Multidimensional Systems and Signal Processing, 2, 421–436 (1991)

Gaussian Wavelet Transform: Two Alternative Fast Implementations for Images

RAFAEL NAVARRO AND ANTONIO TABERNERO
Instituto de Optica "Daza de Valdés" (CSIC), Serrano 121. 28006 Madrid, Spain

Received February 26, 1991, Accepted April 5, 1991

Abstract. A series of schemes for pyramid multiresolution image coding has been proposed, all of them based on sets of orthogonal functions. Several of them are implementable in the spatial domain (such as wavelets), whereas others are more suitable for Fourier domain implementation (as for instance the cortex transform). Gabor functions have many important advantages, allowing easy and fast implementations in either domain, but are usually discarded by their lack of orthogonality which causes incomplete transforms. In this paper we quantify such effect, showing a Gaussian Wavelet Transform, GWT, with *quasiorthogonal* Gabor functions, which allows robust and efficient coding. Our particular GWT is based on a human visual model. Its incompleteness causes small amounts of reconstruction errors (due to small indentations in the MTF), which, however, are irrelevant under criteria based on visual perception.

Keywords. Gaussian wavelets, Gabor functions, image coding, completeness, space and frequency domains implementations

1. Introduction

Multiresolution signal decomposition is being commonly accepted as an optimal solution to processing, coding, and analyzing signals (see for instance [Rosenfeld 1984]). In this sense, the functional organization of the human visual system is based in this kind of image decomposition since individual cells in the cortex are tuned for specific bands of spatial frequencies and orientations [De Valois, Albrecht and Thorell 1982; Campbell and Kulikowski 1966]. In fact, the visual system performs a multiresolution, and also a *multiorientation*, parallel processing of images. Both neurophysiological [Marcelja 1980] and psychophysical [Daugman 1984] studies agree that the receptive fields, or frequency channels in the visual system are very accurately modelable by Gabor functions.

From a point of view of compact coding design, however, Gabor functions are not commonly accepted. Probably this is due to two reasons: first Gabor, in his early work [Gabor 1946] did not really address multiresolution pyramid coding. Second, and most important, Gaussian functions (wave packets) do not directly yield a complete (exact) mapping of the signal. Although this can be overcome by more or less sophisticated methods [Daugman 1988], many other complete transforms have been proposed, such as the Cortex Transform [Watson 1987]; the generalized Gabor scheme [Bastiaans 1981; Porat and Zeevi 1988]; wavelets [Morlet, Arens, Forgeau and Giard 1982; Mallat 1989]; QMF filters [Woods and Neil 1986; Simoncelli and Adelson 1990] etc. Most of those transforms (wavelets, generalized Gabor, QMF, etc.) have been designed to constitute a compact nonredundant coding, involving the computation of a new set of coefficients in a different orthogonal base. Other schemes are based on human vision models (the Cortex Transform, CT, or the Gaussian Wavelet Transform, GWT, proposed here) and basically consist of image decomposition into a set of frequency and orientation channels. Since some amount of redundancy seems to be an intrinsic feature of the image mapping in the visual system (which is multipurpose with a high performance), these schemes may involve some amount of redundancy. Thus, this last type of mapping yields a less compact, but more robust coding, and also shows other advantages; for instance, in order to recover the image it is not necessary to perform any inverse transform, but just to add the channels. As far as we know, it is not clear which among those schemes is the best, but this important number of different proposals indirectly means that image mapping is still an open problem, which is being crucial, e.g., to define standard codes for digital video and broadcast High Definition TV (see for instance [Watson 1990]).

The aim of this paper is to show that Gabor functions (Gaussian wavelets) are also good candidates for image coding, presenting interesting features such as robustness or the possibility of easy and fast implementations in either spatial or spatial-frequency domains. In what follows, we first propose, in Section 2, a Gaussian Wavelet Transform (GWT), which is based on a visual model and is designed to minimize incompleteness effects. The implementations in the Fourier and spatial domains are described in Sections 3 and 4, respectively. The effect of incompleteness are objectively quantified in terms of either the RMS (root mean square) reconstruction error in the spatial domain, or a Modulation Transfer Function, MTF in the Fourier domain, although subjective perceptual criteria can be more adequate in many applications [Gonzalez and Wintz 1977]. The GWT is also compared with other similar, but complete, transform (the Cortex Transform) in the discussion, Section 5.

2. Gaussian Wavelet Representation of Images

In this section we first describe briefly Gaussian wavelets, or Gabor functions, including the reciprocal representation in both spatial and frequency domains, namely receptive fields and channels respectively. Then, among all possible sets of Gabor functions, we shall make our choice based on visual models and image quality criteria.

2.1. Gaussian Wavelets

The Gaussian wavelet, or Gabor functions, is a complex exponential modulated by a Gaussian function. In two-dimensions (2-D), assuming radial symmetry and polar coordinates in the frequency domain, the analytical form in spatial variables is [Tabernero and Navarro 1991]:

$$g_{x_0,y_0,f_0,\theta_0}(x, y) = W \cdot g_{0,0,f_0,\theta_0}(x, y) * \delta(x - x_0, y - y_0),$$

where

$$g_{0,0,f_0,\theta_0}(x, y) = \exp[-\pi a^2((\text{x}\cos\theta_0 + \text{y}\sin\theta_0)^2 + \Gamma^2(\text{y}\cos\theta_0 - \text{x}\sin\theta_0)^2)].$$

$$\exp[i2\pi f_0(\text{x}\cos\theta_0 + \text{y}\sin\theta_0) + \phi] \quad (1)$$

The shape of the Gabor function depends on the following parameters: W which is the gain or amplitude; a which is related to the radial bandwidth, and Γ which is the aspect ratio, while the labels x_0, y_0, f_0, and θ_0 account for the double localization in both space and frequency domains. The Gabor function is complex, having a symmetric (real) and antisymmetric (imaginary) parts. Following an analogy with the human visual system [Jones and Palmer 1987], we shall decompose each Gabor function into two *Receptive Fields* in the space domain, differing in their parity, $p = 0$ (even) or 1 (odd). This choice is not arbitrary because it implies very important computational advantages. Then each single receptive field is defined by five labels (x_0, y_0, f_0, θ_0 and p).

