# Key-Dependent Random Image Transforms and Their Applications in Image Watermarking

Jiri Fridrich
Center for Intelligent Systems
SUNY Binghamton
Binghamton, NY 13902-6000

**Abstract:** *In this paper, we present a fast algorithm for generating key-dependent random smooth orthogonal image transforms for digital image watermarking. The algorithm first generates a set of k random smooth vectors of length N and then Gram-Schmidt orthogonalization procedure is applied to make them orthogonal. A set of $k^2$ orthogonal N×N matrices is finally obtained using tensor products between all possible pairs of the k vectors. We further investigate the average energy-packing properties of the whole class of random transformation and make a comparison to known image transforms. We explain how the key-dependent transforms can be used for construction of ultra secure image watermarking methods and nonlinear secure public watermark detectors.*

**Keywords:** Digital watermarking, security, key-dependent image transforms

## 1. Introduction

Watermarking techniques can be roughly divided into two categories depending on whether the watermark is inserted into the coefficients of some transform or directly into the grayscale levels of pixels. A fundamental advantage of transform based techniques stems from the fact that image transforms, such as DCT, DFT, Hadamard transform, etc., have good energy compactification properties and most of the image energy can be captured with a relatively small number of coefficients. The transform basis functions corresponding to these coefficients carry the most perceptually important information about the image. In spread spectrum techniques, the watermarking is achieved by modulating the transform coefficients with a pseudo-random key-dependent signal (the watermark). Watermarks inserted into the perceptually most important coefficients are typically extremely robust to a wide spectrum of image deformations [1]. At the same time, it is difficult to blindly remove the watermark without visibly degrading the image. This is the basic principle behind most of today's watermarking techniques [2–12].

Key-dependent random smooth transforms for watermarking have been first proposed in a paper by Fridrich [13]. This new concept was introduced with the intent to develop a watermarking technique for which it would be possible to implement the watermark detector as a black-box that anybody could use to determine whether or not a given image is watermarked but the detector itself should not enable a pirate to remove the watermark from the image or recover the watermarking key wired in the black box. Such a detector would find important applications in control of illegal copying of Digital Video Disks DVDs [14,15]. Kalker recently established [16,17] that if the watermark detector is a linear correlator it is always possible to remove the watermark from an image if the black box watermark detector is available. Linnartz et al. [14,18] present an iterative method for removing the watermark based on the behavior of the watermark detector at the detection threshold. By carefully adjusting the gray levels of individual pixels at the threshold, one can estimate the partial derivatives of the detector function at an image that is "far" from the original watermarked image. The authors suggest to use this information to determine the set of most influential pixels and move around the set of watermarked images to an image that is perceptually close to the original. It is not clear if the local information about the linearized detector at points far from the watermarked

image can lead to a globally convergent behavior. It is possible that if the watermark detector is sufficiently nonlinear, the iterative algorithm will randomly wander in vain and will not converge to the desired image.

Key-dependent random smooth patterns may help us build such a secure public black-box watermark detector. Because the choice of the image transform is unknown (depends on a secret key), one cannot purposely modify the coefficients that enter the watermark detector. This is important and it certainly contributes to the security of the system. The key-dependent transform also gives us more possibilities for modulating the coefficients. Since the transform depends on a key, one can make some obvious modifications to the projections (coefficients), such as quantize coefficients to a fixed set of uniformly spaced values, without the fear of producing an insecure scheme. The watermark detector can then simply test for presence of clusters in the modified coefficients and make decisions about watermark presence based on this numerical evidence for clusters. Such a detector would be highly nonlinear, which is exactly what is needed for secure public watermark detector. So far, the major bottleneck of the methods that use key-dependent random transforms has been the computational complexity of the algorithm that generates the transforms [13]. The Gram-Schmidt orthogonalization procedure is the most time consuming part of the algorithm. In this paper, we present a new fast technique for generating the random transforms. The new technique enables construction of practical and fast watermarking techniques that utilize key-dependent transforms and nonlinear detectors. We investigate the energy-compactification properties of the key-dependent transforms and compare them to other known image transformations.

In the next section, we describe the new fast algorithm for generation of key-dependent random smooth patterns. In Section 3, we study the energy-packing properties of the random transformations and make a comparison with known image transforms, such as DFT, DCT, or the Haar transform. It turns out that typical random transforms have energy-packing properties approximately the same as the DFT. Finally, in Section 4 we describe a watermarking algorithm and a nonlinear detector and outline future research directions.

