

A Gabor Expansion-Based Positive Time-Dependent Power Spectrum

Flemming Pedersen

Abstract—This correspondence presents a novel approach to the synthesis of positive time–frequency distributions with marginals corresponding to the instantaneous power and the power spectrum. This class of distribution has mainly been constructed using an entropy approach, whereas this correspondence presents a noniterative approach based on the Gabor expansion. The resulting distribution has been named the positive Gabor spectrogram and is evaluated using different test signals. Interpretation and applicability of positive distributions in connection with signal analysis are discussed.

Index Terms—Gabor expansion, positive time–frequency distributions.

I. INTRODUCTION

The purpose of joint time–frequency analysis (JTFA) is to map a one-dimensional (1-D) function of time *or* frequency into a two-dimensional (2-D) function of time *and* frequency. This 2-D function illustrates how the spectral content of the signal is changing in time. As the spectral/time resolution of the traditional Fourier spectrogram (FS) is limited because of smoothing, considerable effort has been put into the development of new candidates for the FS.

Two popular approaches are “Wigner–Ville (WV) modified distributions” and “statistical correct positive distributions.” In this correspondence, we will show how to obtain the later type of distribution from a Gabor expansion of the function.

The group of “Wigner–Ville modified distributions” is designed with the aim of reducing the effect from the cross-term (CT) and preserving the auto-term (AT) in the WV. Numerous distributions have been obtained by filtering the WV with different filter functions, and they can all be parameterized by the Cohen’s class of distributions. However, the problem with this filtering approach is that it cannot distinguish completely between CT’s and AT’s. A more powerful technique (not belonging in the Cohen’s class) to control the CT’s has been developed by Qian *et al.* [1], [2] and has been named the time–frequency distribution series (TFDS). They use a Gabor expansion of the signal to decompose the WV into CT’s and AT’s. The problem with negative CT’s in the WV can now be localized and reduced efficiently. By changing the number of decomposing terms, the spectral resolution of the TFDS can be changed in a more advantageous way than known from the FS. However, the result strongly depends on how the Gabor expansion is made.

The work of Qian *et al.* has been the inspiration for the development of a new positive time–frequency (TF) distribution known as the positive Gabor spectrogram (PGS) [3], which belongs to the group of statistical correct positive distributions. The aim of the TFDS is to use the Gabor expansion to reduce the contribution from the CT, in the WV, whereas the PGS uses the Gabor expansion to move the CT’s to the position of the AT’s without changing the marginals.

Manuscript received October 1, 1997; revised July 22, 1998. The associate editor coordinating the review of this paper and approving it for publication was Prof. Moeness Amin.

The author is with the Department of Communication Technology, Aalborg University, Digital Signal Processing Research Group, Aalborg, Denmark (e-mail: fp@kom.auc.dk).

Publisher Item Identifier S 1053-587X(99)00749-7.

II. POSITIVE GABOR SPECTROGRAM

The positive Gabor spectrogram is designed as a positive TF distribution with marginals corresponding to the instantaneous power (IP) $|x(t)|^2$ and power spectrum (PS) $|X(\omega)|^2$

$$\begin{aligned} C_x(t, \omega) &\geq 0 \\ \int C_x(t, \omega) dt &= |X(\omega)|^2 \\ \frac{1}{2\pi} \int C_x(t, \omega) d\omega &= |x(t)|^2. \end{aligned} \quad (1)$$

These constraints are not unique and can be satisfied by many differently looking distributions. This group of distributions has been parameterized by Cohen *et al.* as the Cohen–Posch class [4]. As positive distributions are distributions in a statistical sense, Cohen named this class of distributions *proper distributions*, which have been widely accepted. However, this work by Cohen did not specify a way to synthesize a distribution in this class, and it remained a problem for many years as to how to find a meaningful distribution for a given signal. A solution to this problem came with the work by Loughlin [5], [6]. He used an information theoretic approach to estimate a positive distribution. This entropy optimization problem involved the principle of minimum cross-entropy, and he therefore named his distribution a minimum cross-entropy time–frequency distribution (MCE-TFD). Now, instead of following Loughlin’s approach to a positive distribution, we propose a similar distribution based on the Gabor theory.

It is known from the Wigner–Ville case that an important task in constructing a meaningful TF energy distribution is to control the contributions from the inevitable CT’s. The idea of the PGS is to move the marginal contribution of the CT’s to the position of each AT without changing the marginals and preserving a positive level of each AT. This modification of the WV will result in a proper distribution.