We have been dealing until now with Gabor functions in the spatial domain. However, one important advantageous feature of the Gaussian wavelets is their reciprocity in both domains (see Figure 1). This stems from the fact that the Fourier transform of a Gaussian is another Gaussian, which means that the functional form is basically the same in either the spatial or frequency domains. This will allow (next sections) the two alternative implementations in either domain to be equally easy. Such duality is broken in some way by our visual system, which in any case is ultimately the final receptor of images. It is well established that humans can make a very good job in localizing things in space, while their ability for spatial frequency analysis is relatively poor. Our visual system uses an *implementation* in the spatial domain performing a fine sampling of the space; however, this sampling is coarse in the frequency domain (this is imposed by the uncertainty relation) [De Valois 1982; Daugman 1984]. Implementations in the Fourier domain (where the spatial localization is in some way lost) would be more adequate in the hypothetical contrary case: a coarse

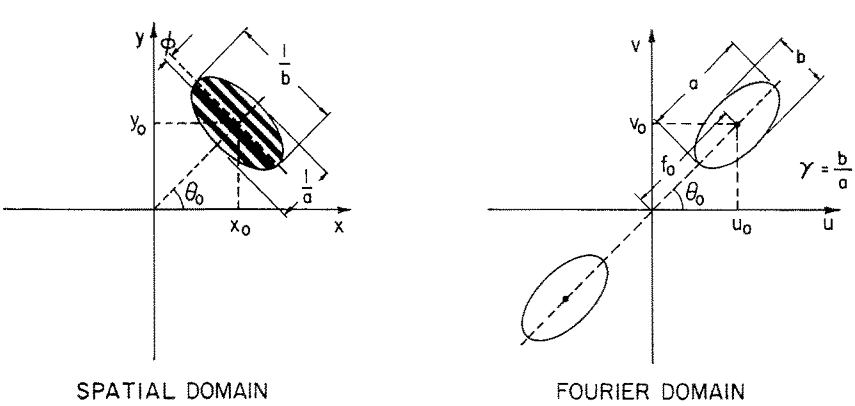

Figure 1. Real component (even receptive field) of a Gaussian wavelet, or Gabor function, (left) and its Fourier transform (right), showing the main parameters of Equation 1.

sampling in space, but a fine Fourier analysis. In any case we shall talk about *Frequency Channels* when dealing with the Fourier domain, and about *Receptive Fields* as regards to the spatial domain.

2.2. *The Set of "Logons"*

Following Gabor [Gabor 1946], in order to represent a signal one can decompose it into a set of elementary quanta of information or logons. Image representations in space (pixels) or by Fourier analysis (spatial frequencies) are just special cases. There is a lot of evidence showing that such representations of images are not the best. Depending on the application one could choose any intermediate case allowing a more or less coarse (or fine) local spectral analysis [Jacobson and Wechsler 1988]. The use of Gabor functions will insure an optimum packing of information (minimizing the uncertainty relation) and the reciprocity in both domains. Unfortunately, Gaussian wavelets present a drawback: due to their shape they do not constitute a complete set. However, by quantifying the reconstruction error, it is possible to find a *quasicomplete* set of Gabor functions with a minimum error, which as we will show is perceptually irrelevant.

Like most of the proposed transforms, our scheme is based on a linear scale in the spatial domain; i.e., a uniform distribution of receptive fields, and a logarithmic sampling in the Fourier domain; i.e., frequency channels distributed in octaves ($\log_2$). This permits multiscale pyramid implementations [Burt and Adelson 1983] with self-similar wavelets: all the wavelets can be obtained by translating, rotating or scaling one of them. The octave distribution of channels has shown to be optimal since the spectra of natural images tend to show exponential type decays [Navarro, Santamaría and Gómez 1987; Field 1987]. This means that the energy per frequency channel tends to be constant.

The additional criteria to design the set of *logons* (the shape and number of wavelets) have been in our case [Tabernero and Navarro 1991]: known features of the human visual system, optimum coverage of both domains and simplicity. The resulting set is composed by four orientation channels ($\theta_0 = 0^o, 45^o, 90^o$ and 135^o) and four frequency channels—and a low-pass residual [Burt and Adelson 1983]. The peak frequencies (f_0) of the four channels are: half the Nyquist frequency ($f_N/2$) for the highest frequency channel, and then halving for the next lower octave channel ($f_N/4$, $f_N/8$ and $f_N/16$). The radial bandwidth has been set to one octave, $a = 0.71 \cdot f_0$ in Equation 1, which yields a nearly optimum coverage of the radial frequency axis. The aspect ratio $\Gamma = 1$ allows circular symmetry, while the resulting angular bandwidth (0.71 radians) is not far from the optimum angular covering of the Fourier domain (for 4 channels is $\pi/4 = 0.785$). Finally, the localization of the receptive fields in space will be given by sampling requirements; sampling in a square grid, with an interval given by the Nyquist frequency of the corresponding channel. Figure 2a shows the basic 4×4 Gabor functions in the spatial domain, and Figure 2b the recovering of the Fourier domain by this set. This last figure clearly illustrates the nature of the incompleteness of the GWT, since the recovery of the Fourier domain is not complete. This causes a nonflat MTF: apart from the loss of high frequencies mainly in the corners (high frequency residue; [Burt and Adelson 1983; Watson 1987]), what is characteristic of the Gabor functions are the indentations between adjacent channels (this is further illustrated in

