

RESEARCH ON IMAGE TRANSLATION BETWEEN SAR AND OPTICAL IMAGERY

FU Zhongliang^{a,b}, ZHANG Wenyan^a

^a School of Remote and Sensing, Wuhan University, Wuhan, China - fuzhl@263.net, wenyuanzhang@126.com

^b State Key Laboratory of Information Engineering in Surveying Mapping and Remote Sensing, Wuhan University, China - fuzhl@263.net

Commission VII, WG VII/6

KEY WORDS: Image Translation; SAR; Optical Imagery; Feature Extraction; Classification; Texture Representation.

ABSTRACT:

Compare to optical sensors, Synthetic Aperture Radar (SAR) sensors can work at all time and under all weather conditions. However, SAR images are less intuitive and more difficult to understand. To complement advantages of optical and SAR sensors, a technique of image translation is put forward. Firstly, the concept named as remote sensing image translation is presented, and a set of technology thinking for multi-source remote sensing image translation is also given. Image understanding, object transformation and representation are considered as three key steps for image translation, and some specific solutions for the three steps are proposed later. Finally, an ENVISAT-ASAR image is translated into a TM image by means of feature extraction of grey scale and texture, SVM classification algorithm, seed-fill algorithm and exemplar-based texture synthesis algorithm. The experiment results demonstrate that the translated SAR images are visually similar to related optical images, while the main spatial structure information of original image is preserved, it could fill the blank regions in the incomplete TM image. All the experiments show that the proposed techniques of image translation are rational and effective.

1. INTRODUCTION

As we all know, Synthetic Aperture Radar (SAR) sensors can work at all time and under all weather conditions; while optical sensors have some limitations in time and weather. Compare to optical images, SAR images are less intuitive and more difficult to understand. As different languages in the world are the barriers to communication, the different characteristics of multi-source remote sensing images also become obstacles to information acquisitions and applications. Therefore, according to the idea of language translation beyond languages and cultural barriers, which promote interpersonal communications and information transfers (Nord and Christine, 2001), the translation techniques from SAR images to optical images are studied, so as to understand SAR images more easily, and achieve the conversion between different remote sensing data and spatial information transfer, and develop the value of remote sensing data more effectively. The translation results could compensate for the lack of optical image under certain conditions, and provide assistance and support for disaster and emergency management.

Remote sensing image translation is a new research topic, with few research works in this area (Resales et al., 2003). This paper proposes a new concept of image translation, and explores some techniques about the translation between SAR images and optical images, so as to promote spatial information transfer, and to meet different demands for certain applications.

2. CONCEPT OF REMOTE SENSING IMAGE TRANSLATION

So far, the concept of remote sensing image translation has not yet been defined. On the basis of language translation, it is defined as a procedure that, for a special purpose, transforms a kind of remote sensing image into another kind of remote

sensing image which characters or form of expression are different.

The remote sensing images translation is to meet certain demands for application. For example, during the process of earthquake emergency treatment, the disaster information may not be acquainted from optical images due to poor weather. However, the SAR sensor could work fine, and the latest SAR images may be translated into optical images, so that the helpful information are obtained and understood more easily.

The objects of image translation are two or more different kinds of remote sensing images, and which are different in image characteristics or form of expression. A variety of sensors and different imaging mechanism lead to different forms of remote sensing images. These objective differences between images make the necessity of image translation.

Conversion and representation are necessary for image translation. Image translation is neither image format conversion, nor image interpretation. It not only needs to understand the spatial information from the source image, but also represent the result as much similar as the target image. The translation result image is usually quite different with the source image.

In order to understand the process of remote sensing image translation, based on classical language translation models (Nida and Taber, 2003), a direct translation model for remote sensing images is put forward, which divides the process of image translation into three main steps: image understanding, target transformation and representation (Figure 1).



Figure 1. Remote sensing image translation model

Image understanding refers to analyse and understand the contents of the source image, and extract spatial information from the source image. Target transform means to convert the extracted information from source image to the equivalent or similar information of target image. Target representing is to organize and represent the equivalent or similar information in the form of target image.

Just as language translation need to understand the meaning of the source language, the image translation also need to understand the information of the source image. It is mainly based on the characteristics of the source image and external prior knowledge, then some image interpretation algorithms are taken to explain or identify information of the source image. Image transform is based on the results of image understanding. It establishes the mapping relationship between the objects of the source and target image, and then transform the features in the source image into the features in the target image. The interpreted objects are represented according to the structure and form of the target image.

