

# IMAGE ANALYSIS OF FUSED SAR AND OPTICAL IMAGES DEPLOYING OPEN SOURCE SOFTWARE LIBRARY OTB

J. D. Wegner<sup>a</sup>, J. Inglada<sup>b</sup>, C. Tison<sup>b</sup>

<sup>a</sup>Leibniz University Hanover / CNES - janwegner@stud.uni-hannover.de

<sup>b</sup>Centre National d'Etudes Spatiales (CNES) - (jordi.inglada, celine.tison)@cnes.fr

**KEY WORDS:** space resources, satellite imagery, image fusion, open source software library, feature extraction

## ABSTRACT:

Within recent years, a variety of new high resolution airborne and spaceborne sensors have been put into place. These sensors are either active (SAR) or passive (optical). Further systems will be launched in near future (TerraSARX, Pléiades, CosmoSkyMed) to even extend today's potential of submetric imagery. Hence new possibilities for the combined analysis of multi modality images of very high resolution arise. The combination of both radar and optical imagery enables the use of complementary information of the same scene and various application scenarios may be thought of. For example, one major field calling for enhanced remotely gathered information is rapid change detection in case of natural hazards. However, time critical applications call for easy to operate automated image processing and information extraction. Therefore one focus of CNES' and IPI's current research is on the automated fusion of optical and radar images as well as on information extraction of such fused data. In order to facilitate international cooperation and knowledge sharing, all algorithms that have been developed are integrated into the open source software library ORFEO Toolbox (OTB), introduced and maintained by CNES. This paper will deal with the current state of research and outline future perspectives. In order to successfully fuse optical and radar imagery, the first step is to achieve an accurate registration of both data sets. The following paragraphs will give an insight into the comprehensive modelling that has to be conducted.

## 1 INTRODUCTION

This paper will deal with the issue of automatic fine registration of high resolution images, captured with SAR and optical sensors. Since SAR and optical sensors deploy different physical principles for the acquisition of a ground scene, the image properties are different as well (both geometry and radiometry). Hence, deploying the correlation coefficient on the pixel level does not always lead to satisfying registration results since this similarity measure only takes into account similarities up to an affine transformation of the radiometry. Therefore, another strategy is proposed in this paper.

First, both radar and optical images have to be resampled to ground geometry, taking into account carefully the different viewing geometries of the sensors and their potential error sources. An external DTM is used to eliminate distortions within the images due to the terrain. The direct 3D localisation functions (image to object) and their inversion (object to image) are introduced. For the optical image the collinearity equations were implemented whereas the equations derived in (Toutin et al., 1992) were applied to the radar image. This first stage has already been accomplished and implemented into OTB. The following steps still have to be implemented and thus just a future perspective of the theoretical approach will be outlined.

The second step is to register the images on ground level. Since the physics of optical and radar sensors are completely different, standard image correlation matching techniques cannot be applied. Two approaches may be thought of to deal with this issue: adaption of correlation techniques on pixel level as well as working at the image feature level. The first approach is useful for areas in the countryside, where a lack of man made structures usually limits the number of detectable features. The algorithm of choice makes use of mutual information techniques in combination with the joint histogram of both images (Inglada and Giros, 2004). However, this algorithm was originally invented for the fusion of rather low resolution SPOT and ERS data. Hence, intensive testing still has to be conducted in order to adjust the

algorithm to high resolution imagery. The second approach will be useful in dense – urban – areas (Tupin and Roux, 2003). A reduction of the noise level is imperative in order to achieve optimum extraction results. In particular the SAR images demand careful smoothing due to severe speckle effects in built up areas. The application of the current state of the art line (Tupin et al., 1998) and point detection tools (Lopès et al., 1993) follows up.

Once the images are accurately registered, the third step consists of information extraction which will allow for automatic and semi-automatic scene interpretation. Segmentation and classification of the joint SAR and optical data will be facilitated.

The developed strategy is tested on simulated data since satellite images with a ground resolution of less than 0.5 m are not available, yet. Therefore, a set of optical (fig. 1) and SAR (fig. 2) aerial images is used. The optical image was taken by the French National Geographic Institute (IGN). The SAR image was taken in X-Band by the aerial sensor RAMSES. It was provided by courtesy of the French DGA (Defence Procurement Agency).