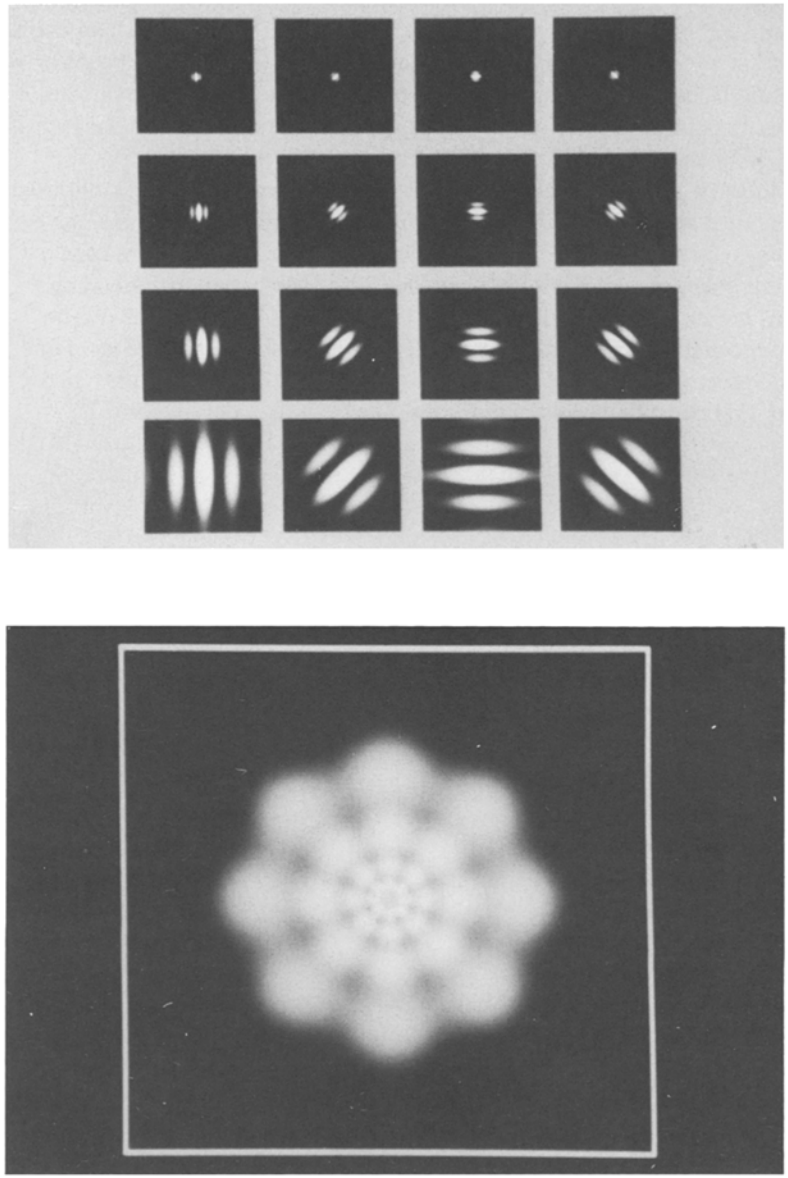

Figure 2. The 4 × 4 basic Gabor functions of the GWT: (a) The (even) Receptive Fields in the space domain; (b) Coverage of the Fourier domain by the Frequency Channels.

Figure 4). This makes the GWT be incomplete since a complete scheme would provide a uniform recovery of both domains.

Once the set of logons is defined, the image representation is obtained by the expansion into these elementary signals. Any implementation will involve the computation of the inner product of the signal with every Gabor function. The resulting numbers are samples of the image in a joint spatial/spatial-frequency domain [Jacobson and Wechsler 1988]. The logarithmic Gaussian mapping has the important feature that the joint size of logons is constant and optimal (minimum uncertainty relation means optimum packing of information) in the joint domain (see Figure 3).

An important aspect of this kind of scheme is that one only needs the even real part of the Gabor functions to encode the image. However, in image analysis, the performance can be highly improved by using the complete transform, including even and odd receptive fields [Tabernero and Navarro 1991]. In fact, it is well known that the simple neurons in the visual cortex are arranged in couples, differing just because their receptive fields are in phase quadrature [Pollen and Ronner 1983]. This *redundancy* seems to be inherent to the visual process: the even receptive fields are bar detectors, while the odd ones are edge detectors [Watson 1990]; (see also Figure 7).