Language translation replaces words with synonyms and reform new sentences in accordance with grammatical structure, the transformation of the image is different with language translation. Due to sensor parameters, weather or seasonal differences, shapes, colors and textures of similar feature classes is different in different kinds of image, even in the same kind image, it is difficult to build the relationship between the same class features of two kinds of image directly. The complexity and variability of features increase the difficulty of image translation. Without equative targets, we can simulate the translated objects by means of regional filling, which based on the priori knowledge or small sample images. Therefore, there are close link between target transformation and representation in image translation. It is difficult to transform and represent the features in the source image into the target image with the same or similar color and texture structure.

The information contained in remote sensing images is very rich; however, loss of information is inevitable in course of the information transfer from image understanding to target transformation and representation. Therefore, how to reduce information losses in the process of image translation is focused on. A good translation result should maintain the features in the spatial structure as faithful to the source image, while the spectrum, color and texture features are much similar with the target image.

3. METHOD OF IMAGE TRANSLATION

3.1 Image Translation Flow

Based on the concept and the model of remote sensing image translation, the general flow of multi-source image translation is summarized as follows.

Image understanding, target transformation and representation are three main steps for image translation, and image understanding is the basis of target transformation and expression.

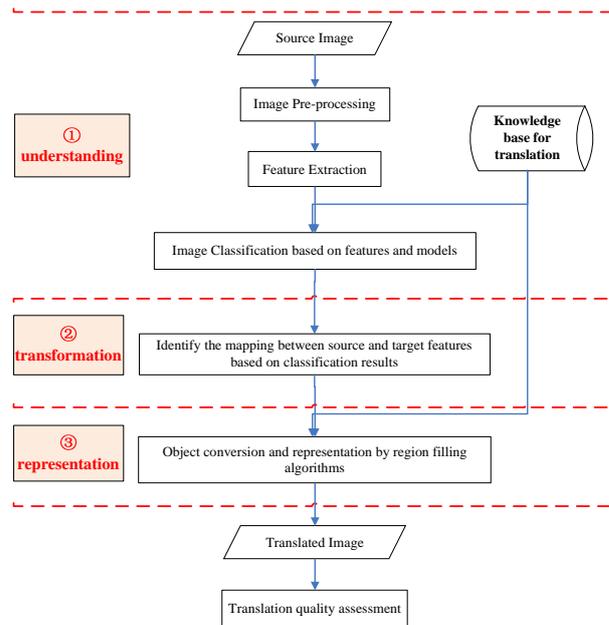


Figure 2. The general flow of image translation

Dictionary is an important tool for language translation, which provides rhyme, explanation, sample, usage, etc. Dictionaries, rule databases and the division of the marked corpus are also helpful to machine translation systems. In remote sensing image translation, grey and texture features, classification models, translation rules, and sample images are usually necessary. Therefore, it is essential to construct a knowledge base, which contains some prior knowledge, rules and sample images. A rich knowledge base would promote automation or intelligence of image translation.

3.2 Image Understanding

The SAR image is interpreted with the method of image classification. Here, the objects are not only extracted, but also the mapping relationships between the SAR and optical image are created by means of the information of class, and it will build a solid base for the following transformation and representation of the objects.

Support Vector Machine (SVM) algorithm is selected to classify the SAR image for the following two reasons. One is that the number of SAR image samples is limited, and SVM algorithm has special advantages on solving the problem about small sample, non-linear and high dimensional pattern recognition. The other is that it has a distinguishing learning performance and nice generalization ability (Waske and Benediktsson, 2007).

The performance of SVM algorithm depends on image features mostly. In order to improve the classification precision of single band and single-polarization SAR image, the combination of grey scale statistic features, GLCM(Gray Level Co-occurrence Matrices) features and Gabor filtering texture features is adopted for classification. The grey scale statistic feature of each pixel is the mean value and grey scale variance value of its adjacent pixels, which are regarded as the basic features for the analysis of SAR images. Five frequently-used and well-performed texture features are extracted from grey-scale co-occurrence matrix, which are angle second order moment, contrast, relativity, entropy and homogeneity (Dell and Gamba, 2003). Two-dimensional Gabor wavelet transform perform well

in processing multi-scale and directional characteristics, and local detail features with different scales and directions could be extracted by Gabor filters, which are benefit to texture image segmentation and classification (Bianconi and Fernandez, 2007). Here, the mean and standard deviation of coefficient image derived from Gabor filters are adopted as Gabor texture features.

3.3 Target Transformation and Representation

The transformation and representation of image objects are more difficult than words or phrase, because it is difficult to find the equivalent targets to replace objects of the source image. Region filling techniques are adopted to realize target transformation between two kinds of images and representation in this paper. The regions of different classes of the source image are filling with the colour or texture information which is from prior knowledge or samples from the target image.