## 2. General key-dependent transforms

The idea to use a whole class of image transforms that depend on a secret key for watermarking has been proposed by Fridrich in [13]. The advantage if this approach is that an attacker does not know the choice of the transform used to watermark a specific image and has practically nothing to start from. The transform is calculated from a user-defined key that is kept secret. In [13], the author proposes the following technique for generating random smooth orthogonal patterns (basis functions). First, a Pseudo-Random Number Generator (PRNG) with uniform distribution on [0,1] is seeded with a secret key and a set of $N$ pseudo-random patterns is generated. The patterns are then smoothened using a low-pass filter and orthogonalized using the Gram-Schmidt orthogonalization procedure. The projections of the patterns onto the image (inner products) play the role of coefficients that can be modulated or modified to embed a watermark. Since the patterns are secret, we have much more freedom for modulating the projections. For example, we could use a simple quantization to predetermined values. The detection would then decide about watermark presence based on numerical evidence that the projections for a particular image have been quantized (clustered). This naturally leads to an interesting nonlinear watermark detector.

Using key-dependent random smooth patterns has at least two potential drawbacks: computational complexity and energy-packing optimality. The computational complexity of the algorithm proposed in [13] is $O(P^2N^2)$, where $P$ is the number of orthogonal patterns, and $N$ is the image dimension. The Gram-Schmidt procedure is computationally very expensive, and it appears that it is not possible to develop a fast computational routine tantanamous to the fast Fourier transform. To avoid large memory requirements for storage

of the random patterns, the patterns need to be generated each time detection is required. This further increases computational complexity of the algorithm.

In this paper, we propose a new technique for fast generation of random smooth orthogonal patterns. Using the new technique, we generate $k^2$ two-dimensional $N \times N$ patterns from an orthogonal system of $k$ random smooth vectors of length $N$ via tensor products of all vector pairs. This is in fact quite similar to how two-dimensional DCT or DFT are constructed. The computational complexity of this algorithm is $O(PN)$, which is a very significant savings when compared to the old technique. It takes less than a second on a 333MHz Pentium II processor to generate 1000 orthogonal random smooth patterns, calculate the projections, modify them, and reassemble a watermarked image with 256×256 pixels.

The algorithm for an $N \times N$ grayscale image is presented below. It consists of four steps:

Step 1. Seed a PRNG with a secret key and generate $k$ random vectors $v_1, \ldots, v_k$ of length $N$.

Step 2. Smooth the vectors using a low-pass filter (keeping the same notation).

Step 3. Set $v_1 = 1$, a vector consisting of all ones.

Step 4. Apply Gram-Schmidt OG procedure to $v_1, \ldots, v_k$ (keeping the same notation).

At this point, the OG random smooth patterns (random DCT basis) can be obtained using tensor products of vectors $v_i$. Assuming $v_i$'s are column vectors, we obtain $k^2$ random smooth OG matrices with $ij$-th matrix equal to $v_i \cdot v_j'$.

The watermarking scheme that uses key-dependent basis functions will have four more steps:

Step 5. (Calculate the projections of the image $I$ onto all $k^2$ matrices). Arrange vectors $v_1, \ldots, v_k$ into one $N \times k$ matrix $V$. The $k^2$ projections $P$ arranged into a $k \times k$ matrix are obtained as $P = V \cdot I \cdot V'$.

Step 6. Modulate the projections $P$ to $Pn$.

Step 7. Reassemble the watermarked image $Iw$ from the new projections: $Iw = I + V' \cdot (Pn - P) \cdot V$.

Step 8. (Take care of boundary effects) if $Iw(p)>1$ for pixel $p$, set $Iw(p)=1$, if $Iw(p)<0$ for pixel $p$, set $Iw(p)=0$.

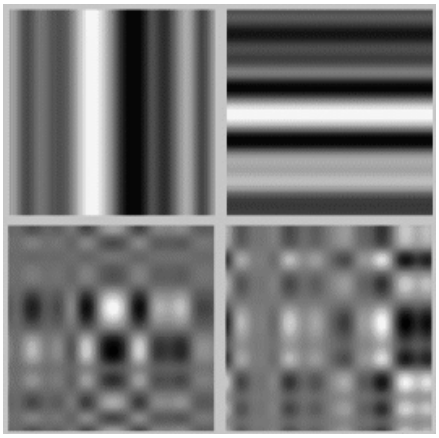

**Figure 1 Examples of four random smooth orthogonal basis functions (patterns).**

As can be seen, the complexity of the process is proportional to $PN$. Examples of random smooth orthogonal patterns are given in Figure 1. One can clearly see a strong resemblance to the DCT. Actually, the basis functions look like randomized DCT modes. This is caused by the construction of the random basis using tensor products from random, smooth, orthogonal one-dimensional vectors.

