Advances in Knowledge-Based Technologies

Proceedings of the
Master and PhD Seminar
Summer term 2010, part 2

Softwarepark Hagenberg
SCCH, Room 0/2
6 July 2010

Software Competence Center Hagenberg
Softwarepark 21
A-4232 Hagenberg
Tel. +43 7236 3343 800
Fax +43 7236 3343 888
www.scch.at

Fuzzy Logic Laboratorium Linz
Softwarepark 21
A-4232 Hagenberg
Tel. +43 7236 3343 431
Fax +43 7236 3343 434
www.flll.jku.at

Program

Session 1. Chair: Roland Richter

9:00	Henrike Stephani: Comparing Feature Selection and Choice of Distances in Hierarchical Clustering
9:30	Verena Schlager: Coherence Probe Microscopy 3D Image Processing – Structural Analysis
10:00	Wolfgang Heidl: Classifier-based analysis of visual inspection: Significance subsumes stability

Comparing Feature Selection Methods and Choice of Distances in Hierarchical Clustering

Henrike Stephani

henrike.stephani@itwm.fraunhofer.de

Abstract

Comparing results of a hierarchical clustering with a known ground truth is a non-trivial problem as the result of such a clustering is not just a partition of the input data in classes but a whole class hierarchy of arbitrary coarseness. The most popular method for such a comparison is the so called FScore (Larsen and Aone 1999). Here, the dendrogram is used to find the branch that contains most members of each original class in relation to members of other classes. The problem with this method is that it generally does not provide a class partition that would be found automatically. Automatic extractions are usually based on the distance levels within the dendrogram. We therefore present a cluster evaluation scheme that finds the distance levels that provide the most correspondences with the ground truth and validate the cluster quality by this number of correspondences.

With this method we are able to evaluate the appropriateness of different feature sets as well as different clustering parameters such as the chosen distance measure and the chosen linkage function. In this work we compare three different types of features sets, namely the full spectra, wavelet coefficients and certain modes of the Empirical Mode Decomposition (EMD). These sets are evaluated with three different linkage functions – complete, single and average linkage – and two different distance measures – Euclidean distance and cosine distance.

Throughout all feature sets it becomes clear that the cosine distance yields better results than the Euclidean distance while the question which linkage function to choose is not as easily answered as it strongly depends on the respective feature set. The feature set that yields the best results are the wavelet level 5 coefficients. The result of the EMD-based clustering is very sensitive to noise within the data and to preprocessing steps applied beforehand.

Coherence Probe Microscopy 3D Image Processing

Structural Analysis

Verena Schlager
Johannes Kepler Universität Linz
FLLL / CDL-MS-MACH

July 06, 2010

Abstract

In *Coherence Probe Microscopy (CPM)* one big problem is to detect noisy respectively failure slices of the recorded image stack. Also a main question is to group slices into different clusters (for example into different directions of the probe structure).
Before applying an analysing algorithm it is necessary to get rid of the radial illumination of the images. Therefor the so-called *rolling ball* algorithm is really suitable (available in ImageJ).
After preprocessing the single slices, different approaches for analysing them and the inner structures are available.
For detecting failure images, *Correlation* of each slice with its neighbouring slice will be computed and then the the resulting values are compared. An other approach is to compute the l^2-*norm* of the pixel value difference of a slice to its neighbour slice or, a more robust version, of a slice to all the other slices and sum the results up. The disadvantage of the second version is the considerably higher runtime. The norm analysis gives a similary result to the correlation approach.
To get some information about the direction of the inner structure of a slice, one can do a simple kind of *Radon Transform*. An other possibility is to compute the *maximum average chord lengths* of an image. This approach needs more preprocessing steps and also a good thresholding algorithm.
Also extracting *feature vectors* out of the images is very usefull. Afterwards one can cluster the images by a *k-Mean algorithm* to group the slices into different direction groups. Further research with other algorithms and features will be done.
Present problems in structural analysis are finding usefull features, optimal thresholding algorithms and analysing algorithms with accurate runtime.

Classifier-based analysis of visual inspection: Significance subsumes stability

Wolfgang Heidl, Christian Eitzinger WOLFGANG.HEIDL@PROFACTOR.AT
Profactor GmbH, 4407 Steyr-Gleink, AUSTRIA

Edwin Lughofer EDWIN.LUGHOFER@JKU.AT
Johannes Kepler University, 4020 Linz, AUSTRIA

1. Introduction

Quality control typically involves the visual inspection of products at the end of a production line. This task is quite often done by women. Their job is to make a quick accept/reject decision and to sort out the bad products. Manufacturing companies often argue that women have more endurance in performing this task and also make decisions with better reproducibility. To our knowledge gender differences in visual inspection decision-making have not been thoroughly investigated.

There is vast literature on sex differences in the behavioral sciences concerning numerous standardized tests on physical strength, spatial orientation, verbal and navigation abilities to name but a few. In the ergonomics and human factor engineering community visual inspection has been extensively studied (Drury & Watson, 2002; Harris & Chaney, 1969). Gender differences have been reported in reaction time changes due to acoustic noise (Gramopadhye & Wilson, 1997), however no gender difference has been found on inspection performance (Lehto & Buck, 2007).

We utilize ML classifiers (Hastie et al., 2009) as a mathematical model of the decision behavior of individual subjects (Figure 1). The goal of model identification, also known as learning, is to generalize from subject responses to specific stimuli to all stimuli stemming from the same random process. Once a ML classifier has been trained that generalizes well, the classifier parameters convey the task-relevant information on subject decision behavior. Further analysis can then be based on these parameters.