CT’s are not a phenomenon limited to bilinear distributions but are due to the nonlinear nature of an energy representation. The CT in the FS are an acknowledged fact [7], [8]. As CT’s exist in all energy distributions, they are also present in 1-D distributions like the IP and PS. In order to construct a 2-D distribution, it will be advantageous first to consider 1-D distributions. This will reveal part of the nature of CT’s and enable us to construct 2-D positive energy distribution with IP/PS marginals.

The following simple example of the CT phenomenon illustrates the principle of how the CT’s can be divided between the AT’s. The absolute square of three arbitrary numbers is

$$|1 + 4 - 2|^2 = \underbrace{|1|^2 + |4|^2 + |2|^2}_{\text{AT}} - \underbrace{12}_{\text{CT}}. \quad (2)$$

The negative level of the sum of CT’s will never exceed the sum of the AT’s. Therefore, it will be possible to divide the contribution from the CT’s between the three AT’s and preserve the positive level of each of the AT’s as

$$\begin{aligned} |1 + 4 - 2|^2 &= \left(|1|^2 + \frac{|1|^2}{|1|^2 + |4|^2 + |2|^2} \cdot (-12) \right) \\ &+ \left(|4|^2 + \frac{|4|^2}{|1|^2 + |4|^2 + |2|^2} \cdot (-12) \right) \\ &+ \left(|2|^2 + \frac{|2|^2}{|1|^2 + |4|^2 + |2|^2} \cdot (-12) \right) \\ &= 0.43 + 6.86 + 1.71. \end{aligned} \quad (3)$$

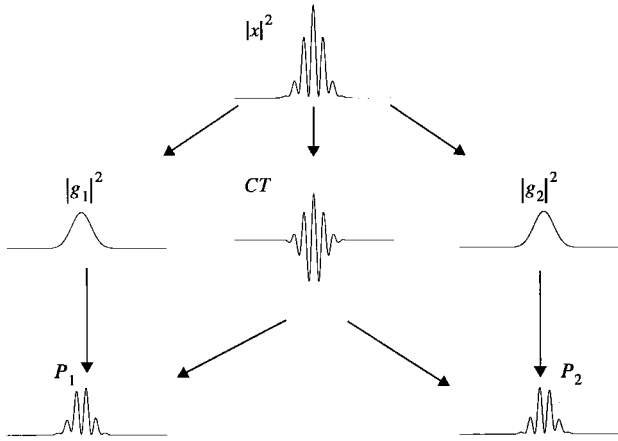


Fig. 1. Procedure for splitting up the cross-term between two Gaussian functions.

Each of the three positive numbers contains its own AT and part of the CT's.

This example will serve as the recipe for the general case. An arbitrary function can be represented by a Gabor expansion in the following way:

$$\begin{aligned} x(t) &= \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} a_{m,n} g(t-mT) e^{-jtn\Omega} \\ &= \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} a_{m,n} g_{m,n}(t) \end{aligned} \quad (4)$$

where

$$g_{m,n}(t) = g(t-mT) e^{-jtn\Omega}.$$

When calculating the IP of $x(t)$, the AT and CT part can now be separated

$$|x(t)|^2 = AT(t) + CT(t)$$

where

$$\begin{aligned} AT(t) &= \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |a_{m,n} g_{m,n}(t)|^2 \\ CT(t) &= |x(t)|^2 - AT(t). \end{aligned} \quad (5)$$

Following the recipe from (3) and (5), we define

$$|x(t)|^2 = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} p_{m,n}(t)$$

where

$$p_{m,n}(t) = \frac{|x(t)|^2}{\sum_{m'=0}^{M-1} \sum_{n'=0}^{N-1} |a_{m',n'} g_{m',n'}(t)|^2} \cdot |a_{m,n} g_{m,n}(t)|^2. \quad (6)$$

We have now expressed the IP of an $M \cdot N$ component function as the summation of $M \cdot N$ $p_{m,n}(t)$ functions, each with its own AT and part of the CT. It is obvious that $p_{m,n}(t)$ is non-negative. The procedure for only two Gaussian functions is illustrated in Fig. 1.

The same decomposition can be made for the PS. First, the Gabor expansion is Fourier transformed

$$X(\omega) = FT\{x(t)\} = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} a_{m,n} G_{m,n}(\omega) \quad (7)$$

where $G_{m,n}(\omega)$ denotes the Fourier transform of $g_{m,n}(t)$, which is also a Gaussian function.

