

# Full Spectrum Coding: A Machine Readable Marking Method for Security Document Images

Hans Oltmans

*Joh. Enschedé Security Print, THE NETHERLANDS*

A method of providing a digital image with a unique, machine readable, code image is presented. It is called "Full Spectrum," because the method uses Fourier transform techniques to embed the code in a wide range of spatial frequencies. The changes made to the original image by the encoding process are meaningless to the observer, and, by proper choice of the embedding parameters and the resolution (reproduction size) of the marked image, they can be made imperceptible. Full Spectrum has some intrinsic advantages for application to printed security structures in a document authentication and identification environment. Using its mathematical properties, the method is shown to be invariant to shifting and robust to cropping, which enables the code image to be reconstructed from a recording of the document with arbitrary position and size. Techniques to deal with possible rotation and scaling of the recorded image with respect to the (printed) original are elaborated. Finally, a general method is developed to match the reconstructed code image to the reference image (expected code image). Experiments show that the code image can survive various graphical transfer processes, such as halftone screening, printing and digitizing. An actual document containing a printed Full Spectrum structure, the *Security ID*, is used as a practical example.

Journal of Imaging Science and Technology 49: 451–463 (2005)

## Introduction

As a consequence of considerable improvements and common availability of digital graphical systems, the authenticity, (copyright) ownership and integrity of digital data are of major concern. This is not only true for image data, but also for video, sound, text, technical drawings etc. Therefore protection of all this *digital content* is an important and rapidly growing field of research. Many of the techniques which have been developed in this field can be comprised by the collective term *digital watermarking*.<sup>1–6</sup> A possible definition is:<sup>6</sup> "Digital watermarking is the practice of (imperceptibly) altering a *work* to embed a *message* about that *work*." Here: a *work* pertains to any kind of content (in the graphical case an image) and the *message* contains some annotation about (the owner, status, etc. of) the content. In the graphical case it will mostly be a mark of authenticity or an ID confirmation. Notice the similarity with a "real" banknote watermark: there also the watermark is thematically related to the printed image and it should be checked (in this case visually) to demonstrate that they correspond. There is a wide range of conceivable applications, e.g.,

- Proof of ownership and copyright enforcement: of vast importance for digital documents, ranging from artwork to scientific papers, for example on the internet.
- Copy control: preventing content to be copied by equipping copying devices with a *watermark detector*.
- Tracking: distribution of different copies to be able to identify the source of piracy in a later stage.

The high quality and resolution which is achievable by modern printing and scanning devices enables the application of digital watermarking techniques in printed structures. This development gives new opportunities for protection of security documents such as IDs, drivers licences etc. Instead of employing special optical devices or measurement systems, authentication or identification can be performed on the basis of a recording or scan of the document, i.e., a digital picture. Features can be extracted and verified from this picture by image processing techniques, which are invisible to, or unrecognizable by the observer. Applying such "covert image coding" in security documents also has some advantages for the producer:

- No special inks or printing technologies are needed.
- Embedding into an arbitrary (raster) design is possible as the visual appearance is not (or is hardly) affected.

In this article, a coding method is developed that is specially tailored to graphical applications. Using general principles of spectral signal analysis, the method is made, to some extent, resistant to the kind of information loss occurring in practical processes.

## Definitions

### Encoding Method

A frequency coded signal is composed of two real signals (time series, images etc.)

$$f(x) = a(x) + b(x), \quad (1)$$

Original manuscript received May 13, 2004

where  $a(x)$  is the original signal containing the “overt information” and  $b(x)$  is the added signal containing “hidden information,” which may depend on  $a(x)$ . The description is given in one dimension here, but can directly be extended to 2D signals. The signal  $b(x)$  should be detectable but not disturbing. Following *information hiding* terminology<sup>1–8</sup>  $a(x)$  is the *cover* signal,  $f(x)$  is the *coded* or *marked* signal and  $b(x)$  is the *stego* signal. Many methods exist to hide information into signals (images) using *transform domain embedding*. The general formalism<sup>9</sup> is

$$f(x) = \mathcal{T}^{-1} \mathcal{C} \mathcal{T} a(x), \quad (2)$$

where  $\mathcal{T}$  is a signal transform and  $\mathcal{C}$  represents the operation that embeds the code.

In the present implementation, the following approach is used: A signal  $c$  is constructed of which the Fourier transform  $C(\omega)$  is a real-valued signal, which is itself regarded as a detectable “pattern.” Because  $C(\omega)$  is real, its inverse transform  $c(x)$  is even. Conversely, we demand  $c(x)$  to be real, which makes  $C(\omega)$  even. Then the modulus of the Fourier transform of the original signal is modulated by  $C(\omega)$ :

$$F(\omega) = [|A(\omega)| + \mu C(\omega)] e^{i\phi(\omega)} = \left[ 1 + \mu \frac{C(\omega)}{|A(\omega)|} \right] A(\omega), \quad (3)$$

where  $\phi(\omega) := \angle A(\omega)$  (the phase) and  $\mu$  is a real positive constant and the symbol  $:=$  denotes as asymmetric operator. Note that the phase factor is undetermined when  $A(\omega)$  vanishes. Comparing with Eq. (1) we see that

$$B(\omega) = \mu C(\omega) e^{i\phi(\omega)} = \mu \frac{A(\omega)}{|A(\omega)|} C(\omega). \quad (4)$$

The Fourier transform of the stego signal  $b(x) = f(x) - a(x)$  has modulus  $\mu C(\omega)$ , whereas its phase is equal to the phase of the Fourier transform of the original signal  $a(x)$ . In terms of the formalism of Eq. (2), the transform  $\mathcal{T}$  is the Fourier transform  $\mathcal{F}$  and the code operation  $\mathcal{C}$  is adding  $\mu C$  to the modulus of  $A$ .

### Alternative Schemes

Sometimes, instead of Eq. (3), the “additive magnitude modulation” (AMM), a “multiplicative modulation” (MMM) is proposed<sup>10,11</sup>:  $|F(\omega)| = |A(\omega)| [1 + \nu C(\omega)]$ . In MMM notation, we observe that  $B(\omega) = \nu C(\omega) A(\omega)$ . This means that  $b(x) = \nu c(x) * a(x)$ , i.e., the stego signal equals the convolution of  $c(x)$ , the inverse transform of the code  $C(\omega)$ , with the cover signal  $a(x)$ . Hence, for Fourier based encoding, the operation performed by MMM can be considered as a convolution filter  $f = h * a$  where the impulse response is the possibly the very complicated signal,  $h(x) = \delta(x) + \nu c(x)$ , which is independent of  $a(x)$ . This method may be vulnerable to deconvolution techniques.

Similar reasoning can be applied to a second alternative proposed in literature, “exponential magnitude modulation” (EMM). Here  $|F(\omega)| = |A(\omega)| \exp[\nu C(\omega)]$ . Using the series expansion of the exponential function, it can be seen that, for  $\nu C \ll 1$ , EMM is equivalent to MMM. EMM also has an interpretation in the space domain, using the the *cepstrum*, by definition<sup>9</sup>

$$\tilde{a}(x) \xrightarrow{\mathbf{F}} \log A(\omega).$$

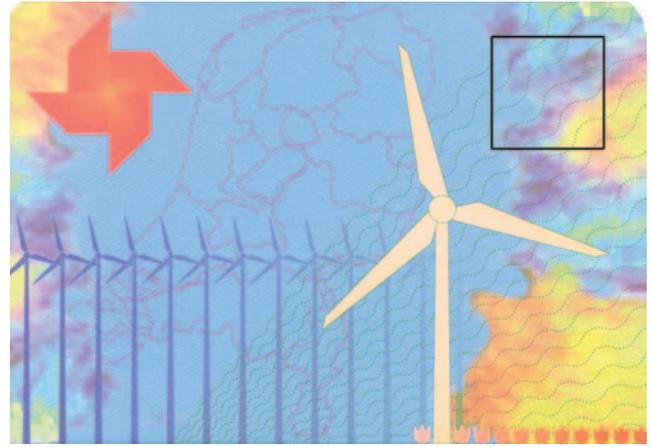


Figure 1. Security ID

This operation has special theoretical application in signal processing because it transforms a convolution into an addition. For EMM, the cepstrum of  $f$  is

$$\tilde{f} = \tilde{a} + \nu c.$$

In the example below, and in the remainder of this article, only the AMM case will be considered.

### Example

The *Security ID* is an example document containing a Full Spectrum (FS) code image. A page of it is shown in Fig. 1. The background of the map of the Netherlands (blue in a color reproduction) carries the FS structure.

The cyan printing phase of the four-color separation, the gray value image *Nederland-C*, has been used to embed the code. A  $256 \times 256$  sub-image thereof is the cover image, see Fig. 2 (left panel). Its location in the document is demarcated by a square frame in Fig. 1.

