Carl-Cranz-Gesellschaft e.V.
Gesellschaft für technisch-wissenschaftliche Weiterbildung

Sensor-Datenfusion

Gintautas Palubinskas
German Aerospace Center (DLR)
Remote Sensing Technology Institute
Oberpfaffenhofen, Germany

7-9.06.2011
Fusion: Optical/SAR
Motivation: Need for new and high quality products

- Why? Complementary nature of sources
- How?
  - Learn (understand) semantic relationships between objects
  - Develop methodology
Data fusion is a formal framework in which are expressed the means and tools for the alliance of data originating from different sources. It aims at obtaining information of greater quality; the exact definition of 'greater quality' will depend upon the application (Wald, L., 1999)

Data fusion is application dependent

Content

- Data fusion concept
- Data pre-processing
  - General fusion framework GFF
  - Orthogonal acquisition geometry
- Data fusion framework INFOFUSE
- Examples for optical/SAR classification
- Simulation of images
- Conclusions and outlook
Data Fusion: Concept

- **Sensors**
  - Source1
  - Source2
  - Source3

- **Methods**
  - Alignment (geo-coding, radiometric normalization)
  - Fusion (levels, methods, tools)

- **Applications**
  - Result1
  - Result2
  - Quality assessment

**Application**
Optical/SAR data fusion

Pre-processing

Orthogonal acquisition geometry
  to minimize displacement effects

Ortho-rectification
  DEM used
  optical data enhancement using TS-X GCPs [1]

Co-registration of optical/SAR data
  using mutual information [2]

Pan-sharpening of multi-spectral and panchromatic optical data

General Fusion Framework

De-speckling of SAR imagery

Feature extraction

Classification

Simulation


Pansharpening

Pre-processing
Orthogonal acquisition geometry
to minimize displacement effects
Ortho-rectification
DEM used
optical data enhancement using TS-X GCPs [1]
Co-registration of optical/SAR data
using mutual information [2]

Pan-sharpening of multi-spectral and panchromatic optical data
General Fusion Framework

De-speckling of SAR imagery

Feature extraction
Classification
Simulation
Pan-sharpening

Input

Low resolution image (multispectral, hyperspectral, …)
High resolution image (panchromatic, SAR, …)

Aim

Include spatial information from high resolution image while preserving spectral properties of low resolution image
General Fusion Framework GFF

Method

Interpolation \[ msi = I(ms) \]

Fusion \[ msf = F(msi, pan) \]

Histogram matching \[ msf = H(msf, ms) \]
Spectral Fusion SF

How?

Add **only** high frequency information from high resolution image to low resolution image

Solution

Fourier/spectral domain

Interpolation and fusion in one step

Spectral Fusion SF

- $\text{PBW}_{HR}/2$
- $\text{PBW}_{LR}/2$
- $f_{\text{cutoff}}$
- $\text{PBW}_{HR}/2$
Spectral Fusion SF

Fourier transform

\[ MS = \text{FFT}(ms) \]
\[ PAN = \text{FFT}(pan) \]

Zero padding and windowing

\[ MSI = \text{ZP}(W \cdot MS) \]

Frequency addition

\[ MSF = MSI + PAN \cdot HPF \]
\[ MSF = MSI + PAN \cdot (1 - LPF) \]

Inverse Fourier transform

\[ msf = \text{FFT}^{-1}(MSF) \]
Spectral Fusion SF

Signal domain

\[ msf = msi + pan \times hpf \]

where

\[ hpf = FFT^{-1}(HF) \]

or

\[ msf = msi + pan - pan \times lpf \]

WV-2 pan-sharpening (München)
TS-X Radar-sharpening (München)
Orthogonal acquisition geometry

Pre-processing

Orthogonal acquisition geometry
to minimize displacement effects

Ortho-rectification
DEM used
optical data enhancement using TS-X GCPs [1]

Co-registration of optical/SAR data
using mutual information [2]

Pan-sharpening of multi-spectral and panchromatic optical data
General Fusion Framework