For example, in order to transform a SAR image to the specific optical image, in the case of lacking the correspond optical image covering the same area with SAR image, optical samples can be collected from the established knowledge base or optical images of neighbouring areas with same feature class. Then, the region of classified SAR image is filled with colour and texture information of optical samples. For the classification result of SAR image, some homogeneous areas (such as rivers, lakes) can be filled with a specified color from the target image, and the classes with rich texture would be represented by texture synthesis based on the samples.

Water objects are usually homogeneous areas in SAR image, and each region is approximately composed by single color. Then, the water target in the SAR image is translated into optical images by filling these areas with the mean value of pixels in water sample of optical images.

Except the homogeneous classes, most objects have both colour information and texture structure, these classes in SAR images could be translated by example-based texture synthesis algorithm. Texture synthesis technique is generated in researching texture mapping in the field of virtual simulation, which can simulate artificial or natural texture by algorithms. The basic idea of sample-based texture synthesis algorithm is to construct a mathematical model with giving a small example image by analysing the texture structure, and large regions with the similar texture are synthesised based on the examples and mathematical models (Efros and Leung, 1999). Sample-based texture synthesis can generate similar and continuous texture, the algorithm is relatively simple, thus is paid attention by more researcher. This algorithm can also be applied to image inpainting or object removing (Criminisi et al., 2004), for example, to fill the missing data or remove the uninterested objects in an image.

In representing SAR regions with texture, the patch-based sample synthesis algorithm (Efros and Freeman, 2001; Liang et al., 2001) search the most matching patch from the input sample and match with the synthesized texture image. Compare to one-pixel-at-a-time synthesis algorithms, pixel-based algorithm, the patch-base sample synthesis algorithm is faster and it makes high-quality texture synthesis, because it synthesizes an image block with texture structure every time.

Efros and Freeman(2001) presented a classic patch-based texture synthesis algorithm. On the basis of Markov random field, the algorithm searches a matched block from samples and

make that the overlapping pixels have the minimum squares of errors. Then, the whole matched blocks is adjoined with the combined image. The steps of this algorithm are described as follows:

- (1) Synthesize image blocks step by step in accordance with scan order.
- (2) For every location, search the input texture for a set of blocks that satisfy the overlap constraints (above and left) within some error tolerance. One is selected randomly.
- (3) Compute the error surface between the new chosen block and the old blocks in the overlap region. Find the minimum cost path along this surface and make it as the boundary of the new block. Paste the block onto the texture.
- (4) Repeat until the whole image is synthesized.

In order to make the cut between two overlapping blocks on the pixels where the two textures match best (that is, where the overlap error is low), dynamic programming was taken to do this.

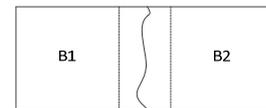


Figure 3. Minimum error boundary cut

The minimal cost path through the error surface is computed in the following mode. If B_1 and B_2 are two blocks that overlap along their vertical edge with the overlap regions B_1^{ov} and B_2^{ov} respectively, as Figure 3, the error surface is defined as $e = (B_1^{ov} - B_2^{ov})^2$. To get the minimal vertical cut through this surface, the cumulative minimum error E_{ij} is computed in every path:

$$E_{i,j} = e_{ij} + \min(E_{i-1,j}, E_{i-1,j-1}, E_{i-1,j+1}) \quad (1)$$

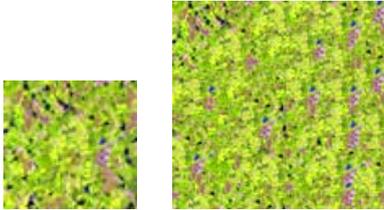
The minimum value of E in the last row will indicate the end of the minimal vertical path though the surface and the best cut path will be found by traced back. The similar procedure can be applied to horizontal overlaps. When there is both a vertical and a horizontal overlap, the best paths in tow direction have the intersection and the final cut boundary is determined by tow paths.

Efros & Freeman's algorithm is not only easy to implement, but also is quite fast. The synthesis results is fine. It is applied to various types of texture. So this algorithm is adopted to transfer and represent the texture targets. In order to synthesize different texture targets, the texture synthesis algorithm is developed, and some parameters, such as block size, percentage of overlap along X or Y dimensions, block reduction factor in each iteration and so on, are adjustable in the algorithm.