The code image is the  $256 \times 256$  image *Eye*, see Fig. 2 (center panel). This image is already even in the DFT sense. Using AMM, it has been embedded with strength  $\mu = 1/4$ , yielding the marked image, see Fig. 2 (right panel). This value is based on the DFT notation with “symmetric normalization” as in Ref. 13 for example. In the graphic display, the *optical notation* is used, meaning that the origin of the coordinate system is in the center of the picture.

The marked CMYK image *Nederland* has been printed, in various techniques, at 120 pixels/cm (approximately 305 dpi), which means that the actual size of the images shown here is  $21.3 \times 21.3$  mm.

### Properties

Full Spectrum coding can be seen as a digital watermarking method, possessing some useful mathematical properties with respect to the two basic characteristics<sup>6,12</sup> for any such coding scheme:

- *fidelity/transparency*, the “imperceptibility” of the code,
- *invariance/robustness*, the extent in which the code survives alterations of its “carrier.”

Using only general properties of the Fourier transform (FT), it will be shown that the (amplitude of the) “noise”



**Figure 2.** Left-to-right: Cover image ( $a$ ); code image ( $C$ ); and marked image ( $f$ ).

introduced by the FS coding, as represented by the fluctuation of the stego signal, can be expressed in, and hence controlled by, the input parameters, i.e., the embedding strength and the code signal. This relationship follows from the direct correspondence between the lowest statistical moments (the mean and the variance) of the stego signal and the input. Furthermore, the characteristic frequency of the stego signal is closely related to the (scale of the) code signal. The noise amplitude and frequency determine the transparency of the code.

Besides these properties pertaining to the signal values (gray or luminance levels for images), the theory of Fourier transformation also provides expressions for the influence on the FS code signal of “geometrical” operations performed on the marked signal. It is an advantage for a feature that is intended to be automatically detectable, that the code turns out to be invariant to the basic operations of shifting (translation) and cropping (cutting), and that the influence of many other operations is predictable from the theory. The behavior of the code with respect to these operations determines the (detection) robustness.

### Statistical Properties of Stego Signal

It follows from the definition of the Fourier transform that

$$B(0) = \int b(x) dx =: \bar{b}, \quad (5)$$

the “average value” of the stego signal, using Eq. (4) this equals  $\mu C(0)$ . Again, the symbol  $:=$  denotes an asymmetric operator. Hence, if we choose the central value of the code  $C(0)$  to be zero, which we will preferably do in order not to affect the low frequencies in  $f(x)$ , the added stego signal has zero mean. Parseval’s relation<sup>14</sup> gives

$$\overline{b^2} := \int |b(x)|^2 dx = \frac{1}{2\pi} \int |B(\omega)|^2 d\omega. \quad (6)$$

Using Eqs. (4) and (5), we obtain an expression for the “variance” or “energy” of the stego signal:

$$E := \overline{b^2} - \bar{b}^2 = \mu^2 \left\{ \frac{1}{2\pi} \int [C(\omega)]^2 d\omega - [C(0)]^2 \right\}. \quad (7)$$

For the special case that  $C(\omega)$  is a binary signal with finite support which does not include the origin, i.e.,  $C(0) = 0$ , we get  $E = \mu^2 S / (2\pi)$ , where  $S$  is the measure of the support of  $C(\cdot)$ , “the white area” for a b/w image.

The quantity  $\sqrt{E}$  can be considered a measure of the “code strength”.

The general form of the stego signal follows from the Fourier transform relationship  $B(\omega) \stackrel{F}{\leftarrow} b(x)$ . Using Eq. (4) and the properties of  $C(\cdot)$ , this can be written as

$$\begin{aligned} b(x) &= \frac{\mu}{2\pi} \int_{-\infty}^{\infty} C(\omega) e^{i[\phi(\omega) + \alpha x]} d\omega \\ &= \frac{\mu}{\pi} \int_0^{\infty} C(\omega) \cos[\phi(\omega) + \alpha x] d\omega. \end{aligned} \quad (8)$$

It must be borne in mind that the description here in terms of continuous, infinite extent and infinite range signals is an idealization. In the practice of (digital) signal processing, the range of values that can be attained will be finite and discrete, typically  $f(\bar{x}) \in \{0, 1, \dots, Q\}$ , where  $\bar{x}$  denotes a discrete coordinate pair and  $Q = 2^k - 1$ . Hence the signal that is in fact digitally stored is clipped and quantized:  $f_d(\cdot) = \mathcal{Q}(f(\cdot))$ , where

$$\mathcal{Q}(z) = \begin{cases} 0 & (z < 0), \\ \lfloor z \rfloor & (0 \leq z < Q), \\ Q & (z \geq Q). \end{cases} \quad (9)$$

Of course  $F_{f_d}$ , the DFT of  $f_d$ , will differ somewhat from  $F$ , the DFT of  $f$ .

It can be concluded that the *amplitude* of the stego signal is only governed by the *volume*, or integral, of the squared code image. Moreover, the *shape* of the code image, i.e., the location of its nonzero values, will determine the characteristic *frequency* of the stego signal fluctuations. We can define the (mean square) radial frequency  $\omega_c$  as

$$\omega_c^2 := \frac{\int |\bar{\omega}|^2 C(\bar{\omega}) d\bar{\omega}}{\int C(\bar{\omega}) d\bar{\omega}} = \frac{[-\nabla^2 c](\bar{0})}{c(\bar{0})}, \quad (10)$$

where vector notation is used to express the validity in 2D. (Note that the mean square radius can be seen as the “moment of inertia” of the code image).

### Example

The stego image for the *Security ID* is shown in Fig. 3. It has been calculated by subtracting the cover image

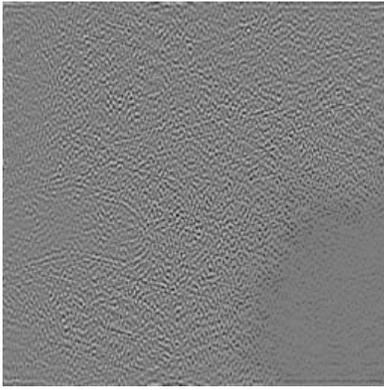


Figure 3. Stego image, i.e., stego signal (b) with +128 offset

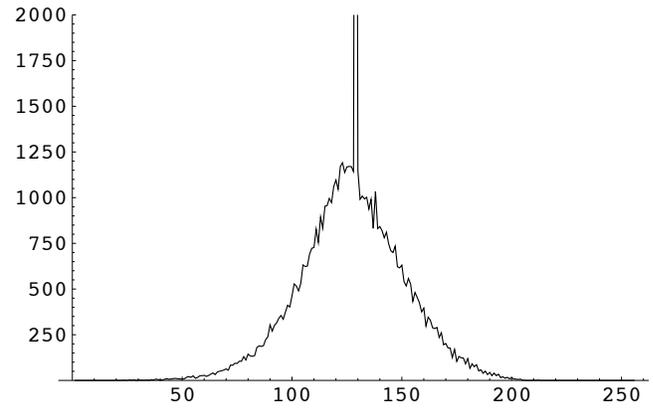


Figure 4. Gray value histogram of the stego image

Fig. 2 (left) from the marked image, Fig. 2 (right). As  $b$  has both positive and negative values, an offset of 128 has been added to each pixel value, to keep the gray values in the range  $\{0, \dots, 255\}$ .

The lower right part of the cover image is saturated. In this white area only negative modulation is possible, the positive values of the stego signal are clipped, which causes many zeroes (pixels with gray value 128) in this area. This is reflected by the large spike in the gray value histogram, Fig. 4. The average gray value is approximately 126, indicating that the mean of the stego image is  $-2$ , slightly smaller than the expected value 0, following from the code image. The standard deviation is 23. The expected value can be calculated from Eq. (7). For binary digital images this becomes  $\sqrt{E} = \mu Q \sqrt{S}$  where  $S \in (0, 1)$  is the relative area of the foreground (the ratio between the number of pixels where  $C = Q = 255$  and the total number of pixels). For the code image of Fig. 2 (center),  $S \approx 0.216$  and hence  $\sqrt{E} \approx 30$ . The error in the estimation of the mean and the variance is a consequence of the digitization, Eq. (9). Figure 5 shows the magnitude of the DFT of the marked image (a contrast stretch was applied to enhance visibility). The digitization effects can be clearly discerned. Finally, Eq. (10) gives the characteristic frequency of the stego signal generated by the code image  $C$ . Expressed in cycles/pixel, it equals  $\omega_c/(2\pi) \approx 0.317$ ; using a print resolution of 120 pixels/cm (see the first example above), this means that the actual mean square frequency in Fig. 3 is approximately 38 pixels/cm.

### Geometrical Properties

Assume that some digital recording  $f'$  of the image  $f$  is obtained. Because of the way (AMM) Full Spectrum is built up, it should be possible to reconstruct the code image by taking the modulus of the DFT of the input image,  $I := |F'|$  and performing pattern recognition on this “image.” However, the reconstruction will be hampered by three causes:

- “Pollution” of  $I$  by the FT of the cover image. Since the code extraction is required to be an *oblivious* (or *blind*)<sup>6</sup> procedure, subtraction of  $a$  is not an option.
- Quantization errors introduced by the digitization, Eq. (9).
- Information loss due to D/A and A/D conversions, i.e., “process noise.”

