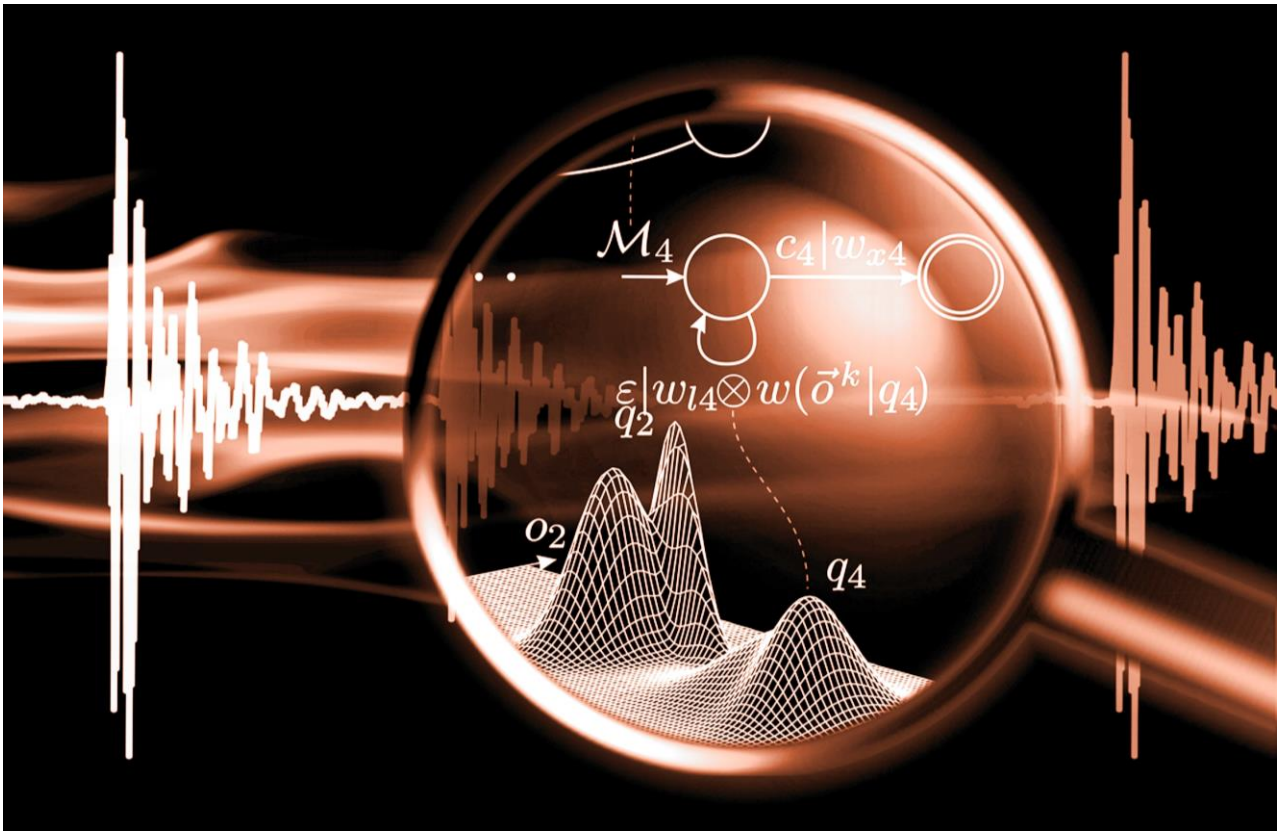
Brandenburgische
Technische Universität
Cottbus - Senftenberg



MiSIS Model Training /M1

Short Manual

Matthias Wolff
2016-05-31

1. Introduction

The MiSIS Model Training Tool provides an easy way to create model files for the MiSIS acoustic pattern recognizer. It wraps the `FEA`, `HMM`, and `REC_PACKDATA` command line tools of the UASR toolkit [UASR] into an Eclipse-based graphical user interface.

2. Installation

2.1. Prerequisites

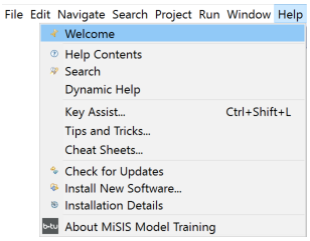
The MiSIS Model Training Tool is a Java-based 64-bit Windows application. Before installing the MiSIS Model Training Tool please install a current 64-bit Oracle Java Runtime Environment.

