WARSAW UNIVERSITY TECHNOLOGY

FACULTY OF ELECTRONICS

AND INFORMATION TECHNOLOGY

INSTITUTE OF RADIOELECTRONICS

Report No.1, 2007

Audio watermarking using two anti-slope chirps and windowed double-dimensional Wigner distributions

Stefan L. Hahn and Kajetana M. Snopek

Warsaw, December 2007

Institute of Radioelectronics, Nowowiejska 15/19, 00-665 Warsaw, tel./fax: (4822) 8255248

e-mail: hahn@ire.pw.edu.pl. K.Snopek@ire.pw.edu.pl

1 Introduction

In the last decade, we observe the intense developement of the process of embedding in the original signal additional (host) data in the form of a watermark signal [1]-[5]. For example, watermarking may serve for the identification of the copyright ownership. The embedded data can carry various information. The audio watermark should be undetectable by a listener and the video watermark invisible for a viewer. However, the efficient watermarking is possible if methods of safe extraction of the watermark from the host are available. It is not our intension to present here a review of watermarking procedures. The paper presents a novel and specific procedure of audio watermarking using two embedded linear chirp signals of opposite slopes and its extraction using the double-dimensional Wigner distribution. The last one has been defined by the authors a few years ago [6].

2 Theoretical background

2.1 Double-dimensional Wigner distributions

The Wigner distribution (*WD*) [6] of a real or complex signal $\psi(t)$ is defined by the Fourier transform of the correlation product $r(t,\tau)=\psi(t+0.5\tau)\psi^*(t-0.5\tau)$:

$$W(t,f)=F\left[r(t,\tau)\right]=\int_{-\infty}^{\infty}\psi(t+0.5\tau)\psi^*(t-0.5\tau)e^{-j2\pi f\tau}d\tau\,. \qquad (1)$$

The double-dimensional *WD,* i.e., the *WD* of the *WD*, is defined by the 2D Fourier transform of the product $W\left(t+0.5\chi_t,f+0.5\mu_f\right)W\left(t-0.5\chi_t,f-0.5\mu_f\right)$:

$$W_W^{(2)}(t,f,\mu,\tau)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(t+0.5\chi_t,f+0.5\mu_f\right)W\left(t-0.5\chi_t,f-0.5\mu_f\right)e^{-j2\pi\left(\mu\chi_t+\tau\mu_f\right)}d\chi_t d\mu_f\,. \qquad (2)$$

However, there is no need to calculate the double integral since the so called (corrected) O'Neil-Flandrin formula can be applied [6]

$$W_W^{(2)}(t,f,\mu,\tau)=W\left(t+\frac{\tau}{2},f-\frac{\mu}{2}\right)W\left(t-\frac{\tau}{2},f+\frac{\mu}{2}\right). \qquad (3)$$

2.2 The Wigner distribution of a single chirp signal

Consider a chirp signal of the form

$$\psi_1(t) = A_1 e^{-c(t-t_1)^2} e^{j2\pi\left(a_1 t + b_1 t^2\right)}. \tag{4}$$

In the time-frequency plane, its "instantaneous frequency" is $f(t) = a_1 + 2b_1 t$. Remark: this is not the instantaneous frequency defined using the Gabor analytic signal since the chirp signal is not analytic (Its real and imaginary parts are not Hilbert transforms). The Wigner distribution of (4) is

$$W_1(t,f) = \sqrt{\frac{\pi}{2c}} A_1^2 e^{-2c(t-t_1)^2} e^{-\frac{\pi^2}{2c}(f-a_1-b_1 t)^2}. \tag{5}$$

The corresponding double dimensional Wigner distribution is defined using (3). Here, our interest is focused only on slices $W_W^{(2)}(t,f,0,0) = \left[W_1(t,f)\right]^2$. Fig.1 shows two examples, the first one with "increasing" and the second one with "decreasing" "instantaneous frequencies".

