Color-Image Watermarking using Multivariate Power-Exponential Distribution

Roland Kwitt, Peter Meerwald and Andreas Uhl
Department of Computer Sciences, University of Salzburg, A-5020 Salzburg, Austria
Email: {rkwitt,pmeerw,uhl}@cosy.sbg.ac.at; Web: http://www.wavelab.at

Abstract

- present a novel watermark detector for additive spread-spectrum watermarking in the DWT domain of color images
- model correlated DWT subbands of RGB color channels by multivariate power-exponential (MPE) distributions
- derive a likelihood ratio test (LRT) for watermark detection based on joint model for DWT detail color subbands

Introduction and Prior Work

Watermarking has been proposed as a technology to ensure copyright protection by embedding an imperceptible, yet detectable signal in digital multimedia content. Most watermarking research focuses on grayscale images. Color image watermarking is usually accomplished by marking only the luminance channel or by processing each color channel separately [1].

Blind detection performance can be improved by accurately modeling the host signal noise [2]. Expressing the joint statistical distribution of transform coefficients across correlated color channels is tedious and has been proposed for the Gaussian host signal case only [3].

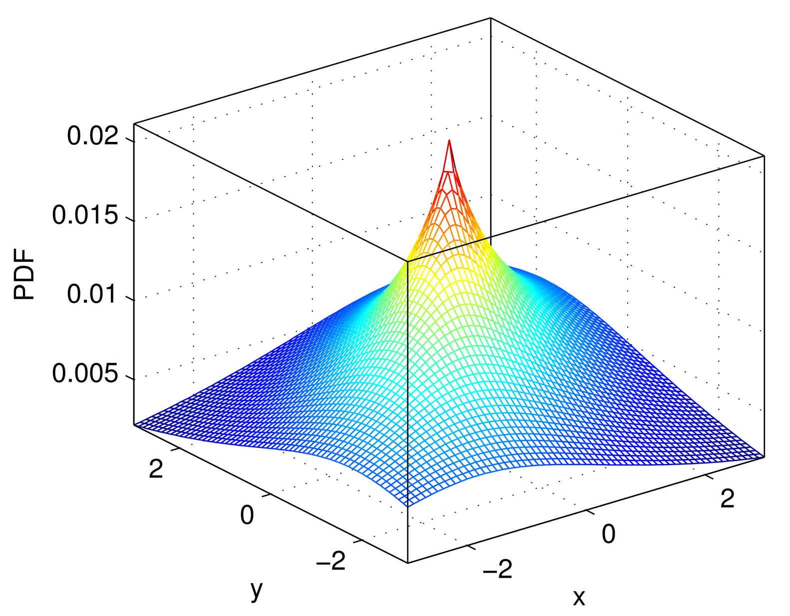
We derive a detector based on the multivariate power-exponential **(MPE)** distribution jointly modeling the DWT subband coefficients of color images. To compare, we implement watermarking approaches in the DWT domain based on linear correlation detection (LC) on the luminance channel **(LC-L)**, on a joint Gaussian model **(LC-J)** [3], on decorrelated channels **(LC-KLT)** [1] as well as a LRT conditioned on a GGD model of the luminance coefficients **(LRT-GGD-L)** [2].

MPE Host Signal Model

We introduce a Likelihood-Ratio Test for watermark detection in host signal noise which follows a multivariate power-exponential (MPE) distribution and discuss threshold determination as well as parameter estimation issues. This noise model for wavelet detail subband coefficients has already been successfully applied in the context of statistical color image retrieval [4] for example. The probability density function (PDF) of the multivariate power-exponential distribution with dimensionality n (MPE_n) is given by

$$p(\mathbf{x};\boldsymbol{\mu},\boldsymbol{\Sigma},\beta) = \frac{n\Gamma\left(\frac{n}{2}\right)}{\pi^{n/2}\Gamma\left(1+\frac{n}{2\beta}\right)2^{1+\frac{n}{2\beta}}} \cdot |\boldsymbol{\Sigma}|^{-1/2}\exp\left\{-\frac{1}{2}\left((\mathbf{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right)^{\beta}\right\} \quad (1)$$

