The pulsed to tonal strength parameter and its importance in characterizing and classifying Beluga whale sounds

Ramón Miralles[a] and Guillermo Lara
Instituto de Telecomunicación y Aplicaciones Multimedia (iTEAM), Universitat Politècnica de València, Valencia, Spain

Jose Antonio Esteban
Parques Reunidos Valencia S. A. L'Oceanogràfic, Ciudad de las Artes y las Ciencias, Valencia, Spain

Alberto Rodriguez
Departamento de Ingeniería de Comunicaciones, Universidad Miguel Hernandez, Elche, Spain

(Received 27 October 2011; revised 11 January 2012; accepted 13 January 2012)

A large number of the vocalization studies on mammals are based on time–frequency analysis of the produced sounds. The patterns, which are extracted from the time–frequency representations, determine the classification in the different sound categories. However, there are situations where this pattern related recognition does not allow a precise characterization of the vocalization to be obtained. In these situations, a feasible alternative, which can help by giving the dominant component of the sound, is to measure the strength of the tonal and pulsed constituent units. In this work, the use of a ratio of pulsed to tonal strength is proposed to objectively measure the distribution of energy between these two components. This pulsed to tonal ratio (*PTR*) can be computed with the aid of the discrete cosine transform. It is demonstrated that the *PTR* can be obtained with a relatively simple expression without having to go through the time–frequency representation. This work presents examples that show how the *PTR* can be used to distinguish between two very similar Beluga whale sounds and how to dynamically track the power distribution between the pulsed and tonal components in non-stationary signals.
 [DOI: 10.1121/1.3682056]

PACS number(s): 43.60.Hj, 43.80.Ka, 43.30.Sf [SAF] Pages: 2173–2179

I. INTRODUCTION

The time–frequency analysis of the underwater sounds of the Beluga whale is a valuable research technique for marine biologists. The way the energy is distributed in time and frequency is used to classify the different sounds that are emitted by the animals. Typical cetacean sound classifications include pulsed and tonal sounds as well as a large variety of combinations and alterations of these two main categories. The process of categorizing whale songs in discrete sets is complex and typically emphasizes the distinctive properties of prototypical units that are heard from a distance.[1]

The classification process, as with many other maritime mammals, involves recognizing the patterns that the sounds produce in the time–frequency representations (TFRs)[2,3] and how the energy is distributed in the different parts of the whale song. The process, which is normally done with the aid of software tools, is frequently supervised by a trained researcher after listening and analyzing the TFR of a whale sound. However, there are some situations where the software algorithms based on geometric parameter extraction from the TFR patterns do not guarantee a correct recognition of the vocalization category. In these situations, even having an expert carefully listen to the vocalizations does not ensure a precise and objective classification of the sound. An example of this is what happens when trying to classify Beluga whale sounds that mix several elements [whistles, regular clicks, and rapid-click buzzes (creaks)] in the same vocalization. The examination of the TFR of these sounds using false color images does not allow the researchers to precisely determine if the sound has a dominant pulsed or tonal component. Figure 1 illustrates this problem. Establishing a threshold to classify sound units as pulsed or tonal just by listening to it is also rather arbitrary and subjective. It would be more interesting if we could extract objective information about how the energy (or mean power) is distributed in the TFR.

A feasible alternative that is closely related to some of the characteristics that are taken into account for sound classification is to obtain the ratio of the pulsed to tonal strength (*PTR*). The *PTR* could be obtained as follows:

$$PTR(dB) = 10 \cdot \log\left[\frac{P_{pulsed}}{P_{tonal}}\right], \quad (1)$$

where P_{pulsed} is the mean power of the pulsed components of the sound and P_{tonal} is the mean power of tonal components of the sound.

The separation of tonal components from impulses can be done using the reassigned spectrogram magnitude as it was presented in Fulop and Fitz.[4] An alternative approach can also be derived by decomposing the TFR using discrete cosine transforms. This new approach is computationally efficient and may complement the Fulop and Fitz technique.

[a]Author to whom correspondence should be addressed. Electronic mail: rmiralle@dcom.upv.es

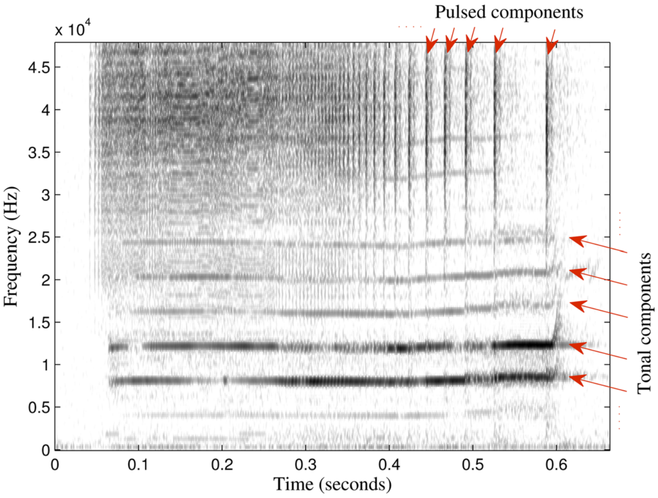

FIG. 1. (Color online) TFR of a Beluga sound (called "Creak whistle" by the biologists). The arrows indicate the location of the pulsed and tonal components.