2.2. Installation and Start

Unpack the `MiSIS-M1-win32-x86_64.zip` archive to an arbitrary folder (in the following named `MiSIS-Folder`). You will find the following directory structure:

```
MiSIS-Folder
`- MiSIS
    | …
    | MiSIS.exe
    ` …
```

To start the application just run `MiSIS.exe` in `MiSIS-Folder/MiSIS`. On the first start, you will see a welcome screen containing an electronic version of this manual. You can go back to the welcome screen at any time by selecting Help/Welcome:



3. Basis Concepts

Acoustic pattern recognition is the process of automatically assigning a class label—e.g. “good” or “bad”—to an acoustic signal. The decision is based on acoustic models, one for each class, which are obtained by machine learning. The MiSIS Model Training Tool provides a convenient mean to compile and structure the data required for machine learning, to perform the learning procedure, and to pack model files for the MiSIS acoustic pattern recognition device.

As shown in Fig. 1, acoustic pattern recognition is performed in two major stages: feature analysis and classification.

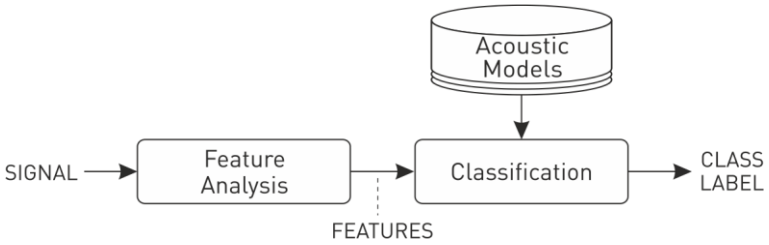


Fig. 1: Basic flowchart of an acoustic pattern recognizer.

Acoustic models are learned from a training set of pre-classified signals as illustrated by Fig. 2. During the iterative training procedure, the performance of the recognizer is evaluated using a test set of pre-classified signals. Training and test sets must be disjoint.

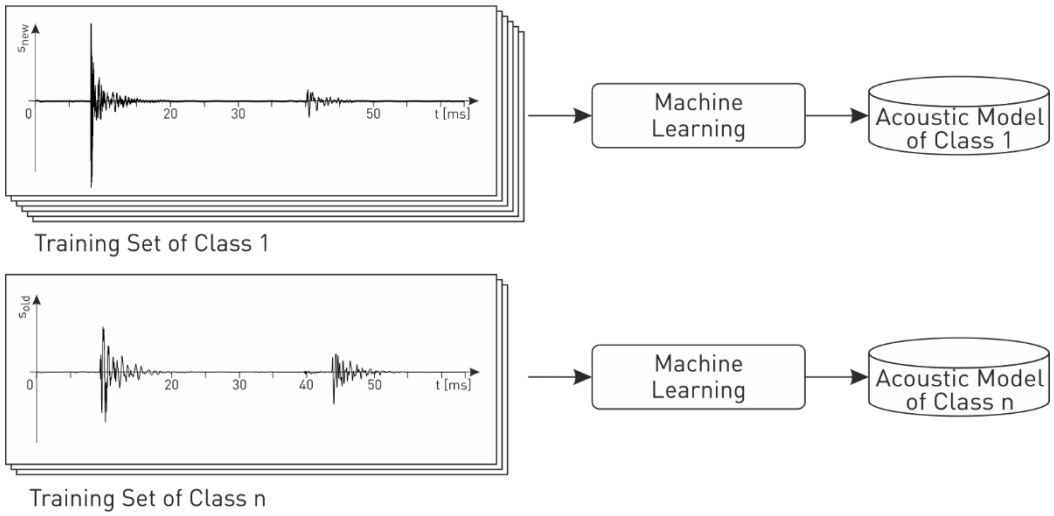


Fig. 2: Illustration of learning acoustic models from pre-classified training sets of signals.

To learn more about acoustic pattern recognition, please read [Tsc12] or [HW15b].