Below, the effects of signal operations that can occur in the reproduction and recording process will be investigated from a theoretical point of view.

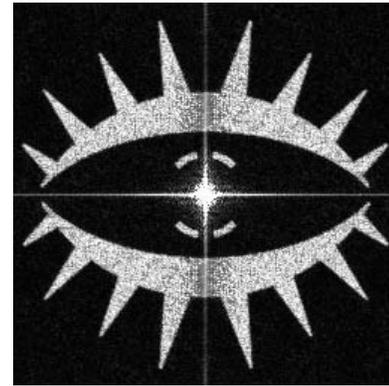


Figure 5. Modulus of the DFT of the marked image

### Shift Invariance

An advantage of the use of (Fourier based) magnitude modulation as described here is the translation invariance. Theoretically, the FT of a shifted image only differs from that of the original image in phase.<sup>14</sup> The magnitude is the same:

$$f(x - x_0) \xrightarrow{\mathbf{F}} e^{-i\alpha x_0} F(\omega). \quad (11)$$

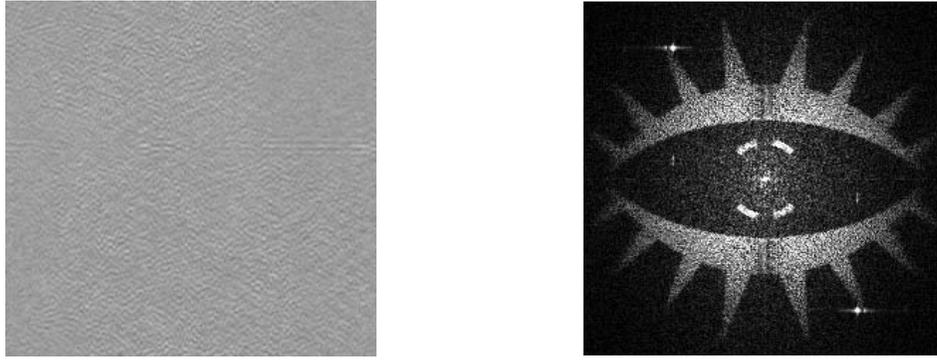
This means that in the reconstructed “image”  $I$ , no searching is necessary. The image, although noisy, will be at the same position.

### Crop Robustness

A quite remarkable property of FS coding is that it exhibits, to some extent, cropping invariance. Due to the properties of the Fourier transform, even a small part of the FS image contains a complete, although blurred, version of the code image. Cropping can be modeled as multiplication of the signal by a window  $p_\kappa(x)$ , which is unity for  $|x| \leq \kappa$  and zero otherwise. The modulation property of the FT gives<sup>14</sup>

$$f(x)p_\kappa(x) \xrightarrow{\mathbf{F}} \frac{\sin \kappa\omega}{\pi\omega} * F(\omega), \quad (12)$$

where the convolution kernel on the RHS “tends to  $\delta(\omega)$ ” for  $\kappa \rightarrow \infty$ . Although detail of the code image will be lost due to this filtering, the global shape is contained in the



**Figure 6.** Sub-image of 305 dpi scan (red channel) of the *Security ID* (left) and its DFT modulus (right).

image fragment. In information hiding terms, the code is *robust* to cropping. If confronted with limited data capacity or computation speed, one could better scan a part of the FS image at high resolution (crop) than the whole image at low resolution (interpolate). The reason is that the FT of the image fragment will yield a blurred code image, whereas the interpolated image will reveal only the central part of the code image.

#### Scaling Property

Scale invariance is usually not an important issue in document processing. The size of a coded image is generally known in advance, hence the size of the recorded image  $f$  in pixels only depends on the scan resolution, say  $R_s$ . The magnification factor is  $R_s/R_o$ , where  $R_o$  is the resolution with which the image is printed.

An FS code image, however, exists in the frequency domain. This means that scale in the code image is related to frequency in the marked image and vice versa. A dot at (relative) position  $\xi$  in the code image results in a component with (normalized) frequency  $\zeta := \omega/(2\pi)$  in the FS image. When the FS image is printed with resolution  $R_o$ , this component will be represented by a spatial frequency  $\zeta R_o$ , a constant. After scanning with resolution  $R_s$ , satisfying the Nyquist criterion, i.e.,  $R_s \geq 2|\zeta|R_o$ , this component will be observed as having a normalized frequency  $\zeta R_o/R_s$ . Hence the “magnification factor” is inverted with respect to spatial coding, which is intuitively understandable. Formally the above follows from the reciprocal scaling property of the FT<sup>14</sup>

$$f(\lambda x) \xrightarrow{\mathbf{F}} F(\omega/\lambda)/|\lambda|, \quad (13)$$

where  $\lambda$  is the ratio of the resolutions here. In practice one works with  $N$ -point DFTs (where  $N$  is usually but not necessarily a power of 2) and the scaling formula becomes

$$\frac{p_s}{p_o} \approx \frac{N_s}{N_o} \frac{R_o}{R_s} = \frac{L_s}{L_o}, \quad (14)$$

where  $p$  is the discrete frequency  $\lfloor N\xi \rfloor$  as used in the DFT formalism and  $L = N/R$  is the physical size of the recorded signal. It follows that the pixel position of a point in the code image does not change if we scan the same area with another resolution, i.e., change  $N$  and  $R$  by the same factor.

#### Example

A scan of the *Security ID* will serve as a qualitative example for the invariance properties. The document has been recorded on a flatbed scanner with 305 dpi, equal to the original printing resolution  $R_o$ . The red channel of the RGB image is taken, as this contains the information of the cyan color phase; see the first example above. Shift invariance can be illustrated by taking an *arbitrary*  $N \times N$  ( $N = 256$ ) sub-image, see Fig. 6 (left panel), with physical size  $21.3 \times 21.3$  mm.

Its DFT magnitude, shown in Fig. 6 (right panel), clearly resembles the DFT magnitude of the original marked image (Fig. 5). (A numerical way to compare these to the original code image of Fig. 2 will be treated in the following sections.)

Scanning the document with a resolution different from  $R_o$ , for example  $R_s = 350$  dpi, and considering again an  $N \times N$  sub-image as in Fig. 7 (left panel), produces an example of scaling. As  $N$  is fixed, the image has also been cropped by a factor  $R_o/R_s$  with respect to the original size.

It follows from Eq. (14) that the associated code image will scale (in this case shrink) by this same factor. This can be verified from the right panel of Fig. 7, which shows the DFT magnitude.

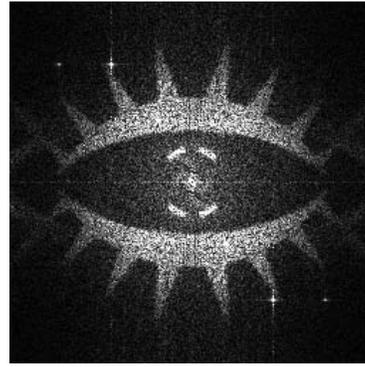
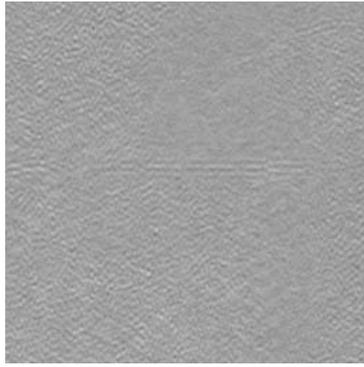
After further cropping by a factor of two in both dimensions, resulting in the  $128 \times 128$  sub-image shown in Fig. 8 (left), the appearance of the code image is still reasonably preserved; see the DFT on the right.