A digital photo and a small TM image were used to test the performance of the texture synthesis algorithm. The size of sample images are 64×64 pixels and the matching blocks is 27×27 pixels, the target image size is 128×128. The synthetic results of the two experiments are shown in figure 4. From the results we know that the algorithm has good texture synthesis effects and it can keep the texture structure of samples effectively.



(a) A digital photo sample and synthesis result



(b) TM image sample and synthesis result

Figure 4. Texture synthesis of different samples

4. EXPERIMENT

4.1 Experiment Data

An ENVISAT-ASAR image and a TM image were selected as the source and target of image translation. The ENVISAT-ASAR data is a single band and HH polarized image, which was imaged on August 17th, 2005, with 30 meter resolution. The image covered East Lake of Wuhan city, contained mainly three classes: lake, vegetation and artificial buildings, and the width and height of the image are 382 and 364 pixels. The intensity image with 256 grey levels of the origin complex SAR data is used, and the speckle of the SAR image had been reduced by enhanced Frost filter algorithm, see Figure 5.



Figure 5. The SAR image

A Landsat TM image covering the area around East Lake was chosen as the optical image. The space resolution of TM image is 28.5 m, and is very close to the SAR image. Band 5, 4, 3 are selected to combine a RGB image. Due to shelter by clouds or other factors, the area around East Lake in the TM image is missed, as shown in figure 6. The goal is to translate the SAR image to a false color TM image, and fill the missed area with the translated SAR image.

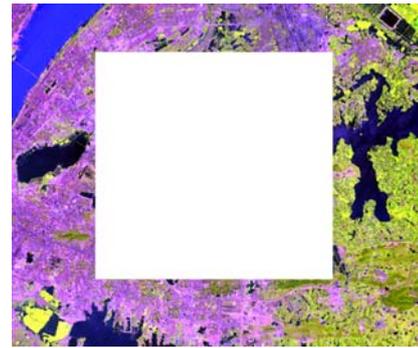


Figure 6. TM image

4.2 Experiment Result

According to the translation procedure and methods described above, the SVM algorithm with fusion of grey scale and texture features was used to classify the SAR image. SVM classification algorithm was developed with LibSVM 3.0 (an Open Source Library) (Chang and Lin, 2011). Radial Basis Function (RBF) was chosen as the kernel Function of SVM, and the parameters of RBF are obtained automatically through the techniques of cross validation and grid searching. The SAR image was classified into three classes: water, vegetation and artificial buildings (shown in Figure 7).

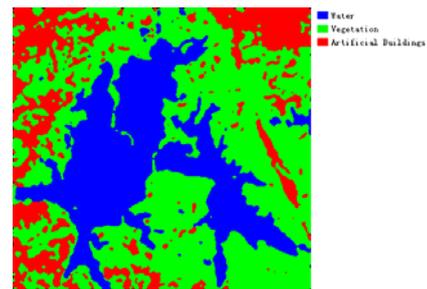
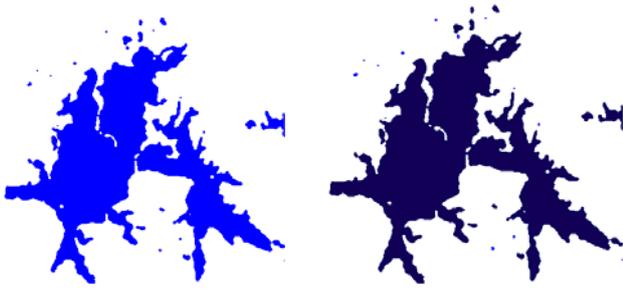


Figure 7. The classification result of SAR image

A certain number of Region of Interests (ROI) were selected precisely as verification samples, the general accuracy of the SVM classification algorithm is 93.13% and the Kappa coefficient is 0.89. This proved that the SVM classification algorithm with the fusion of grey scale and texture features can effectively extract three kinds of typical targets from the SAR image.

Next, we transformed and represented the classes of SAR image with different region filling methods. Firstly, the lake regions was separated from the SAR classification image (Figure 8(a)), and then, the seed filling algorithm was applied to fill these regions by the specific color (RGB = (20, 10, 94)), which was calculated from the lake samples of TM images. The filling result was shown in Figure 8.

