A SYSTEM FOR ADAPTIVE IMAGE REGISTRATION
BASED ON SUPERVISED LEARNING

Line Eikvil, Marit Holden and Ragnar B. Huseby
Norwegian Computing Center
E-mail corresponding author: Line.Eikvil@nr.no

Introduction
The study of time series of satellite images is an important task in many remote sensing applications, for instance for observing different environmental phenomena. For such applications a co-registration of the satellite images acquired at different times is important. This co-registration is often performed using a combination of manual and automatic registration techniques. However, for a multi-temporal problem where the number of images becomes large, manual correction of images is not feasible. Hence, a fully automatic procedure is desirable.

An automatic image registration process tries to determine the most accurate match between the images and will typically combine similarity metrics and matching strategies to achieve this. There exists a range of different algorithms for each step in this process, but there is generally no registration technique that works equally well for all types of images, and selection of the appropriate combination depends on the application and the image specifics. Hence, a single registration technique will generally not be sufficient when handling a range of images. For a user that needs to work on different types of time series, it would therefore be useful to have a more general tool for image registration.

Based on this we have developed a novel registration approach, which aims at offering the user such a tool. The idea is to have a system with a library of different registration algorithms, and to provide a tool on top of this which makes the system able to intelligently choose, at run-time, an appropriate algorithm based on image characteristics. This is achieved by using supervised learning, where the system learns the correspondence between image characteristics and algorithm performance. In addition, we use a region-based strategy to make the approach locally adaptive. This allows for use of different registration algorithms for different regions and permits regions not suited for registration to be discarded.

A first version of a software tool based on this approach was developed through the MIR-project for ESA (2004-2005). Through a follow-up project (Mir-Extension, 2006-2009) the tool has been further developed, validated and integrated as a component in ESA’s KEO framework. In the following we will describe the registration approach in some more detail and present results from the validation.

Methods
The approach that we have developed works as follows. In a training phase, the system learns the correspondence between image features and performance for a set of registration algorithms. At run-time, this enables the system to predict the performance for each algorithm from the image
features. By applying this strategy to image regions rather than the whole image, the approach is made locally adaptive. Features are extracted from regions, and the performance of the available algorithms is predicted for each region. For this prediction we have chosen to use a neural net. Based on the performance prediction, regions and algorithms are selected. For each selected region, a local registration can then be performed with the appropriate algorithm, and from the resulting set of local transforms, a smooth global transform can be estimated. A summary of each step in the process is given below.

Region based strategy
As the characteristics of a remote sensing image may vary across the scene, it is useful to permit the use of different algorithms for different regions. At the same time, it is desirable to be able to identify and discard regions that are not suited for registration. Hence, we have chosen to use a region-based strategy in our approach. As corresponding regions from the images to be registered need to be comparable in shape and size, we have chosen a simple and robust strategy where the images are simply divided into smaller rectangular sub-images of equal size.

Feature extraction
From each region we extract features intended to convey information about image characteristics important in an image registration process. Such important characteristics were identified as (i) the amount of information contained in an image patch, (ii) the level of difference between the master and the slave image and (iii) the relationship between the master and slave image. Features that could say something about this were selected from a set based on texture, image statistics and image differences.

Supervised learning
For learning, and later prediction, we have chosen to use a neural network. Through a separate learning phase, the system is trained on a large number of examples for which the distortion is known. This is performed by first running all registration algorithms on all examples, extracting features and computing the algorithm performance in each case. Afterwards, the neural network is trained to learn the correspondence between image characteristics and algorithm performance.

Performance prediction
The trained neural network will, during run-time, use extracted image features to automatically predict the performance for the set of available registration algorithms. This is done for each region in the image, resulting in a list of scores corresponding to the predicted performance for each algorithm for that region. These scores are then used to select both the regions that are best suited for registration and the registration algorithm to be applied to each region.

Local registration
Local image registration is performed for the regions yielding the best overall score in the performance prediction, by using the registration algorithm resulting in the highest score for each region. The algorithms are fetched from a library offering a selection of registration algorithms with different characteristics (ITK/Insight library, Ibanez et al., 2005), and the selected algorithm for each pair of regions is used to estimate the transform needed to co-register that pair. The result is a set of local transforms, one for each co-registered region.

Outlier removal
At times, some of the local transform estimations may fail. For a best possible result, these should be excluded from the estimation of the global transform. Hence, we have included a procedure based on robust statistics for detecting and removing potential erroneous local transforms prior to the global transform estimation.
**Global transform estimation**
The estimation of the final global transform is performed based on the set of local transforms that remain after the removal of the erroneous ones. First, a set of point coordinates are selected from regions in the slave image, and then the target points are determined by applying the estimated transform for each region to the selected points. This set of points is used as tie points in the final global estimation.

**Multiresolution registration**
To allow for handling of larger distortions, an optional initial step for coarse registration is included. In this step a multiresolution strategy is used, where robust performance without failure is focused rather than sub-pixel accuracy.

**Results**
As we wanted to evaluate the performance for several degrees of difficulty in terms of changes in scene appearance, we selected three classes of images with no differences in content, moderate differences and large differences. The images with no differences were mainly included to verify that the approach works properly under perfect conditions, i.e. when there are no differences in content between the images, and a mean RMS error close to zero for these cases, verified this.

Our set of images with moderate differences in content was selected from time series of Envisat ASAR, Landsat and NOAA-AVHRR images. For these image pairs the mean RMS error was less than one pixel for each pair of images. However, we do not know the exact sub-pixel accuracy of the manual co-registration used as ground truth for these images as this is difficult to verify. Quite consistent deviations from what we believed to be the correct alignment, indicate that we actually get better sub-pixel accuracy with our automatic methods, than what was the case for the initial manual co-registration.

For the image pair with large differences, the performance was more sensitive to the choice of region size and local transform. Analyzing the images in more detail, we did however find that quite large areas were covered by fragmented clouds producing a special pattern. Hence, the main problems here were probably not so much due to the differences between the images, as the large areas that were less suited for registration due to clouds and homogeneous sea regions.

**Conclusion**
The output from the project is a software tool that offers a general tool for co-registration of remote sensing images. The software uses a learning-based strategy where the system learns the relationship between image characteristics and performance for different registration algorithms. By applying this scheme to sub-images, the approach is also made locally adaptive. This enables selection of the best registration algorithm for each region in the image, while regions unsuited for registration can be discarded.

The results from our experiments have demonstrated that the adaptive approach works well and that the same approach can be applied to different types of time series with different types of contents without tedious testing and tuning. The approach is also able to handle images with at least moderate differences in contents, selecting different registration algorithms for different regions and discarding regions that are not suited for registration. In conclusion, the experiments show that both the learning-based and the region-based approaches are fruitful.