Continental Automotive Selected *STATISTICA Data Miner* to Classify Image Data for Quality Monitoring

KEY POINTS

Business
Continental Automotive

Overview
- Continental Automotive Group is present in more than 170 locations worldwide
- Develops innovative and efficient system solutions for vehicle powertrains
- Performs testing on 100% of production steps to ensure quality

Challenge
Needed to develop an automated process that would simulate manual inspection of the physical characteristics of semiconductor solder contacts—to identify voids, weld width, flow pattern, and other parameters whose electronic measurements would be insufficient to assess quality within an acceptable margin of error.

StatSoft Solution
STATISTICA Data Miner

Results
- Has produced "significant savings" over manual review
- Has reduced incorrect quality assessments to 79 ppm

"Thanks to the wide variety of data mining methods in STATISTICA, we have been able to find answers to our questions. STATISTICA's flexibility is well-suited to mapping our manufacturing processes quickly and effectively."

— Dr. Udo Kreißig
Technology Development Manager, Continental Automotive

BACKGROUND

Continental Automotive produces diesel and petrol fuel injectors in which active semiconductor elements, so-called piezo-actuators, are used to control injections. Wires must be inspected, then connected to these semiconductor elements with solder, which is controlled by an automated process. Approximately 700 million soldering connections are made each year.

CHALLENGE

In semiconductor manufacturing, it is necessary for qualitative purposes to perform testing at 100% of the production steps. Standard measurements for current, voltage, and resistance are commonplace, but some quality-related characteristics cannot be sufficiently assessed by the measurement of electrical or geometrical parameters. Especially in the area of contact formation between two components, the human eye is currently essential. For a purely mechanical bond between two devices, the electric measurement of quality parameters is generally sufficient. But in the automotive industry, durability of all parts is imperative, so it is necessary to enhance the quality check beyond standard parameters. Through manual visual assessment of welded and brazed joints, it is possible to assess their quality (based on voids, weld width, flow pattern, etc.). However, for high-volume applications, this manual visual control is very labor intensive and is associated with a known error rate of 2%. Given these conditions, the task was to develop an automated control solution that could incorporate manual assessments.

HOW *STATISTICA* HELPED

STATISTICA Data Miner offers a wide range of classification methods that can be used to automate almost any classification task. For our challenge, we selected *Support Vector Machines* (SVM), a powerful data mining algorithm that can be used for both regression and classification tasks. *STATISTICA*'s implementation of support vector machines is very flexible and allowed us to achieve the desired accuracy without overfitting the data.

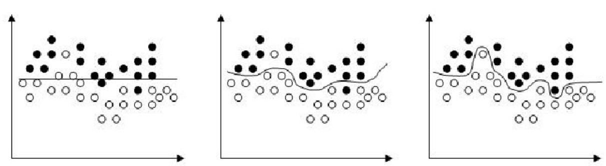

Figure 1: Separation of SVM data using two-dimensional example

Figure 1 (above) displays the decision boundaries generated by different SVMs on a simple two-dimensional example. Notice that as the decision boundaries become more highly nonlinear, an SVM can more accurately classify the different observations, at least on the training data.

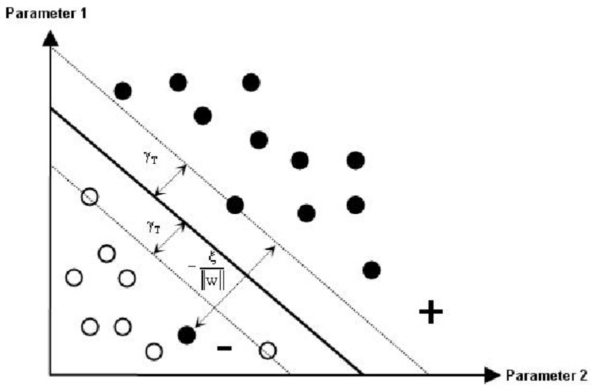

Figure 2: Linear separation of a two-dimensional data set with a misclassification

It was important to build an SVM that captured the pattern or signal in the data without modeling the noise. In general, if a classification model captures not only the signal but also the noise, then the model will struggle to correctly classify new cases it has not yet seen before. *STATISTICA* automates the search process of finding the best SVM, one that would not overfit the training data and would still generalize well to new data.

We were pleased that *STATISTICA* is not just a black box. The software allowed us to directly manipulate algorithm parameters in order to find the best model. By targeted variations of those parameters, we were able to build a successful model.

Model Implementation

The calculations are based on individual images of solder joints that need to be converted into a defined number matrix. In order to enable digital processing efficiently and minimize the learning curve required, the image sizes should be sensibly reduced, even by the removal of areas deemed unimportant. Once images have been reduced appropriately, the quality of single wire soldering quality can be assessed.

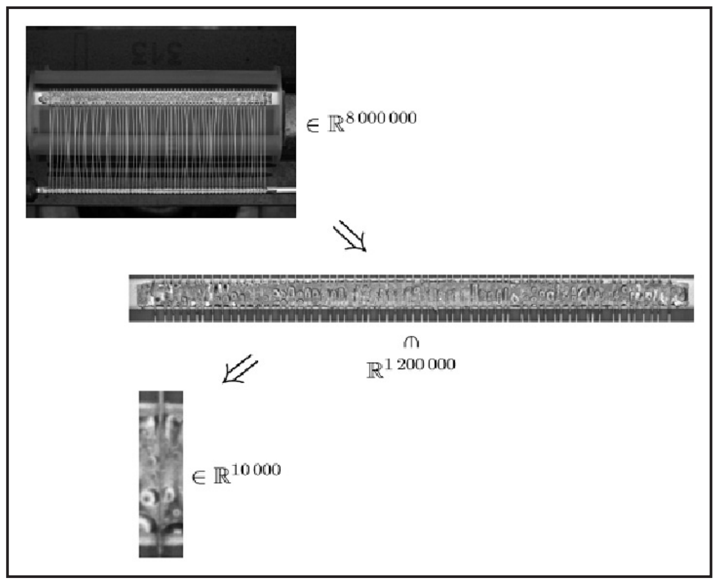

Figure 3: Minimizing referential alignment and reducing the dimensions of images to be analyzed.