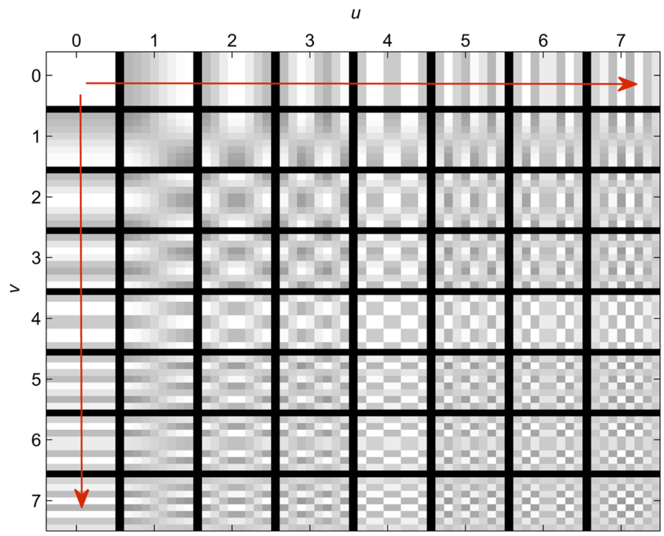

FIG. 2. (Color online) The DCT basis for a transformation of size 8×8. The indexes along the top and left side of the image represent the vertical and horizontal spatial frequency coefficient indexes. The arrows indicate the horizontal and vertical coefficients used for textural analysis.

The work is structured as follows. In Sec. II, we show how to compute the *PTR* parameter using discrete cosine transforms, and we present simplified expressions of the parameter. In Sec. III, the proposed parameter is validated through simulations. A comparison of the proposed method with results obtained using the technique described by Fulop and Fitz is also shown in this section. Section IV presents an application of the proposed parameter to illustrate its importance in characterizing and classifying complex Beluga whale sounds. Finally, we present our conclusions.

II. PULSED TO TONAL RATIO CALCULATION

Let us assume that we have computed the TFR of the underwater Beluga sound that is under analysis (or a fragment of it) using one of the many available algorithms (spectrogram, Wigner–Ville, Cohen's class, etc.). We call the discrete TFR $P(n, k)$, where $n = \{0,1,\ldots, N_1 - 1\}$. The analysis parameter N_1 is related to sound duration or a window data analysis length. The discrete frequency k is obtained by sampling $P(n,w)$ over the unit circle with N_2 points ($k = \{0,1,\ldots, N_2 - 1\}$). In the TFR representation, part of the energy is distributed in the pulsed component of the sound (vertical lines in TFR), and another part of the energy is distributed in the tonal component (horizontal bands in TFR). Figure 1 shows an example of the pulsed and tonal components. An intuitive method for obtaining the relationship of the pulsed to tonal strength can be devised if the TFR is processed as an image. The bidimensional discrete cosine transform (DCT_{2D}) can be used to extract the texture degree of the energy image.[5,6] The horizontal and vertical coefficients of the DCT_{2D} can be used to obtain the mean power in the tonal and pulsed components, which is illustrated in Fig. 2. We are interested in obtaining the ratio of the mean power that is present in the pulsed component to the mean power that is present in the tonal component. We use the sum of the squared DCT_{2D} coefficients since this sum can be used to derive shift-insensitive spatial frequency descriptor according to Frye and Ledley.[7] The ratio can be calculated as

$$PTR(dB) = 10 \log \left[\frac{\sum_{u=0}^{N_1-1} F(u,0)^2}{\sum_{v=0}^{N_2-1} F(0,v)^2} \right], \quad (2)$$

where $F(u, v)$ is the DCT_{2D} of the TFR.

The proposed *PTR* calculation can be simplified if it is taken into account that the DCT_{2D} can be computed using the DCT-II as follows:

$$F(u,v) = \sum_{n=0}^{N_1-1} \sum_{k=0}^{N_2-1} P(n,k) \cos\left[\frac{\pi}{N_1}\left(n+\frac{1}{2}\right)u\right] \cos\left[\frac{\pi}{N_2}\left(k+\frac{1}{2}\right)v\right],$$
$$u = 0,1,\ldots,N_1-1,$$
$$v = 0,1,\ldots,N_2-1. \quad (3)$$