#### Influence of Image Operations on Magnitude Spectrum

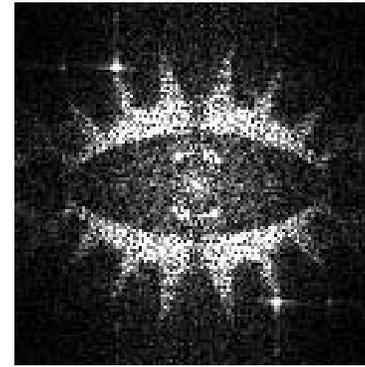
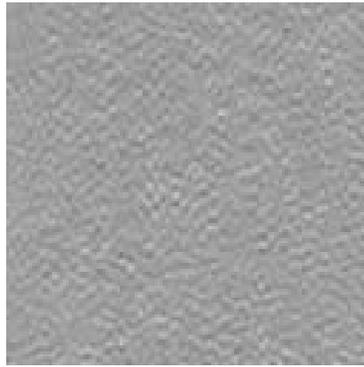
We have seen that, theoretically, an FS code image is shift and, more or less, crop invariant. Furthermore, linear scaling (resampling) of the image results in scaling of the FT and hence of the code image by the inverse factor. To address the last basic operation in image processing, rotation, we have to drop the description in 1D and consider 2D functions. If  $g(x, y)$  is the same image as  $f(x, y)$ , but rotated by an angle  $\alpha$  and scaled by a positive constant factor  $\lambda$ ,  $g(x, y) = f(\lambda(x \cos \alpha + y \sin \alpha), \lambda(-x \sin \alpha + y \cos \alpha))$ , then the Fourier transforms  $F(u, v)$  and  $G(u, v)$  are related by<sup>15</sup>

$$G(u, v) = F((u \cos \alpha + v \sin \alpha)/\lambda, (-u \sin \alpha + v \cos \alpha)/\lambda)/|\lambda|^2. \quad (15)$$

In most cases, when scanning the image to be detected, these operations do not need to be considered because



**Figure 7.** Sub-image of 350 dpi scan (red channel) of the *Security ID* (left) and its DFT modulus (right).



**Figure 8.**  $128 \times 128$  pixel sub-image of 350 dpi scan (left) and its DFT modulus (right).

the orientation of the document and the resolution are fixed, i.e., there is no rotation and the scaling factor follows from the ratio of  $R_s$  and  $R_o$ , see Eq. (14). It is possible, however, that these parameters are not known, for example when inspecting a document positioned obliquely with respect to the image recording device or when recording an image with unknown magnification. Many articles have been written on the subject of recognition with invariance to one or more of these (and many other) operations. In practice some of these operations will depend on each other: a trivial example is that translation and scaling will induce cropping when the recorded area remains unchanged. For the sake of simplicity, however, they are treated separately here. In practice, image operations or alterations will occur frequently, whether deliberately or inadvertently. Table I lists the theoretical influence of commonly used image operations on the modulus of the Fourier transform of the image  $f$ . Assuming the image  $f$  contains a Full Spectrum code image, this determines also the influence on this code image  $C$ . A question mark in the second column of the table indicates that no easy expression for this influence can be found.

The following notation is used: As before,  $f$  is the original image and  $F$  its Fourier transform,  $f \xrightarrow{F} F$  and  $g$  is the image after undergoing the operation. The argument of  $f$ , “time” or “position”, is regarded as a 1D variable  $x$  and, equivalently, the argument of  $F$ , the circular frequency  $\omega$ . Only when it is relevant, vector notation  $\vec{x}, \vec{\omega}$  is used, to denote higher dimensional

variables. In 2D, the components are denoted by  $(x, y)$ ,  $(u, v)$  respectively. Mathematically,  $f$  can be considered as a map from the set of pixel positions  $P$  to the set of (gray) values  $Q$ , so  $f: P \rightarrow Q$ . Roughly, the common image operations can be categorized as:

- *geometrical operation*: a map  $v: P \rightarrow P$ . It involves only the position of the pixels. Because  $y = v(x)$ ,  $g(y)$  is only well-determined if  $v$  is invertible:  $g(y) = f(x) = f(v^{-1}(y))$ . For an actual pixel image, values are only available for a discrete set of positions. In most cases interpolation will be necessary to obtain  $g(y)$  from the values of  $f$  in the neighborhood of  $x = v^{-1}(y)$ .
- *histogram or gray value operation*: a map  $w: Q \rightarrow Q$ . It involves only the pixel values. The function  $w$  is not necessarily invertible.

General operations are maps on the whole set of pixel values “ $Q^P$ ”.

### Detection Method

Full Spectrum coding aims at automatic detection, i.e., machine readability. This means that the reconstructed image  $I = |F'|$  should be compared to the reference image (the code image to be found) in a numerical way, without human intervention (visual judgment). Usually such a comparison is carried out by evaluating a *match score*, a statistic which is large if the signals are “close” and small if they are “distinct.” If, in a specific detection case, the score exceeds a certain value, the threshold, the signals are assumed to match.

**Table I. Influence of Image Operations on FS Code Image**

operation	influence on FSCI	formulas and remarks
<b>Geometrical operations</b>		
shift (translate)	none	$g(x) = f(x - x_0), G(\omega) = e^{-i\omega x_0} F(\omega)$
crop	blur with “sinc kernel”	$g(x) = f(x)p_\kappa(x), G(\omega) = \frac{\sin \kappa\omega}{\pi\omega} * F(\omega)$
scale (zoom)	scale by inverse factor	$g(x) = f(\lambda x), G(\omega) = F(\omega/\lambda)/ \lambda $
rotate	rotate by same angle	$g(x, y) = f(x \cos \alpha + y \sin \alpha, -x \sin \alpha + y \cos \alpha),$ $G(u, v) = F(u \cos \alpha + v \sin \alpha, -u \sin \alpha + v \cos \alpha);$ in FS code image rotation is modulo $\pi$
rotate by $\pi$	none	$g(x, y) = f(-x, -y)$ (is reflection in origin), $ G(u, v)  =  F(-u, -v)  =  F(u, v) $ by the conjugate symmetry of $ F $ .
mirror	mirror in same axis	for example mirror in $y$ -axis: $g(x, y) = f(-x, y),$ $G(u, v) = F(-u, v)$
general linear	inverse transpose	$g(\vec{x}) = f(\mathbf{V}^{-1}\vec{x})$ ( $\mathbf{V}$ is invertible matrix), $G(\vec{\omega}) =  \mathbf{V} F(\mathbf{V}^T\vec{\omega});$ scale, rotate, mirror are special cases.
nonlinear	?	$g(\vec{x}) = f(\vec{v}^{-1}(\vec{x}))$ ( $\vec{v}$ is invertible vector field), $G(\vec{\omega}) = \int f(\vec{x}) e^{-i\vec{\omega}\cdot\vec{v}(\vec{x})}  (\partial\vec{v}/\partial\vec{x})  d\vec{x}$
<b>Histogram operations</b>		
shift (add constant)	none (except for $\omega = 0$ )	$g(x) = f(x) + a$ ( $a$ is real constant), $G(\omega) = F(\omega) + a\delta(\omega),  G(\omega)  =  F(\omega) $ ( $\forall \omega \neq 0$ )
stretch, shrink	same	$g(x) = bf(x)$ ( $b$ is real constant), $G(\omega) = bF(\omega),  G(\omega)  =  b  F(\omega) $
invert	none (except for $\omega = 0$ )	$g(x) = Q - f(x)$ ( $Q$ is maximum grey value), $G(\omega) = Q\delta(\omega) - F(\omega),  G(\omega)  =  F(\omega) $ ( $\forall \omega \neq 0$ )
general linear	stretch / shrink	$g(x) = bf(x) + a,  G(\omega)  =  b  F(\omega) $ ( $\forall \omega \neq 0$ ); shift, stretch, shrink, invert are special cases.
nonlinear	?	$g(x) = w(f(x))$ ( $w$ is “curve”), $G(\omega) \stackrel{\mathcal{F}}{\leftarrow} w \circ f(x)$
<b>Filter operations</b>		
convolution	filter multiplication	$g(x) = h(x) * f(x)$ ( $h$ is impulse response), $G(\omega) = H(\omega)F(\omega), h \stackrel{\mathcal{F}}{\rightarrow} H$
special cases:		
Gaussian blur	Gaussian window	$h(\vec{x}) = e^{-\vec{x}\cdot\mathbf{G}^{-1}\vec{x}/2}/(2\pi\sqrt{ \mathbf{G} })$ ( $\mathbf{G}$ is Cov matrix), $H(\vec{\omega}) = e^{-\vec{\omega}\cdot\mathbf{G}\vec{\omega}/2}$
(2D) uniform filter $\odot$	circular window	$h(\vec{x}) = \omega_0 J_1(\omega_0 \vec{x} )/(2\pi \vec{x} )$ (“Airy pattern”), $H(\vec{\omega}) = 1$ for $ \vec{\omega}  \leq \omega_0$ and 0 otherwise
(2D) uniform filter $\square$	square window	$h(x, y) = \sin(\omega_0 x) \sin(\omega_0 y)/(\pi^2 xy)$ (“2D sinc”), $H(u, v) = 1$ for $ u ,  v  \leq \omega_0$ and 0 otherwise
Laplace	parabolic HPF	$g(\vec{x}) = -\nabla^2 f(\vec{x}), G(\vec{\omega}) =  \vec{\omega} ^2 F(\vec{\omega})$
unsharp mask	enhance high freqs.	$g(\vec{x}) = f(\vec{x}) - \epsilon \nabla^2 f(\vec{x}), G(\vec{\omega}) = (1 + \epsilon \vec{\omega} ^2)F(\vec{\omega})$

The invariance properties of the FS code considered above are very well suited for a machine readable feature, as, at least theoretically, the operations for which  $I$  is invariant should have no influence on the match score. For other operations, however, the changes made to  $I$  have to be found and compensated for before

the matching can take place. For geometrical operations, this process is called *registration*.

**Registration**

To enable comparison of the reconstructed image  $I := |F'|$  with the reference image, the problem of registration has

to be solved first. Here it is assumed that the possible operations are limited to scaling and the basic rigid ones of rotation, translation, and possibly, cropping. As explained above, registration is required only with respect to the first two operations, because, for FS coding, the problems of translation and cropping are more or less solved implicitly by the coding method. Of course this is only true in practice for a limited range of values: large shifts will take the coded image outside the field of view of the recording device and extreme small portions of the image will not contain enough information to reconstruct the code image.

