SAR Image Content Retrieval by Speckle Robust Compression-based Methods

Daniela Espinoza-Molina, Jayashree Chadalawada, Mihai Datcu
Remote Sensing Technology Institute
German Aerospace Center (DLR)
Oberpfaffenhofen, Germany

INTRODUCTION

This poster presents a study of content-based image retrieval using compression-based methods with original and despeckled TerraSAR-X images. This study aims at analyzing the behavior of SAR image retrieval regarding speckle noise. Our study method is based on the Lempel-Ziv-Welch (LZW) compression algorithm for feature extraction and Fast Compression Distance (FCD) as a similarity metric. The intention of employing compression-based image retrieval techniques is to exploit the compression properties of objects and to estimate the information being shared between them. In particular, Fast Compression Distance can be applied to large datasets. FCD between any two objects can be computed using the sizes of their dictionaries (sequences of recurring patterns) extracted through compression with the LZW algorithm and the intersection of their dictionaries. The experimental results demonstrate that the method is immune to speckle noise.

IMAGE RETRIEVAL BY FAST COMPRESSION DISTANCE

- Normalized Compression Distance (NCD) [1] measures the similarity between two images based on their compression rate. The NCD can be directly computed between any two strings or files **x** and **y** and it represents how different they are.
- Pattern Recognition based on Data Compression (PRDC) was introduced by [2] as a classification methodology for general data. The main idea of PRDC is to extract dictionaries by applying a compressor (e.g., the LZW algorithm), directly from each object represented by a string. The data have to be previously encoded into a string.
- Fast Compression Distance (FCD) is a similarity measure based on data compression following a parameter-free approach. It was proposed in [3]. This approach combines the robustness of NCD with the speed of PRDC by establishing a link between both concepts.
- The FCD between two images **x** and **y** is given by the following equation[3]:

$$FCD(x,y) = \frac{|D(x)| - \cap(D(x),D(y))}{|D(x)|}$$

where $|D(x)|$ and $|D(y)|$ are the dictionary sizes of **x** and **y** respectively and $\cap(D(x),D(y))$ is the intersection between the dictionaries of **x** and **y**. Fig. 1 displays a graphical representation of this concept.

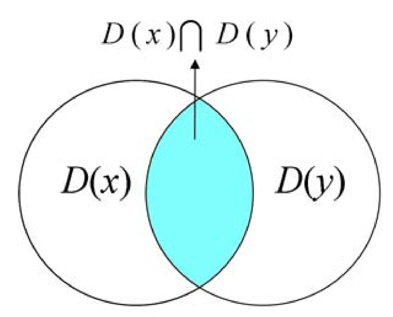

Fig. 1: Graphical representation of the intersection of two dictionaries (x,y).

- The LZW compressor [4] linearly scans the data and encodes the data into strings that are incorporated and continuously updated in a data structure called "Dynamic Dictionary". This method is used for the compression of images and extracts their dictionaries.

DATASET AND METHODOLOGY

Test dataset

The data used in this experiment stems from a Spatially Enhanced (SE) Spotlight mode (SL) Single polarization (S) Multilook Ground range Detected (MGD) TerraSAR-X image of Toronto City. We created two test datasets: the first dataset contains the original TerraSAR-X image, while the second dataset is composed of TerraSAR-X sub-scenes after despeckling. An Enhanced Lee Filter was used for despeckling.

We selected TerraSAR-X sub-scenes representing 6 major land use classes (High rise buildings, Industry, Residential area, Roads, Vegetation, and Water bodies) of the Toronto image. An example of each land use class is shown in Fig. 2.

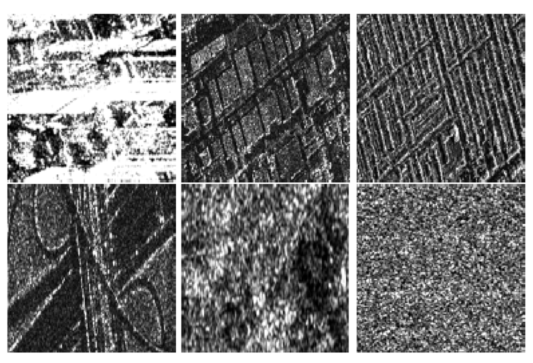

Fig. 2: TerraSAR-X sub-scenes representing (left to right): high rise buildings, industry, residential area, roads, vegetation and water bodies.