De-speckling of SAR imagery

Feature extraction
Classification
Simulation
3D Object height

Optical sensor

Radar sensor

DEM

$\Delta h$

$\Delta x$

$\Delta h$

$\Delta x$
Orthogonal acquisition geometry

\[ \theta_{\text{opt}} + \theta_{\text{sar}} = 90^\circ \]

Two pairs of Optical/SAR acquisition

IKONOS2

TS-X ascending

TS-X descending

WV-1
Classification

Pre-processing
- Orthogonal acquisition geometry
to minimize displacement effects
- Ortho-rectification
  - DEM used
  - optical data enhancement using TS-X GCPs [1]
- Co-registration of optical/SAR data
  - using mutual information [2]
- Pan-sharpening of multi-spectral and panchromatic optical data
- General Fusion Framework
- De-speckling of SAR imagery

Feature extraction

Classification

Simulation
Data fusion framework INFOFUSE

1. Input data
   - Input multisensory data: Multispectral, SAR, GIS, DEM, etc.

2. Feature extraction
   - Information fission: extraction of exhaustive set of features (e.g. Gabor, Co-Occurrence)

3. Feature representation
   - Feature representation on predefined discrete domain using unsupervised clustering

4. Aggregation Method
   - Neural network, Bayesian network, Graphical model

   \[ \text{Selected classes} \]

## Data fusion requirements and solutions

<table>
<thead>
<tr>
<th>Requirements</th>
<th>Solutions</th>
</tr>
</thead>
<tbody>
<tr>
<td>• to follow consensus theory*</td>
<td>• Data fission is employed (to calculate exhaustive feature set). Quasi-full characterization of classes</td>
</tr>
<tr>
<td>• Data non-commensurability: to work with different input data (e.g. nature and statistics of optical and SAR)</td>
<td>• Feature representation on a finite predefined domain (e.g. numbers of a range). Reduce of data size</td>
</tr>
<tr>
<td>• Arbitrary number of data sources</td>
<td>• Neural Networks, Bayesian networks, or Graphical models</td>
</tr>
<tr>
<td>• Aggregation method to be able to use arbitrary number of calculated input features</td>
<td></td>
</tr>
<tr>
<td>• Acceptable complexity</td>
<td></td>
</tr>
</tbody>
</table>

## Feature aggregator selection

<table>
<thead>
<tr>
<th>Feature Aggregator</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neural Network</td>
<td>Acceptable training time</td>
<td>Overtraining</td>
</tr>
<tr>
<td></td>
<td>Variety of methods for learning</td>
<td></td>
</tr>
<tr>
<td></td>
<td>High accuracy of classification</td>
<td></td>
</tr>
<tr>
<td>Bayesian Network</td>
<td>Knowledge representation in probabilistic way</td>
<td>High training/classification time</td>
</tr>
<tr>
<td></td>
<td>Proper configuration using expert knowledge</td>
<td></td>
</tr>
<tr>
<td>Graphical model</td>
<td>Knowledge representation in probabilistic way (assuming multi-nominal distributions)</td>
<td>High training/classification time</td>
</tr>
<tr>
<td></td>
<td>Proper configuration using expert knowledge (compare to neural network)</td>
<td>No practical methods for structure learning</td>
</tr>
</tbody>
</table>
## Data characteristics

<table>
<thead>
<tr>
<th>Sensor Parameter</th>
<th>TerraSAR-X</th>
<th>WorldView-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image time (local)</td>
<td>7-Jun-2008 07:17:48</td>
<td>12-Jul-2010 10:30:17</td>
</tr>
<tr>
<td>Mode</td>
<td>Spotlight HS</td>
<td>Pan-sharpened VNIR bands</td>
</tr>
<tr>
<td>Look angle</td>
<td>49.45° Right</td>
<td>5.2° Left</td>
</tr>
<tr>
<td>Orbit</td>
<td>Descending</td>
<td>Descending</td>
</tr>
<tr>
<td>Polarization</td>
<td>Single, VV</td>
<td></td>
</tr>
<tr>
<td>Product</td>
<td>EEC</td>
<td>L2A</td>
</tr>
<tr>
<td>Resolution gr x az (m)</td>
<td>1.0 x 1.14</td>
<td>0.5 x 0.5</td>
</tr>
<tr>
<td>Pixel spacing (m)</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>Radiometric resolution</td>
<td>16 bit</td>
<td>11 bit</td>
</tr>
</tbody>
</table>
Experiment 1