Second, following the recipe, the PS is divided into an AT and a CT part

$$|X(\omega)|^2 = AT(\omega) + CT(\omega)$$

where

$$\begin{aligned} AT(\omega) &= \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |a_{m,n} G_{m,n}(\omega)|^2 \\ CT(\omega) &= |X(\omega)|^2 - AT(\omega). \end{aligned} \quad (8)$$

Finally, following the recipe from (3) and (8), we define

$$|X(\omega)|^2 = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} P_{m,n}(\omega)$$

where

$$P_{m,n}(\omega) = \frac{|X(\omega)|^2}{\sum_{m'=0}^{M-1} \sum_{n'=0}^{N-1} |a_{m',n'} G_{m',n'}(\omega)|^2} \cdot |a_{m,n} G_{m,n}(\omega)|^2. \quad (9)$$

We have now likewise expressed the PS as a summation of $M \cdot N$ positive functions $P_{m,n}(\omega)$. The link between the two functions $p_{m,n}(t)$ and $P_{m,n}(\omega)$ is a common AT, which is related to the same TF shifted Gaussian function $g_{m,n}$ in the Gabor expansion. The key to use this fragmentation of the IP and PS into $p_{m,n}(t)$ and $P_{m,n}(\omega)$ relates to the fact that the WV of a Gaussian function can be expressed as the product between the IP and PS divided by the energy of the Gaussian function

$$WV_g(t, \omega) = \frac{|g(t)|^2 \cdot |G(\omega)|^2}{\int |g(t)|^2 dt}. \quad (10)$$

Now, we have not only one Gaussian function but $M \cdot N$ Gaussian functions given by the Gabor expansion. The resulting distribution is, therefore, obtained by summing up these products between the time functions $p_{m,n}(t)$ and the frequency functions $P_{m,n}(\omega)$ normalized by the area of one of them. The resulting distribution is referred to as the positive Gabor spectrogram

$$PGS(t, \omega) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \frac{p_{m,n}(t) \cdot P_{m,n}(\omega)}{\frac{1}{2\pi} \int P_{m,n}(\omega') d\omega'}. \quad (11)$$

It is clearly seen that the PGS is positive, as each function $p_{m,n}(t)$ and $P_{m,n}(\omega)$ is positive per definition. A test for the time marginal gives

$$\begin{aligned} \frac{1}{2\pi} \int PGS(t, \omega) d\omega &= \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} p_{m,n}(t) \frac{\frac{1}{2\pi} \int P_{m,n}(\omega) d\omega}{\frac{1}{2\pi} \int P_{m,n}(\omega') d\omega'} \\ &= \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} p_{m,n}(t) = |x(t)|^2. \end{aligned} \quad (12)$$

In the case of the frequency marginals, we have

$$\int PGS(t, \omega) dt = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} P_{m,n}(\omega) \frac{\int p_{m,n}(t) dt}{\frac{1}{2\pi} \int P_{m,n}(\omega') d\omega'}. \quad (13)$$