## 2 ORFEO TOOLBOX (OTB)

The algorithm proposed in this paper will be implemented in C++ as a part of the open source software library ORFEO Toolbox (OTB). While the a priori model-based geometric registration is already accomplished, the second step, feature-based fine registration, is ongoing and has not been integrated into OTB, yet. Furthermore, joint SAR and optical segmentation and classification tools are to be implemented in near future.

OTB is part of the ORFEO Accompaniement Programme ([http://smsc.cnes.fr/PLEIADES/FR/A\\_prog\\_accomp.htm](http://smsc.cnes.fr/PLEIADES/FR/A_prog_accomp.htm)). This programme was set up in order to prepare for the exploitation of high resolution images, derived from the new sensors introduced by ORFEO. ORFEO is a French-Italian high resolution earth observation satellite programme, incorporating the optical system Pléiades (France) and the SAR system CosmoSkyMed (Italy).



Figure 1: Optical image taken with an aerial sensor



Figure 2: SAR image taken by an aerial sensor

OTB contains a set of algorithms for the operational exploitation of the future submetric radar and optical images e.g. tridimensional aspects, change detection, texture analysis, pattern matching and optical radar complementarities. It is mainly build around the Insight Toolkit (ITK), an open source software library facilitating the analysis of medical images. The approach presented in this paper relies on already developed algorithms of OTB and adds new tools. The proposed image fusion algorithm will be integrated into OTB after severe testing.

### 3 GEOMETRIC DEFORMATION MODELING

Modelling has to be conducted in order to account for the different viewing geometries of SAR and optical sensors. Distorsions due to the terrain have to be taken care of, too. Both images are projected from sensor space to the ground as a first geometric reg-

istration step. A general a priori model-based approach was chosen, capable of rectifying the images without in depth knowledge of sensor parameters. However, current work also comprises the integration of the software library OSSIM into OTB. Once integrated, OSSIM will enable the usage of the precise sensor model for each sensor. An external DTM is used to reduce distorsions introduced by rough terrain.

In case of optical imagery, the inverse 3D collinearity equations (object to image) are used in order to project the image to the ground.

$$x = x_0 - f * \frac{r_{11} * (X - X_C) + r_{12} * (Y - Y_C) + r_{13} * (Z - Z_C)}{r_{31} * (X - X_C) + r_{32} * (Y - Y_C) + r_{33} * (Z - Z_C)} \quad (1)$$

$$y = y_0 - f * \frac{r_{21} * (X - X_C) + r_{22} * (Y - Y_C) + r_{23} * (Z - Z_C)}{r_{31} * (X - X_C) + r_{32} * (Y - Y_C) + r_{33} * (Z - Z_C)} \quad (2)$$

- $X, Y$  and  $Z$  are the ground coordinates,
- $x_0, y_0$  and  $f$  are the parameters of the interior orientation of the sensor,
- $X_C, Y_C$  and  $Z_C$  are the coordinates of the sensor's principle point,
- $r_{ij}$  are the elements of the rotation matrix at line  $i$  and column  $j$ .

Each pixel of the image on the ground is transformed via the previously displayed equations into the original image. The entire geometric modelling process is conducted in physical coordinates. Within the original image in sensor geometry, an interpolation takes place, in order to achieve the appropriate pixel value for the rectified image on the ground. In this case, a simple linear interpolation technique was chosen for the testing of the entire algorithm in order to decrease computation time. As soon as the entire registration process will be fully developed and tested, more advanced interpolation techniques, e.g. B-Spline interpolation, will be introduced. Displacement effects rest because buildings are not included in the DTM. These effect could be treated by the introduction of a DSM, provided from an external source e.g. LIDAR. However, LIDAR data is not widely available, in particular not in development countries. Further research conducted at CNES and IPI will show if an internal DSM, derived with InSAR techniques from the same data set where the SAR was derived from, will lead to sufficient results in urban areas.

