Colour Enhanced Time/Pressure Envelope (CETPE), a novel on-screen rendering of digital sound

Cesare Brizio

World Biodiversity Association,
Museo Civico di Storia Naturale di Verona,
Lungadige Porta Vittoria, 9, 37129 Verona
briziocesare@gmail.com

Ricevuto: 27/8/2023
Accettato: 2/11/2023

DOI: 10.3280/ria2-2023oa16390
ISSNe: 2385-2615

The most widespread forms of digital sound rendering in audio-management software applications include Time/Pressure Envelopes (TPÈs aka "oscillograms"), Frequency/Pressure Analyses (FPA's), and Time/Frequency Spectrographic Images (TFSI, often referred to as "spectrograms"). Only the latter, thanks to the colour mapping of pressure values, are capable to deliver simultaneously visual information about the three numerical domains (time, horizontal axis; frequency, vertical axis; pressure, colour) digitally defining the acoustic phenomenon.
Here the Colour Enhanced Time/Pressure Envelope (CETPE) is proposed, a 24-bit RGB colour mapped form of the commonplace TPE available on-screen in any digital audio software: this novel Fast Fourier Transform (FFT)-based rendering, controlled by user-defined parameters, is capable to deliver partial but potentially relevant information about the presence and the intensity of interesting frequency bands in an audio file. A proof-of-concept demonstrator of the 2-color CETPE was developed in Python language, and a sample output screen, consistent with the expected look of the rendering, is included in this paper.
Keywords: FFT, oscillogram, spectrogram, screen, colour, digital

Colour Enhanced Time/Pressure Envelope (CETPE), una nuova rappresentazione del suono digitale su schermo
Le più diffuse forme di visualizzazione del suono digitale comprendono gli Inviluppi Tempo/Pressione (TPE, anche detti "oscillogrammi"), le Analisi Frequenza/Pressione (FPA), e le Immagini Spettrografiche Tempo/Frequenza (TFSI, spesso citate come "spettrogrammi"). Solo queste ultime, grazie alla mappatura a colori dei valori di pressione, sono in grado di fornire simultaneamente informazioni visuali sui tre domini numerici (tempo, asse orizzontale; frequenza, asse verticale; pressione: colore) che definiscono digitalmente il fenomeno acustico. Qui viene proposto il Colour Enhanced Time/Pressure Envelope (CETPE), una forma mappata su colore RGB a 24-bit RGB del comune TPE disponibile a schermo in qualsiasi software per l'audio digitale: questa nuova rappresentazione basata sulle Fast Fourier Transform (FFT), controllata da parametri definiti dall'utente, è capace di fornire informazioni parziali ma potenzialmente rilevanti sulla presenza e sull'intensità di interessanti bande di frequenza in un file audio. Un dimostratore proof-of-concept è stato sviluppato in linguaggio Python, e un esempio di schermo di output, coerente con l'aspetto atteso della rappresentazione, è incluso in questo articolo.
Parole chiave: FFT, oscillogramma, spettrogramma, schermo, colore, digitale

1 | Introduction

The process of analog/digital conversion inherent in digital audio recording has a discrete nature: the continuum of an analogic signal is fractioned in as many discrete units, as the sampling rate allows following the Whittaker–Nyquist–Shannon cardinal theorem of interpolation (Nyquist [1], Shannon [2]). In the recording phase, each sample is separately assigned its specific value, directly proportional to the sound pressure collected by the microphone/recorder equipment during the sampling tempuscule and depending on the available bit depth (Boulanger & Lazzarini [3]).

Besides the instantaneous sound pressure at any time, frequency is the most relevant physical quantity for a qualitative description of acoustic phenomena. While sound pressure can be easily and immediately translated in a digital quantity by simple analog-to-digital converter (ADC) devices, computation is needed to discover which specific frequencies concur to the formation of sound: particularly in the field of bioacoustics, absolute frequency values and patterns of frequency emission are often the main, or the only, subject of study.

Dealing with frequencies, the most extensively adopted algorithm in user-application sound processing is without doubt the direct Fast Fourier Transform (FFT) – an algorithm that computes the discrete Fourier transform of a sequence of sound data, or its Inverse Fast Fourier Transform (IFFT). Fourier analysis converts a signal from its original domain (often time or space) to a representation in the frequency domain and vice versa (Heideman et al. [4], Smith [5]).

The frequency resolution of FFT-based analyses is directly proportional to FFT size (Welch et al. [6], National Instruments Corporation [7]). Several methods to generate Frequency/Pressure analyses by scanning a continuous interval of an audio file have been available for decades, e.g., the Blackman-Harris method (Blackman & Tukey [8], Harris [9], Nuttall [10]).

Whichever its purpose (research in fields including bioacoustics, recreational purposes such as personal music recording and production, or other), audio processing software must necessarily provide visual user interfaces, delivering all the information needed for interactions that include recording, playback, post-processing activities (selective deletion of soundtracks, tracks merging, digital filtering techniques…) and sound analyses.