If Eq. (3) is particularized for $F(u, 0)$ and $F(0, v)$ which are, respectively, used to calculate the numerator and the denominator of Eq. (2) we obtain

$$F(u,0) = \sum_{n=0}^{N_1-1} \sum_{k=0}^{N_2-1} P(n,k) \cos\left[\frac{\pi}{N_1}\left(n+\frac{1}{2}\right)u\right],$$
$$u = 0, 1, \ldots, N_1 - 1$$

$$F(0,v) = \sum_{k=0}^{N_2-1} \sum_{n=0}^{N_1-1} P(n,k) \cos\left[\frac{\pi}{N_2}\left(k+\frac{1}{2}\right)v\right],$$
$$v = 0, 1, \ldots, N_2 - 1. \quad (4)$$

It can be observed from Eq. (4) that only TFR marginals need to be calculated. The discrete time and frequency marginals are given by (5) and (6), respectively,

$$P(n) = \sum_{k=0}^{N_2-1} P(n,k), \quad (5)$$

$$P(k) = \sum_{n=0}^{N_1-1} P(n,k). \quad (6)$$

Note that now, in the computation of Eq. (4), only one-dimensional DCTs are needed. Depending on the TFR associated with the computed marginals, Eqs. (5) and (6) can be simplified even more, and optimum and fast algorithms can be devised to calculate the *PTR*.

A. Computation of the *PTR* with marginals of the spectrogram

As a first example, we can obtain the marginals assuming that $P(n,k)$ is the spectrogram computed using the short-time Fourier transform. We define DFT_{N_2} $[\cdot]$ as an operator to compute the discrete-time, discrete-frequency Fourier transform with N_2 being the number of equally spaced samples from $w = 0$ to 2π. We use $x(n)$ to design the sound fragment of interest where we want to compute the *PTR*. We also call the N_2 analysis window $w(n)$, which is assumed to be non-zero only in the interval $n = [0, 1, \ldots, N_2 - 1]$. The marginals (5) and (6) can then be written as

$$P(n) = \sum_{k=0}^{N_2-1} |DFT_{N_2}[w(m) \cdot x(n-m)]|^2$$
$$= N_2 \sum_{m=0}^{N_2-1} |w(m) \cdot x(n-m)|^2 = N_2 \cdot \hat{x}(n), \quad (7)$$

$$P(k) = \sum_{n=0}^{N_1-1} |DFT_{N_2}[w(m) \cdot x(n-m)]|^2 = \widehat{X}(k). \quad (8)$$

Equation (7) has been simplified using Parseval's theorem, and it shows that the time marginal is proportional to the smoothed energy calculation of the sound fragment using the sliding window $w(n)$. We use $\hat{x}(n)$ to refer to this smoothed energy estimation. Equation (8) shows that the frequency marginal is equivalent (with a scale factor) to smoothed spectra estimation using Bartlett's method.[8] We use $\widehat{X}(k)$ to refer to the smoothed spectra estimation. Taking all this into account, Eq. (2) can be computed using the results of the following:

$$F(u,0) = DCT_{1D}[N_2 \cdot \hat{x}(n)], \quad (9)$$

$$F(0,v) = DCT_{1D}[\widehat{X}(k)], \quad (10)$$

with $\hat{x}(n)$ and $\widehat{X}(k)$ previously defined. In this case, only one-dimensional discrete cosine transforms are required ($DCT_{1D}[\cdot]$). This greatly simplifies the computation.

B. Computation of the *PTR* with distributions that satisfy marginal properties

The computation of the *PTR* can be simplified even more if the associated discrete TFR satisfies marginal properties such as discrete Wigner–Ville distribution[9,10] or discrete positive distributions (Coehn's class type II,[11] etc.). For these distributions, Eqs. (5) and (6) result in

$$P(n) = K_1 |x(n)|^2, \quad (11)$$

$$P(k) = K_2 |X(k)|^2, \quad (12)$$

where $X(k) = DFT_{N_2}[x(n)]$. The constants K_1 and K_2 depend on how the particular positive distribution is extended to discrete-time signals. Table I illustrates the possible values for K_1 and K_2, which are extracted from some of the alternatives for computing the discrete-time, discrete-frequency TFRs.[9,10,12]

Thus Eq. (2) can also be computed using

$$F(u,0) = DCT_{1D}[K_1 |x(n)|^2], \quad (13)$$

$$F(0,v) = DCT_{1D}[K_2 |X(k)|^2], \quad (14)$$

with K_1 and K_2, which have already been given in Table I.

The proposed computation of the *PTR* presented in Secs. II A and II B, which only involves marginals, is not affected by inherent problems of the TFR computation such as cross-terms. Additionally, resolution and parameter setup becomes an easy task, and as a result classical spectral density estimation concepts can be used.[8]

C. Application to non-stationary audio signals with slow variation of the *PTR*

In the analysis of bioacustic signals, it is quite frequent to find slow variations of the tonal or pulsed components. The above-proposed simplifications allow the *PTR* parameter to be easily computed and used to dynamically track the pulsed to tonal strength ratio.