The SAR image is projected to the ground deploying the inverse equations from (Toutin et al., 1992), originally derived from the collinearity equations. This approach models the residual errors still present after the image has been generated from raw data, e.g. residuals in the estimation of slant range, Doppler frequency, ephemeris and the ellipsoid. It incorporates three different models: the motion, sensor and earth models. Hence three coordinate systems are used: the image coordinate system, the intermediate coordinate system and the ground cartographic coordinate system. The first step is a transformation of the ground

coordinates into the intermediate coordinate system. It simply applies one translation and one rotation. Furthermore, the resulting coordinates of the intermediate system  $(x, y, h)$  are transformed to the image coordinates  $(p, q)$  with the equations shown below. Image coordinate  $p$  corresponds to the azimuth while  $q$  corresponds to the distance.

$$p = \frac{-y * (1 + \delta\gamma * X) + \tau * h}{P} \quad (3)$$

$$q = \frac{-X - \frac{\theta * h}{\cos\chi}}{\alpha * \left(Q + \theta * X - \frac{h}{\cos\chi}\right)} \quad (4)$$

$$X = (x - a * y) * \left(1 + \frac{h}{N_0}\right) + b * y^2 + c * x * y + \delta h * h \quad (5)$$

- $N_0$  is the normal distance between the sensor and the ellipsoid,
- $a$  a function of the non-perpendicularity of axes,
- $\alpha$  the field-of-view of an image pixel,
- $P$  a scaling factor in along-track direction,
- $Q$  a scaling factor in across-track direction,
- $\tau$  and  $\theta$  are functions of the leveling angles in along-track and across-track direction
- $b, c, \chi, \delta\gamma$  and  $\delta h$  are second order parameters.

This approach was found appropriate because it implies several desirable properties. It models the complete viewing geometry of the sensor and works with both ground control points or a DTM. In contrast to polynomial rectification methods, all variables and factors are directly related to physical values like e.g. the leveling angles of the satellite. As a matter of fact, parameters, which translate directly to physical properties of the sensor, make the equations somehow easier understandable and interpretable. Appropriate parameter values may be derived in a relatively simple way.

## 4 IMAGE REGISTRATION

### 4.1 Registration Framework

The image registration process is embedded into a registration framework, originally provided by ITK (fig. 3). The inputs to this framework are two images. While one image is called the fixed image, the other one is called the moving image. The goal is to find the optimum spatial mapping that aligns the moving image with the fixed image. The framework treats image registration as an iterative optimization problem and consists of several components: the transform component, the interpolator, the metric component and the optimizer. The transform component applies a geometric transformation to the points in fixed image space in order to map them to moving image space. The interpolator evaluates intensities in the moving image at non-grid positions. The metric component measures the similarity between the deformed moving image and the fixed image. This similarity value is maximized by the optimizer and a new set of parameters for the transformation of the following iteration is determined. The major advantage of this modular conception of the registration framework is easy compatibility of a large variety of optimizers, geometric transformations and similarity measure techniques.

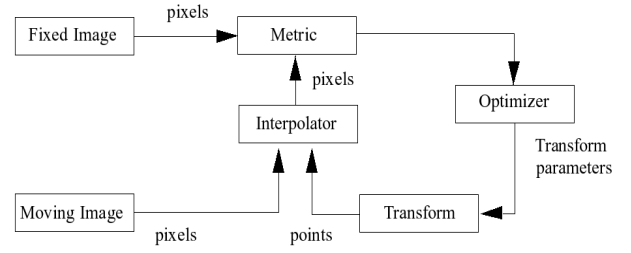


Figure 3: The ITK registration framework

### 4.2 Image Preprocessing

In order to reduce the noise level and to facilitate the extraction of lines, contours and regions, SAR and optical images first have to be processed with edge preserving smoothing filters.

For the optical image, an anisotropic diffusion filter, already existing in OTB, was chosen. This filter implements an N-dimensional version of the anisotropic diffusion equation proposed in for scalar-valued images proposed by (Perona and Malik, 1990). The anisotropic diffusion algorithm includes a variable conductance term which depends on the differential structure of the image. It is deployed in order to adjust the appropriate smoothing of edges, i.e. by limiting the strength of diffusion at edge pixels. In this case, the conductance term was chosen as a function of the gradient magnitude of the optical image.

In SAR imagery, noise reduction is even more critical than in optical imagery. In particular, the speckle effect causes severe perturbations within SAR images. Hence, edge preserving speckle reduction was conducted with the Lee filter (Lee, 1981). This filter applies a linear regression, minimizing the mean-square error in the context of a multiplicative speckle model. Although more advanced anti speckle filters already exist, the Lee filter was found sufficient in this case. In addition, the filtered SAR image was filtered with a median filter in order to eliminate the brilliant lines induced by power lines. The median filter was adjusted in such a way that it does not consider the brightest 20% of the pixels within the filter kernel for the determination of the median.