To perform the desired operations, users need data on all the three domains of the acoustic phenomenon, namely time, pressure and frequency. The structure of any visual user interface is a compromise among physical constraints (such as the 2D nature of the flat video display), the physiology of human perception and the desire to provide the highest level of information. Among the on-screen renderings, one of such compromises takes the shape of Time/Pressure Envelopes[1] (TPÈs), also cited as oscillograms or sonograms.

TPÈs emerge as the most widespread, intuitive, and practical representation of sound. They are usually displayed on screen in the form of an outlined or single-colour filled area, with time as the horizontal axis, whose vertical span (in other word, the length of the vertical segment representing each sample) is directly proportional to the sample's overall sound pressure, according to the graduation displayed in the relative vertical scale.

Despite its practicality and intuitiveness, ordinary TPÈs provide just two information items: time (horizontal axis) and volume (vertical axis), contrary to the Time/Frequency Spectrographic Images (TFSI's) that – by mapping of relative pressure values to a colour scale – provide three-dimensional data (time, horizontal axis; frequency, vertical axis; pressure, colour).

Inspired by the TFSI, the novel rendering proposed here, Colour Enhanced Time/Pressure Envelope (CETPE), is a TPE that includes a third, frequency-related information item: colour. In that respect, it can be described as a TFSI under disguise.

In fact, it's another way of presenting the same data that appear in a TFSI, with the radical difference that:

[1] The term "envelope", used in the signal processing field with reference to the output of a Hilbert transform, is here used in the acceptation widespread in the bioacoustic community to refer to the time-history of a pressure signal.

- the rendering shape and its horizontal scale will coincide with an ordinary TPE, at the price of a reduced resolution in time (while the standard TPE resolution depends only on the sampling frequency, the CETPE time resolution will be determined by FFT parameters such as FFT size);
- colour information may be restricted to a user-defined range of frequencies;
- colour will appear only when the sound pressure in the desired frequency range exceeds a user-defined threshold.

Optionally, colour information may be restricted to a user-defined range of overall sound pressures, thus excluding the loudest or feeblest portions of the audio file.

The enhancement provided by colour is a very effective diagnostic indicator: it may allow to identify the portions of an audio file, that contain or lack specific frequencies, interesting for a special research purpose.

2 | Materials and methods

2.1 | TFSI, TPE and screen – dealing with different resolutions

It is known that – in FFT – time and frequency resolution are inversely related: if FFT resolution is increased (high FFT size, high number of discrete "frequency bins"), then time resolution will decrease. FFT size is expressed in "points" or unitary samples, while the number of bins, by default, is equal to half the FFT size + 1, e.g., 8192 points correspond to 4097 bins. For any meaningful scientific use of a time/frequency analysis, FFT sizes of 8192 or above are commonplace. The duration of the analysis window, or time resolution, is inversely proportionate to the frequency resolution. The longer the window, the fewer "images" we get of the signal evolution in time. As an example, according to the general equation:

$$t = nfft / fs$$

where t is the time resolution, nfft is the FFT size and fs is the sampling rate, with a sampling rate of 48 kHz and a FFT size of 1024 points, the spectrum is equally split into images representing a 21 ms duration. If we choose a 4096 points FFT, the spectrum is equally split into images representing a longer 85 ms duration that provides a less precise temporal resolution.

On such premises, it's very obvious that the intrinsic time resolution of an FFT-based rendering, including the TFSI on which the CETPE is based, will be up to hundreds of times lower than the native resolution of pressure data as acquired at the native sampling rate of the recording equipment. Operationally, with audio data translated into data arrays, this will result in radically different dimensions of, on one side, the one- (for monophonic recordings) or bi-dimensional (for stereophonic recordings) amplitude array, that will contain as many columns as recorded samples and, on the other side, the one- or bi-dimensional (idem) array containing the results of the FFT, with the latter containing as many rows as the frequency bins and as may columns as dictated by the FFT size.

At the array level, the search of an interesting frequency above a given threshold pressure, that may involve one (monophonic) or two (stereophonic) array dimensions, is performed by sweeping the full breadth of a row in the spectrogram array, corresponding to the frequency bin that includes the interesting frequency, and locating the spectrogram array cells whose pressure value exceeds the threshold. Each cell in the spectrogram array subsumes the data from hundreds of adjacent samples in the one-dimensional amplitude array, that correspond to as many timestamps. The list of the spectrogram array cells can be stored in a separate results array for successive use.

To generate a coloured on-screen CETPE, two alternative strategies can be applied:

- colour highlights can be overlaid on a previously generated ordinary TPE as shown by the proof-of-concept Python application (see the Supporting Material section) – a result that can be obtained by calls to existing programming language libraries;
- a full-resolution TPE may be rendered de novo, checking sample by sample whether the pressure data fall into a TFSI column appearing in the results array: all the adjacent pressure data that belong to such a column will be rendered in colour – a result that probably requires programming from scratch.

The CETPE will be rendered in a screen window of a given pixel width. Talking about screen resolution, unless the number of spectrogram array columns in the un-zoomed (full breadth) CETPE exceeds the screen window pixel width, any spectrogram column (and thus any colour enhancement) shall not exceed one pixel of width.