To do so, we refer to $r(n)$ as the audio signal where we want to compute the evolution of the *PTR*. We also use a temporal rectangular window $v(n)$ of length $N = N_1 + N_2$. With the algorithm described in Sec. II B we can compute the evolution of the *PTR* as the moving window $v(n)$ slides over the audio signal $r(n)$. The evolution of the *PTR* in a given sound can be a valuable tool for examining and comparing sounds. It can also be used to establish thresholds to determine if the cetacean vocalization is more pulsed than tonal or vice versa.

III. SIMULATION

We have used MATLAB to create synthetic sounds with controlled pulsed and tonal mean power.[13] Equation (15) models how the synthetic sounds were obtained,

$$x_i(t) = \tilde{w}(t) + \sum_m \Pi\left(\frac{t - m\rho_{pulsed}^{-1}}{T_{pulsed}}\right)\tilde{m}(t)$$
$$+ \sum_{m=1}^{N_h} A\,sin(2\pi f_0 mt). \quad (15)$$

TABLE I. Marginal constants K_1 and K_2 for some discrete-time, discrete-frequency TFRs found in the bibliography.

Positive distribution	K_1	K_2	Defined in:
Discrete Wigner–Ville	N_2/f_s	f_s	As formulated in Ref. 9
Discrete Wigner–Ville (WS)	N_2	N_1	As formulated in Ref. 10
Type II Cohen's Class	1	1	With the kernel constraints in Ref. 12

In Eq. (15), $\{\tilde{w}(t)\}$ and $\{\tilde{m}(t)\}$ are zero mean Gaussian stochastic processes that respectively model the observation noise (mean power σ_w^2) and the pulsed bursts (mean power σ_m^2). The function $\Pi(t/T)$ is the rectangular function of duration T. Some other variables appear in the model in Eq. (15): the pulsed burst length (T_{pulsed}), the density of pulses per second (ρ_{pulsed}), the amplitude of the tonal component (A), the central frequency of the tonal component (f_0), and the number of harmonics (N_h). Parameters in Eq. (15) ensure that $\rho_{\text{pulsed}}^{-1} > T_{\text{pulsed}}$ to avoid overlapping of pulsed components. In all the simulations presented in this section, the signals have been sampled at $f_s = 96$ kHz.

Theoretical *PTR* was computed as stated in the following:

$$PTR_{theo}(dB) = 10\ log\left[\frac{\rho_{pulsed} T_{pulsed} \sigma_m^2}{N_h A^2/2}\right]. \quad (16)$$

Figure 3 shows simulations of Eqs. (15), (2), and (16) for the proposed DCT-based estimator in two cases—when the amplitude of the tonal component (A) increases and the rest of the variables remain fixed; and when ρ_{pulsed} increases and the rest of the variables remain fixed. Figure 3 also shows a comparison of the *PTR* when the tonal and pulsed components are separated using the reassigned spectrogram through the channelized instantaneous frequency with mixed partial derivative (CIFderiv) thresholding.[4] The thresholds were selected based on the recommendations by Fulop and Fitz. A threshold value of 0.01 was used for isolating tonal components, whereas threshold values of between 0.75 and 1.25 were used for isolating pulsed components. The reassigned spectrogram was computed using frames of 256 points and a frame advance of 20 points. A Hanning window was employed in the spectrogram computation.

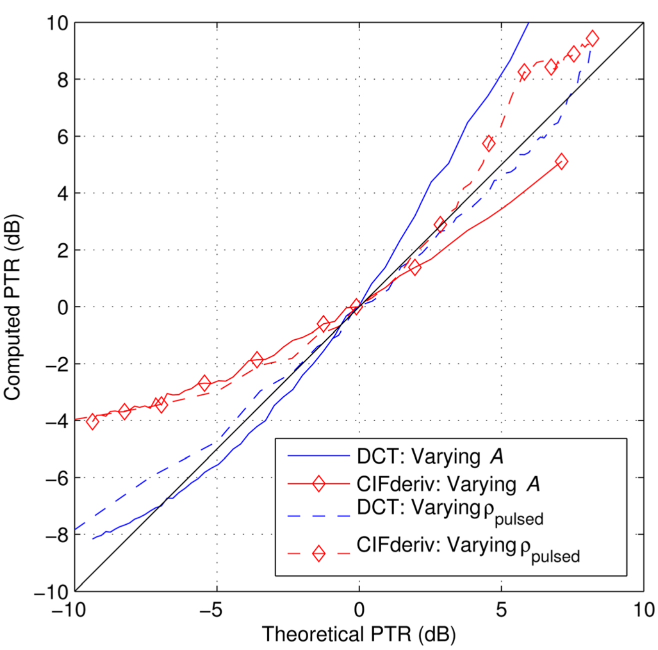

FIG. 3. (Color online) Evolution of the estimated *PTR* when the theoretical *PTR* increases. *PTR* estimates with the proposed method in the case where TFR satisfies marginal properties (DCT based), and with the mixed derivative thresholding (CIFderiv based) are shown. Signal length: 0.5 s, $T_{\text{pulsed}} = 5$ ms, $\sigma_m^2 = 4$, $N_h = 1$. The value changes in the range $A = [1.5,\ldots,10]$ and $\rho_{pulsed} = 70$ in one case and in the range $A = 2.5$ and $\rho_{\text{pulsed}} = [2,\ldots,250]$ in the other.