Urban area classification:
- Munich city
- 6 classes

Features:
- WV-2 VNIR (4 bands)
- TS-X Texture (Co-occurrence)

Fusion strategies:
- VNIR
- VNIR+SAR Texture
Urban area classification

Input data:
- WorldView-2 multispectral (VNIR range)
- TerraSAR-X single polarization band
- Haralick texture features

INFOFUSE classification

<table>
<thead>
<tr>
<th>Label</th>
<th>Classes/subclasses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buildings</td>
<td>8 subclasses</td>
</tr>
<tr>
<td>Roads</td>
<td>2 subclasses</td>
</tr>
<tr>
<td>Water</td>
<td>1 class</td>
</tr>
<tr>
<td>Forest/Trees</td>
<td>1 class</td>
</tr>
<tr>
<td>Grass</td>
<td>1 class</td>
</tr>
<tr>
<td>Shadows</td>
<td>1 class</td>
</tr>
</tbody>
</table>
Urban area classification zoom

SAR data allow to reduce errors in building/road separation
### Confusion Matrices

#### INFOFUSE Multispectral (VNIR)
- **Overall Accuracy:** 70.95%
- **Kappa Coefficient:** 0.6059

#### INFOFUSE Multispectral (VNIR)+SAR
- **Overall Accuracy:** 97.19%
- **Kappa Coefficient:** 0.9613

<table>
<thead>
<tr>
<th>Class</th>
<th>Water</th>
<th>Grass</th>
<th>Trees</th>
<th>Buildings</th>
<th>Road</th>
<th>Shadow</th>
<th>Water</th>
<th>Grass</th>
<th>Trees</th>
<th>Buildings</th>
<th>Road</th>
<th>Shadow</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water</td>
<td>100.00</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>Water</td>
<td>100.00</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Grass</td>
<td>0</td>
<td>85.47</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>Grass</td>
<td>0</td>
<td>98.20</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Trees</td>
<td>0</td>
<td>14.53</td>
<td>100.00</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>Trees</td>
<td>0</td>
<td>1.80</td>
<td>98.90</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Buildings</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>60.05</td>
<td>42.25</td>
<td>4.72</td>
<td>Buildings</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>96.36</td>
<td>3.91</td>
</tr>
<tr>
<td>Road</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>39.95</td>
<td>57.75</td>
<td>0</td>
<td>Road</td>
<td>0</td>
<td>0</td>
<td>1.10</td>
<td>3.64</td>
<td>96.06</td>
</tr>
<tr>
<td>Shadow</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>95.28</td>
<td>Shadow</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>99.53</td>
</tr>
<tr>
<td>Total</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>Total</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
</tbody>
</table>
Experiment 2

Urban area classification:
Munich city
23 classes

Features:
WV-2 (8 bands), Red Band Texture-Gabor (6 orientations, 4 sine modulations, 2 sigma sizes)
TS-X Texture-Gabor (6 orientations, 4 sine modulations, 2 sigma sizes)
DSM (Generated from WV-2 panchromatic stereo pair)

Fusion strategies:
Multispectral, DSM (9 features),
Texture SAR, Texture optical, DSM (97 features),
Multispectral, Texture SAR, Texture optical (104 features),
Multispectral, Texture SAR, Texture optical and DSM (105 features)
Classes

23 classes were defined:

- 1. Water
- 2. Forest/Trees
- 3. Grass/Low vegetation
- 4. Bare soil
- 5. Construction site
- 6. Swimming pool
- 7. Asphalt road
- 8. Concrete road
- 9. Football field
- 10. Tennis field
- 11. Green house
- 12. Rail road
- 13. Tram line
- 14. Cemetery
- 15. Parking/car
- 16. Shadow
- 17. Red roofing tiles
- 18. Grey roofing tiles
- 19. Dark roofing tiles
- 20. Roofing concrete
- 21. Vegetation roof
- 22. Zinc roof
- 23. Roofing copper

Ground truth

- ATKIS vector map (Landesamt für Vermessung und Geoinformation)
- Material vector data (Dr. Wieke Heldens)
## Classification accuracy

<table>
<thead>
<tr>
<th>Method</th>
<th>OVA, %</th>
<th>Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum likelihood (does not follow consensus theory)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multispectral, DSM (9 features)</td>
<td>85.4841</td>
<td>0.8409</td>
</tr>
<tr>
<td>Texture, DSM (97 features)</td>
<td>60.5719</td>
<td>0.5666</td>
</tr>
<tr>
<td>Multispectral, Texture (104 features)</td>
<td>81.4288</td>
<td>0.7932</td>
</tr>
<tr>
<td>Multispectral, Texture, and DSM (105 features)</td>
<td>82.1923</td>
<td>0.8019</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Neural Network (1 hidden layer, 40 neurons for 97, 104, or 105 features, 8 neurons for 9 features)</th>
<th>OVA, %</th>
<th>Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multispectral, DSM (9 features)</td>
<td>85.6575</td>
<td>0.8426</td>
</tr>
<tr>
<td>Texture, DSM (97 features)</td>
<td>60.8644</td>
<td>0.5643</td>
</tr>
<tr>
<td>Multispectral, Texture (104 features)</td>
<td>82.6471</td>
<td>0.8076</td>
</tr>
<tr>
<td>Multispectral, Texture, and DSM (105 features)</td>
<td>87.0697</td>
<td>0.8566</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>INFOFUSE (based on Neural Network, 1 hidden layer, 40 neurons for 97, 104, or 105 features, 9 neurons for 8 features) (50 clusters for each feature)</th>
<th>OVA, %</th>
<th>Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multispectral, DSM (9 features)</td>
<td>85.1835</td>
<td>0.8360</td>
</tr>
<tr>
<td>Texture, DSM (97 features)</td>
<td>71.8699</td>
<td>0.6906</td>
</tr>
<tr>
<td>Multispectral, Texture (104 features)</td>
<td>88.8692</td>
<td>0.8768</td>
</tr>
<tr>
<td><strong>Multispectral, Texture, and DSM (105 features)</strong></td>
<td><strong>90.1092</strong></td>
<td><strong>0.8907</strong></td>
</tr>
</tbody>
</table>
Classification accuracy
Classification map 1

- water
- forest
- grass
- bare_soil
- construction
- pool
- road_asphalt
- football
- tennis
- green_house
- rail
- tram
- cemetery
- parking/car
- shadow
- concrete
- red_roofing_tiles
- concrete_roof
- vegetation_roof
- dark_roofing_tiles
- zink_roof
- roofing_copper
- grey_roofing_tiles
Classification map 2

water
forest
grass
bare_soil
construction
pool
road_asphalt
football
tennis
green_house
rail
tram
cemetery
parking/car
shadow
concrete
red_roofing_tiles
concrete_roof
vegetation_roof
dark_roofing_tiles
zink_roof
roofing_copper
grey_roofing_tiles
Classification map 3

- water
- forest
- grass
- bare_soil
- construction
- pool
- road_asphalt
- football
- tennis
- green_house
- rail
- tram
- cemetery
- parking/car
- shadow
- concrete
- red_roofing_tiles
- concrete_roof
- vegetation_roof
- dark_roofing_tiles
- zink_roof
- roofing_copper
- grey_roofing_tiles
## Influence of data sources on classification