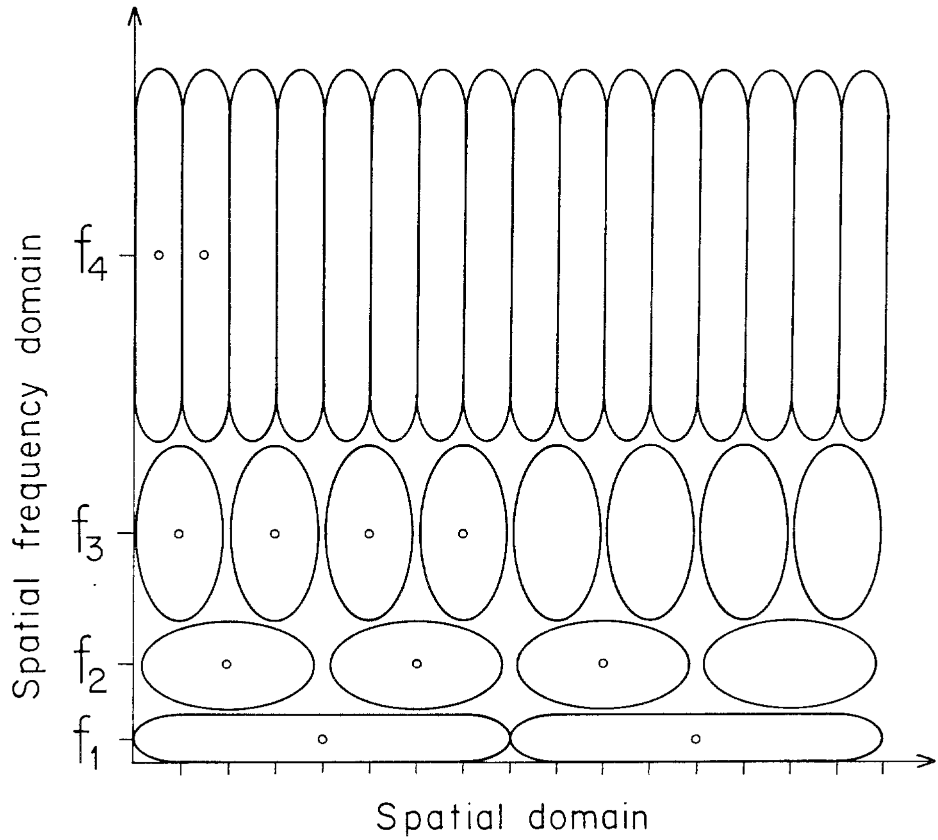

Figure 3. GWT coverage of the joint spatial/spatial frequency domain, with the minimum number of Gabor functions (logons), all of them having the same area (constant and maximum uncertainty relation).

3. Implementation in the Fourier Domain

As far as we know, the implementation in the Fourier domain has been common in this kind of transform [Watson 1987], since on the one hand the frequency channels are straightforwardly implementable as linear Fourier filters (in our case Gaussian windows in the frequency domain, Figure 2b), and on the other hand, by using a Fast Fourier Transform algorithm, and a pyramid implementation [Watson 1987], the computer time is quite reasonable.

Most of the details of this implementation are the same as in similar schemes [Watson 1987]. The unique but important question is that the Gaussian channels do not allow a complete and uniform covering of the Fourier domain. This can be appreciated in Figure 4. This figure shows the modulation transfer function, MTF of the Gaussian Wavelet

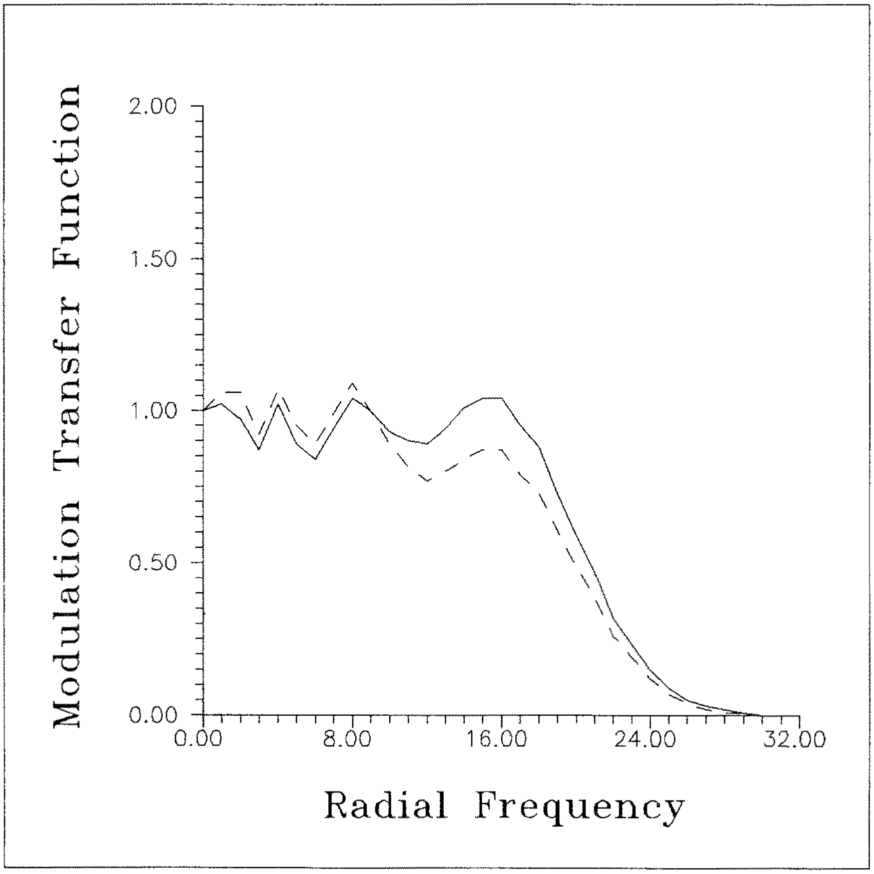

Figure 4. Modulation Transfer Function of the Fourier domain GWT implementation. The dashed line corresponds to the case of constant gain for the different channels, and the continuous line is after equalization by assigning different gains to the frequency channels (gains are 0.93, 0.93, 0.93 and 1.15 respectively).