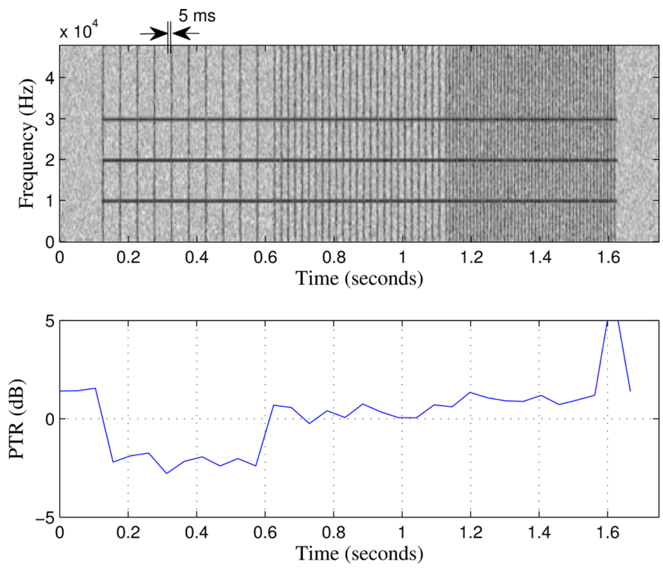

FIG. 4. (Color online) Simulated sound of 1.5 s. Spectrogram (top); *PTR* evolution (bottom). The first 0.5 s contain a pulse density of 20 pulses/s, the next 0.5 s contain a pulse density of 50 pulses/s, and the third 0.5 s contain a pulse density of 100 pulses/s.

The comparison in Fig. 3 shows that both techniques give quite an accurate estimation of the real *PTR*. However, the proposed DCT-based estimation gives a higher sensitivity to small changes in low *PTR* signals. Figure 3 also shows that, even though the proposed estimators of the *PTR* have some bias and are not linear, they are monotonically increasing in the simulation range. The maximum bias observed in the [–10,10] dB range is obtained at the ends of the curve, when the sound is mainly tonal or pulsed. In these situations, the measured error is approximately 4 dB for the DCT-based estimator and 6 dB for the CIFderiv-based estimator. Despite the observed bias,

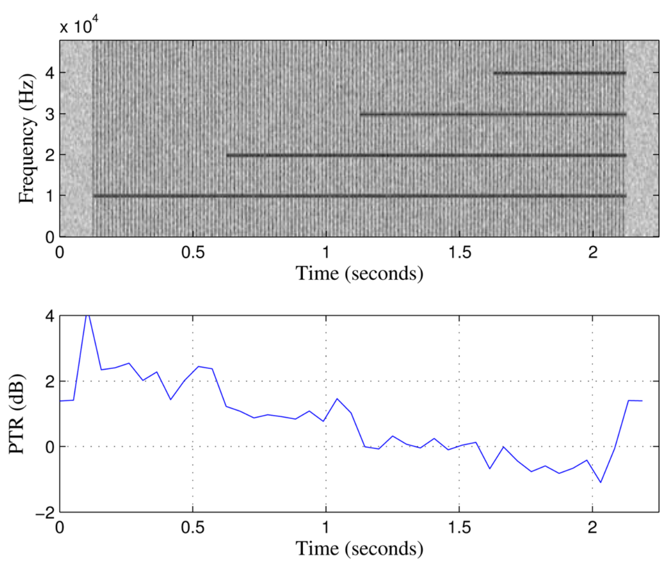

FIG. 5. (Color online) Simulated sound of 2 s. Spectrogram (top); *PTR* evolution (bottom). The number of 10 kHz harmonics increases by one every 0.5 s.