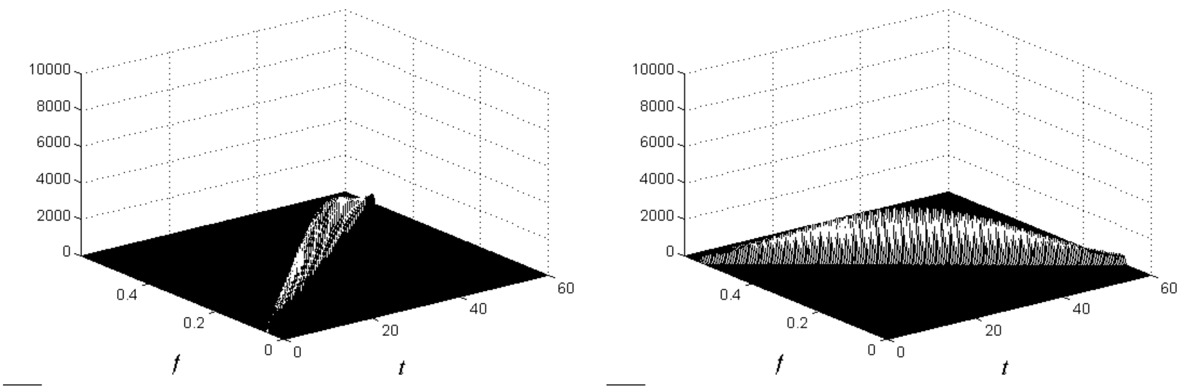

Fig.1. Left: The slice $W_W^{(2)}(t.f,0.0)$ of a chirp (Gaussian envelope) with linearly increasing frequency. Right: The same with linearly decreasing frequency

3 Watermarking using "chirp" signals

3.1 Watermarking using the sum of two anti-slope chirp signals

Consider the time-frequency plane with the following range of variables: $t \in (0, t_{\max})$ and $f \in (0, f_{\max})$ and a sum of two chirp signals:

$$\psi(t) = e^{-c(t-t_0)^2}\left(A_1 e^{j2\pi\left(a_1 t + b_1 t^2\right)} + A_2 e^{j2\pi\left(a_2 t + b_2 t^2\right)}\right). \tag{6}$$

We assume $a_1 = \alpha f_{\max}$, $a_2 = (1-\alpha) f_{\max}$, $b_1 = f_{\max} / 2t_{\max}$ and $b_2 = -b_1$. This yields chirps with the following "instantaneous frequencies:": The increasing frequency $f_1(t) = \alpha f_{\max} + 2bt$ and the decreasing frequency $f_2(t) = (1-\alpha) f_{\max} - 2bt$ (see Fig. 2).

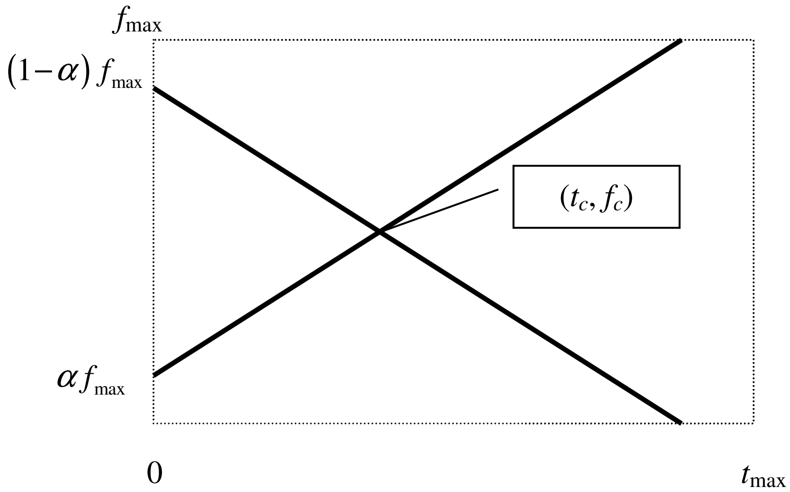

Fig.2. The lines $f_1(t) = \alpha f_{\max} + 2bt$ and $f_2(t) = (1-\alpha) f_{\max} - 2bt$.

The coordinates of the crossing point given by the equation $f_1(t) = f_2(t)$ are

$$t_c = \frac{(1-2\alpha) f_{\max}}{4b} \quad ; \quad f_c = 0.5 f_{\max} \tag{7}$$

The double-dimensional *WD* of this signal is displayed in Fig. 3, left. We observe a large peak at the crossing point of the instantaneous frequencies. A sequence of such peaks may be used to encode any information, i.e., serve as a watermark. It is well known that the Wigner distribution of a sum of two signals contains so called cross-terms. However, the cross-terms generated by the above described chirp signals are negligible, as displayed in Fig. 3, right.