Several approaches are conceivable. It is possible to use visible marking symbols in the document and perform pattern recognition on the recorded image to find their mutual position/orientation. Here, however, we propose to start from the FS code itself and carry out the registration in the frequency domain.

### Log-Polar form of Spectrum

There exists a well-known coordinate transformation which is rotation and scale invariant in the sense that a rotation or scaling of an image  $f$  causes a *translation* of the transformed image. Consider the expression for the FT of a scaled and rotated image Eq. (15). In *log-polar* coordinates,<sup>16</sup>

$$F_{LP}(\tau, \varphi) = F(e^\tau \cos \varphi, e^\tau \sin \varphi), \quad (16)$$

scaling of the image (and thus its FT) transforms into a linear shift of  $F_{LP}$  along the  $\tau$ -axis and rotation into a circular shift, i.e., modulo  $2\pi$ , along the  $\varphi$ -direction:

$$G_{LP}(\tau, \varphi) \propto F_{LP}(\tau - \log \lambda, \varphi - \alpha). \quad (17)$$

The polar angle  $\varphi$  naturally varies between 0 and  $2\pi$ . Because of the conjugate symmetry, however  $F_{LP}(\tau, \varphi + \pi) = F(-u, -v) = F^*(u, v) = F_{LP}^*(\tau, \varphi)$ , so  $|F_{LP}|$  is  $\pi$ -periodic in  $\varphi$ , which means that we only have to consider one half of the frequency plane.

An elegant notation is obtained by regarding the frequency vector as a complex number  $w := u + iv$ . Then the log-polar transformation is equivalent to taking the complex logarithm  $\chi := \tau + i\varphi = \log w$ . Defining  $\bar{F}(u + iv) := F(u, v)$ , Eq. (16) can be written as  $\bar{F}_{LP}(\chi) = \bar{F}(e^\chi)$ . In this way, rotation and scaling is just a multiplication by  $\lambda e^{i\alpha}$  in the complex  $w$  plane and addition of  $\log \lambda + i\alpha$  (modulo  $2\pi i$ ) in the complex  $\chi$  plane (strip).

Sometimes it is proposed<sup>6,12</sup> to convert this property into an actual invariance by taking a second FT of the log-polar image. As for the translation invariance of the  $|F|$ , the translation of the  $|F_{LP}|$  caused by the rotation and scaling will only affect the phase of this second FT. The magnitude should be invariant. A transformation that combines the coordinate and Fourier transform is the so-called *Fourier-Mellin Transform*, see Ref. 17, for example. The combined invariance of a code with respect to the three basic operations is denoted as *RST invariance*.<sup>15,17</sup>

### Register Mark

The log-polar transformation provides a method of determining the rotation and scaling undergone by an FS image. Because these operations are transformed into simple translations, the values  $\lambda$  and  $\alpha$  can theoretically be found by linear cross-correlation.<sup>18,19</sup>

Imagine that a “marking signal”  $W_{LP} := Z(\tau, \varphi)$  is incorporated into the log-polar transformed code image  $C$ . We do not exactly specify the functional shape of  $Z$

yet, but make it understood that it can “easily” be located by correlation. For example, a separable pattern can be implemented by taking  $Z(\tau, \varphi) = v\zeta_1(\tau)\zeta_2(\varphi)$  where  $v$  is some constant and the functions  $\zeta_i(\cdot)$  are, possibly identical, 1D “perfect correlation signals”:  $\zeta(x) \otimes \zeta(x) := \zeta(x) * \zeta^*(-x) \rightarrow \delta(x)$ . After the reverse transformation according to Eq. (16), the mark appears as a signal  $W(u, v)$ , located in an preferably “empty” part of both  $C(u, v)$  and  $A(u, v)$ , the spectrum of  $a$  (meaning that the reverse transformed mark  $W$  is disjoint with  $F$ ). The embedding of the mark can be done by addition.

Consequently, the mark should be present, possibly in rotated and scaled form, in the magnitude spectrum of the recorded image  $I = |F'(u, v)|$ . It can be found by matched filtering.<sup>9</sup> The cross-correlation of  $I_{LP}$  with  $Z$

$$\begin{aligned} \Psi(\tau, \varphi) &:= \int_0^\pi \int_{-\infty}^\infty I_{LP}(v, \psi) Z(v - \tau, \psi - \varphi) dv d\psi \\ &= \int |F'(\bar{\omega})| \frac{Z(\log|\bar{\omega}| - \tau, \angle\bar{\omega} - \varphi)}{|\bar{\omega}|^2} d\bar{\omega} \end{aligned} \quad (18)$$

should exhibit a peak at the relative position  $(\log \lambda, \alpha)$ , see Eq. (17). The second expression in this equation allows direct computation of  $\Psi(\tau, \varphi)$  from the original DFT of  $f$  (without interpolation). The integral is over a semiplane, excluding the origin.

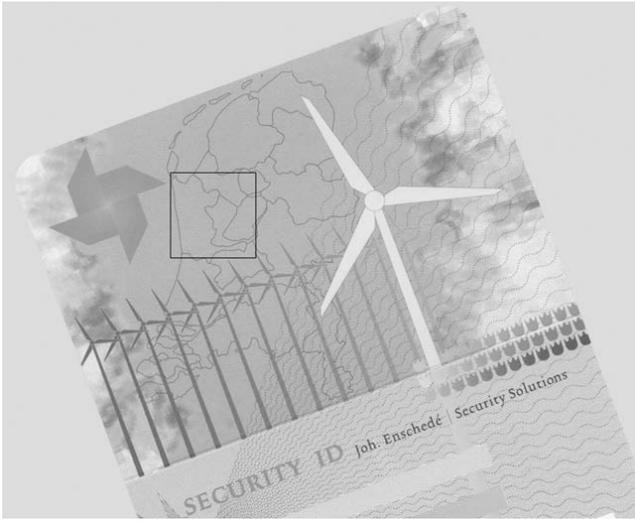
### Example

The *Security ID* contains a “rudimentary” example of a register mark, in the shape of the two (4 by symmetry) small concentric annular segments, see Fig. 2 (center). These segments become rectangles under the log-polar transformation.

To illustrate the effect of rotation and scaling, the *Security ID* was placed obliquely (with random orientation) on the glass platen of a flatbed scanner and scanned with a resolution  $R_s = 600$  dpi. The result is shown in Fig. 9.

A  $512 \times 512$  sub-image thereof is analyzed, (see the square frame in Fig. 9), and the red channel of this sub-image is shown in Fig. 10(a). Its DFT magnitude, shown in Fig. 10(b), is clearly rotated and scaled compared to the digital original, Fig. 5. Figure 10(c) shows the log-polar form of the DFT magnitude. In this picture,  $\tau$  runs from left to right. At the right border of the picture  $\tau := \log(N/2)$ , which corresponds to the maximal “physical” spatial frequency  $R/2$ , the axis stretches over four octaves (frequency doubling intervals) here, so at the left border  $\tau = \log(N/2) - 4 \log 2$ , corresponding to a frequency  $R/32$ . The polar angle  $\varphi$  runs from top ( $\varphi = 0$ ) to bottom ( $\varphi = \pi$ ), where the positive direction is clockwise. The diagram was discretized to  $256 \times 256$  pixels (a rather arbitrary choice), such that the “units” are  $256/4 = 64$  pixels/octave in the  $\tau$ -direction and  $256/180 \approx 1.42$  pixels/degree in the  $\varphi$ -direction. Figure 10(d) shows the log-polar form of the original code image *Eye* of Fig. 2 (center) and Fig. 10(e) that of the DFT magnitude of the marked image, Fig. 5; these two are geometrically identical.

In the diagram of the 600 dpi scan, Fig. 10(c), the rectangular “marks”, although deteriorated by various sources of information loss, are still discernible, shifted approximately 62 pixels = 0.97 octaves to the left, compared to the original. This corresponds to a scaling factor  $\lambda \approx 2^{-0.97} \approx 0.51 \approx R_o/R_s$ . The vertical shift (which is cyclic due to the  $\pi$ -periodicity of  $\varphi$ ) is approximately 31 pixels upwards corresponding to a rotation  $\alpha \approx -22^\circ$ , i.e., counter-clockwise.