<table>
<thead>
<tr>
<th>Confused classes</th>
<th>Sensor or feature influence for proper classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class 1</td>
<td>Class 2</td>
</tr>
<tr>
<td>Road</td>
<td>Building</td>
</tr>
<tr>
<td>Rail road/Tram road</td>
<td>Road</td>
</tr>
<tr>
<td>Rail road</td>
<td>Tram road</td>
</tr>
<tr>
<td>Bare soil</td>
<td>Construction site</td>
</tr>
<tr>
<td>Football field</td>
<td>Grass/Low vegetation</td>
</tr>
<tr>
<td>Parking/car</td>
<td>Road</td>
</tr>
<tr>
<td>Cemetery</td>
<td>Grass/Low vegetation</td>
</tr>
<tr>
<td>Green house</td>
<td>Road</td>
</tr>
<tr>
<td></td>
<td>Sensor/Feature</td>
</tr>
<tr>
<td></td>
<td>DSM</td>
</tr>
<tr>
<td></td>
<td>SAR Texture</td>
</tr>
<tr>
<td></td>
<td>SAR Texture</td>
</tr>
<tr>
<td></td>
<td>SAR Texture</td>
</tr>
<tr>
<td></td>
<td>SAR Texture, Multispectral</td>
</tr>
<tr>
<td></td>
<td>Texture on optical data</td>
</tr>
<tr>
<td></td>
<td>SAR Texture</td>
</tr>
<tr>
<td></td>
<td>SAR Texture, Multispectral</td>
</tr>
</tbody>
</table>
Pre-processing

Orthogonal acquisition geometry
to minimize displacement effects

Ortho-rectification

DEM used
optical data enhancement using TS-X GCPs [1]

Co-registration of optical/SAR data
using mutual information [2]

Pan-sharpening of multi-spectral and panchromatic optical data

General Fusion Framework

De-speckling of SAR imagery

Feature extraction

Classification

Simulation
Laser DSM

Horizontal resolution: 1m
vertical resolution: 0.1m
Simulation optical/SAR images

Simulated optical image
WV-1 image
TS-X image
Simulated SAR image

7-9.06.2011
Single building extraction from DSM

1. FK_LS
2. FK_haus
3. FK_road
4. FK_road_median
5. FK_road_median_idw
6. FK_haus_G1m
7. FK_haus_G1
8. FK_haus_G1r
Single building model

DEM + single building model
Single building simulation in optical image
Single building simulation in SAR image

Simulated SAR-Images
Line extraction in TS-X image
Line extraction in simulated image
Line extraction for single building
Extracted lines

TSX image

Simulated image

Extracted line features
Lines matching

Lines from TSX (highlighted in colour)
Lines from simulated image (white)
Scene interpretation

TSX (cyan), Simulated image (red)
Single building interpretation

TSX (cyan), reflection and shadow area of single building model (red)
Conclusions

- Data fusion concept is introduced
- Acquisition geometry should be considered
- Data preparation is very important
- General fusion framework GFF introduced
- Data fusion framework INFOFUSE introduced
  - Suitable for multi-sensor data classification
  - Separate feature processing and representation on a finite discrete domain allows to reduce storage and processor requirements
  - Classification process is not restricted by data size
- Examples for WV-2 and TS-X joint urban area classification are presented
  - SAR features help to resolve road/building confusion
  - Increase number of urban classes
  - Increase classification accuracy
Conclusions II

Problems

- Influence of shadowing (mostly self shadows) on classification accuracy and material confusion
- Time gap of sensors acquisition – possible confusion of classes but prospects for class-specific change detection
- Dependency on acquisition time for different imaging sensors

Future work

- Orthogonal acquisition model for optical and SAR data
- Employment of the model for class-specific change detection on single or multisensory data
- A proper validation methodology is required
Conclusions III

- Simulated optical image for direct and quick identification of objects in the SAR image
- Simulated SAR images present single or multiple scattering in the SAR images, useful for building recognition and reconstruction
- Orthorectification enables a direct comparison with the SAR images
- Automatic matching with SAR Image
  - Line group matching

Future work
- Change detection between DSM and SAR
References


References


References