## 3. Energy compactification efficiency

In this section, we study the energy-packing efficiency of the new transform. There is a good reason why DCT is so frequently used for watermarking. This has to do with the performance of DCT, which is very close to the optimal Karhunen-Loève transform. The coefficients of the DCT pack the energy of typical images very efficiently (see, for example, [19]). This implies that watermarks inserted into DCT coefficients will have good robustness properties. Replacing the DCT with

a key-dependent random transform may lead to a less efficient transform. Obviously, we need to investigate the energy-packing efficiency of such transforms and somehow compare their performance to known image transforms.

We performed an analysis similar to the one described by Clarke in [19]. The images were modeled as first order Markov processes with covariance *cov* between pixels $(i_1, j_1)$ and $(i_2, j_2)$ dependent only on the difference $\Delta = \sqrt{(i_1 - i_2)^2 + (j_1 - j_2)^2}$, $cov = \rho^{\Delta}$. The covariance matrix of the original image is a Toeplitz $k \times k$ matrix $C_{ij} = \rho^{|i-j|}$. Given an orthonormal system of $k$ basis vectors of length $k$ arranged into a $k \times k$ matrix $V$, the covariance between the transform coefficients is the matrix $T = V \cdot C \cdot V'$. The energy-packing efficiency is evaluated using two numerical measures (for details, see [19]): The decorrelation efficiency $\eta_C$ and the relative amount of energy in the first $d$ diagonal components $\eta_E(d)$ as a function of $d$:

$$\eta_C = 1 - \frac{\sum_{\substack{i,j=1 \\ i \neq j}}^{k} C_{ij}}{\sum_{\substack{i,j=1 \\ i \neq j}}^{k} T_{ij}}$$

The decorrelation efficiency is the ratio between the sum of off-diagonal elements of the covariance matrix $C$ and the corresponding sum for matrix $T$. For a transform that achieves a complete decorrelation, the decorrelation efficiency is 100%. This is the case for the Karhunen-Loève transform. The relative amount of energy $\eta_E(d)$ tells us how much energy is contained in the first $d$ diagonal elements of the transform. Higher values correspond to higher energy-packing efficiency. The abbreviations used in the table are: DCT=Discrete Cosine Transform, KLT= Karhunen-Loève Transform, DFT=Discrete Fourier Transform, Haar=Discrete Haar Transform, DST=Discrete Sine Transform, WHT=Walsh-Hadamard Transform, HCT=High Correlation Transform, Slant=Slant Transform, Random=Random key-dependent Transform, Std=standard deviation of the entries in the row Random. Both numerical measures were calculated for an 8×8 square.

$$\eta_E(d) = 1 - \frac{\sum_{i=1}^{d} C_{ii}}{\sum_{i=1}^{d} T_{ii}}$$

**Table 1 Decorrelation efficiency and energy packing efficiency for various image transformations.**

| Transform | $\eta_C$ | $\eta_E(1)$ | $\eta_E(2)$ | $\eta_E(3)$ | $\eta_E(4)$ | $\eta_E(5)$ | $\eta_E(6)$ | $\eta_E(7)$ |
|---|---|---|---|---|---|---|---|---|
| DCT | 98.05 | 79.3 | 90.9 | 94.8 | 96.7 | 97.9 | 98.7 | 99.4 |
| KLT | 100.00 | 79.5 | 91.1 | 94.8 | 96.7 | 97.9 | 98.7 | 99.4 |
| DFT | 89.48 | 79.3 | 86.0 | 92.7 | 94.7 | 96.7 | 97.8 | 99.0 |
| Haar | 92.70 | 79.3 | 89.3 | 92.4 | 95.5 | 96.6 | 97.8 | 98.9 |
| DST | 84.97 | 73.6 | 84.3 | 92.5 | 95.0 | 97.4 | 98.4 | 99.4 |
| WHT | 94.86 | 79.3 | 89.3 | 92.7 | 95.5 | 96.7 | 97.9 | 99.0 |
| HCT | 96.72 | 79.3 | 90.5 | 94.4 | 96.0 | 97.1 | 98.3 | 99.3 |
| Slant | 97.16 | 79.3 | 90.7 | 94.6 | 96.3 | 97.4 | 98.6 | 99.3 |
| Random | 91.48 | 79.3 | 86.0 | 92.5 | 94.3 | 96.0 | 97.6 | 98.8 |
| Std | ±1.1 | ±0.0 | ±2.0 | ±0.8 | ±0.7 | ±0.7 | ±0.2 | ±0.3 |