**Figure 9.** 600 dpi scan of the *Security ID*, skewed.

### Transformation Invariants

The register mark is a very general tool for registration. By determining its position in the log-polar plane, it is possible to match  $I = |F'|$  to the correct reference image, which is a rotated and scaled version of the code image. Other approaches to rotation and scaling independent recognition are feasible. For example, a code image can be used which is itself rotation invariant. In this case only the scaling problem has to be dealt with. However, instead of sacrificing the generality of the code image, it is also possible to extract invariant feature signals from the code image  $C$  and compare these to the corresponding signals computed from  $I$ . For example the circular profile of a signal  $I$  is defined as

$$\bar{I}_c(w) := \frac{1}{2\pi w} \oint_{|\bar{\omega}|=w} I(\bar{\omega}) ds = \frac{1}{\pi} \int_0^\pi I(w\bar{e}_\varphi) d\varphi \quad (19)$$

(using the point symmetry in the origin). Here  $\bar{e}_\varphi = (\cos \varphi, \sin \varphi)$ , the unit vector in the direction  $\varphi$ . Clearly  $\bar{I}_c(w)$  does not change if  $f$  (and hence  $I$ ) is rotated. Scaling of  $f$  causes scaling of  $I$  and hence  $\bar{I}_c(w)$  by the inverse factor. This scaling can be transformed into a translation by taking  $\tau = \log w$ , as before. Similarly, we can define the radial profile

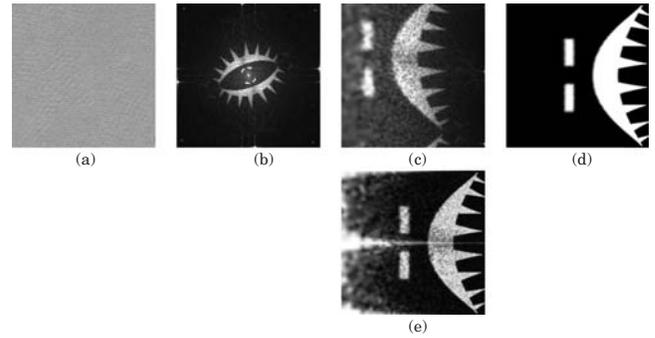
$$\bar{I}_r(\varphi) := \int_{\omega_0}^\infty I(w\bar{e}_\varphi) dw. \quad (20)$$

Here the lower integration limit is  $\omega_0$  rather than 0 to eliminate the influence of the cover image, this will be explained in the next section. Due to the symmetry, the signal is  $\pi$ -periodic. Rotation of  $f$  by  $\alpha$  causes a (circular) shift  $\alpha$  in  $\bar{I}_r(\varphi)$ . Scaling theoretically does not change  $\bar{I}_r(\varphi)$ . Observe that the signals and are just the projections along the coordinate axes of the (log-)polar form of  $I$ .

### Matching

#### The Normalized Inner Product (NIP)

To verify if an input image  $f'$  contains the FS code image  $C$ , we have to find the “correspondence” between



**Figure 10.** Comparison of log-polar magnitude spectra; see text.

its Fourier magnitude  $|F'|$  and  $C$ . Here it is assumed that  $C$  has already been scaled (and possibly rotated) to the proper scanning resolution (orientation). In the case that it differs from the “original” code image, it is better to refer to  $C$  as the *reference image*. A possible way to assess such a correspondence is (linear) *correlation*. Because of the shift invariance, however, we only have to consider the zero lag cross correlation which can be expressed as a normalized inner product

$$\text{NIP} = \frac{(|F'|, C)}{\sqrt{(|F'|, |F'|)(C, C)}}. \quad (21)$$

The inner product is defined as

$$(G(\omega), H(\omega)) := \int G(\omega)H(\omega)d\omega, \quad (22)$$

where the integral is replaced by a normalized sum if the signals are discrete. Because the signals are real in this case, we can also use the “complex inner product notation”

$$\langle G(\omega), H(\omega) \rangle := \int G(\omega)H^*(\omega)d\omega = (G(\omega), H^*(\omega)). \quad (23)$$

Confusion can only arise when complex signals are involved. Note that  $(|F'|, |F'|) = \int |F'(\omega)|^2 d\omega = \langle F', F' \rangle$ , but  $(|F'|, C) = \int |F'(\omega)| C(\omega) d\omega = \langle |F'|, C \rangle \neq \langle F', C \rangle$ . The “auto inner product”  $\langle G, G \rangle$  is sometimes denoted as  $\|G\|^2$ , the (Euclidean) norm of  $G$ . In the definition of Eq. (21),  $\|C\|$  is a constant for a given code image. The quantity that has to be calculated for each input image is  $\text{SIM} := \|C\| \times \text{NIP} = (|F'|, C)/\|F'\|$ , the *similarity*, see Ref. 10 for example. Computation efficiency can be gained by utilizing the symmetry of the signals. The NIP equals unity if and only if  $|F'| \propto C$ . Multiplication of the signals by a constant does not affect the value.

Usually, in the definition of correlation, the signals entering the inner product are first shifted to zero mean, by subtracting the average:  $\hat{f} := f - \bar{f}$ . This is also a good idea for matching, because, if this were not done, matching errors, i.e., deviations of  $|F'|$  from a constant times  $C$ , on the foreground (the part of the code image where  $C > 0$ ) are “penalized,” whereas errors on the background ( $C = 0$ ) are not. The subtraction can be performed by replacing  $C$  in Eq. (21) by  $\hat{C} = C - \bar{C}$  and likewise for  $|F'|$ :

$$\begin{aligned} \text{NIP} &= \frac{(|\hat{F}'|, \hat{C})}{\sqrt{(|\hat{F}'|, |\hat{F}'|)(\hat{C}, \hat{C})}} \\ &= \frac{(|\hat{F}'|, C) - |\bar{F}'| |\bar{C}|}{\sqrt{[(|F'|, |F'|) - |\bar{F}'|^2][(C, C) - \bar{C}^2]}} \end{aligned} \quad (24)$$

The problem with this technique is that the central peak in the spectrum of  $f'$ , which is caused by the cover image, is very strong. As a consequence, a low value of the inner product is obtained even if the similarity is “optically” good. To solve this, the input can be high-pass filtered before computing the correlation, this amounts to replacing  $F'(\omega)$  in Eq. (24) by

$$F'_h(\omega) := [1 - H(\omega)]F'(\omega), \quad (25)$$

where  $H(\omega)$  is the frequency domain representation of a low-pass filter. The filtering can be done already in the pixel domain by taking  $f'_h(x) = (\delta(x) - h(x)) * f'(x)$ . In most cases a simple cut-off window will be used, i.e.,  $H(\omega) = 1$  for  $|\omega| \leq \omega_0$  and 0 otherwise. It can be advantageous, however, to use another kind of filter. If, for example, a Gaussian  $H(\omega) := e^{-\sigma^2 \omega^2 / 2}$  is chosen, the low-pass filter impulse response is also Gaussian,  $h(x) = e^{-x^2 / (2\sigma^2)} / \sqrt{2\pi} \sigma$ , enabling simple implementation (see Table I).

The standard deviation  $\sigma$  is inversely proportional to the filter bandwidth. It can be calculated that for the equal-pass situation, i.e., fixed  $\int |H(\omega)| d\omega$ , the cut-off frequency  $\omega_0$  for the (spherical) uniform filter in  $d$  dimensions and the standard deviation  $\sigma$  are related by  $\omega_0 \sigma = [\Gamma(d/2 + 1)]^{1/d}$ . This parameter should be chosen such that the central peak is suppressed while the code image “survives.” If  $F'(\omega)$  is replaced by  $F'_h(\omega)$  in Eq. (24), we get a match value  $\text{NIP}(\cdot)$  that depends on  $\sigma$  (or  $\omega_0$ ). It is expected that this optimizes for a good choice of the filter parameter.

### Numerical Considerations

Consider Eq. (21) again. As pointed out above, although the code image is clearly recognizable in  $I = |F'|$ , the value of the inner product is mainly influenced by the cover image  $a$ . To see this, let us calculate this value in the ideal case  $|F'| = |A| + |B|$ , where  $A$  is an unperturbed but unknown cover image and  $|B| = \mu C$ . Substituting into Eq. (21), we obtain

$$\text{NIP} = \frac{(|A|, C) + \mu(C, C)}{\sqrt{[(|A|, |A|) + 2\mu(|A|, C) + \mu^2(C, C)](C, C)}}, \quad (26)$$

which becomes unity when  $A \equiv 0$ . Introducing

$$\begin{aligned} \gamma^2 &:= \frac{(|A|, |A|)}{\mu^2(C, C)}, \\ \delta &:= \frac{(|A|, C)}{\sqrt{(|A|, |A|)(C, C)}}, \end{aligned} \quad (27)$$

we can express in dimensionless quantities:

$$\text{NIP} = \frac{1 + \gamma\delta}{\sqrt{1 + \gamma(\gamma + 2\delta)}}. \quad (28)$$

The parameter  $\gamma^2$  is the ratio between the cover image and the stego image “energy”,  $\delta$  is just the NIP of  $|A|$  with  $C$ ; it can be interpreted as the “relative spectral mixing” between cover and code image. Usually  $\delta \ll 1$  by choice of  $C$ . Hence we can approximate the RHS of Eq. (28) by the first terms of its series expansion

$$\text{NIP} = \frac{1}{\sqrt{1 + \gamma^2}} \left[ 1 + \frac{\gamma^2}{1 + \gamma^2} (\gamma\delta) + \frac{1 - 2\gamma^2}{2(1 + \gamma^2)^2} (\gamma\delta)^2 + \dots \right]. \quad (29)$$

Generally  $\gamma$  cannot be made arbitrarily small without influencing the visibility of  $b$ . With the notation of Eq. (23), and using Eqs. (6) and (7), we have

$$\gamma^2 = \frac{\langle A, A \rangle}{\mu^2 \langle C, C \rangle} = \frac{\langle a, a \rangle}{\langle b, b \rangle} = \frac{\bar{a}^2}{\bar{b}^2} = \frac{\bar{a}^2 + \text{Var}(a)}{\bar{b}^2 + E}. \quad (30)$$

A similar calculation can be made for the zero mean variables as in Eq. (24), yielding again Eq. (28) with  $\gamma, \delta$  replaced by  $\hat{\gamma}, \hat{\delta}$ , defined analogous to Eq. (27) in terms of  $|\hat{A}|$  and  $\hat{C}$ . In this case, however, the correspondence to the pixel domain quantities is less clear. The desired effect of the filtering of Eq. (25) will be a strong decrease of  $\gamma$ . The value of  $\gamma\delta$  will not be affected very much.