### 4.3 Feature Extraction in Optical Imagery

Feature extraction algorithms are applied after the noise level has been significantly reduced. The most promising results so far have been achieved with contour detecting algorithms. The image was filtered with the Canny-Deriche detector. The result is a binary image, displaying all detected contours as white lines (fig. 4). Its quality highly depends on the previously conducted noise reduction.

### 4.4 Feature Extraction in SAR Imagery

Three different categories of features visible in SAR imagery may be useful for the fusion of SAR and the optical imagery: linear features, regions and point targets. A point detection filter proposed by (Lopès et al., 1993) was implemented. Results are promising but still need further refinement. The best extraction results for the SAR image are achieved with linear feature extraction. Visible linear features in SAR images of urban areas are:

- Very strong reflections occur at the foundation of buildings, where the building's walls meet the ground (indicated in red box in fig. 2).

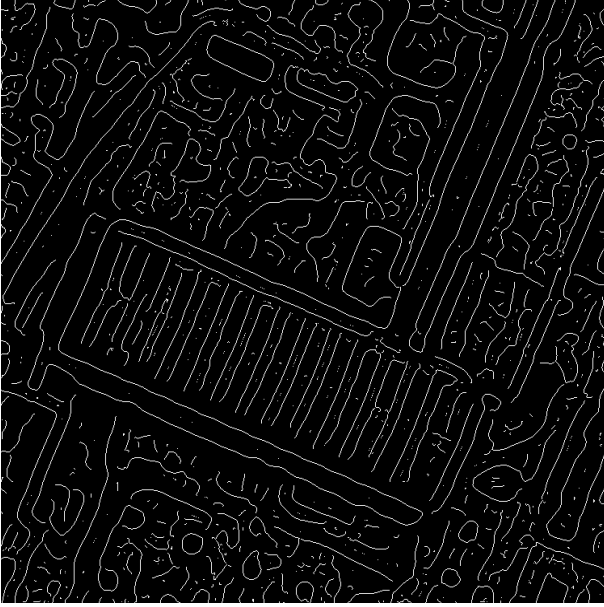


Figure 4: Contours detected with the Canny-Deriche algorithm

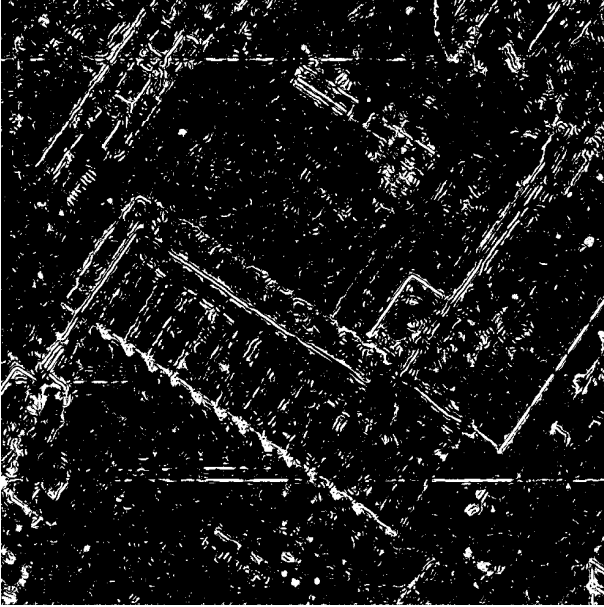


Figure 5: Contours detected with the asymmetric fusion of lines algorithm

- The regularly structured surface of factory hall roofs, if oriented perpendicularly to the sensor's flight direction, is visible in the image.
- Power lines and any ground feature charged with electricity appears bright in the SAR image.
- Very often the contrast between roofs and soil is clearly visible.

However, not all of these detectable features are useful for the image fusion process. In fact, only linear features visible in both SAR and optical imagery contribute to the fusion process in a positive sense. All extracted features detectable in only one of the images will lead to ambiguities and misalignments.