The test database is composed of:

- ✓ Residential area- 351 images
- ✓ High rise buildings - 148 images
- ✓ Industry – 93 images
- ✓ Water bodies – 45 images
- ✓ Roads – 27 images and
- ✓ Vegetation – 16 images

Methodology

The evaluation methodology is composed of the following steps:

1. Test database creation by visual indexing and manual annotation of image patches. The database comprises image patches of 64x64 pixels being representative of typical land use classes.
2. Dictionaries are extracted from each patch using the LZW algorithm.
3. One patch per land use class is chosen for querying and similar patches are retrieved based on Fast Compression Distance.
4. Computation of evaluation metrics (e.g., precision and recall, confusion matrix).

In order to create the second dataset, we applied an Enhanced Lee Filter to the same database resulting in a second dataset with despeckled images.

EVALUATION RESULTS

The evaluation was carried out using a TerraSAR-X image of Toronto City with 1.2 m spatial resolution.

Retrieval experiment using high resolution TerraSAR-X image patches

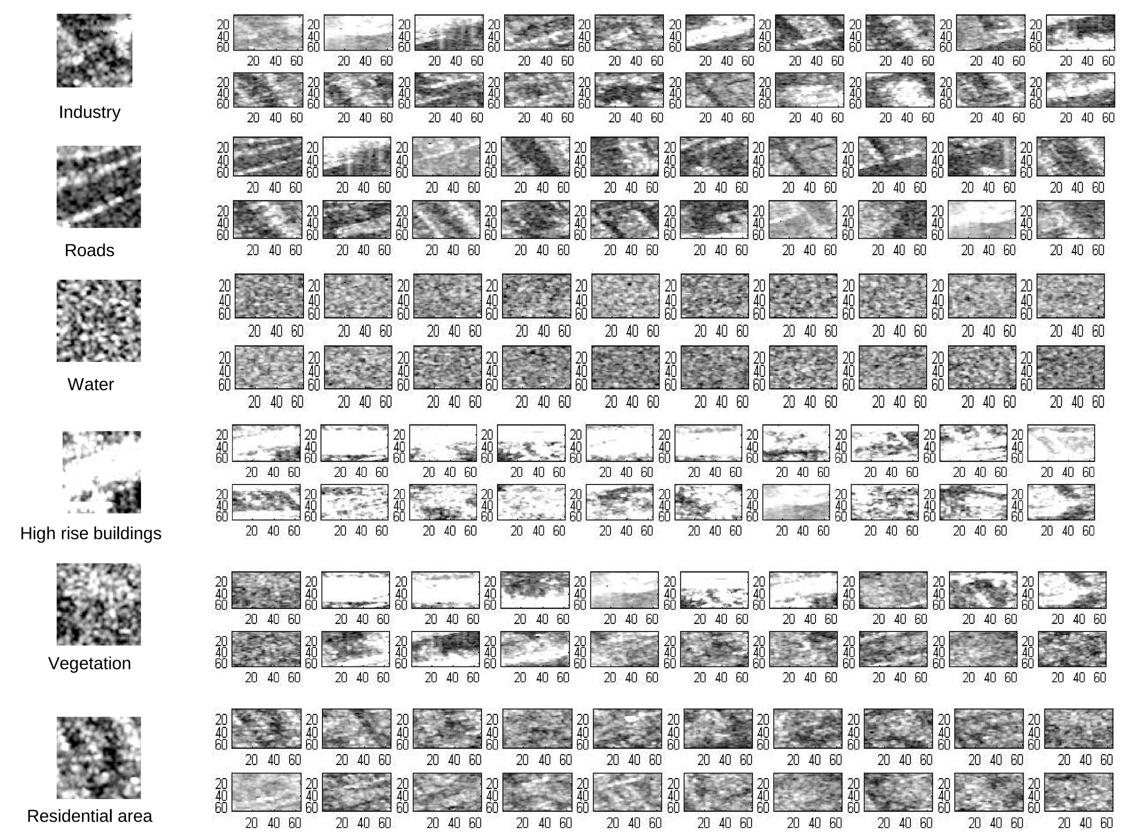

Fig. 3: Retrieval results of TerraSAR-X patches of Toronto City showing the top ranking 20 patches.

The confusion matrix corresponding to Fig. 3 is presented below:

Category	High rise	Resident.	Water	Industry	Vegetation	Roads	Total
High rise	90	20	0	11	0	3	**124**
Resident.	46	251	0	37	7	0	**341**
Water	0	0	45	0	0	13	**58**
Industry	12	54	0	24	2	1	**93**
Vegetation	0	15	0	7	7	1	**30**
Roads	0	11	0	14	0	9	**34**
Total	**148**	**351**	**45**	**93**	**16**	**27**	**680**

- Water has the best performance followed by Residential area, High rise buildings, Vegetation, Roads and Industry.
- Water often interferes (about 50%) with Roads.
- Vegetation often interferes with Residential area and the least with Industry.
- The overall accuracy is 62%.
- The overall retrieval accuracy of TerraSAR-X patches with a size of 64x64 pixels is 62.65% in spite of the interference between urban subclasses.