### Fourier Transform Duality

In the calculation of  $\gamma$  in Eq. (30), Parseval’s relation, Eq. (6) was used to express (auto) inner products of frequency domain variables in terms of space (“pixel”) domain variables. This is also possible for cross inner products by using the *general form* of Parseval’s relation<sup>20,21</sup>

$$\langle A, B \rangle = (2\pi)^d \langle a, b \rangle. \quad (31)$$

Here  $a, b$  are functions of a  $d$ -dimensional variable with  $a \xrightarrow{\mathcal{F}} A, b \xrightarrow{\mathcal{F}} B$ . For  $a = b$  and  $d = 1$ , Eq. (31) reduces to Eq. (6). For example, the mixing parameter can be written as  $\delta = (\psi_a, c) / \sqrt{\langle a, a \rangle \langle c, c \rangle}$  where, by definition,  $\psi_a \xrightarrow{\mathcal{F}} |A|$  and  $c$  is the inverse FT of the code image ( $c \xrightarrow{\mathcal{F}} C$ ) as before.

In the same way, Eq. (31) can be used to write the general formula for the NIP, Eq. (21) as

$$\text{NIP}(|F'|, C) = \frac{(\psi_{f'}, c)}{\sqrt{\langle f', f' \rangle \langle c, c \rangle}} = \frac{\int \psi_{f'}(x) c(x) dx}{\sqrt{\int [f'(x)]^2 dx \int [c(x)]^2 dx}} \quad (32)$$

In this expression  $\psi_{f'}(x) \xrightarrow{\mathcal{F}} I$  is the only signal that cannot be derived directly from the pixel image. The formula for the inverse FT gives  $\psi_{f'}(x) = \int |F'(\omega)| e^{i\omega x} d\omega / (2\pi)$ . If  $|F'|$  in the integrand were replaced by its square, the well-known identity<sup>20</sup> that the squared modulus of the FT, sometimes called the *energy spectrum* of the signal, equals the FT of the *autocovariance function*;  $\psi_{f'}(x) := \int f'^*(y) f'(y + x) dy$  can be invoked to write the NIP in terms of known pixel domain signals. (Referring to this analogy,  $\psi_{f'}(x)$  could be

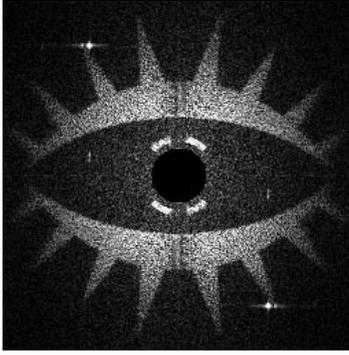


Figure 11. DFT modulus of 305 dpi scan, filtered.

TABLE II. Inner Products for 305 dpi Scan

Inner product	“Plain”	Filtered Uniform $r_0 = 20$	Filtered Gaussian $\sigma = 3$
$ \overline{F'} $	6.644	5.805	5.807
$\overline{C}$	55.028	same	same
$( F' C)$	807.317	807.317	801.063
$( F'   F' )$	24733.368	80.528	79.428
$(C, C)$	13311.728	same	same
NIP	0.0445	0.7797	0.7790
$\widehat{NIP}$	0.0277	0.7030	0.7023

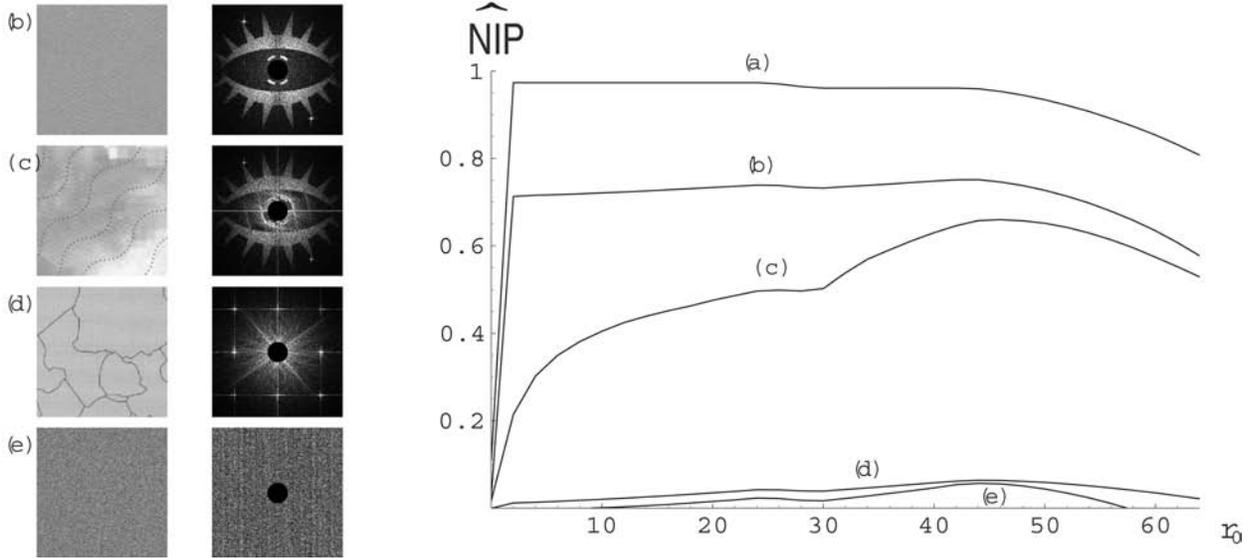


Figure 12. Graph of  $NIP(r_0)$  for (a) Security ID, sub-image of the digital original; (b) scan of the Security ID, “clean” part; (c) scan of the Security ID, “overprinted” part; (d) scan of a document containing the cover image of the Security ID, without FS coding; (e) Gaussian noise image, computer generated. On the left, the corresponding sub-image and its DFT (filtered with  $r_0 = 20$ ) is shown for each case.

called the *rooted* autocovariance function). For a binary code image,  $C^2 \propto C$  and thus:

$$NIP(|F'|^2, C^2) = \frac{(|F'|^2, C)}{\sqrt{(|F'|^2, |F'|^2)(C, C)}} = \frac{(\psi_{f'f'}, c)}{\sqrt{(\psi_{f'f'}, \psi_{f'f'})(c, c)}}. \quad (33)$$

Although it is an interesting result that all factors in the match score have a “pixel domain counterpart,” i.e., can in principle be computed without explicitly evaluating the FFT, we have found no straightforward application of this identity yet. The reason for this is that matching  $I^2$  to  $C^2$ , instead of  $I$  to  $C$ , gives poor results in practice due to the process noise.

#### Example

The scanned image, Fig. 6 (left), which was used before to illustrate properties of FS coding, serves as a numerical example for matching. The size of the

recorded image  $f'(x, y)$  is  $256 \times 256$  pixels. No rescaling of the code image is necessary, because  $N_s = N_o$  and  $R_s = R_o$ , see Eq. (14). Hence, the image *Eye* from Fig. 2 is the reference image  $C(p, q)$ . (The magnitude of  $F'(p, q)$ , the DFT of  $f'(x, y)$ , is shown in Fig. 6 (right). In Table II, we list the values of the means and the discrete inner products  $(G, H) := \frac{1}{N^2} \sum_{p=0}^{N-1} \sum_{q=0}^{N-1} G(p, q)H(p, q)$ . The normalization by  $N^2$  is convenient, because it avoids the values becoming awkwardly big and, with this definition,  $\overline{G} = (G, 1)$ , so we can use Eq. (24), rather than calculating with the zero-mean variables.