Classifiers have recently been successfully applied to

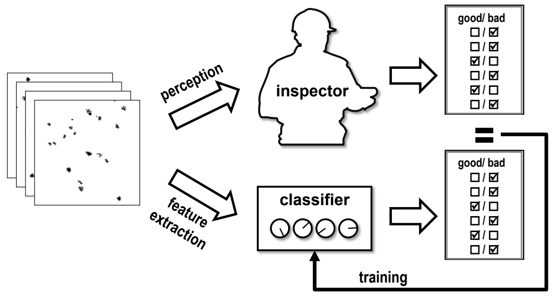

Figure 1. Classifiers modeling human decisions in visual inspection. During training the classier parameters (displayed as tuning knobs) are adjusted to minimize the discrepancy between human and classier decisions.

complex visual inspection tasks (e.g. (Eitzinger et al., 2009; Lughofer et al., 2009)). However, up to now they are mainly based on considerations of statistics and probability distributions of the features that they classify (Vapnik, 1999; Hastie et al., 2009). The fact that the decision is originally made by a human is often neglected. Consequently, there has been no investigation of differences in male and female decision-making from the viewpoint of classifiers.

Unlike classical parametric and latent variable models (Skrondal & Rabe-Hesketh, 2004) the model parameter and structure identification process during classifier training includes some form of randomization, sometimes already within the actual learning algorithm, but most often when the expected classification error is Monte-Carlo estimated for validation and complexity tuning(Stone, 1974). Depending on *classifier stability* (Bousquet & Elisseeff, 2002) those random effects propagate to the identified parameters and hence to our model of subject decision-behavior. This gives rise to the question of whether lack of classifier stability could result in observed gender differences that are merely spurious.

This work has been supported by the FEMtech program of The Federal Ministry for Transport, Innovation and Technology under FFG grant No. 318113. It reflects only the authors' views.

Statistical hypothesis testing (Neyman & Pearson, 1933; Fisher, 1966) is precisely concerned with assessing the likelihood of effects being observed by "luck of the draw". Following the line of (Ludbrook & Dudley, 1998) we adopt the Fisherian model of statistical inference and assess statistical significance with permutation tests (Good, 1994). We show that the errors resulting from randomized model identification are — with respect to statistical significance — no different from the omnipresent measurement errors and therefore do not change validity of permutation tests over so-identified measures. Classifier instability can not increase significance, therefore statistical significance of observed effects implies sufficient stability for that purpose.

Permutation tests confirm the significance of previously reported results (Heidl et al., 2010) that have been tested with classical t-tests. We extend those results by significant differences identified using CART (Breiman et al., 1993) decision trees. Additionally, we show that significant model-free gender differences exist, based solely on the analysis of binary subject responses.

Acknowledgment

The authors would like to thank Daniela Kirchberger who organized recruitment of the subjects and carried out the majority of experiment sessions.

References

Bousquet, O., & Elisseeff, A. (2002). Stability and generalization. *The Journal of Machine Learning Research*, *2*, 499–526.

Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1993). *Classification and regression trees.* Boca Raton: Chapman and Hall.

Drury, C. G., & Watson, J. (2002). *Good practices in visual inspection* (Technical Report). Federal Aviation Administration/Office of Aviation Medicine, Washington DC.

Eitzinger, C., Heidl, W., Lughofer, E., Raiser, S., Smith, J., Tahir, M., Sannen, D., & Van Brussel, H. (2009). Assessment of the influence of adaptive components in trainable surface inspection systems. *Machine Vision and Applications.*

Fisher, R. A. (1966). *The design of experiments.* New York: Hafner Publishing Company. Eighth edition edition.

Good, P. (1994). *Permutation tests.* Springer-Verlag Berlin and Heidelberg.

Gramopadhye, A., & Wilson, K. (1997). Noise, feedback training, and visual inspection performance. *International Journal of Industrial Ergonomics*, *20*, 223–230.

Harris, D. H., & Chaney, F. B. (1969). *Human factors in quality assurance.* New York: Wiley.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction, second edition (springer series in statistics).* Springer. 2nd ed. 2009. corr. 3rd printing edition.

Heidl, W., Thumfart, S., Lughofer, E., Eitzinger, C., & Klement, E. P. (2010). Classifier-based analysis of visual inspection: Gender differences in decision-making. *Proceedings of SMC2010, IEEE Conference on Systems, Man and Cybernetics.*

Lehto, M. R., & Buck, J. (2007). *Introduction to human factors and ergonomics for engineers.* CRC Press. 1 edition.

Ludbrook, J., & Dudley, H. (1998). Why permutation tests are superior to t and f tests in biomedical research. *The American Statistician*, *52*, 127–132.

Lughofer, E., Smith, J., Caleb-Solly, P., Tahir, M., Eitzinger, C., Sannen, D., & Nuttin, M. (2009). On human-machine interaction during on-line image classifier training. *IEEE Transactions on Systems, Man and Cybernetics, part A - Systems and Humans*, *39*, 960–971.

Neyman, J., & Pearson, E. S. (1933). On the problem of the most efficient tests of statistical hypotheses. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, *231*, 289–337.

Skrondal, A., & Rabe-Hesketh, S. (2004). *Generalized latent variable modeling: Multilevel, longitudinal, and structural equation models.* Chapman and Hall. 1 edition.

Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society*, *36*, 111–147.

Vapnik, V. N. (1999). *The nature of statistical learning theory.* Statistics for Engineering and Information Science. Springer, Berlin.