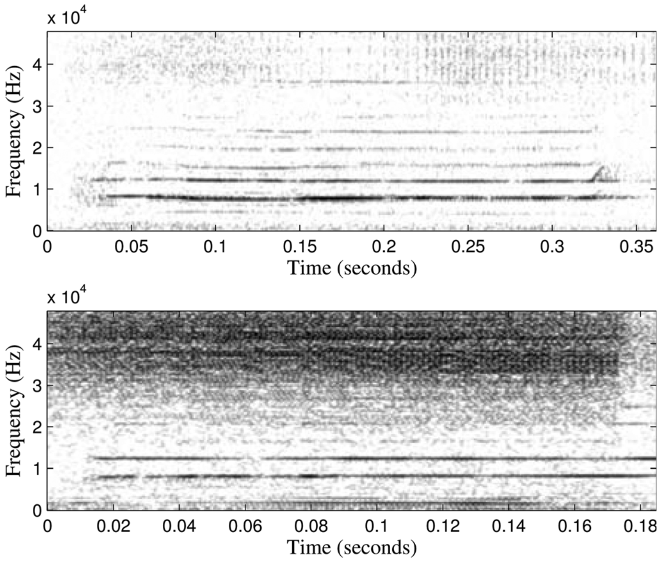

FIG. 6. Whistle creak (top) and Creak whistle (bottom).

the proposed estimator can be a valuable technique to measure slight variations in the way power changes between the tonal or pulsed components as we discussed in Sec. IV.

It is interesting to observe that, even though the Fulop and Fitz method allows high quality TFR to be obtained, it has a high computational cost. Thus, in situations where we are only interested in obtaining an estimation of the amount of power distributed between the pulsed and tonal components, the proposed DCT-based computation of the *PTR* is a better alternative. Also, the low computational cost allows this technique to be used to track the *PTR* in non-stationary signals.

Let us now observe the behavior when there are non-stationary audio signals with slow variation of the *PTR*. Observation noise has been modeled as described earlier. The pulsed bursts have been simulated with zero mean Gaussian noise $T_{\mathrm{pulsed}} = 5\,\mathrm{ms}$ and $\sigma_m^2 = 4$. The number of harmonics has been set to $N_h = 3$. The signal has been obtained by joining three fragments of 0.5 s. The first 0.5 s contain

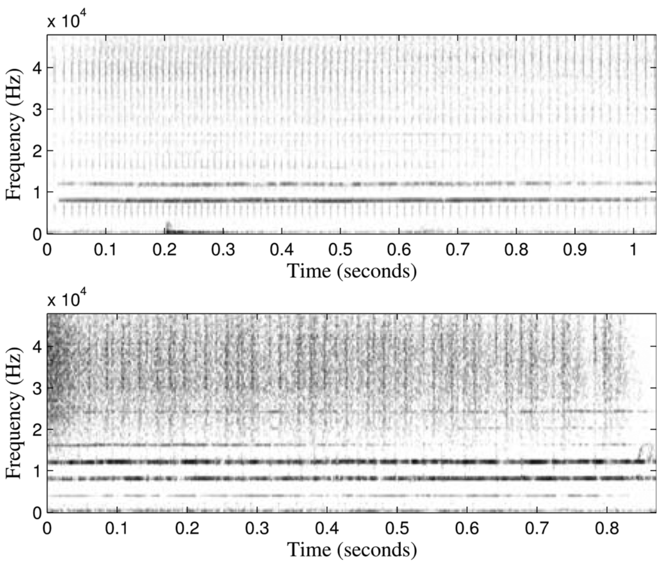

FIG. 7. TFR of Beluga sounds s_1 (top) and s_2 (bottom).

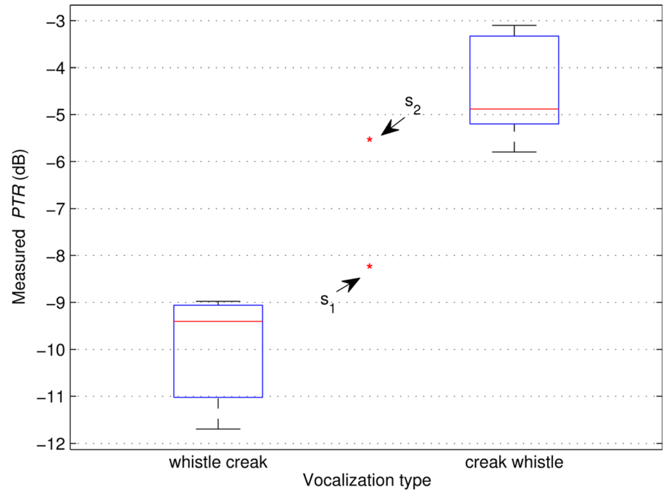

FIG. 8. (Color online) Box and Whisker plot (median ±25th and 75th percentiles) for the Whistle creak and Creak whistle vocalizations.

a pulse density of 20 bursts/s, the second 0.5 s contains a pulse density of 50 bursts/s, and, finally, the third 0.5 s contains a pulse density of 100 bursts/s. Figure 4 illustrates the spectrogram (top) of one of these synthetic sounds. Figure 4 (bottom) shows the estimated *PTR* evolution. The window length used to obtain the *PTR* evolution was $N = 5000$ samples ($\approx 52\,\mathrm{ms}$ at $f_s = 96\,\mathrm{kHz}$). This graph shows that as the density of bursts increases, the *PTR* reflects this behavior.