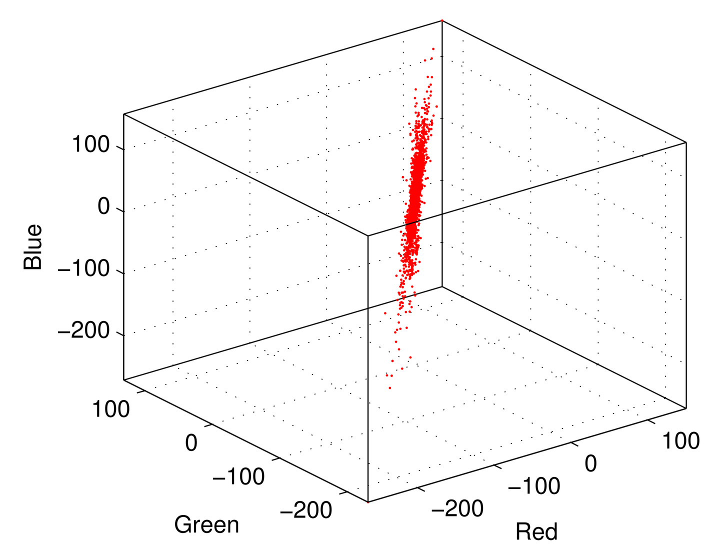
where $\boldsymbol{\Sigma}$ is a positive-definite symmetric $n\times n$ matrix, $\beta\in(0,\infty)$ denotes the shape parameter and $\boldsymbol{\mu}\in\mathbb{R}^n$ denotes the location vector.



PDF of a $\text{MPE}_2(\mathbf{0},\boldsymbol{\Sigma},0.4)$ with $\boldsymbol{\Sigma}=\left(\begin{smallmatrix}1 & 0.6\\ 0.6 & 1\end{smallmatrix}\right)$.

Since we aim at modeling wavelet coefficient distributions, it is reasonable to assume zero mean to reduce free parameters. We assume RGB images, thus we have three color bands and $n=3$.

The main reason for choosing a multivariate statistical model is the high correlation which can be observed between the same wavelet detail subbands of different color image channels. The statistical model of Eq. (1) is a special case of the Kotz-type family of distributions [5].



Scatter-plot of the HL subband coefficients of the R,G and B channel of the *Island* image at decomposition level two.

Parameter Estimation

Parameter estimation of the MPE_n noise model is accomplished using the method of moments by matching the variance and Mardia's multivariate Kurtosis coefficient [6, 7] to their empirical estimates. Let $X\sim\text{MPE}_n(\boldsymbol{\Sigma},\beta)$, then the variance of X is given by

$$\mathbb{V}(X) = \frac{2^{\frac{1}{\beta}}\Gamma\left(\frac{n+2}{2\beta}\right)}{n\Gamma\left(\frac{n}{2\beta}\right)}\boldsymbol{\Sigma}. \quad (2)$$

Mardia's multivariate Kurtosis coefficient is defined as

$$\gamma_2(X) = \mathbb{E}\left[\left((X-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(X-\boldsymbol{\mu})\right)^2\right] - n(n+2) \quad (3)$$

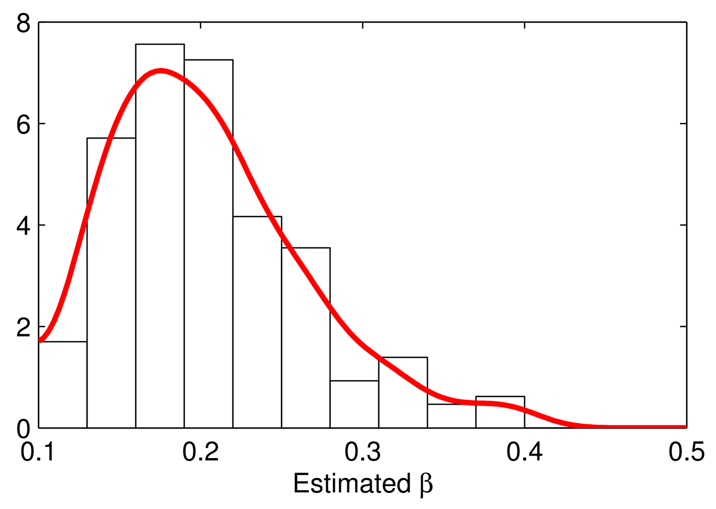
which leads to

$$\gamma_2(X) = \frac{n^2\Gamma\left(\frac{n}{2\beta}\right)\Gamma\left(\frac{n+4}{2\beta}\right)}{\Gamma^2\left(\frac{n+2}{2\beta}\right)} - n(n+2) \quad (4)$$

in case of a $\text{MPE}_n(\boldsymbol{\Sigma},\beta)$ distribution. Given that $\mathbf{S}$ denotes the classic sample variance (setting $\boldsymbol{\mu}=\mathbf{0}$) we can estimate $\gamma_2(X)$ by

$$\hat{\gamma}_2(\mathbf{x}_1,\ldots,\mathbf{x}_m) = \frac{1}{m}\sum_{i=1}^{m}\left(\mathbf{x}_i^T\mathbf{S}^{-1}\mathbf{x}_i\right)^2 - n(n+2). \quad (5)$$

from our data $\mathbf{x}_1,\ldots,\mathbf{x}_m$ where m denotes the number of wavelet coefficients in each target subband. We use $\hat{\beta}$ to obtain an estimate for $\boldsymbol{\Sigma}$. $\beta=1$ corresponds to a multivariate Gaussian distribution.