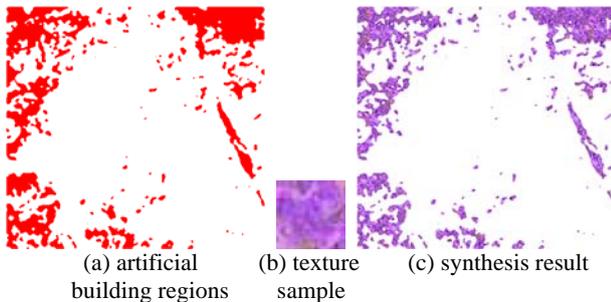


(a) the extracted lake regions (b) seed filling result
 Figure 8. Translation of lake class

Secondly, patch-based texture synthesis algorithm is applied to represent the regions of vegetation and artificial buildings from SAR classification. The sample images are searched from the translation knowledge base or TM images, and a small vegetation sample (Figure 9(b)) with the size of 16×16 pixels was selected to synthesize the regions of vegetation (Figure 9(c)). The input texture image size of artificial buildings is 27×27 pixels (Figure 10(b)), and the synthesis result based on the input sample was shown as Figure 10 (c).



(a) vegetation regions (b) sample image (c) synthesis result
 Figure 9. Translation of vegetation class



(a) artificial building regions (b) texture sample (c) synthesis result
 Figure 10. Translation of artificial building class

Finally, the three kinds of translation results were combined as a complete image, as shown in Figure 11.

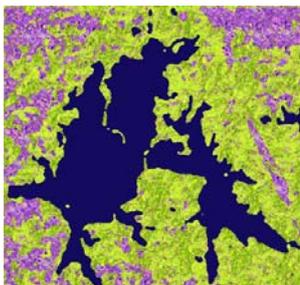


Figure 11. The final result of SAR image translation



Figure 12. Filling the blanks of TM image with translated result

Filling the missed areas of TM image with the translated SAR image, we can get a complete optical image (shown in Figure 12). The filling result shows that the translated image is visually similar to the original TM image, while as much as spatial structure information of the origin SAR image is preserved. There is little difference in the global colors and textures between the translated image and the origin TM image, but some nuances exist at the edge of two images. Since the color and texture of vegetation in different regions of the TM image are not always the same, and the single band SAR image is difficult to be classified as different kinds of vegetation, all the regions of vegetation in the translated image could only be synthesized by one kind of small sample from the TM image, so a small amount of difference between the translation result and the TM image is inevitable.

5. CONCLUSION

Remote sensing image translation is a new research topic with practical significance. So far, there are few research theories and results. This paper puts forward the new concept of remote sensing image translation and explores the translation technologies based on image understanding, target transform and representation. Grey scale, GLCM features and texture features by Gabor filtering are extracted from the single-band SAR image, and the fusion of the three kinds of features is input to SVM, water, vegetation and artificial buildings are classified from the SAR image based on the SVM algorithm. Then, these classes are represented by the seed filling algorithm or sample-based texture synthesis algorithm. Finally, the translated SAR image could be filled the blank areas in the TM image. The experiments of image translation demonstrate that the translated SAR images are visually similar to the related optical images, while the original spatial structure information is preserved. The filling result based on image translation also performs well. The proposed methods of remote sensing image translation are feasible and effective.

Although the ENVISAT-ASAR image is successfully translated into a TM image, there may be some challenges in translating those remote sensing images with high resolution.

REFERENCES

Bianconi F., Fernandez A., 2007. Evaluation of the effects of Gabor filter parameters on texture classification. *Pattern Recognition*, 40, pp. 3325-3335.

Chang C. C., Lin C. J., 2011. LIBSVM: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2, pp. 27.

Criminisi A., Perez P., Toyama K., 2004. Region filling and object removal by exemplar-based image inpainting. *IEEE Transactions on Image Processing*, 13, pp. 1200-1212.

Dell'Acqua F., Gamba P., 2003. Texture-based characterization of urban environments on satellite SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 41, pp. 153-159.

Efros A. A., Freeman W. T., 2001. Image quilting for texture synthesis and transfer, Citeseer; City. pp. 341-346.

Efros A. A., Leung T. K., 1999. Texture synthesis by non-parametric sampling, Published by the IEEE Computer Society; City. pp. 1033-1038.

Liang L., Liu C., XU Y. Q., et al., 2001. Real-time texture synthesis by patch-based sampling. *ACM Transactions on Graphics (ToG)*, 20, pp.127-150.

Nida E. A., Taber C. R., 1969. The theory and practice of translation. Brill Academic Pub.218 pages

Nord, Christine, 2001. Translating as a Purposeful Activity: Functionalist Approaches Explained. Shanghai Foreign Language Education Press; Shanghai.

Resales R., Achan K., Frey B., 2003. Unsupervised Image Translation, IEEE; City, pp. 472-478.

Waske B, BENEDIKTSSON J A, 2007. Fusion of Support Vector Machines for Classification of Multisensor Data. *IEEE Transactions on Geoscience and Remote Sensing*, 45, pp. 3858-3866.