Control Engineering
Hof University
Alfons Goppel Platz 1
95028 Hof

Hof University
University of Applied Sciences

Lab Sessions Condition Monitoring and Predictive Maintenance

Lab PMCM01
Data Cleaning, Feature Reduction

Prof. Dr.-Ing. V.Plenk

Version 0.3

1 Project "INTEUM"

Hof University's Institute of Information Systems (iisys) is a partner in the project INTEUM. The grant application for the project names the following project goal:

> *Medium-sized manufacturing companies often have a heterogeneous machine park consisting of existing machines, for the monitoring of which there are usually only very product- and process-specific retrofit solutions. Within the framework of the INTEUM project, a system is to be developed that makes it possible to predict wear and tear (predictive maintenance) and product quality independently of the machine and across all processes. The system works with structure-borne sound data, which can cover a very high frequency range and thus a large number of different machine types. By enhancing the sensor data with additional meta-information (ERP, product and process data), the robustness of the developed plug-and-play solution is increased in productive use.*

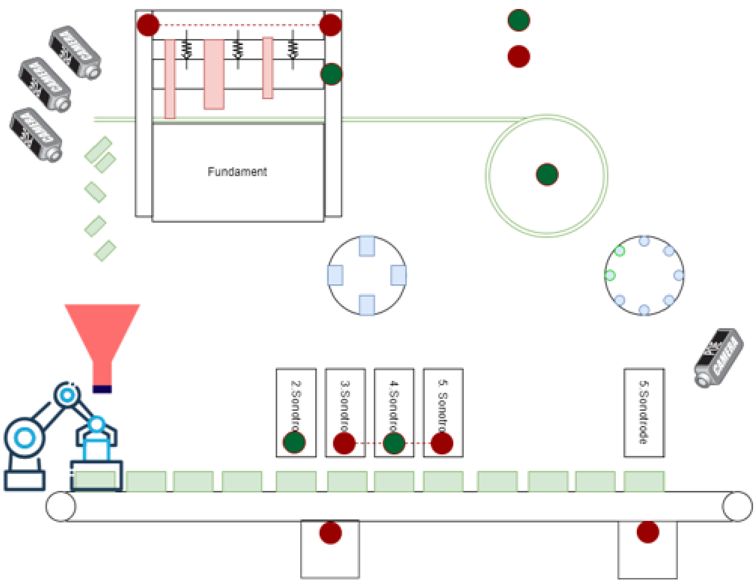

Figure 1.1: Schematic of the production machine

In this lab session you will work with data from a complex production machine as shown in figure 1.1. Copper sheet is unrolled from a coil and punched out (top). The resulting parts are inspected by a camera system. Good parts are assembled and welded on a plastic carrier. This partly finished product is again inspected by a camera system. Four carriers are welded together. The resulting product is inspected by a camera system.

In addition to the video cameras already present in the machine a high frequency structure-borne sound pickup and some "sensor beacons" were retrofitted to the machine.

Figure 1.2 shows the features/dimensions measured by the camera system. With the exception of the characteristics marked with numbers 1 and 2, these are standard length and width measurements. The dimension lines of characteristics 1 and 2 delimit areas in which the minimum and maximum dimensions of the corresponding dimension (0.604 or 0.74) are recorded.

The ridge on the left side of the illustration serves as a reference for the symmetry line of the SKV contours (`Sym␣aussen␣SKV#1`, `Sym␣Einstich␣SKV#1`).

The two symmetry features are measured in the drawn measuring areas in relation to the symmetry line. In the case of the feature `Licht-Einstich␣frei␣SKV#1`, only the drawn-in area (elongated rectangle) is measured for the absence or presence of punching waste.

The sound data is sampled on two channels and typically shows three distinct phases of sound emission per punching-cycle (figure 1.3).

The sensor beacons provide "slow" data – e.g. the temperature (which is not expected to change from part to part. In the future this data will be used to detect a change of the copper sheet/coil.

Our project partner has performed data fusion for all the different sensors and supplies us with a CSV-file containing the following features:

1. Timestamps:
   The primary timestamp is `timestamp_skv`. This refers to the first punching step for one finished part in the timezone UTC. [1]

2. Sound features:
   The sound-processing software calculates different features either for the whole duration of one punching-cycle or for each of the three phases.

   The features are named according to the following schema:

   - `ST_CH1_P_F0`
     - `ST` for punching system (german: Stanze)
     - `CH1` for channel 1: this sensor is on the side of the punching tool that is closer to the output coil.
     - `P` for the entire punching cycle

[1] The other time-related features originate from the different data sources and can be ignored for now: `timediff_skv`, `timestamp_sensor`, `timediff_sensor`, `SynchroDiff`, `timediff_sensor`, `ST_CH1_P0_Time`, `ST_CH1_P_Time`, `ST_CH1_P1_Time`, `ST_CH2_P_Time`, `ST_CH2_P0_Time`, `ST_CH2_P1_Time`, `ST_CH2_P2_Time`, `timestamp` (3 columns with identical names; the timestamp refers to the following feature).

    - F0 for feature 0, 1, 2 . . .

      The features only use anonymous names to protect the intellectual property of our partner.

  - ST_CH1_P0_F0
    - P0 First process step
    - F0 for feature 0, 1, 2 . . .
  - ST_CH1_P1_F0
    - P1 Second process step
    - F0 for feature 0, 1, 2 . . .
  - ST_CH1_P2_F0
    - P2 Third process step
    - F0 for feature 0, 1, 2 . . .
  - For channel 2, which is on the quality monitoring side of the camera, the designations are analogous.

3. The cameras measure different dimensions and return these dimensions as features. The features are named in the schema 4.713␣SKV#1, 4.713␣SKV#1 . . .