To transform the individual image sections into numeric values, the gray level of each pixel gets translated, in order, from top left to bottom right, thus producing a number matrix having one row per image. The resulting number of columns is defined by the number of pixels to be evaluated, in this case 10,000. During evaluation of these high-resolution images via database analysis, it was assumed that the distribution of gray levels in the image could explicitly correlate to an expression of solder quality.

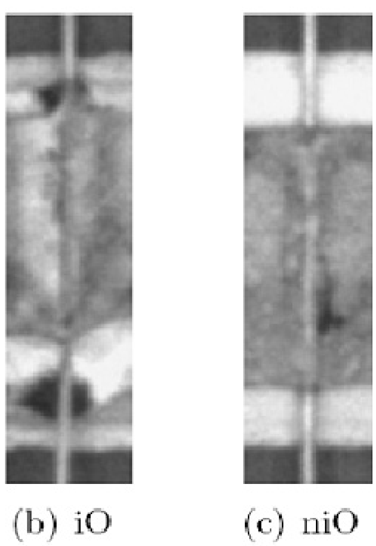

Figure 4: 10,000-pixel images of two soldered wires (b = Good, c = Bad). Question: has the solder flowed, or are the components only mechanically connected?

The mathematical conversion of visual data into numeric data was achieved with *STATISTICA*. Many visual images were evaluated beforehand for quality-relevant features and classified with a separate values column indicating good = 1 or bad = 0.

While creating the training model, care was taken to ensure that the ratio of good to bad was balanced (in order to offset the impact of the current quality situation) through the use of *STATISTICA Data Miner*'s SVM functionality, which was carefully cross-validated. Since the process is highly dependent on the established parameters Sigma and C, and these in turn are adapted to each specific application area, a parameter variation becomes very useful. This takes a lot of time and computing power to complete due to the high complexity in the modeling and the determination of the ideal parameters for the separation function.

RESULTS

To assess the quality of the SVM algorithm's output, ROC curves were used. This image shows the quality of the process. The farther away the curve from the lower left to upper right diagonal runs, the better the separation efficiency of the underlying algorithm.

In this case, 10% of the good parts are classified incorrectly to ensure a sufficiently large safety margin in relation to any slip of bad components; this is performed according to the required safety procedures regarding misclassifications. Any bad components are subjected to a second, manual visual inspection. Since *STATISTICA*'s SVM offers the opportunity to constantly adapt from new errors, the partition function is trained with new fault patterns continuously.

The model was integrated directly into the production line and automatically started analyzing solder joints. This has led to significant savings in resources over manual review. Long-term studies to verify the system's assessments are currently underway and, so far, have demonstrated that incorrect classifications are currently at only 79 ppm, a value that is excellent for a statistical model.

Dr. Udo Kreißig
Technology Development Manager, Piezo-Actuators
Continental Automotive

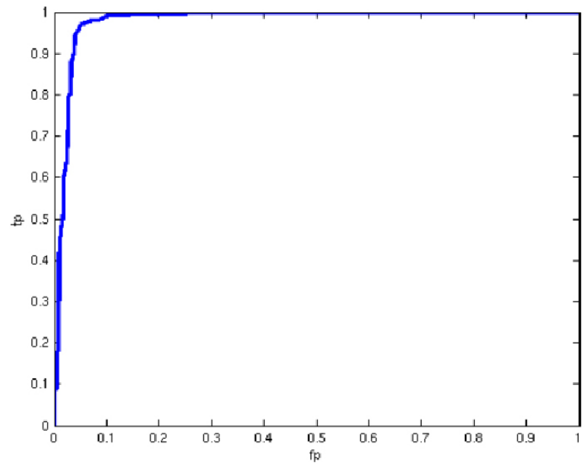

Figure 5: ROC curve of the SVM separation function.
Sigma = 2, C = 1 in this object space → test error 4.01%

Headquarters: StatSoft, Inc. • 2300 East 14 Street • Tulsa, OK 74104 • USA • +1 (918) 749-1119 • info@statsoft.com • www.statsoft.com

Australia: StatSoft Pacific Pty Ltd.
Brazil: StatSoft South America Ltda.
Bulgaria: StatSoft Bulgaria Ltd.
Chile: StatSoft South America Ltda.
China: StatSoft China
Czech Rep.: StatSoft Czech Rep. s.r.o.
France: StatSoft France
Germany: StatSoft GmbH
Hungary: StatSoft Hungary Ltd.
India: StatSoft India Pvt. Ltd.
Israel: StatSoft Israel Ltd.
Italy: StatSoft Italia srl
Japan: StatSoft Japan Inc.
Korea: StatSoft South Korea
Netherlands: StatSoft Benelux
Norway: StatSoft Norway AS
Poland: StatSoft Polska Sp. z o.o.
Portugal: StatSoft Iberica Lda
Russia: StatSoft Russia
S. Africa: StatSoft S. Africa (Pty) Ltd.
Spain: StatSoft Iberica Lda
Sweden: StatSoft Scandinavia AB
Taiwan: StatSoft Taiwan
UAE/Egypt: StatSoft Middle East
United Kingdom: StatSoft Ltd.