**Table 2 Decorrelation efficiency as a function of image block size (see [19] pp. 132).**

| Block size | 4 | 8 | 16 |
|---|---|---|---|
| $\eta_C$ | 95.3±1.1 | 91.3±1.3 | 87.5±1.3 |

**Table 3 Decorrelation efficiency as a function of ρ (Block = 8) (see [19] pp. 131).**

| ρ | 0.99 | 0.96 | 0.91 | 0.84 | 0.75 |
|---|---|---|---|---|---|
| $\eta_C$ | 99.0±0.2 | 96.0±0.6 | 91.3±1.3 | 85.0±2.2 | 77.7±3.2 |

Table 1 shows that the energy-packing ability of a general random smooth transform is on average very similar to the discrete Fourier transform, although the decorrelation efficiency is somewhat better. Overall, the performance of the random transform does not appear to be significantly worse than the performance of most well known image transforms. This suggests that watermarking schemes based on random smooth transforms will likely achieve comparable robustness properties.

## 4. Conclusions and Future Effort

In this paper, we have developed a new technique for fast generation of key-dependent random smooth orthogonal patterns (basis functions). The patterns are tensor products of pairs of random smooth orthogonal vectors obtained from a PRNG, smoothened using a low-pass filter, and orthogonalized using the Gram-Schmidt procedure. The new algorithm can generate over 1000 random smooth patterns with 256×256 pixels in a fraction of a second (on a 333MHz Pentium II computer in Matlab 5.1). Actually, the performance of the algorithm appeared to be even faster than techniques based on the DCT. The random patterns were further analyzed for their energy-compactification properties. We concluded that their performance measured using decorrelation efficiency coefficient and energy-packing coefficient is on average very close to the discrete Fourier transform. This implies that techniques based on key-dependent transforms will have good robustness properties similar to techniques based on the DCT or other common transforms.

As discussed in the introduction, one of the most important arguments for using key-dependent basis functions is the fact that this concept may enable construction of a secure public watermark detector that is implemented as a black-box in a tamper-proof hardware. Such watermark detectors find important applications in copy control of DVD [14,15]. The box accepts integer matrices on its input and outputs one bit of information. It is assumed that the complete design of the detector and the corresponding watermarking scheme are known except a secret key, and that an attacker has one watermarked image at his disposal. The latest attacks on public watermark detectors [14–18] indicate that it is not clear if a secure public watermark detector can be built at all. It has been proven that all watermark detectors that are thresholded linear correlators can be attacked using a variety of techniques [16–18]. Kalker [16,17] describes a simple statistical technique using which the secret key can be recovered in $O(N^2)$ operations, where $N$ is the image dimension. The main culprit seems to be the fact that the correlation function is linear and that the quantities $c_i$, which are correlated with the watermark sequence $w_i$, can be directly modified through the pixel values. Linnartz and Cox [14,15] attack public detectors by investigating the sensitivity of the watermark detector to individual pixels for a critical image – the image at the detection threshold. Once the most influential set of pixels is found, its gray levels are scaled and subtracted from the watermarked image. They repeat the process in a hope to converge to an image that does not have the watermark. The assumption here is that we can actually learn the sensitivity of the detection function *at the watermarked image* from its sensitivity *at the critical image* that will generally be far from the watermarked image.

In order to design a watermarking method with a detector that would not be vulnerable to those attacks, we need to mask the quantities that are being correlated so that we cannot purposely change them through pixel values. Also, we must introduce nonlinearity into the scheme to prevent the attack by Linnartz et al. [14,15,18].

Towards this purpose, we propose to use key-dependent basis functions and a watermarking algorithm that clusters (quantizes) the projections $c_i$ onto $k^2$ random smooth orthogonal patterns to a set of points $x_i$ that follow geometric sequence with quotient $q$. By selecting

$$q = \frac{1+\alpha}{1-\alpha},$$

$$x_i = x_0 q^i,$$

it is easy to see that $| |x| - round_q(|x|) | \le \alpha |x|$ for any real $x$. In the last expression, $round_q(x)$ denotes the operation of rounding $x$ to the closets point $x_i$. Therefore, by rounding the projections $c_i$, one will never change $c_i$ by more than $\alpha$ %.