Histogram and kernel density fit for the shape parameter $\hat{\beta}$ over all 24 Kodak test images.

Embedding and Detection Problem

We create a bipolar, pseudo-random watermark sequence $\mathbf{w}=[w_1,\ldots,w_m]^T$ with $w_i\in\{+1,-1\}$ (depending on a secret key K). According to the rule of additive spread-spectrum watermarking, $\mathbf{W}=[\mathbf{w}^T\ \mathbf{w}^T\ \mathbf{w}^T]$ is added to the signal matrix by $\mathbf{Y}=\mathbf{X}+\alpha\mathbf{W}$ where α denotes embedding strength. The hypothesis of the signal detection problem are

$$\begin{aligned}\mathcal{H}_0 &: \mathbf{Y}=\mathbf{X} \quad \text{(not watermarked)}\\ \mathcal{H}_1 &: \mathbf{Y}=\mathbf{X}+\alpha\mathbf{W} \quad \text{(watermarked)}\end{aligned} \quad (6)$$

Assuming independence of the observations $\mathbf{x}_1,\ldots,\mathbf{x}_m$, the statistic of the LRT is

$$l(\mathbf{Y}) = \frac{\prod_{i=1}^{m}p(\mathbf{y}_i-\alpha\mathbf{w}_i;\boldsymbol{\Sigma},\beta)}{\prod_{i=1}^{m}p(\mathbf{y}_i;\boldsymbol{\Sigma},\beta)}. \quad (7)$$

Taking the logarithm and inserting the PDF of Eq. (1) we obtain

$$L(\mathbf{Y}) = -\frac{1}{2}\sum_{i=1}^{m}\left((\mathbf{y}_i-\alpha\mathbf{w}_i)^T\boldsymbol{\Sigma}^{-1}(\mathbf{y}_i-\alpha\mathbf{w}_i)\right)^{\beta} + \frac{1}{2}\sum_{i=1}^{m}\left(\mathbf{y}_i^T\boldsymbol{\Sigma}^{-1}\mathbf{y}_i\right)^{\beta}. \quad (8)$$

$L(\mathbf{Y})$ follows a Normal distribution under $\mathcal{H}_0$ and $\mathcal{H}_1$ with parameters (μ_0,σ_0^2) and (μ_1,σ_1^2), resp. If we consider $\mathbf{y}_i$ fixed, the only variable term is $\mathbf{w}_i$ and the expected value μ_0 under $\mathcal{H}_0$ (note that $\mathbf{y}_i=\mathbf{x}_i$) can be calculated as

$$\mu_0 = -\frac{1}{4}\sum_{i=1}^{m}\left((\mathbf{x}_i-\alpha\mathbf{w}_i)^T\boldsymbol{\Sigma}^{-1}(\mathbf{x}_i-\alpha\mathbf{w}_i)\right)^{\beta} + \left((\mathbf{x}_i+\alpha\mathbf{w}_i)^T\boldsymbol{\Sigma}^{-1}(\mathbf{x}_i+\alpha\mathbf{w}_i)\right)^{\beta} + \frac{1}{2}\sum_{i=1}^{m}\left(\mathbf{x}_i^T\boldsymbol{\Sigma}^{-1}\mathbf{x}_i\right)^{\beta} \quad (9)$$

and the variance σ_0^2 of the detection statistic $L(\mathbf{Y})$ is given by

$$\sigma_0^2 = \frac{1}{16}\sum_{i=1}^{m}\left(\left((\mathbf{x}_i+\alpha\mathbf{w}_i)^T\boldsymbol{\Sigma}^{-1}(\mathbf{x}_i+\alpha\mathbf{w}_i)\right)^{\beta} - \left((\mathbf{x}_i-\alpha\mathbf{w}_i)^T\boldsymbol{\Sigma}^{-1}(\mathbf{x}_i-\alpha\mathbf{w}_i)\right)^{\beta}\right)^2. \quad (10)$$

Having obtained both parameters of the Gaussian under $\mathcal{H}_0$ allows the determination of the detection threshold T in a Neyman-Pearson sense as $T=\text{erfc}^{-1}(2P_f)\sqrt{2\sigma_0^2}+\mu_0$ where P_f denotes the desired probability of false alarm. The detection statistic parameters (μ_1,σ_1^2) under $\mathcal{H}_1$ are $\mu_1=-\mu_0$ and $\sigma_1^2=\sigma_0^2$ as $\forall i:\mathbf{y}_i=\mathbf{x}_i+\alpha\mathbf{w}_i$.