Transform GWT, with, and without, assigning different weights (gains) to the different channels. The reason to assign different weights to the channels is to equalize (flatten) the MTF as much as possible; i.e., to minimize the reconstruction errors. The method used to find the gains, in both spatial and Fourier implementations, was empirical. The MTF was obtained using a random white noise as test image, as the ratio of the Fourier spectra: that of the recovered image after the GWT, versus that of the original. As can be appreciated, it is not possible with Gausian functions to obtain a completely flat MTF. This explains why the GWT is not complete, in the sense that we do not exactly recover the original. However, we will show that while objective criteria (RMS error in space, or MTF in frequency), give noticeable differences with respect to the original, under subjective visual criteria it is difficult to distinguish between the original and the reconstruction.

4. Implementation in the Spatial Domain

The self-similarity of the Gabor functions, along with their Gaussian shape, allows easy pyramid implementations in either spatial or frequency domains. Although any filtering operator can be equally performed in both domains by the convolution theorem, the Fast Fourier Transform (FFT) algorithms usually are more appropriate to implement the filtering in the Fourier domain. However, adequate filter design, with small filter convolution masks, can match, or even improve, the computing time when compared with a FFT implementation. For instance, 3×3 filter masks are widely used in real time image processing, including low-pass Gaussian filters. For this purpose we started designing a couple (even and odd) of filter masks corresponding to the highest frequency, horizontal $\theta = 0^0$ channel. This consisted of finding the minimum number of samples needed to reasonably represent the shape of that Gabor function. As a result of this tradeoff, the receptive fields were implemented in 7×7 convolution masks [Tabernero and Navarro 1991]. The resulting filter masks, shown in Figure 5, permit a fast implementation. The computing time for

$$
\begin{pmatrix}
0 & -3 & 0 & 13 & 0 & -3 & 0 \\
0 & -8 & 0 & 16 & 0 & -8 & 0 \\
0 & -13 & 0 & 22 & 0 & -13 & 0 \\
0 & -16 & 0 & 26 & 0 & -16 & 0 \\
0 & -13 & 0 & 22 & 0 & -13 & 0 \\
0 & -8 & 0 & 16 & 0 & -8 & 0 \\
0 & -3 & 0 & 13 & 0 & -3 & 0
\end{pmatrix}
\qquad
\begin{pmatrix}
3 & 0 & -8 & 0 & 8 & 0 & -3 \\
6 & 0 & -11 & 0 & 11 & 0 & -6 \\
8 & 0 & -15 & 0 & 15 & 0 & -8 \\
10 & 0 & -16 & 0 & 16 & 0 & -10 \\
8 & 0 & -15 & 0 & 15 & 0 & -8 \\
6 & 0 & -11 & 0 & 11 & 0 & -6 \\
3 & 0 & -8 & 0 & 8 & 0 & -3
\end{pmatrix}
$$

EVEN FILTER ODD FILTER

Figure 5. Filter masks of 7×7 pixels used in the spatial domain implementation. They correspond to the maximum frequency $f_N/2$, vertical $\theta = 0^0$, even and odd receptive fields.

the spatial implementation is roughly proportional to $7^2 \cdot N^2$ (N being the number of rows in a square image), versus $k \cdot N^2 \cdot \log_2 N^2$ (typical values for k are between 2 and 3) of the fast Fourier implementation. In consequence, for typical image sizes both implementations are approximately equivalent in computing time.

Once the basic couple of receptive filters is designed, the complete set is obtained by rotating and/or scaling operations. While the four orientations are in fact obtained by rotating the originals, a pyramid implementation is more adequate for scaling: instead of magnifying the filters, which would imply much more computing time, the pyramid implementation applies the same set of filters to successive compressed versions of the image. Each compressed version is obtained by first applying a low-pass Gaussian filter (5×5 mask) to the image in order to avoid aliasing, and then a decimation by a factor of 2 in both dimensions [Burt and Adelson 1983]. This scheme is typical in multiresolution methods [Rosenfeld 1984]. Figure 6 shows the result of the GWT expansion. The image is recovered by adding only the even logons (a), while the odd part (b) is equivalent to a set of *edge detectors*.

As in the Fourier domain implementation, the effect of incompleteness is that the GWT is not exact. Furthermore, undersampling of the receptive fields causes a broadening of the Fourier channel, which affects the MTF. This can be appreciated in Figure 7. This figure shows the MTF of the spatial domain GWT (computed from a white noise image), with and without assigning different weights, or gains, to the different channels. The equalized version is obtained with gains 0.45, 0.60, 0.65, and 0.90 for the four frequency channels, f_1 to f_4 respectively (the gain of the low frequency residual, 1, is unchanged). The assignment of different weights to the channels allows the equalization of the two implementations.

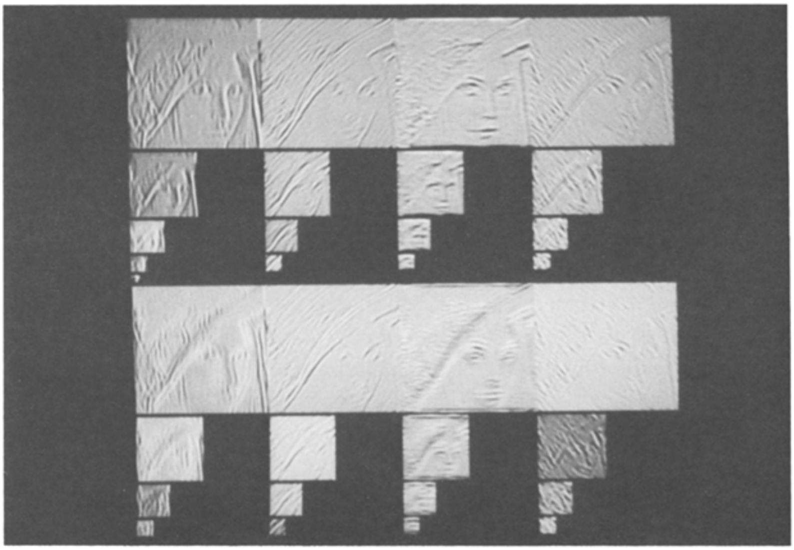

Figure 6. Result of the complete GWT. The grey scale is proportional to the value of each logon. Note that the even receptive fields (up), behave as bar detectors, and are enough to code and reconstruct the image. The odd receptive fields are edge detectors (bottom).