The chosen algorithm for the extraction of linear features in SAR imagery is an asymmetric fusion of lines approach. It was originally developed for the detection of road networks (Tupin et al., 1998). The result of this algorithm applied to the SAR image (after thresholding) may be seen in fig. 5. Its strategy is the fusion of the outcome of two separate line detectors D1 and D2, both of them consisting of two edge detectors. D1 is based on the ratio edge detector proposed in (Touzi et al., 1988) while D2 uses the normalized centered correlation of two pixel regions. The following two paragraphs will explain these two line detectors in further detail.

Line detector D1 consists of two ratio edge detectors, one on each side of the region of interest. The ratio edge detector is a filter with a fixed false alarm rate. It is appropriate for SAR images because speckle noise is considered as multiplicative (in contrast to optical images where noise is considered additive) and calculates the ratio of local means  $\mu$  of the regions  $i, j$  on both sides of an edge. The edge detectors response  $r_{i,j}$  is defined as

$$r_{i,j} = 1 - \min\left(\frac{\mu_i}{\mu_j}, \frac{\mu_j}{\mu_i}\right). \quad (6)$$

The line detector D1 minimizes the responses of two edge detectors. Considering three regions 1, 2 and 3, where region 2 is the probable line and 1 and 3 are the neighboring regions, the response  $r$  to the line detector D1 is

$$r = \min(r_{12}, r_{23}). \quad (7)$$

Diverse widths of region 2 are tried since the width of linear features may vary. Additionally, eight directions are tested for each pixel and only the best response is kept. A pixel is considered a line pixel, as soon as the response  $r$  exceeds a previously chosen threshold. Lowering the threshold results in more detected lines but also in a higher false-alarm rate. Therefore, this decision threshold is a compromise between the chosen false-alarm rate and the minimum detectable contrast. However, the false-alarm rate may also be decreased by increasing the regions size. The more pixels contribute to the empirical mean of the region, the less false-alarms occur. It has to be considered that larger regions increase computation time.

The second line detector D2 is based on two edge detectors. The edge detector is based on the normalized-centered cross correlation coefficient  $\rho_{ij}^2$ . An edge consists of two regions  $i$  and  $j$  with their corresponding means  $\mu$ , the pixel number inside the region  $n$  and the ratio of standard deviation and mean  $\gamma$ . In the following equation the ratio of the regions' means  $\bar{c}_{ij}$  is interpreted as the empirical contrast between the two regions  $i$  and  $j$ . The closer to one this variation coefficient is, the more homogenous is the area.

$$\rho_{ij}^2 = \frac{1}{1 + (n_i + n_j) * \frac{n_i * \gamma_i^2 * \bar{c}_{ij}^2 + n_j * \gamma_j^2}{n_i * n_j * (\bar{c}_{ij} - 1)^2}} \quad (8)$$

The advantage of this edge detector is its dependence on both the contrast between the two regions  $\bar{c}_{ij}$  and inside each region  $\gamma$ . In contrast, the ratio of means detector may be influenced by outliers contained in the regions. Again, the line detector strives for the minimum response  $\rho = \min(\rho_{12}, \rho_{23})$  of the edge detectors neighboring the potential line.

#### 4.5 Image Registration Strategy

The approach presented in this paper uses two distance images as input to the previously outlined registration framework: one distance map of extracted features of the optical image (fig. 6) and

one of the extracted features of the SAR image (fig. 7). The classical Danielsson approach is deployed to compute the distance maps. Using distance maps instead of the feature images themselves has the advantage that the features' absolute spatial positions are avoided as error sources. Absolute feature positions may cause errors due to the different viewing geometries of optical and SAR imagery. Certain building fragments may be detected quite well in both distance images (see coloured boxes in fig. 6 and fig. 7).

