. The training samples are 512 positive samples and 512 negative samples, and the test samples are 512 positive samples and 793 negative samples.

Table 1. The features and the D-Score values

Feature Code	F1	F2	F3	F4	F5
D-Score value	166.55	151.29	139.04	82.056	41.19
Feature Code	F6	F7	F8	F9	F10
D-Score value	34.98	30.55	28.18	2.41	0.56

Notes: F1 - elevation variance, F2 - image gray variance, F3 – the difference of the highest and secondary peak of the correlation plane, F4 – autocorrelation coefficient, F5 – information entropy, F6 - edge density extracted by Sobel operator, F7 - surface density extracted by OTSU, F8 – Harris corner response value, F9 - edge density extracted by Canny operator, F10 – the maximum height difference

Tab.2 The feature subsets and the corresponding classification accuracy

Feature subset	F1	F1F2	F1F2F4
Overall classification accuracy (%)	69.58	71.72	78.93
Feature subset	F1F2F4F6	F1F2F4F6F7	F1F2F4F6F7F10
Overall classification accuracy (%)	80.77	81.04	81.92

The results selected by experts and by the prediction model are visualized separately in Figure 4, the suitability areas are shown in light grey, the unsuitability areas are shown in the dark grey, the areas in white are undetermined due to lacking of DEM. Most of both are identical. Several small areas are taken out for detail analysis. Figure 5 shows the examples of correct classification. The suitability area shown as Figure 5(a) is flat and its image is rich in feature information. The mountain and hill area shown as Figure 5(b) is an unsuitability area although its image is informative. If an image is short of feature information like Figure 5(c), the area covered by the image must be an unsuitability area. Figure 6 shows an area that the expert labels it an unsuitability area, but the class labeled by the model is on the contrary. This is because half of the image is rich in information, the other is poor in information, the suitability features used in this paper are all statistics based on the whole area, which can’t describe the spatial arrangement of the image features.

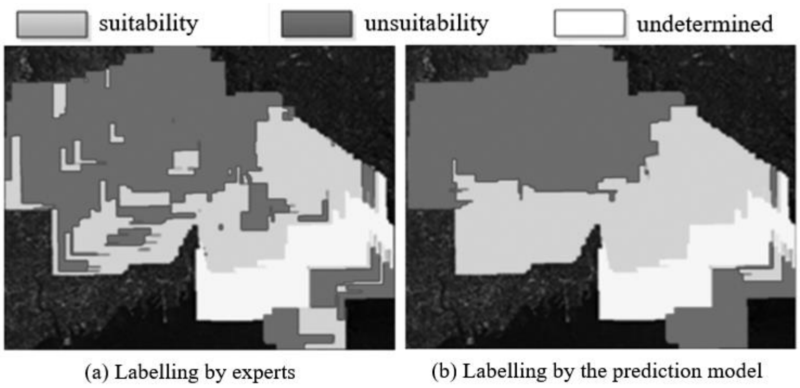


Fig.4 Labelling by experts compared with the prediction model

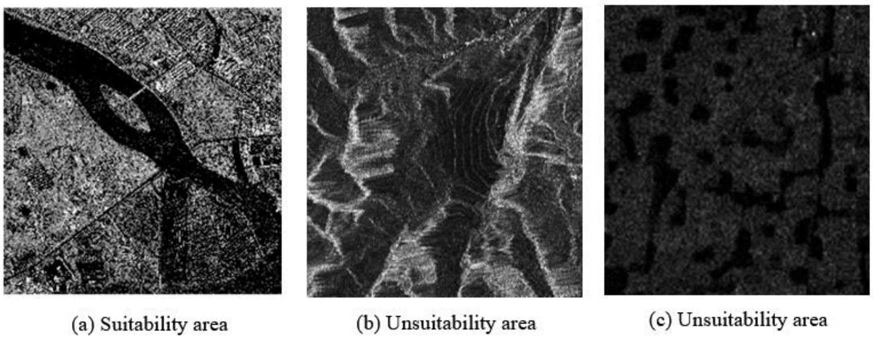


Fig.5 Examples of correct classification

Fig.6 Example of misclassification

4. CONLUSIONS

In this paper, a novel method is proposed to build the sample dataset of suitability analysis. The real-time images taken by aircraft at the stage of test flight are used to match the reference image, then the scene areas are labeled as "suitability" or "unsuitability" according to the matching results. Experimental results verify the prediction model trained based on D-score and SVM by supervised learning can select the suitability area effectively. Now the suitability features used in this paper are all statistics, which can't describe the structure information and spatial arrangement of the scene areas, this leads to some errors. In the future study, deep learning will be used to extract the better suitability features to improve the prediction accuracy.

ACKOWLEDGEMENT

The work is sponsored by The National Key Research and Development Program of China (2018YFB0505402).

REFERENCES

Bu, Yan-Long. 2009. Ph.D. dissertation: Matching suitability of SAR imaging areaes for INS/SAR integrated navigation. National University of Defense Technology, Changsha, China

Chen, Xueling, Zhao Chunhui, Li Yaojun, Cheng Yongmei. 2013. Multi-feature suitability analysis of matching area based on D-S theory. Journal of Computer Applications, 33(6): 1665-1669.

Dai, Kun, Yu, Hong-yi, Ma, Xue-gang, Li, Qing. 2014. Feature selection algorithms based on support vector machine. Journal of Information Engineering University, 15(1): 85-91.

Fu, WenXing, Wang, JianMin, Jin, ShangLiang. 2003. A practical method for selecting scene matching area. Journal of Astronautics, 24(4):348-353.

Jiang, Biaochu. 2007. Ph.D. Dissertation: Study on matching probability of radar scene and relative problems. Tongji University, Shanghai, China.

Mao, Yong, Zhou, XiaoBo, Xia, Zheng, Yin, Zheng, Sun, YouXian. 2007. A survey for study of feature selection algorithms. Pattern Recognition and Artificial Intelligence, 20(2):211-218.

Shen, Lincheng, Bu, Yanlong, Xu, Xin, Pan, Liang. 2010. Research on matching-area suitability for scene matching aided navigation. Acta Aeronautica Et Astronautica Sinica, 31(03): 553-563.

Weston J., Mukherjee S., Chapelle O.， et al． 2001. Feature selection for SVM． Advances in Neural Information Processing Systems，13: 668-674.

Xie, Juanying, Lei, Jinhu, Xie Weixin, Gao Xinbo. 2011. Hybrid feature selection methods based on D-score and support vector machine. Journal of Computer Applications, 31(12):3292-3296.

Yang, Zhaohui, Chen Yingying. 2009. Support Vector machine for scene matching area selection. Joural of Tongji University (Natural Science), 37(5): 690-695.

Ye, Qin, Chen, Hongmin, Zhang, Shaoming. 2015. Analysis of the matching suitability of remote sensing radar images of different sensors based on Sketch Token. Joural of Tongji University (Natural Science), 43(12): 1888-1894.

Yin, Decheng. 2009. Review of development of missile-borne SAR guidance technology. Modern Radar, 31(11):20-24.

Zhang, Hui, Su, Xiaojing, Zhang cong, Hao, Mengqian. 2018. Research on SAR scene matching route planning based on fitness evaluation. Journal of Test and Measurement Technology, 32(1):60-64.