Performance evaluation using original and despeckled TerraSAR-X data

The left part of Fig. 4 shows the precision and recall curves of original TerraSAR-X data, while the precision-recall curves summarizing the performance of despeckled patches of six land use classes are shown on the right side of Fig. 4.

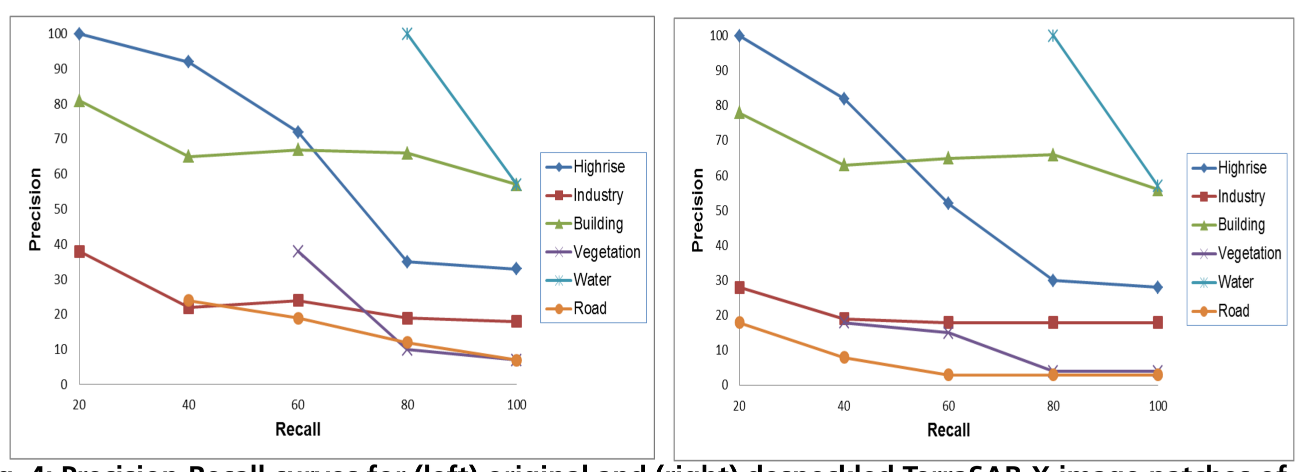

Fig. 4: Precision-Recall curves for (left) original and (right) despeckled TerraSAR-X image patches of 64x64 pixels for six land use classes.

The following table compares the precision and recall results of the first 40 retrieved patches. For High rise buildings, Residential area and Water, the precision of the land use classes is the same for original and despeckled patches showing an independence of the method regarding speckle.

Class	No. of patches	Original patches		Despeckled patches	
		Precision (%)	Recall (%)	Precision (%)	Recall (%)
High rise buildings	148	100	27.02	100	27.02
Residential area	351	80	9.88	80	9.88
Water bodies	45	100	88.88	100	88.88
Industry	93	35	15.05	25	10.75
Roads	27	25	37.03	12.5	18.51
Vegetation	16	27.5	68.75	12.5	31.25

CONCLUSION

- ✓ We could demonstrate the capacity of FCD in effectively understanding SAR images that vary in content and resolution.
- ✓ FCD can be used for urban scene recognition and gives good retrieval results for high resolution optical and SAR image patches with a size as small as 64x64 pixels.
- ✓ Our FCD experiment using a TerraSAR-X scene of Toronto City yielded excellent results for Water, Buildings and High rise buildings and unsatisfactory understanding for Roads and Industry. Despite the confusion, we obtained an overall retrieval accuracy of 62.65%.
- ✓ The retrieval performance of all classes drops significantly for patches with a size of less than 64x64 pixels.

REFERENCES

[1] Li, M., Chen, X., Li, X., Ma, B., Vitanyi, P. M. B. (2004), "The similarity metric", IEEE Transactions on Information Theory, 50(12): 3250–3264.
[2] Watanabe, T., Sugawara, K., Sugihara, H. (2002), "A new pattern representation scheme using data compression", IEEE Transactions on Pattern Analysis and Machine Intelligence 24(5): 579–590.
[3] Cerra, D. & Datcu, M. (2012), "A fast compression-based similarity measure with applications to content-based image retrieval", Journal of Visual Communication and Image Representation, 23: 293-302.
[4] Welch T, (1984), "Technique for high-performance data compression", Computer, 17(6): 8–19.

Remote Sensing Technology Institute