To have correct frequency marginals

$$\int p_{m,n}(t) dt = \frac{1}{2\pi} \int P_{m,n}(\omega) d\omega \quad (14)$$

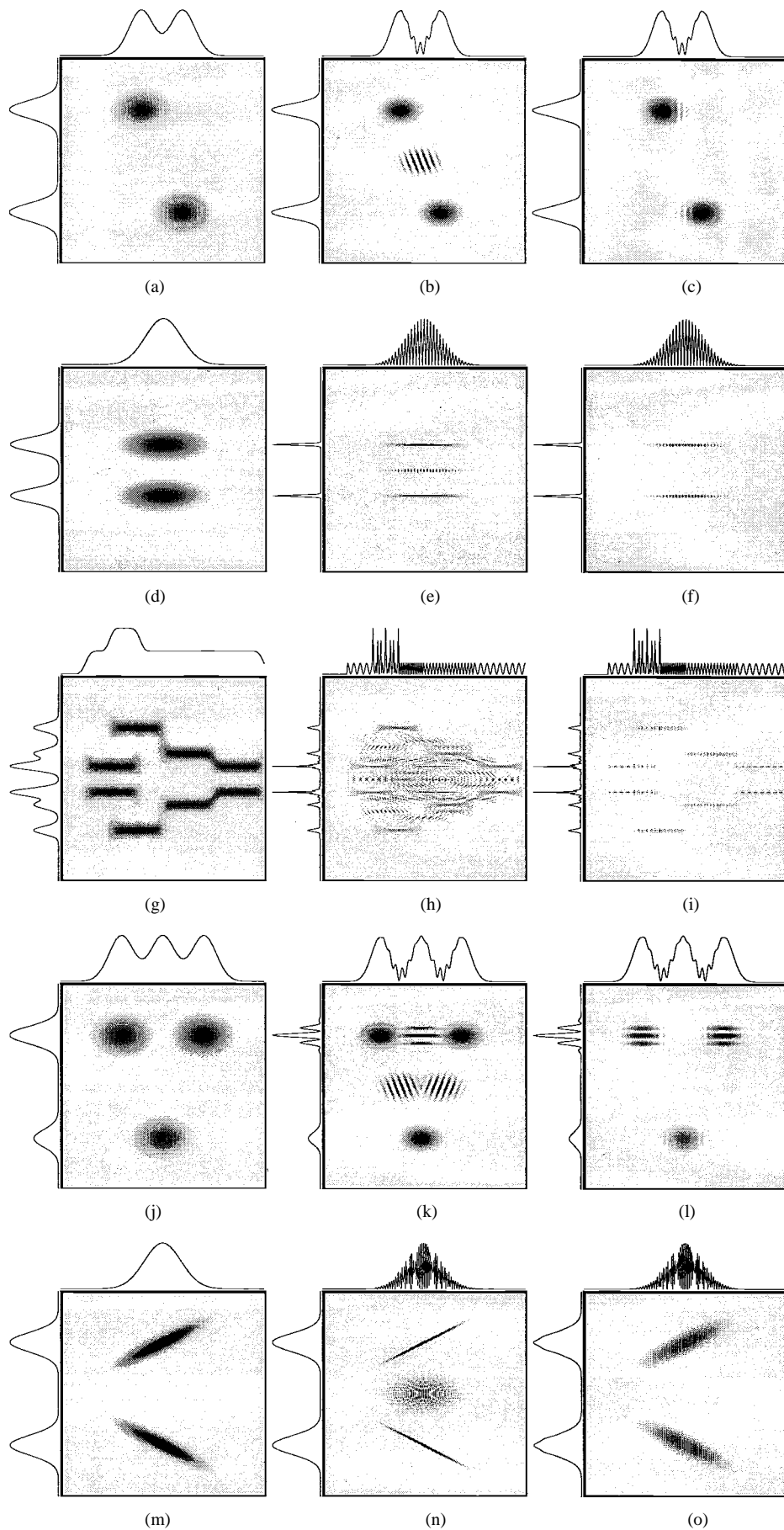


Fig. 2. Fourier spectrogram, Wigner-Ville, and positive Gabor spectrogram of five test signals.

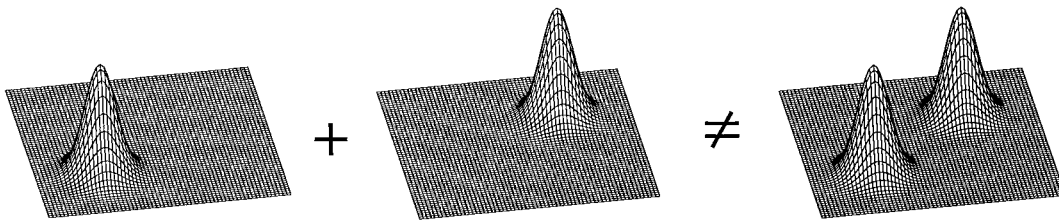


Fig. 3. Illustration of the nonadditivity of energy distributions.

must be satisfied. At present, no analytical proof of (14) has been found, but numerical simulations, in all cases that we have tested, indicate that the differences between right and left sides integrations are negligible.

In order to evaluate the nature of a proper distribution, we will test the PGS with some synthetic test signals (toy signals). The Gabor coefficients are calculated using an orthogonal-like Gabor expansion [9]. For each test signal, we have calculated the FS, the WV distribution, and the PGS. The reason for these choices is that both the WV and FS have properties that we want the PGS to possess, i.e., correct marginals and a positive level. For each distribution, their corresponding marginals have been calculated and placed above (time marginal) and to the left (frequency marginal) for the distribution.

By inspecting the PGS in Fig. 2(a)–(l), it is seen that the AT's are only affected by the CT's when these are contributing to the marginals. The WV always has CT's between the AT's independently of whether they contribute to the marginals or not. The FS gives a traditional low-resolution result. These results of the PGS are in agreement with our expectations. The last test signal *test5* contains a Gaussian shaped chirp signal, and the results are illustrated in Fig. 2(m)–(o). This result is not as convincing as the results from (a)–(l), and the reason for this will be addressed in the Discussion.