4. Getting Started

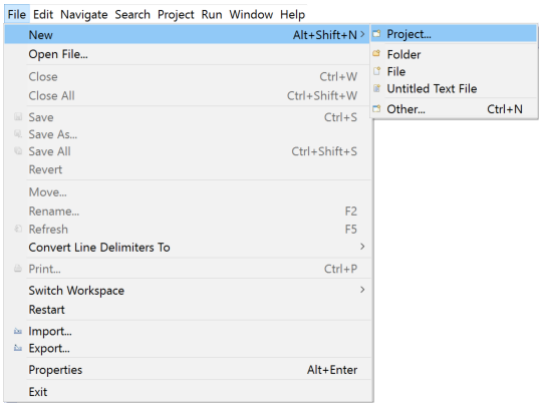
4.1. Decide on Classes and Prepare Training and Test Data

First you need to decide on the classes the acoustic pattern recognizer is going to distinguish. A very simple recognizer may discern “good” and “bad” signals. Then you need to provide signal examples for each of the classes. You will need at least 100 signal examples per class for the machine learning of acoustic models and additional 30 signal examples for testing. However, depending on the difficulty of the classification task you may need much more training examples to achieve a satisfactory classifier performance.

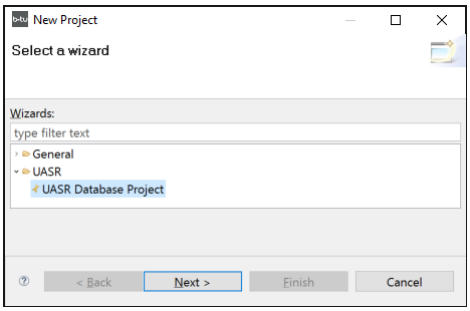
4.2. Create a Training Database

To create model files, you need to setup a database [UASRb] using the “New UASR Database” wizard:

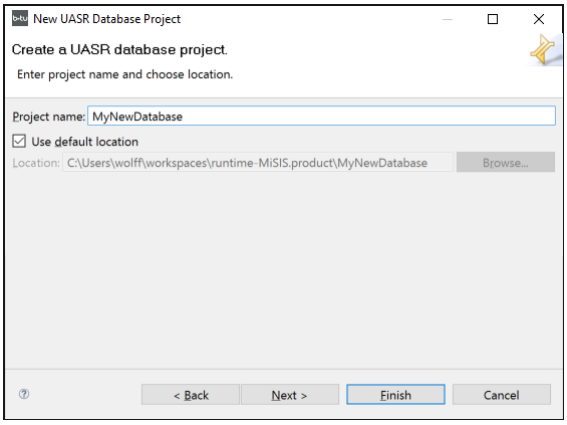
1. Select File/New/Project…



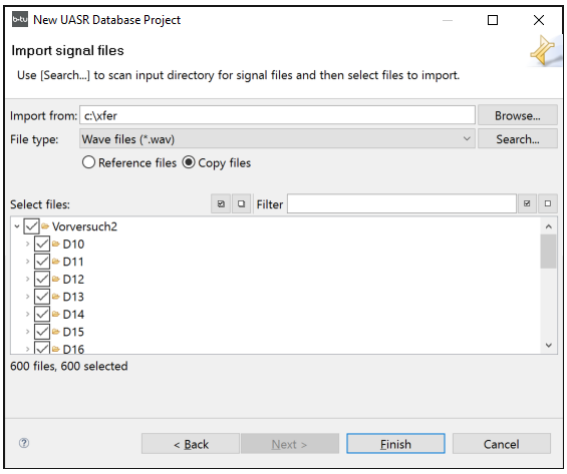
2. Select “UASR/UASR Database Project” and click [Next >].



3. Enter a project name and click [Next >].

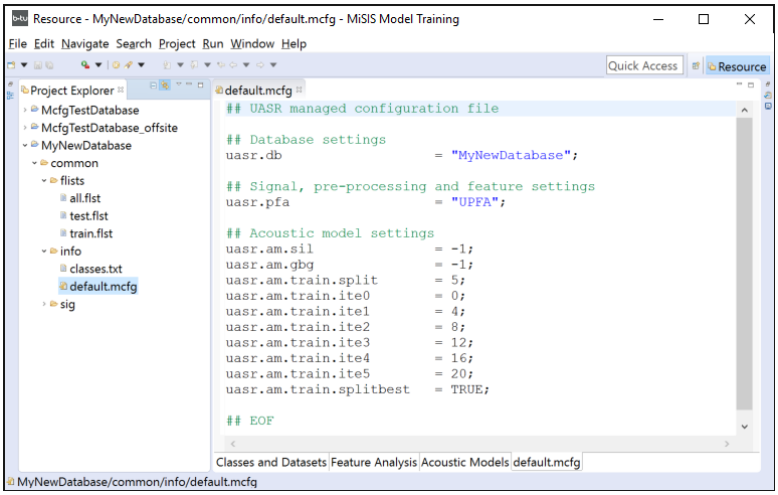


4. Select signal files to import into the database and click [Finish].



The wizard will create a new project in the workspace and initialize the default directory structure of a UASR [UASRb] database comprising:

- file lists for the training set (`common/flists/train.flst`), the test set (`common/flists/test.flst`), and for the union of all data sets (`common/flists/all.flst`),
- a classes definition file (`common/info/classes.txt`),
- a UASR configuration file [UASRc] (`common/info/default.mcfg`), and,
- if you chose to copy signal files into the project, a folder containing the imported signal files (`common/sig`).



The wizard will automatically open the configuration file. The editor includes four tabs:

- **Classes and Datasets**
  graphical editor for the classes definition file (`classes.txt`) and the file lists (`*.flst`),
- **Feature Analysis**
  graphical editor for the feature analysis settings and feature analysis tools,
- **Acoustic Models**
  graphical editor for the model learning settings, model learning and packaging tools, and
- **<filename>.mcfg**
  text editor for the configuration file.

4.3. Setup the Database

Work through the graphical editor tabs from left to right. Define classes, datasets and adjust settings. Section 5 contains detailed instruction for the single pages. These instructions can also be found in a context help section on the right-hand side of each editor page.

4.4. Run Feature Analysis, Model Learning and Model Packing

Feature analysis, model learning and model packing must be run in this order. You find respective buttons on the tool areas of the “Feature Analysis” and “Model Learning” pages of the MCFG editor. For details, see section 5 below.

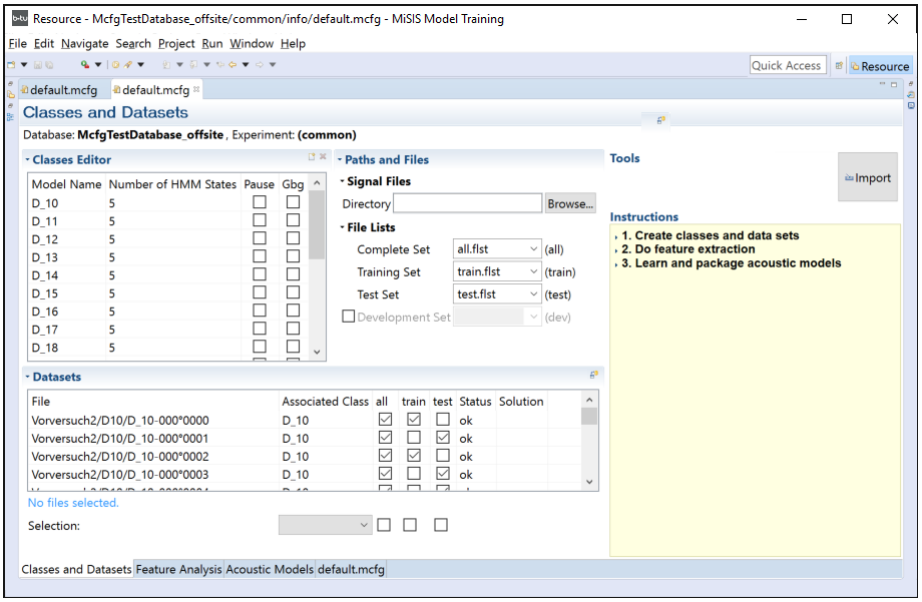
5. MCFG Editor

As outlined above, each training database contains at least one UASR configuration file [UASRc] which controls the model training. You find the configuration file (usually name `default.mcfg`) in the `common/info` folder of your training database:

```
<mydatabase>
  `- common
     |- flists
     |- info
     |   |- classes.txt
     |   `- default.mcfg
     `- sig
```

Double-click this file in the project explorer to open the configuration file in the graphical MCFG editor. The editor contains four pages which are explained in the following.

