Wavelets, Fourier Transforms, and Neural Networks for Ford Classification

Paavo Nieminen

Department of Mathematical Information Technology
University of Jyväskylä
Finland

IEEE World Congress on Computational Intelligence
1–6 June, 2008

UNIVERSITY OF JYVÄSKYLÄ
DEPARTMENT OF MATHEMATICAL
INFORMATION TECHNOLOGY

Outline

UNIVERSITY OF JYVÄSKYLÄ
DEPARTMENT OF MATHEMATICAL
INFORMATION TECHNOLOGY

Percentages for the pre-announced data sets

	Train	**Validation**
Ford A	93.4 % (6.3 % fpr)	93.9 % (7.1 % fpr)
Ford B	96.1 % (3.8 % fpr)	88.1 % (14.8 % fpr)

The above table gives the train and validation accuracy and false positive rate (fpr) for the competition data sets Ford A and Ford B. It's fun to see on July 4th, what the outcome was on the undisclosed testing set.

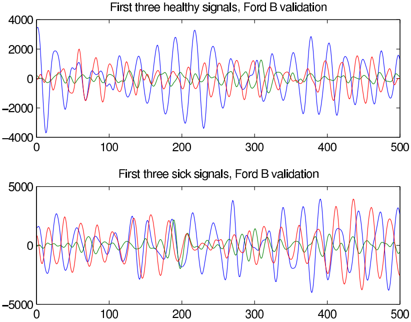
UNIVERSITY OF JYVÄSKYLÄ
DEPARTMENT OF MATHEMATICAL
INFORMATION TECHNOLOGY

I used a quite traditional processing path (Ford B described)

- Tukey window, zero-padding $\rightarrow$ 512 samples, FFT and DWT applicable
- DWT $\rightarrow$512 wavelet coefficients; threshold for de-noising.
- Relative distribution of energy into the different fluctuation levels $\rightarrow$ just a couple of Wavelet-based features.
- Inverse DWT $\rightarrow$ presumably de-noised signal.
- FFT $\rightarrow$ spectrum; normalization $\rightarrow$ spectrum with energy of one.
- Linear scaling of the input variables to the range $[-1, 1]$.
  $\rightarrow$ balanced training of an MLP with *tanh* activation.
- Simple MLP neural network (3 hidden neurons)
  $\rightarrow$ Classification of a signal.
- Train MLP using only training set (Ford A)
  and with validation set included (Ford B).
  Fixed (small) number of iterations.

UNIVERSITY OF JYVÄSKYLÄ
DEPARTMENT OF MATHEMATICAL
INFORMATION TECHNOLOGY

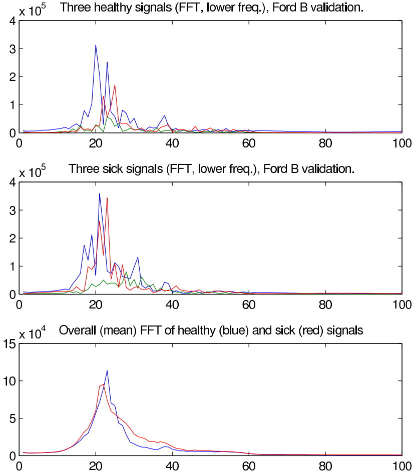
Time domain



- No conspicuous differences between classes.
- Obvious variance in amplitude between signals.
- This seemed like a dead end immediately.

UNIVERSITY OF JYVÄSKYLÄ
DEPARTMENT OF MATHEMATICAL
INFORMATION TECHNOLOGY

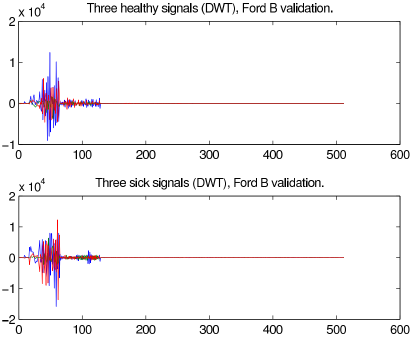
Frequency domain



- Basic discrete fourier transform seems to give some differences.
- Overall, there might be a difference in the peak frequencies.
- Interplay of different bands?

UNIVERSITY OF JYVÄSKYLÄ
DEPARTMENT OF MATHEMATICAL
INFORMATION TECHNOLOGY

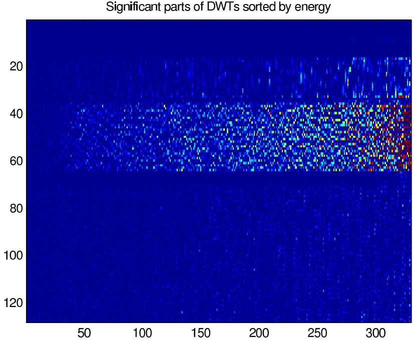
My Trip to Wavelets: first steps



- I took this as an opportunity to learn about the Wavelet transform.
- On the left there is one possible discrete wavelet transform (DWT) of some of the Ford B signals.

