The Influence of the Image Basis on Modeling and Steganalysis Performance

12th Information Hiding Conference

V. Schwamberger[2], P.H.D. Le[1], B. Schölkopf[2], M.O. Franz[1]

[1]University of Applied Sciences at Konstanz, Brauneggerstr. 55, D-78462 Konstanz, Germany

[2]Max Planck Institute for Biological Cybernetics, Spemannstr. 38, D-72076 Tübingen, Germany

Overview

Overview

Image model and image basis

- Main problem of steganalysis: detect a tiny manipulation in a large amplitude signal (here: a suitably transformed image)
- An image model predicts transform coefficients from their local neighborhoods based on the coefficient statistics of the image.
- By analyzing the prediction error instead of the whole image, most of the carrier signal is removed.
- The initial image transform determines the image basis in which the image is modeled and in which the residual is characterized by a suitable set of statistical descriptors for the final classification stage.

The choice of the image basis influences both the modeling process and the statistical image descriptors.

The blind steganalyzer of Lyu and Farid (2006)

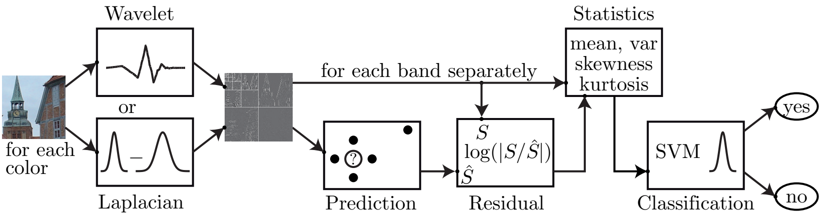


Image model: coefficients are predicted from their neighbourhood with linear regression (originally least squares, here Gaussian Processes).

Image basis: originally a QMF wavelet pyramid (3 layers, 3 orientation subbands), here alternatively a Laplace pyramid (3 layers, no orientation subbands).

Feature selection: statistics are computed only from coefficients above a threshold.

Goal of the study

Hypothesis: The best image basis (or the best associated subband transform) should lead to the image model with the highest prediction performance since this most effectively removes the carrier from a potential stego image.

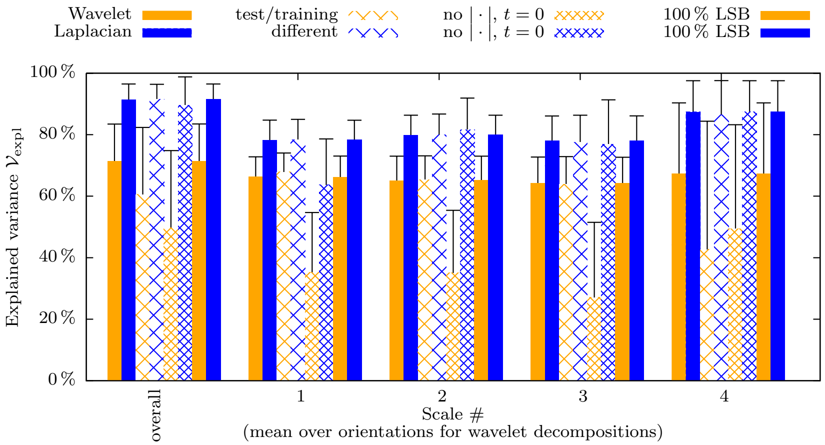
Method: Using the modified steganalyzer of Lyu and Farid, compare the image modeling performance measured in explained variance:

$$\mathcal{V}_{\text{expl}} = \frac{\mathcal{V}_{\text{img}} - \mathcal{V}_{\text{err}}}{\mathcal{V}_{\text{img}}} = 1 - \frac{\sum_{i,j}\left(S(x_i, y_j) - \widehat{S}(x_i, y_j)\right)^2}{\sum_{i,j}\left(S(x_i, y_j) - \bar{S}\right)^2}$$

to the steganalysis performance of the wavelet and Laplace image basis.