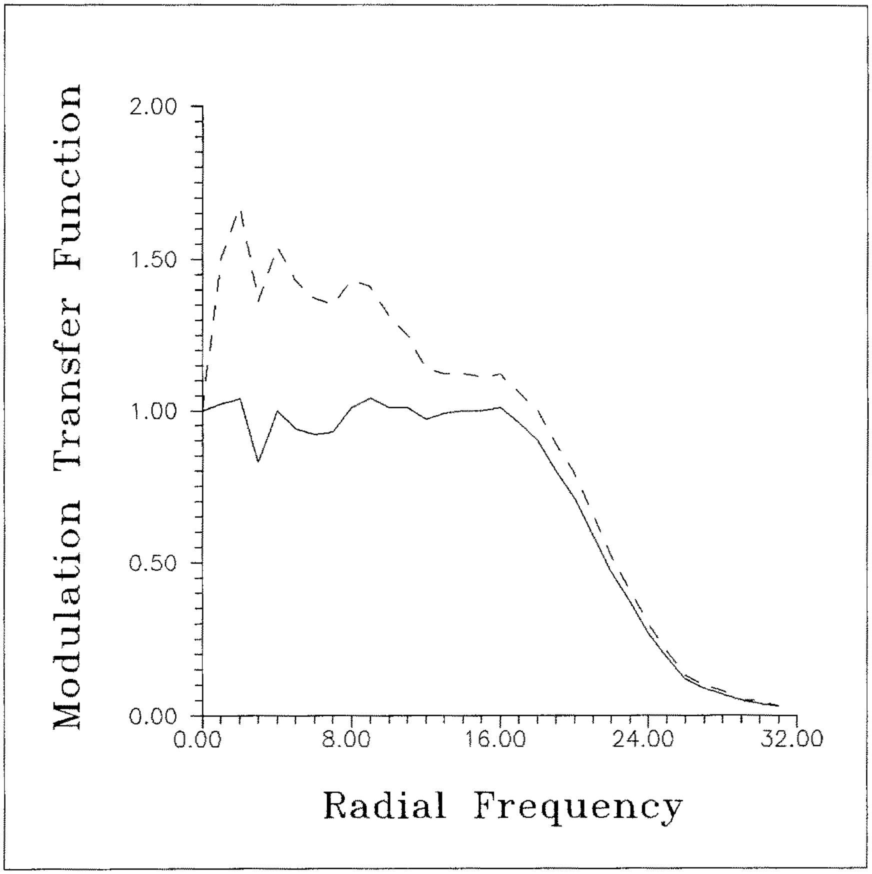

Figure 7. Modulation Transfer Function of the spatial domain GWT implementation. The dashed line corresponds to the case of constant gain for the different channels, and the continuous line is after equalization by assigning different gains to the frequency channels (gains are 0.45, 0.60, 0.65 and 0.90 respectively).

The matching is not total anyway, since the two implementations involve different computing errors, sampling, quantization, etc. Figure 8 shows (a) the original image, and reconstructions: (b) from Fourier domain and (c) from spatial domain implementations of the GWT. The RMS normalized errors of the reconstructions are 13% in both cases. Most of this RMS error (about the 11% for the CT. See Section 5) comes from disregarding the high frequency residual band, corresponding to the energy contained in the corners of the Fourier domain [Watson 1987]. In fact, the visual appearance of the reconstructions resembles a very slightly low-pass filtered version of the original. The high frequency residual can be incorporated to the GWT, but its value is usually going to be negligible. Moreover, this high frequency residual will be zero when the acquisition system uses both conventional optical systems and adequate sampling.

Figure 8. Original image (a) and reconstructions from Fourier domain (b), and spatial domain (c) implementations of the GWT.

5. Discussion

In this section we shall compare our GWT with a similar but complete transform. For this purpose, among the different published schemes, we chose the Cortex Transform, CT [Watson 1987], because this is probably the one presenting more similar features with respect to our GWT (4 × 4 channels, also based on visual models, etc.). The main difference is that the CT is complete; also the CT is difficult to implement in the spatial domain. In what follows, we use a version of the CT with exactly the same frequency, and orientation channels as our GWT.

5.1. Comparison Under Objective and Subjective Visual Criteria

Figure 9 shows two reconstructions from the CT (a) and the GWT (b) respectively. In both cases the high frequency residual has been neglected, which causes most of the error. The RMS error is slightly smaller in the CT case: 11% in the CT (which is only due to neglecting the high frequency residual), versus 13% in the GTW; while the visual appearance is almost identical. The effect of incompleteness is objectively quantizable in the Fourier

Figure 9. Comparison of the reconstructions obtained with the Cortex Transform, CT (a), and the Gaussian Wavelets Transform GWT (b). In both cases, the peak frequencies and orientations are the same, and the high frequency residual has been neglected.

domain as little indentations in the MTF of the transform or as a higher RMS error in the reconstruction. However, since the final receptor of images is usually going to be the human eye, subjective perceptual criteria are more adequate. In Figure 9, the GWT reconstruction could be better classified by several observers, although most of them would find it very difficult to establish a difference. This result points out that although Gabor functions have been previously discarded because of their lack of completeness, the GWT proposed here is *quasicomplete*; i.e., the effect of incompleteness is small in terms of RMS error, and irrelevant under perceptual criteria.