The detection algorithm tests the presence of clusters in the projections. Given $k^2$ projections, the following function can be used to numerically evaluate the evidence for clusters:

$$dist = \sum_{i=1}^{k^2} \left( | c_i | - x_0 q^{round_q(|c_i|)} \right)^2 .$$

This function is the sum of squares of differences between the projections and the centers of the closest clusters.

Preliminary tests on the classical grayscale test image Lenna with 256×256 pixels indicate that this technique is robust with respect to low-pass filtering, JPEG compression, resampling, brightness and contrast adjustment, and gamma correction. On the other hand, techniques based on clustering key-dependent projections are not robust to overwatermarking (embedding multiple watermarks in one image) and the collusion attack. A search for a more sophisticated watermark embedding and detection function will be a part of our future research.

## 5. References


[1] I.J. Cox and M.L. Miller, "A review of watermarking and the importance of perceptual modeling", *Proceedings of Electronic Imaging'97*, February 1997.

[2] I.J. Cox, J. Kilian, T. Leighton, and T. Shamoon, Secure Spread Spectrum Watermarking for Multimedia," NEC Research Institute, *Technical Report* 95-10.

[3] B.O. Comiskey and J.R. Smith, "Modulation and Information Hiding in Images," in: *Information Hiding, First International Workshop*, edited by Ross J. Anderson. Cambridge, U.K., May 30–June 1, 1996, Proceedings. Lecture Notes in Computer Science, Vol. 1174, Springer-Verlag, 1996.

[4] A. Piva, M. Barni, F. Bartolini, V. Cappellini, "DCT-based watermark recovering without resorting to the uncorrupted original image", preprint, 1997.

[5] A. Piva, M. Barni, F. Bartolini, V. Cappellini, "Threshold Selection for Correlation-Based Watermark Detection",

[6] M. Barni, F. Bartolini, V. Cappellini and A. Piva, "A DCT-domain system for robust image watermarking", Signal Processing, Special Issue in Copyright Protection and Access Control for Multimedia Services, (66)3, 1998, pp. 357-372.

[7] J. J. K. Ó Ruanaidh, W. J. Dowling, and F. M. Boland, "Phase Watermarking of Digital Images", *Proc. IEEE Int. Conf. on Image Processing*, vol. 3, pp. 239–242, Lausanne, Switzerland, Sep 1996.

[8] J. J. K. Ó Ruanaidh, F. M. Boland, and O. Sinnen. "Watermarking digital images for copyright protection", *Proc. of the Electronic Imaging and Visual Arts Conference*, Florence, Feb 1996.

[9] J. J. K. Ó Ruanaidh, F. M. Boland, and C. Dautzenberg, "Watermarking Digital Images for Copyright Protection", *Proc. of the IEE Conference on Image Processing and its Applications*, Edinburgh, pp. 326–330, 1995.

[10] J. J. K. Ó Ruanaidh, W. J. Dowling, and F. M. Boland, "Watermarking digital images for copyright protection", *IEE Proc. Vision, Image and Signal Processing*, **143**(4), pp. 250–256, August 1996.

[11] J. J. K. Ó Ruanaidh and T. Pun, "Rotation, scale and translation invariant digital image watermarking", *Proc. of the ICIP*, vol. 1, pp. 536–539, Santa Barbara, California, Oct 1997.

[12] J. Fridrich, "Combining Low-frequency and Spread Spectrum Watermarking", *Proc. SPIE Int. Symposium on Optical Science,*

*Engineering, and Instrumentation*, San Diego, July 19–24, 1998.

[13] J. Fridrich, 2 Lt Arnold C. Baldoza, and Richard J. Simard, "Robust digital watermarking based on key-dependent basis functions", *2nd Information Hiding Workshop*, Portland OR, April 15–17, 1998.

[14] I.J. Cox and Jean-Paul M.G. Linnartz, "Public watermarks and resistance to tampering", *ICIP–97*, October 1997, CD version of Proceedings.

[15] I.J. Cox and Jean-Paul M.G. Linnartz, "Some general methods for tampering with watermarks", *preprint,* 1998.

[16] T. Kalker, "Watermark Estimation Through Detector Observation", Philips Research Eindhoven, Netherland, *preprint* 1998.

[17] T. Kalker, J.P. Linnartz and M. van Dijk, "Watermark Estimation Through Detector Analysis", *ICIP-98*, October 1998.

[18] J.P. Linnartz and M. van Dijk, "Analysis of the sensitivity attack against electronic watermarks in images", in *Proceedings of the 2nd Workshop on Information Hiding,* Portland OR, April 15–17, 1998.

[19] R. J. Clarke, *Transform Coding of Images*. Academic Press, London, 1985.