Similarly, we have created a synthetic sound to illustrate the behavior when the number of tonal components increases. We have used a pulse density of 70 bursts/s in the whole audio register, and we have increased the number of 10 kHz harmonics by one every 0.5 s. The top graph of Fig. 5 shows the spectrogram and the bottom graph shows the *PTR* evolution. Figure 5 shows how the *PTR* decreases as the number of tonal components increases.

IV. APPLICATION TO THE CLASSIFICATION AND CHARACTERIZATION OF BELUGA SOUNDS

In this section, we illustrate how the proposed parameter can be a valuable tool in a real application for comparing sounds that contain energy that is concentrated in the tonal or pulsed components.

As we have already described, the way that the energy is distributed in time and frequency is used to classify the sound emitted by the Beluga whales. The classification categories are related to observed Beluga behavior and include pulsed and tonal sounds as well as a combination of these two main categories. In some situations, simple observation of the TFR or listening to the sounds is not enough to determine if a given Beluga vocalization has a predominant "pulsed" component or a predominant "tonal" component. In these situations, the calculation of the proposed *PTR* can help to decide.

TABLE II. Measured *PTR* of the different sounds.

	Whistle creak	Sound s_1	Sound s_2	Creak whistle
Measured PTR (mean)	−9.6 dB	−8.3 dB	−5.6 dB	−4.5 dB

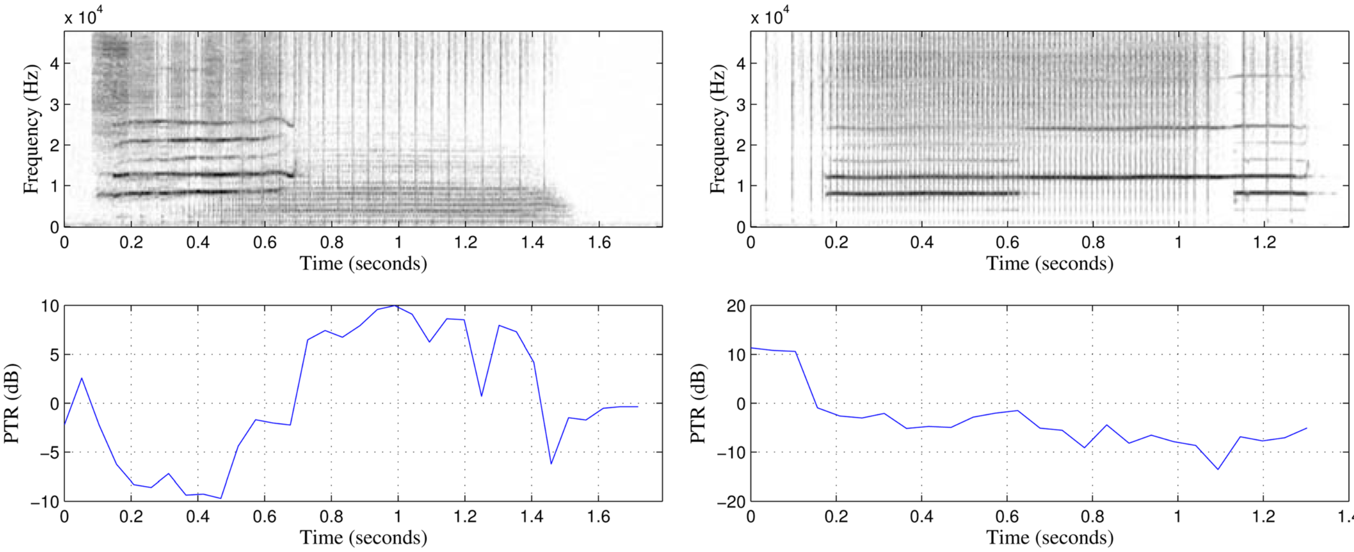

FIG. 9. (Color online) Sound recording composed of a Creak whistle followed by pulsed bursts (top) and sound recording composed of a mix of flat whistles, with different number of harmonics, and pulsed bursts (bottom).

Figure 6 shows the TFR of two Beluga sounds that were recorded at the Oceanographic of Valencia research facilities. Both sounds contain tonal components (whistles) and pulsed components (creaks). The sound at the top of Fig. 6 has a dominant whistle component and is therefore called Whistle creak. The sound at the bottom of Fig. 6 has a dominant creak component and therefore is called Creak whistle. These underwater sounds, as well as the rest of the sounds presented in this section, were recorded using a computer with a Roland (Edirol) FA-101 sound acquisition system (24 bits and frequency samples $f_s = 96$ kHz), a Bruel & Kjaer 8103 hydrophone, and a Bruel & Kjaer 2692 Nexus amplifier.