The redundancy inherent in using an incomplete representation has been also argued against Gabor functions. For image compression applications, orthogonal transforms should give a better result than those nonorthogonal; i.e., a more compact coding with smaller entropy or bit/pixel rate, since minimum entropy also means minimum redundancy [Gonzalez and Wintz 1977]. To assess that redundancy, we have computed and compared the entropy of the different frequency channels of Figure 6, with that equivalently obtained by the CT. The resulting values for the higher frequency channel are 4.4 and 4.3 respectively, which involves only a 2% of increment, while for the other channels there is not a noticeable difference (values are 4.6, 5 and 5.2 respectively. Note that these values of entropy are relatively high because they have been computed using the original quantization in 256 linear levels. By an adequate quantization and coding, the entropy can be highly reduced). In order to complete this comparison we also need to know the 2-D (area) bandwidth ratio between the GWT and CT frequency channels, because the total amount of bits required to code a channel is proportional to the entropy of that channel times its bandwidth. This has been estimated in the following way: first the minimum box in the Fourier domain containing a 99% of the energy of a CT channel has been determined, and second, the percentage of energy of the equivalent GWT channel inside that box has been computed. The result is 90% approximately, which once again does not imply a big difference. Actually this means a small increment in the amount of redundancy in the GWT with respect to the CT, which, as we shall show in Section 5.2, increases the robustness of the coding. Additionally, the reciprocity of Gaussian wavelets under Fourier transformation, makes them specially appropriate as a base of functions in joint representations. Moreover, the smoothness of Gaussian windows is very well suited to avoid edge effects, aliasing and ringing artifacts, etc., which are typical in Fourier filtering. As we shall show, this property, along with some amount of information redundancy, is important in robust coding.

5.2. Comparison in Terms of Robustness of Coding

One of the main applications of these kinds of transforms (Cortex, QMF, wavelets, etc.) is image coding. There is strong evidence suggesting that the human visual system also uses a similar scheme [Marcelja 1989; De Valois et al. 1982; Daugman 1984] and it has been shown that similar transforms allow high rates of data compression [Daugman 1988; Watson 1987]. Also the visual system makes an important use of redundancy: for instance the arrangement of neurons in couples with phase quadrature involves redundancy, but it is very important in shape and edge detection, or texture recognition. Many tasks can be more efficiently performed with this visual mapping of images, but also this mapping is well suited for robust coding.

Figure 10. Image reconstructions with partial losses of frequency channels. The first row (up) corresponds to the CT and the second (down) to the GWT. The first column (left) shows the reconstruction from only three frequency channels (missing the high frequency channel), and the second and third columns to reconstructions from two and only one channels, respectively.

Probably the most important features of good coding are robustness and efficiency. In particular, the robustness under partial information losses (or errors), which could be random or even systematic, is critical [Watson 1990]. Pyramid multiresolution coding is robust under high frequency losses. This has been tested, and the result is shown in Figure 10. The first row presents image reconstructions from the CT, with progressive loss of high frequency channels: with only three channels (first column left, 19%), and with two (second column $E_{RMS} = 26\%$), and just the lowest frequency channel (third, $E_{RMS} = 34\%$). The second row corresponds to the GWT in the same conditions ($E_{RMS} = 20\%$, 26% and 35%, respectively). In spite of these relatively high values of the RMS error, this figure illustrates the robustness of this kind of representation. Our eye can perform a very good job in face recognition, even when just the two lower frequency channels are available (a small subset, 6% approximately, of the total number of logons). This also could exlain to some extent why this coding is convenient in human vision, since people with low visual acuity (myopic, amblyopics, etc.) can recognize objects, and perform most visual tasks.

Figure 10 also illustrates a very important point. In the case of any kind of information losses, or errors, no coding is complete. In consequence, robustness is not at all dependent on completeness. If we compare the first and second rows of Figure 10, we can observe some artifacts, *ringing*, which are more clear in the first row, CT. Once again, although the RMS error is slightly smaller for the CT (11%, 19%, 26% and 34% in the CT versus 13%, 20%, 26% and 35% in the GWT), visual criteria will not make the same difference. Besides, in the case of partial losses (compare the first and second couples of pictures) the GWT presents slightly smoother artifacts, which means that under visual criteria, the GWT would be preferable. This fact is further illustrated in Figure 11, representing orientation channels losses.

Figure 11. Image reconstructions with partial losses of orientation channels. The first row corresponds to the CT and the second to the GWT. The first column shows reconstructions without a single orientation θ_{45} in only one frequency f_4 (high frequency). The second and third columns show reconstructions with one orientation channel missing for all frequencies (θ_0 and θ_{135} respectively).

Figure 11 shows reconstructions with orientation channel losses: the first column (left) corresponds to loosing a single orientation channel ($\theta_0 = 45^0$) only for the frequency band f_3; the second and third columns correspond to orientation losses (θ_0 and θ_{135} respectively) for all frequencies. The first row (top) corresponds to the CT and the second (bottom) to the GWT. Apart from some more or less curious effects (noise removal, etc.), this is a significant example in which the CT is subject (on average) to more visual artifacts, the GWT appearing to be more robust. This is because when an orientation channel is missing in the CT, which is less redundant, a deep and sharp hole appears in the Fourier domain. On the contrary, with Gabor functions, that hole is much less deep since it is partially filled by the adjacent channels, and of course it has much smoother edges. This is a very important advantage, because in these examples, the partial filling of holes in the GWT will allow image restoration (by inverse or Wiener filtering methods, etc.), while the deep holes in the CT are very difficult to restore.