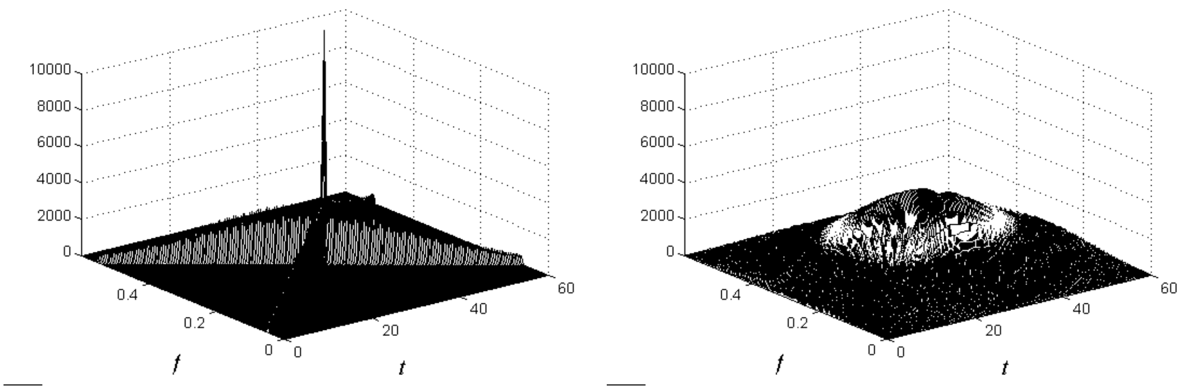

Fig.3. Left: The slice of the double-dimensional *WD* of two chirps defined by (6). Right: The cross term with the amplitude multiplied by 100.

3.2 Windows used to extract the pulse of Fig. 3.

The pulses generated by the watermark chirp signals can be extracted from the sum of the host signal and chirp signals using very narrow windows in the time-frequency plane. Fig. 4 shows examples of such windows.

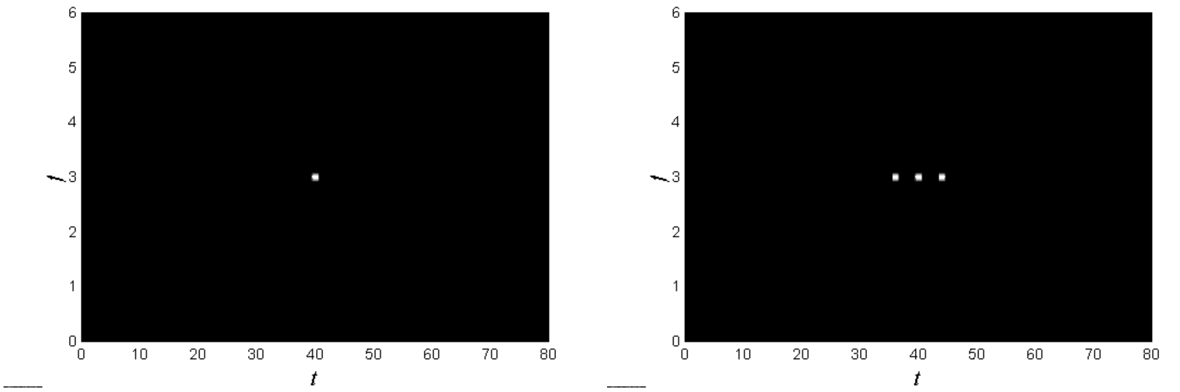

Fig. 4. Windows in the time-frequency plane. Left: A single window. Right: Three windows.

3.3 Watermarking using a set of 3 chirps

Let us introduce to (6) a third chirp signal getting

$$u(t) = e^{-c(t-t_i)^2}\left(A_1\cos\left[2\pi\left(a_1t+b_1t^2\right)\right]+A_2\cos\left[2\pi\left(a_2t+b_2t^2\right)\right]+A_3\cos\left[2\pi\left(a_3t+b_3t^2\right)\right]\right). \tag{8}$$

If we apply $A_1 = A_2 = A_3$ and equal negative slopes $b_2 = b_3 = -b_1$, with different values $a_2 \neq a_3$, we get two points of intersection of "instantaneous frequences" in the time-frequency plane. Notice that the "rising" chirp is

$$u_1(t) = A_1 e^{-c_1(t-t_0)^2} \cos\left[2\pi\left(\alpha f_{\max} t + bt^2\right)\right]. \qquad (9)$$

The two "falling" chirps are

$$u_2(t) = A_2 e^{-c_2(t-t_1)^2} \cos\left\{2\pi\left[\left(\left((1-\alpha) f_{\max} + \beta\right)t - bt^2\right)\right]\right\}, \qquad (10)$$

$$u_3(t) = A_3 e^{-c_3(t-t_2)^2} \cos\left\{2\pi\left[\left(\left((1-\alpha) f_{\max} - \beta\right)t - bt^2\right)\right]\right\}. \qquad (11)$$