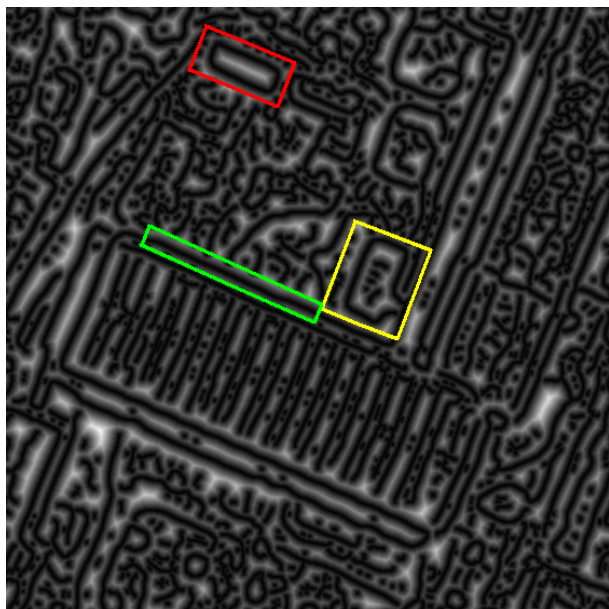


Figure 6: Danielsson distance map of the optical feature image

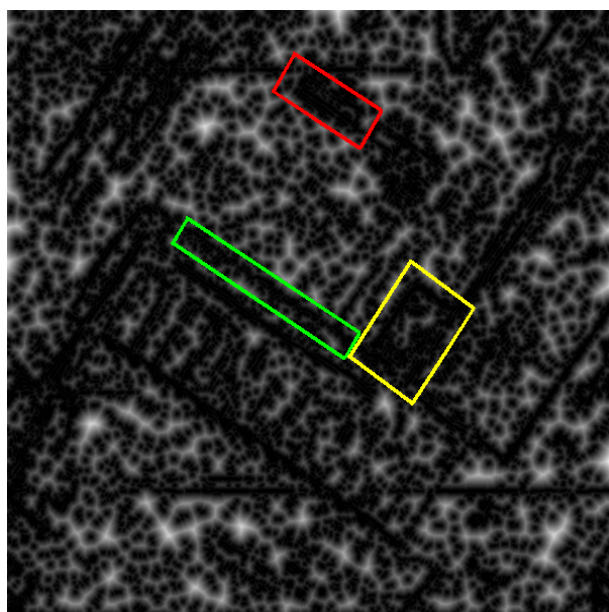


Figure 7: Danielsson distance map of the SAR feature image

## 5 CONCLUSIONS AND FUTURE WORK

The so far implemented OTB proved to be very useful in the context of image fusion. However, the implemented code represents only the very first step. Future work will comprise the implementation of the next steps necessary for a fully automatic fusion of SAR and optical images. Algorithms for image fusion in rural areas are still lacking as well as the integration of segmentation and classification techniques for the fused data. Additionally, further refinements of the parameters and algorithms still have to be made. The fact that all L-shaped brilliant features correspond to ground altitude (meeting of building foundations and soil) may be useful for an enhancement of the extraction results. The appearance of both the ground altitude and the roof altitude in the SAR image may enable building height estimation, in particular, if the optical image shows both altitudes, too (in case the sensor view comes from the same direction for optical and SAR image).

## REFERENCES

- Inglada, J. and Giros, A., 2004. On the possibility of automatic multisensor image registration. *IEEE Transactions on Geoscience and Remote Sensing* 42(10), pp. 2104 – 2120.
- Lee, J.-S., 1981. Speckle analysis and smoothing of synthetic aperture radar images. *Computer Graphics and Image Processing* 17, pp. 24 – 32.
- Lopès, A., Nezry, E., Touzi, R. and Laur, H., 1993. Structure detection and statistical adaptive speckle filtering in sar images. *International Journal of Remote Sensing* 14(9), pp. 1735 – 1758.
- Perona, P. and Malik, J., 1990. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis Machine Intelligence* 12, pp. 629 – 639.
- Toutin, T., Carbonneau, Y. and St-Laurent, L., 1992. An integrated method to rectify airborne radar imagery using dem. *Photogrammetric Engineering and Remote Sensing* 58(4), pp. 417 – 422.
- Touzi, R., Lopes, A. and Bousquet, P., 1988. A statistical and geometrical edge detector for sar images. *IEEE Trans. Geosci. Remote Sensing* 26, pp. 764 – 773.
- Tupin, F. and Roux, M., 2003. Detection of building outlines based on the fusion of sar and optical features. *ISPRS Journal of Photogrammetry and Remote Sensing* 58(1), pp. 71 – 82.
- Tupin, F., Maître, H., Mangin, J.-M., Nicolas, J.-M. and Pechersky, E., 1998. Detection of linear features in sar images: Application to road network extraction. *IEEE Transactions on Geoscience and Remote Sensing* 36(2), pp. 434 – 453.