5.1. Classes and Datasets Page



Classes Editor

Each signal file is assigned a class (e.g. “good”, “fair”, “bad”, etc.) Add or remove signal classes by the and buttons, give each class a unique name and select the number of HMM states (typically 3, reasonable values 1...10). One model can be declared as “pause” (silence model) and one model as “garbage” (noise model).

Important: class names must not contain spaces or any special characters except the underscore.

Paths and Files

Specify the signal files folder and the file lists of the datasets. Normally you should not need to change these settings.

Datasets

Associate the correct classes to each signal file[1] and specify if the signal file is member of the training set (used for model learning) or the test set (used for model testing). Each signal is member of the complete set (“all”).

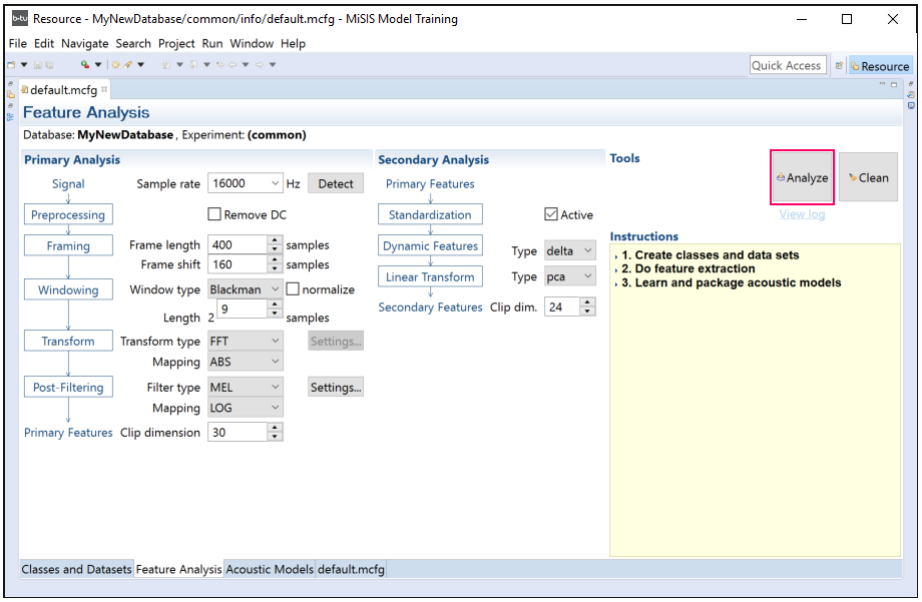
Note: A signal can be either in the training or in the test set, but never in both sets!

Tools

• Import: Import new signal files into the project.

[1] This information is needed for learning and testing

5.2. Feature Analysis Page



Primary Analysis

Settings for the short-term signal transform based on FFT, Wavelets, or the mel-generalized cepstrum. Tool tips contain information for the settings. For more information on the signal transform see documentation of UASR scripts util/fea.itp and util/ufea.itp in the UASR manual [UASR].

Secondary Analysis

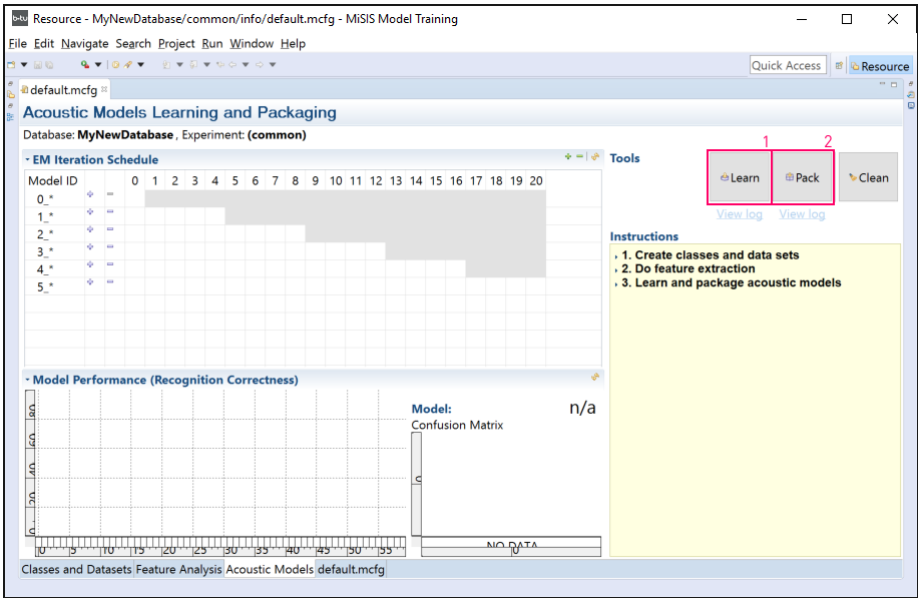
Settings for dynamic features and statistical feature dimension reduction. Tool tips contain information for the settings. See section 8.2 of [HW15b] for details on the secondary transform.