The corresponding „instantaneous frequencies" are:

$$f_1(t) = \alpha f_{\max} + 2bt, \qquad (12)$$

$$f_2(t) = (1-\alpha) f_{\max} + \beta - 2bt, \qquad (13)$$

$$f_3(t) = (1-\alpha) f_{\max} - \beta - 2bt. \qquad (14)$$

We get the following coordinates of crossing points (see Fig. 5): The first one is given by the equation $f_1(t) = f_2(t)$:

$$t_{12} = \frac{(1-2\alpha) f_{\max} + \beta}{4b}; \quad f_{12} = 0.5\left(f_{\max} + \beta\right). \qquad (15)$$

The second one by the equation $f_1(t) = f_3(t)$:

$$t_{13} = \frac{(1-2\alpha) f_{\max} - \beta}{4b}; \quad f_{23} = 0.5\left(f_{\max} - \beta\right) \qquad (16)$$

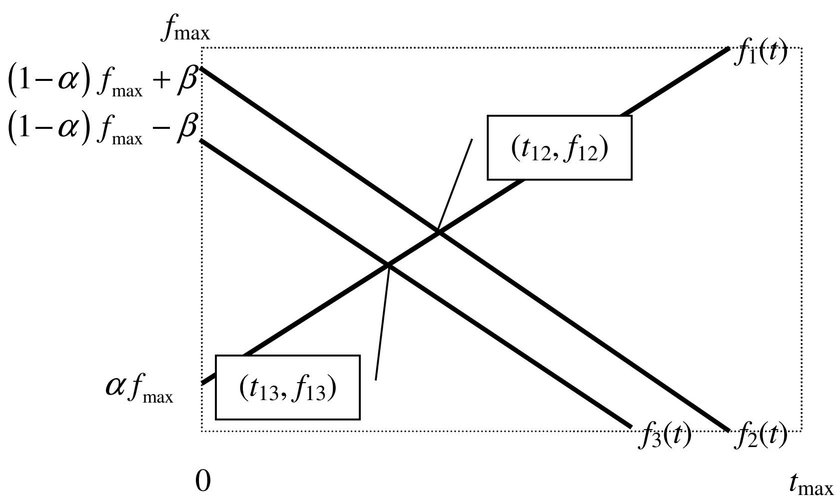

Fig. 5 The lines of "instantaneous frequences" of the sum of three chirps

3. 4 Watermarking using a set of 4 chirps

The four chirp signals have the form

$$u_1(t) = A_1 e^{-c_1(t-t_A)^2} \cos\left\{2\pi\left[\left(\alpha f_{\max} + \beta\right)t + bt^2\right]\right\}, \tag{17}$$

$$u_2(t) = A_2 e^{-c_2(t-t_B)^2} \cos\left\{2\pi\left[\left(\alpha f_{\max} - \beta\right)t + bt^2\right]\right\}, \tag{18}$$

$$u_3(t) = A_3 e^{-c_3(t-t_C)^2} \cos\left\{2\pi\left[\left((1-\alpha) f_{\max} + \beta\right)t - bt^2\right]\right\}, \tag{19}$$

$$u_3(t) = A_4 e^{-c_4(t-t_D)^2} \cos\left\{2\pi\left[\left((1-\alpha) f_{\max} - \beta\right)t - bt^2\right]\right\}. \tag{20}$$

The corresponding "instantaneous frequencies" are

$$f_1(t) = \alpha f_{\max} + \beta + 2bt, \tag{21}$$

$$f_2(t) = \alpha f_{\max} - \beta + 2bt, \tag{22}$$

$$f_3(t) = \left(1 - \alpha f_{\max}\right) + \beta - 2bt, \tag{23}$$

$$f_4(t) = \left(1 - \alpha f_{\max}\right) - \beta - 2bt. \tag{24}$$

We have four points of intersection of the *WD*s defined by the equality of the 'instantaneous frequencies" (see Fig. 6 points denoted with subscripts 13, 23, 14 and 24).