UNIVERSITY OF JYVÄSKYLÄ
DEPARTMENT OF MATHEMATICAL
INFORMATION TECHNOLOGY

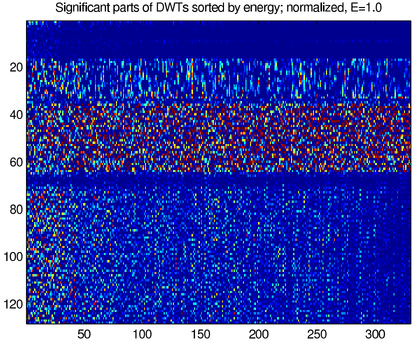
My Trip to Wavelets: fiddling around



- A plot of side-by-side DWTs.
- Some of the signals are very weak compared to others.
- Also, the weak signals seem to contain mostly noise (as guessed from a normalized plot)
- So . . . What, if anything, can I do with those?

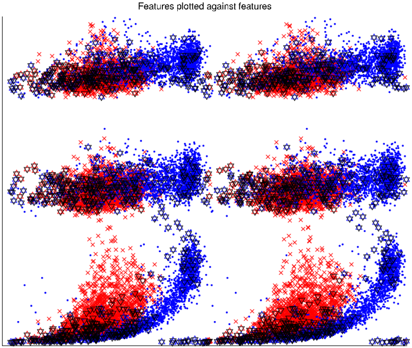
UNIVERSITY OF JYVÄSKYLÄ
DEPARTMENT OF MATHEMATICAL
INFORMATION TECHNOLOGY

The normalized DWT image is here:



UNIVERSITY OF JYVÄSKYLÄ
DEPARTMENT OF MATHEMATICAL
INFORMATION TECHNOLOGY

What do the features tell

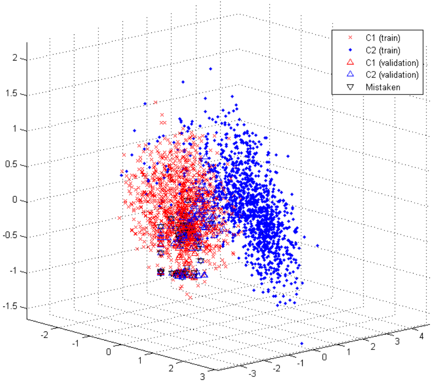


- I wanted to visualize the features somehow, and assess the usability of them.
- Here are some scatter plots of feature-pairs (Ford B).
- At the bottom we have the interplay of wavelet fluctuation distributions.
- Others are DFT bin magnitudes.

UNIVERSITY OF JYVÄSKYLÄ
DEPARTMENT OF MATHEMATICAL
INFORMATION TECHNOLOGY

The data in 3 dimensions (PCA from final features)

Ford B, with DWT and DFT features



- What can we find from a 3D-projection?
- At least zooming and rotating is much fun.
- In reality, this tells that some signals are, in fact, very easy to classify. Problems occur within a well-located minority.
- Unfortunately, my features take the noisy validation set close to this!

UNIVERSITY OF JYVÄSKYLÄ
DEPARTMENT OF MATHEMATICAL
INFORMATION TECHNOLOGY

Shortcomings that I realize only afterwards:

- Why on earth didn't I use a smaller window?
  - The interesting things might happen only sporadically!
  - A sliding window yields translation-invariance!
- I should have selected only the "most discriminant training signals" using some between-class distance metric.
  - What if the training data contains some false labels?
  - What if the sought-after, but elusive, discriminant feature does not occur within each 500-sample batch?
- Should I have excluded the most noisy batches from the training? What can you learn from plain noise!
- I didn't try out classifiers other than the simple MLP.

Summary

- It seems that for this case, traditional and simple methods performed quite well.
- Most of the work was done by trial and error.
- Interactive visualizations of data aid a lot in understanding what is possible and what perhaps is not.
- I thank the organizers and co-competitors. There was fun in computing this time!

Big thanks also to the Department of Mathematical Information Technology, University of Jyväskylä, for providing a travel grant to WCCI 2008, and to prof. Tommi Kärkkäinen for tipping me off about the competition.

UNIVERSITY OF JYVÄSKYLÄ
DEPARTMENT OF MATHEMATICAL
INFORMATION TECHNOLOGY

Greetings from Finland. Please, make a visit some time!



UNIVERSITY OF JYVÄSKYLÄ
DEPARTMENT OF MATHEMATICAL
INFORMATION TECHNOLOGY

Some Introductional Material on Wavelets I

- J. S. Walker,
  A Primer on Wavelets and their Scientific Applications.
  Boca Raton: CRC Press LLC, 1999.
- J. Buckheit, S. Chen, D. Donoho, I. Johnstone, and J. Scargle,
  "About WaveLab" 2005.
  [Online]. Available: `http://www-stat.stanford.edu/~wavelab/Wavelab_850/AboutWaveLab.pdf`

UNIVERSITY OF JYVÄSKYLÄ
DEPARTMENT OF MATHEMATICAL
INFORMATION TECHNOLOGY