Tools

• Analyze: Compute feature files. There is one feature file per signal stored in `common/fea/<fea_type>`. You can open and view feature files from the navigator view. Note: You must re-run the analysis after you have changed settings on this page!

• Clean: Delete feature and/or log files.

5.3. Acoustic Models Page



EM Iteration Schedule

Acoustic model[2] learning is an iterative process of splitting Gaussian probability density functions (lines of schedule) and step-wise parameter optimization (columns of schedule). Learning progress is indicated by green cells. Available models are indicated by the icon . Click on the icon symbol to see model performance details.

• Lines: The more lines in the schedule, the better the recognizer performance and the higher the demand for training data. Use the and buttons in the section tool bar to add or delete lines.

• Columns: The higher the line number, the more optimizations steps are required. Use the and buttons at the beginning of the line to adjust.

Model Performance

• Left: Graph of the test performance (percentage of correctly recognized signal files in test set) of the iteratively improved models with confidence interval. Click on a data point to see performance details.

• Right: Test performance and class confusion matrix (lines: automatically recognized class, columns: correct class) of the selected acoustic model.

Tools

• Learn: Run acoustic model learning. Note: May take very long depending on the size of the datasets and the iteration schedule!

• Pack: Pack one of the learned models (see database symbols in schedule) for usage with the MiSIS hardware recognizer.

• Clean: Delete model and/or log files.

[2] MiSIS uses Hidden Markov Models (HMM) as acoustic models.

5.4. MCFG Text Editor Page

On this page you can edit the actual configuration file as text. See [UASRc] for a documentation of important UASR configuration keys.

6. Advanced Topics

6.1. Installing Eclipse Plug-Ins in MiSIS Model Training

As MiSIS Model Training is an Eclipse product, you can install any Eclipse plug-ins you may want to use. Just select "Help|Install New Software…" and enter http://download.eclipse.org/releases/mars/ (or any other update site) under "Work with". Then select and install plug-ins as usual.

6.2. Installing MiSIS Model Training as an Eclipse-Plug-In

On the delivered medium you find a folder named `UpdateSite` besides the product zip package. Select "Help|Install New Software…" in Eclipse, click the `[Add…]` button in the top-right corner of the "Install" dialog and then the `[Local…]` button in the "Add Repository" dialog being opened. Select the location of the local update site folder an install the two plug-ins contained therein.

Note: You may have to uncheck the "Group items by category" option in the "Install" dialog in order to see the content of the local update site.

References

[UASR] UASR – Unified Approach to Speech Synthesis and Recognition, Manual. Online: https://rawgit.com/matthias-wolff/UASR/master/manual/index.html.

[UASRb] UASR – Unified Approach to Speech Synthesis and Recognition, Manual. Section "Organization of UASR Data Files". Online: https://rawgit.com/matthias-wolff/UASR/master/manual/index.html?automatic/util/cfg.itp.html;index:datadirs.

[UASRc] UASR – Unified Approach to Speech Synthesis and Recognition, Manual. Section "UASR Configuration Files". Online: https://rawgit.com/matthias-wolff/UASR/master/manual/index.html?automatic/util/cfg.itp.html;index:cfgf.

[Tsc12] C. Tschöpe: Akustische zerstörungsfreie Prüfung mit Hidden-Markov-Modellen. Dresden: TUDpress, Studientexte zur Sprachkommunikation, Bd. 60, 2012.

[HW15] R. Hoffmann, M. Wolff: Intelligente Signalverarbeitung 1: Signalanalyse, 2. Auflage. Springer Vieweg, 2015. ISBN 978-3662453223.

[HW15b] R. Hoffmann, M. Wolff: Intelligente Signalverarbeitung 2: Signalerkennung, 2. Auflage. Springer Vieweg, 2015. ISBN 978-3-662-46725-1.