Experimental Results

Nine Kodak color images are resized to 192×128 pixel, decomposed with biorthogonal 7/9 wavelet filters and the HL subband at decomposition level 2 is watermarked with $\alpha=5$ (PSNR $\approx$ 46 dB). We verify that the detector responses follow a Gaussian law by a Lilliefors test [8] at the 5% significance level and check that the theoretical values μ_0 and σ_0^2 are close to the experimental values we obtain from Eq. (9) and Eq. (10) under $\mathcal{H}_0$.

Image	μ_0	$\hat{\mu}_0$	σ_0	$\hat{\sigma}_0$
Barn	−238.71	−239.39	805.58	800.72
Facade	−216.56	−216.51	658.74	637.65
Girl	−133.77	−133.30	372.45	350.54
House	−134.91	−134.23	254.95	244.17
Island	−353.10	−353.55	1609.80	1571.80
Parrots	−287.28	−285.25	1738.73	1657.28
Rafting	−185.94	−186.64	801.52	826.01
Window	−139.83	−140.59	986.03	984.52
Zentime	−349.46	−352.74	1491.52	1461.93

Theoretical and experimental values of the detector statistics under $\mathcal{H}_0$

The mean and variance of the detection statistic are estimated experimentally under $\mathcal{H}_1$ from 1000 test runs for each image and $\hat{P}_m=\frac{1}{2}\text{erfc}\left((\hat{\mu}_1-T)/\sqrt{2\hat{\sigma}_1^2}\right)$ is used to determine the empirical probability of missing the watermark for a false-alarm rate of $P_f=10^{-6}$. The MPE detector performs better than the LRT-GGD-L detector except for one image and significantly outperforms the LC detectors.

Image	LC-L	LC-J	LC-KLT	LRT-GGD-L	MPE
Barn	10^{-6}	10^{-7}	10^{-27}	10^{-31}	$\mathbf{10^{-50}}$
Facade	0.46	0.35	10^{-5}	10^{-5}	$\mathbf{10^{-24}}$
Girl	0.97	0.97	0.43	10^{-95}	$\mathbf{10^{-103}}$
House	0.03	0.02	10^{-12}	10^{-8}	$\mathbf{10^{-22}}$
Island	10^{-9}	10^{-10}	10^{-42}	$\mathbf{10^{-189}}$	10^{-166}
Parrots	10^{-6}	10^{-6}	10^{-20}	10^{-70}	$\mathbf{10^{-86}}$
Rafting	0.21	0.15	10^{-7}	10^{-12}	$\mathbf{10^{-16}}$
Window	0.02	0.03	10^{-9}	10^{-54}	$\mathbf{10^{-79}}$
Zentime	0.12	0.06	10^{-10}	10^{-121}	$\mathbf{10^{-168}}$

Empirical probability of missing the watermark ($\hat{P}_m$) at 46 dB PSNR for different detectors ($P_f=10^{-6}$, 1000 test runs).

Conclusion

We proposed a novel detector for additive, spread-spectrum watermarking of color image DWT subbands based on the multivariate power-exponential distribution. This signal model allows to capture the highly correlated structure of the subbands. The derived likelihood ratio test achieves increased detection performance compared to watermarking the luminance channel only and earlier detectors based on a Gaussian host signal model.

Source code available at http://www.wavelab.at/sources. Research funded by Austrian Science Fund project FWF-P19159-N13.

References

[1] M. Barni, F. Bartolini, A. DeRosa, and A. Piva, "Color image watermarking in the Karhunen-Loeve transform domain," *Journal of Electronic Imaging*, vol. 11, no. 1, pp. 87–95, Jan. 2002.

[2] J. Hernández, M. Amado, and F. Pérez-González, "DCT-domain watermarking techniques for still images: Detector performance analysis and a new structure," *IEEE Transactions on Image Processing*, vol. 9, no. 1, pp. 55–68, Jan. 2000.

[3] M. Barni, F. Bartolini, and A. Piva, "Multichannel watermarking of color images," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 12, no. 3, pp. 142–156, Mar. 2002.

[4] G. Verdoolaege, S. De Backer, and P. Scheunders, "Multiscale colour texture retrieval using the geodesic distance between multivariate generalized Gaussian models," in *Proceedings of the 15th IEEE International Conference on Image Processing, ICIP '08*, San Diego, CA, USA, 2008, pp. 169–172.

[5] S. Nadarajah, "The Kotz-type distribution with applications," *Statistics*, vol. 37, no. 4, pp. 341–358, 2003.

[6] K. Zografos, "On Mardia's and Song's measures of Kurtosis in elliptical distributions," *Journal of Multivariate Analysis*, vol. 99, no. 5, pp. 858–879, May 2008.

[7] K.V. Mardia, "Measures of multivariate Kurtosis and Skewness," *Biometrika*, vol. 57, pp. 519–530, 1970.

[8] H. Lilliefors, "On the Kolmogorov-Smirnov test for normality with mean and variance unknown," *Journal of the American Statistical Association*, vol. 62, pp. 399–402, June 1967.