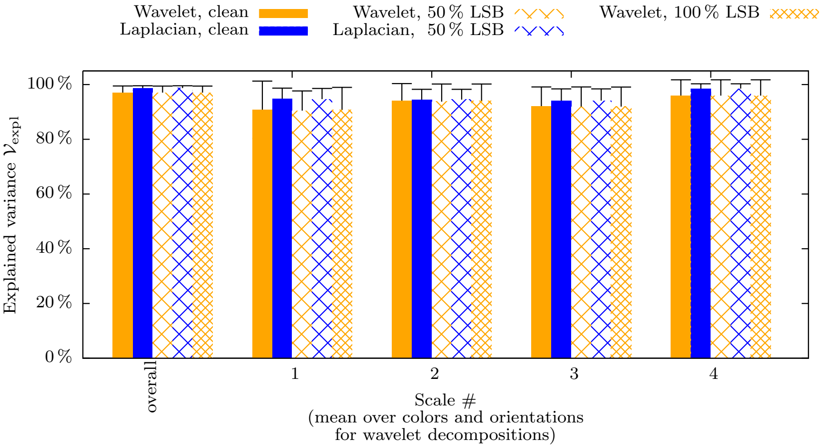
Overview

Explained variance on grayvalue images



Brodatz texture database, 111 images of size 640×640

Explained variance on color images



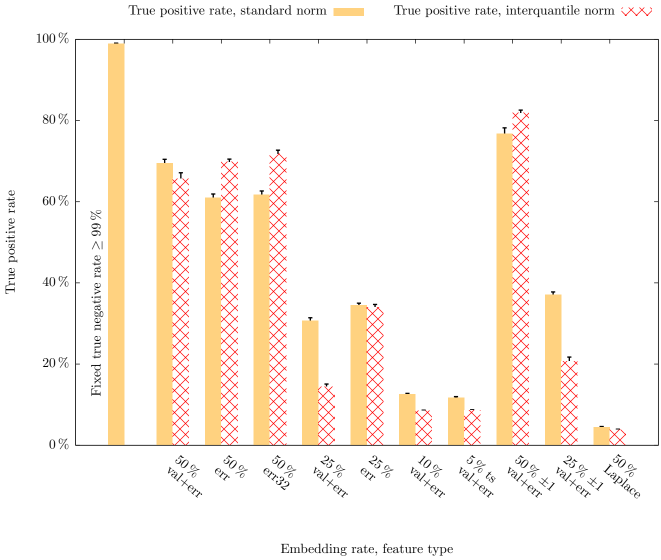
German Federal Office for Information Security database, 1640 never compressed RGB images

Image modeling performance

- Laplacian image basis outperforms the wavelet basis significantly in terms of explained variance, even if training and test region of the images are not the same.
- This happens both in pixel space and across the different scales of the Laplace or wavelet decomposition.
- For RGB images, the advantages of the Laplacian basis are less pronounced but still significant because of the high correlations between the color channels.
- Reasons: (1) Laplace coefficients are higher correlated with their neighbourhood than wavelet coefficient magnitudes; (2) The Laplace pyramid is overcomplete by a factor of $4/3 \Rightarrow$ more finely grained modeling of the local dependency structure.

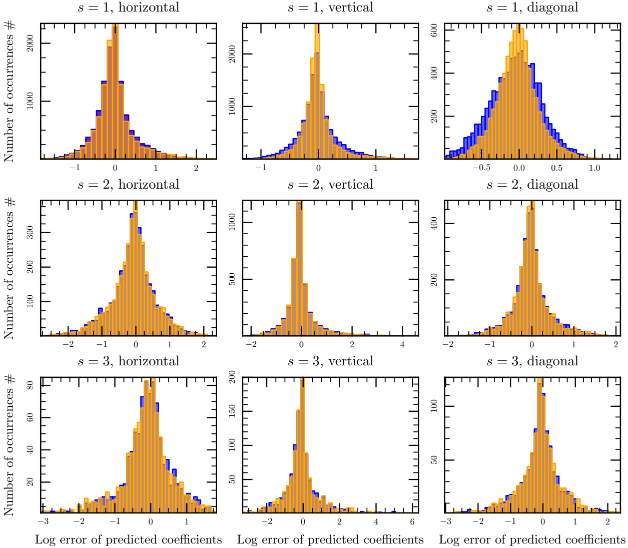
Overview

Steganalysis performance

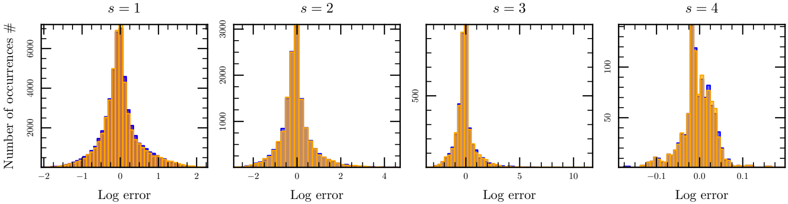


German Federal Office for Information Security database, 1000 training, 600 test image pairs, 100 random splittings

Log wavelet residual of stego and cover image



Log Laplacian residual of stego and cover image



Laplacian image basis performs vastly inferior to wavelets
⇒ a better image model does not automatically lead to better steganalysis performance!

Overview

Dimensionality of feature space

Reason 1

Due to the larger number of subbands, images are represented in 120 dimensions in the wavelet basis, 48 in the Laplacian. The probability of two classes being separable increases with dimensionality.

- Experiment with a hybrid steganalyzer: image is modeled in the Laplace domain, backtransformed in pixel domain, residual is wavelet transformed.
- Result: performance increases from a detection rate of $(4.5 \pm 0.1)\,\%$ to $(37.3 \pm 1.0)\%$, although the performance of the wavelet steganalyzer of $(69.5 \pm 0.9)\,\%$ could not be achieved.
- Reason: the backtransformed Laplace reconstruction is less accurate in the wavelet domain than the direct reconstruction.

Image modeling and denoising

Reason 2

Image modeling can be viewed as denoising: the "true" or denoised cover image is reconstructed from the contaminated or "noisy" stego image. Thresholded wavelets are better for denoising than other, less sparse image bases.

- Subband statistics for the steganalyzer are computed after thresholding the coefficients. This is a standard procedure for image denoising (threshold estimator, e.g. wavelet coring).
- It can be shown (Donohue & Johnson, 1993) that denoising performance is higher if the image basis is sparser.
- Wavelet bases are sparser for natural images $\Rightarrow$ the residual statistics can be estimated with higher accuracy after thresholding.
- Experiment: disable thresholding in wavelet steganalyzer $\Rightarrow$ detection rate drops from $(69.5 \pm 0.9)\,\%$ to $(27.1 \pm 0.6)\,\%$.

Conclusions

- Although the Laplacian representation leads to a better image model, it shows an inferior detection performance.
- High performance in image modeling does not directly transfer to a higher steganalysis performance.
- Sparseness in connection with thresholding and higher dimensionality of the feature vectors turned out to be more important than accurate modeling.
- It seems important to stay in the same transformation space in all steganalysis steps from image modeling, thresholding, to computation of feature vectors.

Thanks to: German Federal Ministry of Education and Research, German Federal Office for Information Security, Medav GmbH, Markus Messmer, Olga Schönemann, Jürgen Keppler.