The sounds presented in Fig. 6 are clear examples of Whistle creak and Creak whistle. However, there are situations where this distinction may not be so obvious (see Beluga sounds s_1 and s_2 in Fig. 7). In these situations, the *PTR* parameter can be used to quantitatively decide between these two categories. We calculated the box and whisker plot of several Whistle creak and Creak whistle sounds as well as the *PTR* of sounds s_1 and s_2. The results are summarized in Table II and Fig. 8. According to this results, sound s_1 should be assigned to Whistle creak, whereas sound s_2 can be considered a Creak whistle.

The proposed *PTR* was also calculated for underwater sounds with variations in how the energy was distributed between the tonal and pulsed components. Two recordings of Beluga whale sounds of approximately 1.5 s were selected. The first one [Fig. 9 (left)] is composed of a Creak whistle followed by pulsed bursts. The second one [Fig. 9 (right)] is composed of a mix of flat whistles, with different number of harmonics, and pulsed bursts. Figure 9 shows the TFR and the *PTR* for both recordings.

It is easy to notice from the observation of the *PTR* of the sound record containing the creak whistle [Fig. 9 (left)] that the proposed parameter clearly demarcates the tonal and the pulsed fragments of the sound if a *PTR* threshold of 0 dB is established. In the case of Fig. 9 (right), the *PTR* helps to observe that the flat whistle has more strength on the tonal component in the interval $t = [0.6\text{–}1.1]$ s, even though some harmonic components disappear in this time interval.

V. CONCLUSIONS

We have presented a new parameter to measure the pulsed to tonal strength *PTR*. Conceptually, the textural information has been extracted by computing the bi-dimensional cosine transform of the TFR. A mathematical demonstration shows that the proposed parameter can be easily obtained without computing the TFR and using only one-dimensional discrete cosine transforms.

Simulations have demonstrated that the proposed *PTR* gives information about the mean power in the pulsed component in relation to the mean power in the tonal component of a given signal. A comparison with previously published techniques, such as the reassigned spectrogram, have shown that the proposed method gives similar results with less computational complexity. A real example has shown the utility of the parameter in helping to classify Beluga sounds with mixed pulsed and tonal elements in the same vocalization, using objective criteria. The *PTR* parameter can also be used in other bioacoustic sounds or in other completely different areas such as the monitoring of cracks in engines or turbines using the nondestructive testing technique of acoustic emission.

ACKNOWLEDGMENTS

This work was supported by the national R + D program under Grant No. TEC2011-23403 (Spain), the *Generalitat Valenciana* PROMETEO 2010/040, and the *Cátedra Telefónica* in the *Unviersitat Politècnica de València*.

[1]E. Mercado, J. Schneider, A. Pack, and L. Herman, "Sound production by singing humpback whales," J. Acoust. Soc. Am. **127**, 2678–2691 (2010).
[2]J. N. Oswald, S. Rankin, J. Barlow, and M. Lammers, "A tool for real-time acoustic species identification of delphinid whistles," J. Acoust. Soc. Am. **122**, 587–595 (2007).
[3]C. Ioana, C. Gervaise, Y. Stephan, and J. I. Mars, "Analysis of underwater mammal vocalisations using time-frequency-phase tracker." Appl. Acoust. **71**, 1070–1080 (2010).

[4] S. Fulop and K. Fitz, "Separation of components from impulses in reassigned spectrograms," J. Acoust. Soc. Am. **121**, 1510–1518 (2007).
[5] Y. Huang and R. Chang, "Texture features for DCT-coded image retrieval and classification," in *Proceedings of the 1999 IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP'99* (IEEE, New York, 1999), Vol. 6, pp. 3013–3016.
[6] D. Sim, H. Kim, and R. Park, "Fast texture description and retrieval of DCT-based compressed images," Electron. Lett. **37**, 18–19 (2001).
[7] R. E. Frye and R. S. Ledley, "Texture discrimination using discrete cosine transformation shift-insensitive (DCTSIS) descriptors," Pattern Recogn. **33**, 1585–1598 (2000).
[8] J. G. Proakis and D. K. Manolakis, *Digital Signal Processing*, 4th ed. (Prentice-Hall, Upper Saddle River, NJ, 2006), pp. 910–911.
[9] E. Chassande-Mottin and A. Pai, "Discrete time and frequency Winger-Ville distribution: Moyal's formula and aliasing," IEEE Signal Process. Lett. **12**, 508–511 (2005).
[10] M. Richman, T. Parks, and R. Shenoy, "Discrete-time, discrete-frequency time-frequency representations," in *1995 IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP-95* (IEEE, New York, 1995), Vol. 2, pp. 1029–1032.
[11] A. Papandreou-Suppappola, *Applications in Time-Frequency Signal Processing*, 1st ed. (CRC Press, Boca Raton, FL, 2002), pp. 273–306.
[12] F. Hlawatsch and T. Twaroch, "Extending the characteristic function method for joint a-b and time-frequency analysis," in *1997 IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP-97* (IEEE, New York, 1997), Vol. 3, pp. 2049–2052.
[13] MATLAB code is available from the first author's web page.