III. DISCUSSION

This correspondence has presented a positive TF distribution based on a Gabor expansion of the signal. Both the TFDS and the PGS utilized the Gabor expansion and the resulting distributions depend strongly on the particular orthogonal Gabor expansion. For both distributions, the Gabor coefficients that best reveal the TF content of the signal have to be used.

Distributions with only positive/marginal constraints have only optimal focus when the signal components are oriented in vertical or horizontal lines in the TF plan (which is not the case of *test5*). This orientation dependency of the marginal constraint was discussed by Fonollosa in [10], and here, he suggested that Loughlin's approach could be extended by including arbitrary directions in the TF plan as constraints in the optimization problem. On the other hand, he also pointed out that it will only be possible to use a finite number of angles because the only distribution that is optimally concentrated in every direction is the WV, and this is not a positive distribution.

A fundamental property of a true energy distribution is the *nonadditivity* property. Because of the nonlinear nature of energy distributions, the components are interacting with each other.

For this reason, the logic in Fig. 3 is true.

It illustrates the fact that the spectral contents of two functions added together are not the sum of the spectral content for each individual function—the spectral content is not additive. The implication of this property will be the subject of the rest of this discussion.

The effect of satisfying the marginals is that every zero-crossing in the signal is fading the distribution to zero (making vertical zero lines). Likewise, every zero in the power spectrum is also fading the distribution to zero (making horizontal zero lines). The

resulting distribution of a real signal will therefore have a very spiky shape. This effect is called *TF fading* and is a consequence of the nonadditivity property. When the distribution was implemented, the corresponding distribution was sampled in discrete points. However, if a sequence of these points falls close to or identical to these zero crossings, a part of the resulting distribution will fade to zero. Another consequence of TF fading is that if two distributions are calculated for the same signal, except that one has all time points shifted one half of the sample, then the two distributions will be quite different. This effect is not known from the FS because of the smoothing effect. TF fading is one reason why correct marginals may not be practicable.

A consequence of the frequency marginal is that one should be very cautious when interpreting, as an example, an analysis of a speech signal. A proper distribution is context dependent, which means that a spectral representation of a particular word in a sentence will depend strongly on what is said before and after this word. Therefore, proper distributions may not be the best tool for speech analysis.

REFERENCES

- [1] S. Qian and J. M. Morris, "Wigner distribution decomposition and cross-terms deleted representation," *Signal Process.*, vol. 25, pp. 125–144, May 1992.
- [2] S. Qian and D. Chen, "Decomposition of the Wigner–Ville distribution and time-frequency distribution series," *IEEE Trans. Signal Processing*, vol. 42, pp. 2836–2842, Oct. 1994.
- [3] F. Pedersen, "Joint time frequency analysis in digital signal processing," Ph.D. dissertation, Aalborg Univ., Aalborg, Denmark, May 1997.
- [4] L. Cohen and T. E. Posch, "Positive time-frequency distribution functions," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, pp. 31–38, Feb. 1985.
- [5] P. J. Loughlin, "Time-frequency energy density functions: Theory and synthesis," Ph.D. dissertation, Univ. Washington, Seattle, 1992.
- [6] P. J. Loughlin, J. W. Pitton, and L. E. Atlas, "Construction of positive time-frequency distributions," *IEEE Trans. Signal Processing*, vol. 42, pp. 2697–2705, Oct. 1994.
- [7] S. Kadambe and G. F. Boudreaux-Bartels, "A comparison of the existence of 'cross terms' in the Wigner distribution and the squared magnitude of the wavelet transform and the short time Fourier transform," *IEEE Trans. Signal Processing*, vol. 40, pp. 2498–2517, Oct. 1992.
- [8] J. Jeong and W. J. Williams, "Mechanism of the cross-terms in spectrograms," *IEEE Trans. Signal Processing*, vol. 40, pp. 2608–2613, Oct. 1992.
- [9] S. Qian and D. Chen, "Discrete Gabor transform," *IEEE Trans. Signal Processing*, vol. 41, pp. 2429–2438, July 1993.
- [10] J. R. Fonollosa, "Positive time-frequency distributions based on joint marginal constraints," *IEEE Trans. Signal Processing*, vol. 44, pp. 2086–2091, Aug. 1996.