   The numerical value corresponds to one of the dimensions shown in figure 1.2 (and also indicates the expected value for this feature). #1 indicates which of the four cameras has measured the feature.

4. xxx_RH refers to the relative humidity, xxx_PRESS to the air pressure and xxx_TEMP to the air temperature.

5. Tool data from Scherdel maintenance:

   - pieces is a counter for the number of punch-actions made with the current tool. This counter is reset once the tool undergoes maintenance.
   - tool␣io a flag indicating that the tool requires maintenance (soon). This flag becomes goes to state "1" up to 48 hours before NIO event. The data preparation process currently requires this (over)large window.

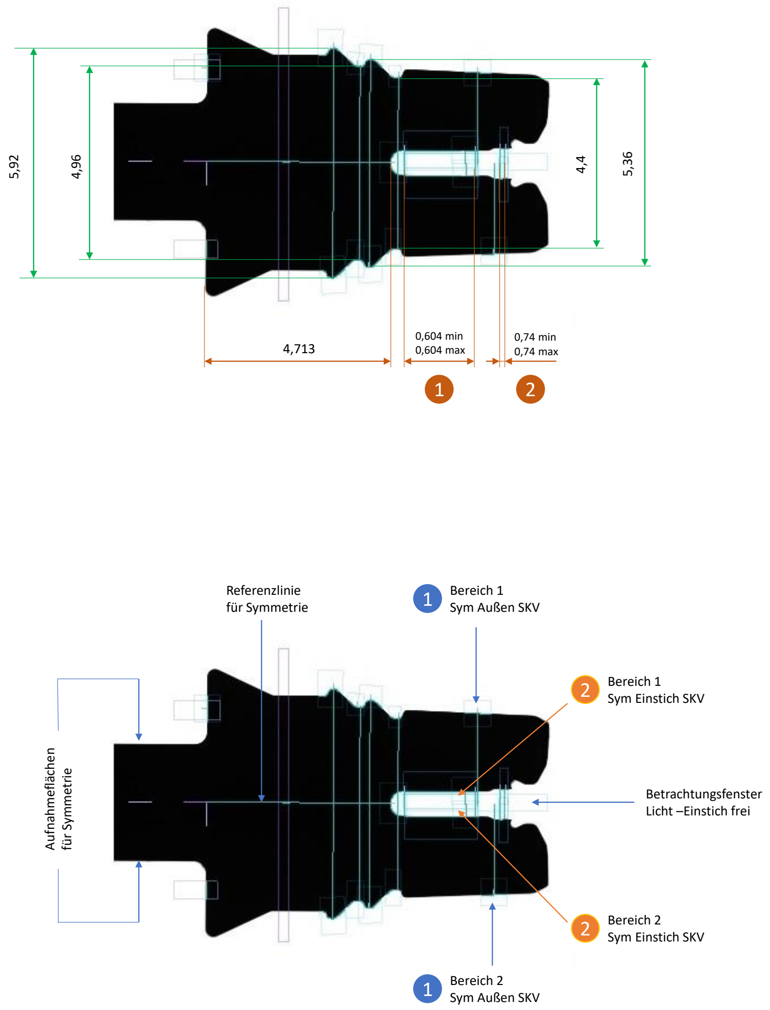

Figure 1.2: Dimensions checked by the camera systems

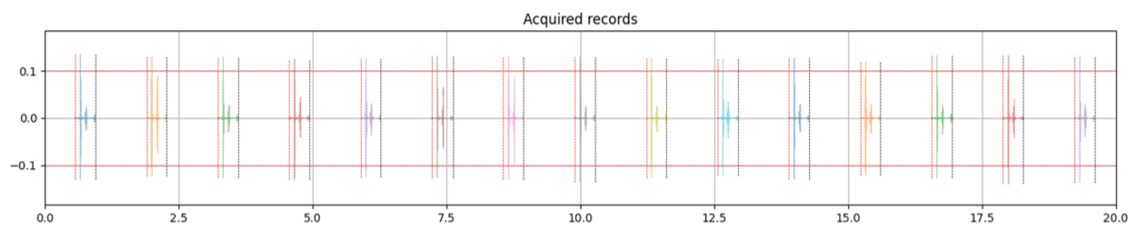

Figure 1.3: Typical time series for structure-borne sound (channel 1)

Task 1: Upload the CSV-file
Upload the CSV-File to your instance of Jupyter-Labs on the iisys-server!

Task 2: Read the CSV-file
Create a Jupyter notebook to load the CSV-File .
Look at the column-names and check whether the features were loaded correctly!

Hints:

- You might want to create a separate notebook for each of the following tasks.
- Feel free to work in groups of two or three students.

Task 3: Check Timestamps
Firstly check whether the timestamps are equidistant – reflecting a constant rate of production – or whether there are "pause"-intervals!
If you find such intervals try to formulate hypotheses to explain these "pauses"!

Repeat the following tasks for each feature set – i.e. for all sound feature groups `ST_CH1_P_*`, `ST_CH1_P0_*`, `ST_CH1_P1_*`, `ST_CH1_P2_*`, `ST_CH2_P_*`, `ST_CH2_P0_*`, `ST_CH2_P1_*`, `ST_CH2_P2_*` and all camera features `*␣*SKV#*`!

Task 4: Drop rows with missing features, normalize
Check whether there are rows with missing features in the current feature group! If so, drop the rows and document how many rows you had to drop.
Normalize the remaining rows column by column!

Task 5: PCA and Visualization
Perform a PCA on the data and visualize the data!
If you find apparent clusters try to formulate and visualize hyptheses to explain the clusters! (E.g: before / after "pause"-intervals, younger vs. older data, . . .)
Repeat the process with different sub-sets of the feature-list to identify features that yield plausible results!