A uniform (circular) high-pass filter was applied in the frequency domain. The cut-off radius  $r_0 = N\omega_0/2\pi = 20$  was chosen by inspection of  $C$  and  $|F'|$ . The filtered DFT modulus is shown in Fig. 11. The “corresponding” Gaussian filter has standard deviation  $\sigma = \sqrt{2} / \omega_0 \approx 2.9$ . We have applied a filter with  $\sigma = 3$  in the pixel domain. The results are very similar to those with the uniform filter. The NIP has a very low value for the plain image. As mentioned, this is due to the huge central peak in  $|F'|$ . We observe that the filtering has only a small effect on the cross-product, while  $(|F'|, |F'|)$  is much smaller than

TABLE III. Inner Products for Digital Original

Inner product	"Plain" value (parameter, $\mu = 1/4$ )	Filtered Uniform $r_0 = 20$
$ \bar{A} $	2.832	0.975
$( A , C)$	43.321 ( $\delta = 0.0018$ )	same ( $\delta_0 = 0.1014$ )
$( A ,  A )$	41397.564 ( $\gamma = 7.054$ )	13.707 ( $\gamma_0 = 0.1284$ )
$ \bar{F} $	16.589	14.732
$( F , C)$	3371.253	same
$( F ,  F )$	42251.210	867.351
$ \bar{F} $	14.782	12.990
$( F , C)$	2568.985	same
$( F ,  F )$	41227.143	536.516
$NIP_i$	0.1422	0.9921
$\bar{NIP}_i$	0.1183	0.9902
$NIP$	0.1097	0.9613
$\bar{NIP}$	0.0855	0.9534

$(|F'|, |F'|)$ . The value of  $\bar{NIP}$  is somewhat smaller than the value obtained without incorporating the zero mean shift. It is expected, however, that the value is lower for unmarked images. As an illustration, in Fig. 12 a graph is shown of  $NIP$  as a function of the cut-off radius  $r_0$  of the uniform circular filter for several 305 dpi images.

Finally we will analyze the original digital signal, to illustrate the fact that, even in the noise-free case, no perfect matching can be achieved because of the quantization in the computation of the FS signal and the spectral mixing of the cover and the stego-signal. The values of the means and the discrete inner products for Fig. 2 are listed in Table III.

Here  $A$  is the DFT of the cover image  $a$ . The subscript  $i$  denotes the "ideal" case  $|F_i| = |A| + \mu C$ , (see Eqs. (26) through (28)), which differs from the "digital"  $|F|$ ; see the remark above following Eq. (9). So  $(|F_i|, C) = (|A|, C) + \mu(C, C)$ , and  $(|F_i|, |F_i|) = (|A|, |A|) + 2\mu(|A|, C) + \mu^2(C, C)$ . Both computed values are larger than the actual values.

## Conclusions

FS coding embodies a useful method of providing a digital image with a code, in such a way that the added information is meaningless or even imperceptible to the common observer.

The cover image can be any gray value image with sufficient modulation, i.e., images containing large even areas are not useful. In principle, the code may contain the full information of a gray value image at the size of one half of the cover image, due to the required symmetry. In practice, some restrictions will be imposed on the code image to obtain better properties for the marked image. For the purpose of document identification, it is expected that code images will be used that are relatively simple to produce and recognize. The code images in the examples considered until now are binary and have a global, coarse shape, i.e., do not contain small details. However, application of actual halftone pictures is also conceivable. Some aspects of the influence of the code image on the properties of the marked image have been analyzed here by signal processing and statistical techniques.

FS coding has some properties that are very useful for document processing applications. Extraction of the code image is supported by its invariance to shifting and cropping and its predictable behavior with respect to other common practical operations such as contrast

stretching, resampling/(re)scaling and rotation. These invariances indicate a strong analogy to holography. An FS coded image can therefore be considered as a kind of "graphical hologram." Ways to perform the detection in the presence of these "disturbances" have been proposed here. In addition, experiments have shown that the code image is able to survive various graphical transfer processes, such as halftone screening, printing and digitizing, etc.

On the other hand, the effects of other practical processes such as

- digital transformations, e.g. compression, embedding of additional watermarks
- printing, writing etc. on the marked areas
- soiling, damaging of the substrate, wear and tear etc.

may be insurmountable in the sense that they can affect the code image to the extent of its becoming unrecognizable.

In cases where these processes are likely to occur special care has to be taken to enable recoverability of the code.

Various applications in security documents are currently the subject of study. Combination of FS with other types of frequency coding techniques<sup>22-25</sup> can produce structures with wide applicability in document protection and copyright enforcement. For example, modern variable printing technology offers the possibility to embed a different code in each document, which opens opportunities for owner identification or tracing.

Moreover, machine authentication is enabled by the recent technological progress of document scanning and sorting systems. Although the detection examples in this article have been produced using an ordinary office (flatbed) scanner, high speed inspection platforms are now capable of recording documents in color, with comparable resolution and quality.  $\blacktriangle$

## References

1. R. Anderson, Ed., "Information Hiding: Proceedings of the 1st International Workshop, Cambridge, May/June 1996", *Lecture Notes in Computer Science* **1174** (1996).
2. D. Aucsmith, Ed., "Information Hiding: Proceedings of the 2nd International Workshop, Portland, OR, April 1998", *Lecture Notes in Computer Science* **1525** (1998).
3. A. Pfitzmann, Ed., "Information Hiding: Proceedings of the 3rd International Workshop, Dresden, Germany, September/October 1999", *Lecture Notes in Computer Science* **1768** (2000).
4. I.S. Moskowitz, Ed., "Information Hiding: Proceedings of the 4th International Workshop, Pittsburgh, PA, April 2001", *Lecture Notes in Computer Science* **2137** (2001).
5. F.A.P. Petitcolas, Ed., "Information Hiding: Proceedings of the 5th International Workshop, Noordwijkerhout, The Netherlands, October 2002", *Lecture Notes in Computer Science* **2578** (2003).
6. I.J. Cox, M.L. Miller, and J.A. Bloom, *Digital Watermarking*, (Morgan-Kaufmann, San Francisco, 2002), pp. 2, 11, 49, 275.
7. R. Anderson, Security Engineering, A Guide to Building Dependable Distributed Systems, (Wiley, New York, 2001).
8. W. Bender, D. Gruhl, N. Morimoto, and A. Lu, "Techniques for Data Hiding", *IBM Systems Journal* **35**(3-4), 313 (1996).
9. W. K. Pratt, *Digital Image Processing*, (Wiley, New York, 1978), p. 279, 327, 553.
10. I.J. Cox, J. Kilian, T. Leighton, and T. Shamon, A Secure, Robust Watermark for Multimedia, in Ref. 1, p. 185.
11. I.J. Cox, J. Kilian, T. Leighton, and T. Shamon, Secure Spread Spectrum Watermarking for Multimedia, *IEEE Trans. Image Proc.* **6**(12), 1673 (1997).
12. G.C. Langelaar, I. Setyawan, and R.L. Lagendijk, "Watermarking Digital Image and Video Data: A State-of-the-Art Overview", *IEEE Signal Processing Magazine*, **17**(5), 20-46 (2000).
13. M. Rabbani and P.W. Jones, *Digital Image Compression Techniques*, (SPIE Optical Engineering Press, Bellingham, WA, 1991), p. 107.
14. A.V. Oppenheim, A.S. Willsky, and I.T. Young, *Signals and Systems*, (Prentice-Hall, Englewood Cliffs, NJ, 1983), pp. 211, 205, 219, 207.

15. D. Zheng, J. Zhao, and A. El Saddik, "RST Invariant Digital Image Watermarking Based on Log-Polar Mapping and Phase Correlation", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. **13**(8), 753–765 (2003).
16. B.J.A. Kröse, "A structure description of visual information: modulation of globality in the feature space", in *First Quinquennial Review of the Dutch Society for Pattern Recognition & Image Processing*, E. Backer et al., Eds., (DEB Publishers, Pijnacker, The Netherlands, 1986) p. 93.
17. P.A. Fletcher and K.G. Larkin, "Direct Embedding and Detection of RST Invariant Watermarks", in Ref. 5, p. 129.
18. A. Herrigel, J. O'Ruanaidh, H. Petersen, S. Pereira, and T. Pun, "Secure Copyright Protection Techniques for Digital Images", in Ref. 2, p. 169.
19. J. O'Ruanaidh and T. Pun, "Rotation, Scale and Translation Invariant Spread Spectrum Digital Image Watermarking", *Signal Processing* **66**(3), 303 (1998).
20. A. Rosenfeld and A.C. Kak, *Digital Picture Processing*, vol. 1, (Academic Press, Orlando, FL, 1982), p. 19.
21. F. Natterer, *The Mathematics of Computerized Tomography*, (Wiley, Chichester, UK, 1986), p. 181.
22. H. Oltmans, *Recognition of a SABIC-code*, Report for the Eindhoven University of Technology and Joh. Enschede Security Printing, Haarlem, The Netherlands, 1995.
23. S. Spannenburg, "Digital Copying Security Elements", in *Optical Document Security*, 2nd ed., R.L. van Renesse, Ed., (Artech House, Boston, London, 1997), p. 169.
24. S. Spannenburg, "Optically- and Machine Detectable Copying Security Elements", *Proc. SPIE* **2659**, 76 (1996).
25. S. Spannenburg, "Developments in Digital Document Security", *Proc. SPIE* **3973**, 88 (2000).