Another important aspect of the GWT is that it is well suited for parallel optical implementations. In this sense, an optical wavelet transform has been recently proposed by [Freysz, Pouligny, Argoul and Arneodo 1990]. With an adequate design of optical filters, it would be amenable to completely parallel optical implementation; i.e., by multiplying the object by a high frequency square array of dots, many replicas of the object spectrum can be simultaneously obtained allowing to apply several filters at the same time.

6. Conclusion

Gaussian wavelets (Gabor functions) present a series of important advantages with respect to other wavelets, at the cost of incompleteness. Many authors have argued against Gabor functions for their lack of orthogonality and completeness, and proposed other alternative complete schemes and orthogonal bases of functions. However, we have shown that by a careful choice of the Gabor functions, in order to optimize the recovery of the frequency domain, the effect of incompleteness of the GWT is small, and perceptually almost insignificant. Moreover, the GWT proposed here can be considered in practice quasicomplete. The advantages are very important (most of them were already shown by Gabor in 1946): minimizing the uncertainty relation (optimum packing of information); formal reciprocity, which makes implementations possible and easy in either domains, and robustness under information losses, minimizing artifacts, etc. A final advantage is that the GWT is also based on human vision. Since the human visual system shows a high performance in most aspects, which has not been improved, or even reached, by artificial systems, image processing environments based on visual models are expected to be well suited for artificial vision.

Acknowledgments

This research has been supported by the Comisión Interministerial de Ciencia y Technologia, Spain, under grant P18-88/88-0198 (PRONTIC).

References

M.J. Bastiaans, "A sampling theorem for the complex spectrogram, and Gabor's expansion of a signal in Gaussian elementary signals," *Opt. Engineer.* 20, pp. 594–598, 1981.

P.L. Burt and E.H. Adelson, "The laplacian pyramid as a compact image code," *IEEE Transactions on Communications* COM-31, pp. 532-540, 1983.

F.W. Campbell and J.J. Kulikowski, "Orientation selectivity of the human visual system" *J. of Physiology* 187, pp. 437–445, 1966.

J.G. Daugman, "Spatial visual channels in the Fourier plane," *Vision Research* 24, pp. 891–910, 1984.

J.G. Daugman, "Complete discrete 2-D Gabor transform by neural networks for image analysis and compression," *IEEE Transactions on Acoustic Speech and Signal Processing* ASSP-36, pp. 1169–1169, 1988.

R.L. De Valois, D.G. Albrecht, and L.G. Thorell, "Spatial frequency selectivity of cells in macaque visual cortex," *Vision Research* 22, pp. 545–559, 1982.

D.J. Field, "Relation between the statistics of natural images and the response properties of cortical cells," *J. of the Optical Society of America A* 4, pp. 2379–2394, 1987.

E. Freysz, B. Pouligny, F. Argoul, and A. Arneodo, "Optical wavelet transform of fractal aggregates," *Phys. Rev. Lett.* 64, pp. 745–748, 1990.

D. Gabor, "Theory of communication," *J. Inst. Elect. Eng.* 93, pp. 429–457, 1946.

R.C. Gonzalez and P. Wintz, *Digital Image Processing*. Addison-Wesley: London, 1977.

L.D. Jacobson and H. Wechsler, "Joint spatial/spatial-frequency representations," *Signal Processing* 14, pp. 37–68, 1988.

J. Jones and L. Palmer, "An evaluation of the two-dimensional Gabor filters model of simple receptive fields in cat striate cortex," *J. of Neurophysiology* 58, pp. 538–539, 1987.

S.G. Mallat, "A theory of multiresolution signal decomposition: the wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI-11, pp. 674–693, 1989.

S. Marcelja, "Mathematical description of the response of simple cortical cells," *J. of the Optical Society of America* 70, pp. 1297–1300, 1980.

J. Morlet, G. Arens, I. Forgeau, and D. Giard, "Wave propagation and sampling theory," *Geophysics* 47, pp. 203-236, 1982.

R. Navarro, J. Santamaría, and R. Gómez, "Automatic log spectrum restoration of atmospheric seeing," *Astronomy and Astrophysics* 174, pp. 334–351, 1987.

D.A. Pollen and S.F. Ronnen, "Visual cortical neurons as localized spatial filters," *IEEE Transactions on Systems, Man and Cybernetics* SMC-13, pp. 284–302, 1983.

M. Porat and Y.Y. Zeevi, "The generalized Gabor scheme for image representation in biological and machine vision," *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI-10, pp. 452–468, 1988.

A. Rosenfeld, *Multiresolution Image Processing and Analysis*. Springer-Verlag: New York/Berlin, 1984.

E.P. Simoncelli and E.H. Adelson, "Nonseparable extensions of quadrature mirror filters to multiple dimensions," *Proc. of the IEEE* 78, pp. 652–664, 1990.

A. Tabernero and R. Navarro, "Performance of Gabor funcitons for texture analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI (submitted).

A.B. Watson, "The cortex transform: rapid computation of simulated neural images," *Computer Vision, Graphics and Image Processing* 39, pp. 311–327, 1987.

A.B. Watson, "Perceptual-components architecture for digital video," *J. of the Optical Society of America A* 7, pp. 1943–1954, 1990.

J.W. Woods and S.D. O'Neal, "Subband coding of images," *IEEE Trans. on Acoustic Speach and Signal Procesing* ASSP-34, pp. 1278–1288, 1986.