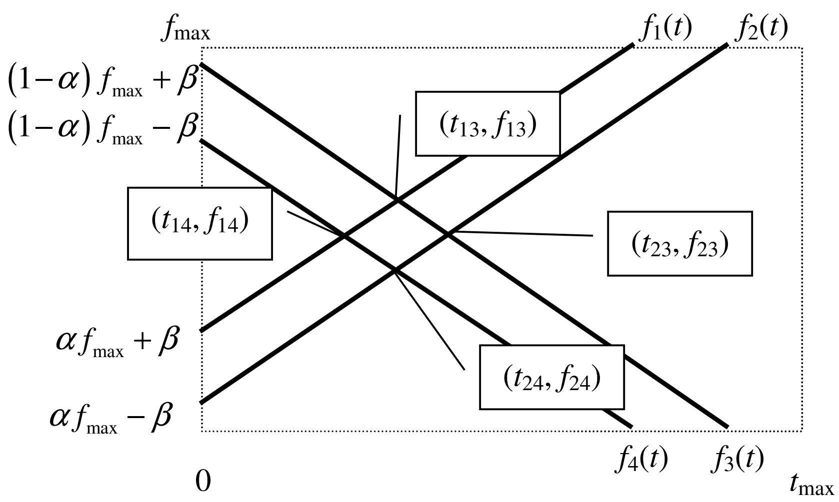

Fig. 6 The lines of "instantaneous frequences" of the sum of four chirps

The coordinates of the point 13 are given by the equation $f_1(t) = f_3(t)$ yielding

$$t_{13}=\frac{(1-2\alpha)f_{\max}}{4b}\ ;\ f_{13}=0.5f_{\max}+\beta\,. \tag{25}$$

Analogously, the equation $f_2(t)=f_3(t)$ yields

$$t_{23}=\frac{(1-2\alpha)f_{\max}+2\beta}{4b}\ ;\ f_{23}=0.5f_{\max}\,. \tag{26}$$

The equation $f_2(t)=f_4(t)$ yields

$$t_{24}=\frac{(1-2\alpha)f_{\max}}{4b}\ ;\ f_{24}=0.5f_{\max}-\beta\,. \tag{27}$$

and the equation $f_1(t)=f_4(t)$ yields

$$t_{14}=\frac{(1-2\alpha)f_{\max}-2\beta}{4b};\ \ f_{14}=f_{23}=0.5f_{\max}\,. \tag{28}$$

4 Transmission of information using the pulses generated in the WD of the watermark signal

4. 1 Binary transmission using a single pulse

We assume that two anti-slope chirp signals generating a single pulse of the double-dimensional *WD* are repeated at regular time intervals. The presence of the pulse is detected as logical ***1*** and the absence of the pulse as logical ***0***. We can have binary transmission corresponding to the on-off keying. Each pulse is detected using a single window in the time-frequency plane shown in Fig. 4.

4.2 M-ary transmission

M-ary transmission can be realized in various ways. For example, if the generated pulses are shifted in time they can be detected by a multi-window, for example the window of Fig. 4 with three time shifted windows. Using two such windows it is possible to realize a robust binary transmission, where the symbol ***0*** corresponds to the detection of a pulse by the first window and the symbol ***1***, by the second window. However, if we allow generation of pulses in both windows, we get 4-ary transmission: ***00, 01, 10*** and ***11***. Using a sequence of three time shifted windows we can realize a robust ternary transmission using ***0, 1*** and ***2*** or a nonrobust 8-ary transmission using ***000, 001, ….,***

110, 111. Other possibilities are given by a larger number of crossing points, for example three crossing points of Fig. 5 or four crossing points of Fig.6

References

[1] S.Disch, J.Herre, J.Kammerl, "Audio watermarking using subband modulation spectra," ICASSP 2007, I-245-248.

[2] I.Djurović, S.Stanković, I.Pitas, L. Stanković, "Generalization of the Fourier domain watermarking to the space/spatial-frequency domain, First Intern. Workshop on Image and Signal Processing, June 14-15,2000, Pula, Croatia.

[3] P.Dymarski, "Watermarking of Audio Signals Using Adaptive Subband Filtering and Manchester Signaling,"

[4] S.Erküclik, S.Krishnan, M.Zeytinoğlou, "A robust watermark representation based on linear chirps," IEEE Trans.on Multimedia, vol.8, No.5, October 2006, pp.925-938

[5] I.Orović, P.Zogović, N.Žarić, S.Stanković, "Speech signal protection via log-watermarking based on time-frequency analysis,"

[6] S.L. Hahn, K.M.Snopek, "The Double-dimensional Distributions. Another Approach to "Quartic"Distributions," IEEE Transactions on Signal Processing, Vol. 50, No.12, December 2002, 2987-2997.

[7] S.Stanković, I.Diurović, I.Pitas, "Watermarking in the spece/spatial-frequency domain using two-dimensional Radon-Wigner distribution," IEEE Transactions on Signal Processing, Vol.10, No.4, April 2001, pp.650-658.