The lower resolution of the multi-sample CETPE enhancements as opposed to the unitary samples will become apparent only as soon as the zooming level expands a section of the CETPE to the point that the adjacent samples coincident with a spectrogram row engage a width of more than one pixel: even in that condition, the CETPE will continue to be more useful and informative than its predecessor TFSI, allowing a quicker focus on the interesting song portions than that provided by its corresponding spectral rendering, as illustrated in the following Application Scenario.

2.2 | Application Scenario and Proof-of-Concept Python Application

A realistic application scenario may be exemplified as follows: two species, “red” and “blue”, that respectively sing with the frequency profile shown in Figure 1, sing at the same time, with subequal intensity, and with similar echemes, undistinguishable in a standard TPE. Figure 1 is based on two recordings taken on an ASUS 1225B Netbook PC, at the sampling frequency of 250 kHz, following the protocols by Brizio, Buzzetti & Pavan [11] in their 2020 investigation of the inaudible components of Orthopteran songs. As usual in bioacoustic studies, the horizontal axis is linear (non-logarithmic) to provide a higher detail on the different frequency peaks, as required for an accurate description of the song. Here, such an illustration will allow to grasp subtle differences between the two songs, on which the example illustrated in the following lines will be based.

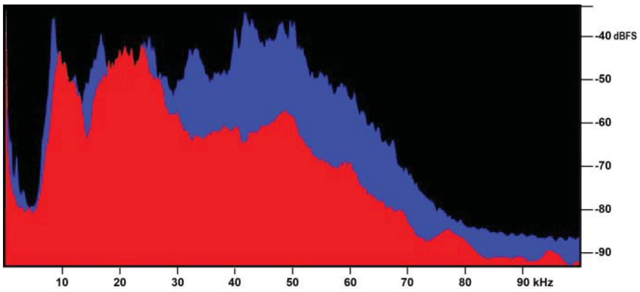

Fig. 1 – The compared FPA’s of two hypothetical species, “red” and “blue” (see text). 0 kHz-100 kHz frequency window from a 250 kHz recording (see text). Horizontal axis is linear (non-logarithmic) to allow an easier understanding of the example provided in the text

Comparazione della FPA (analisi frequenza/pressione) di due specie ipotetiche, “rossa” and “blu” (vedi testo).Finestra di frequenza 0 kHz – 100 kHz da una registrazione a 250 kHz (vedi testo). L’asse orizzontale è lineare (non logaritmico) per consentire una più agevole comprensione dell’esempio fornito nel testo

A researcher needs to measure the following parameters for the red species:

- echeme/syllable duration;
- number of syllables per echeme;
- interval between syllables;
- interval between echemes.

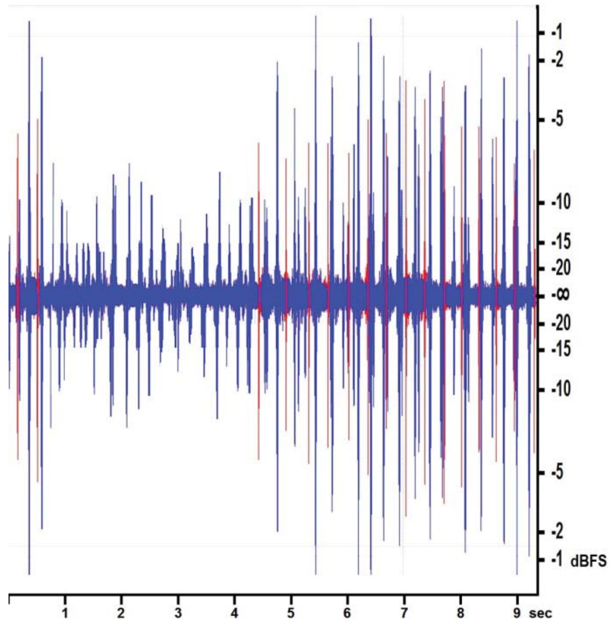

Fig. 2 – A simulated 2-colour CEPTE of a monophonic recording – the red lines correspond to samples whose overall volume falls in the desired Pressure Range (PR) (the lower volume portion is excluded from the analysis), and that contain components in the peculiar Frequency Range (FR) of the red species

Simulazione del CEPTE a 2 colori di una registrazione monofonica – le linee rosse corrispondono a campioni il cui volume cade nell’ambito di pressione (PR) desiderato (la porzione a volume più basso è esclusa dall’analisi), e che contengono componenti nel peculiare ambito di frequenza (FR) della specie rossa

Such values are read manually, on-screen by suitably positioning a visual cursor or the mouse pointer. Unfortunately, for algorithmic reasons the on-screen TFSI, with its blurred peaks, is structurally inadequate for high-precision, time-related measures. But it may be impossible to recognize the red species echemes on a monochrome TPE: the researcher needs simultaneously the optimal time resolution of a TPE, and some discrimination aid that marks only the echemes by the red species, e.g., on the basis that it lacks components above −60 dBFS in the frequency range 35 kHz-60 kHz, or that it includes components above −50 dBFS in the range 20 kHz-25 kHz. Here, a 2-colour CETPE (simulated in Figure 2) can provide a suitable solution.

A simpler scenario would occur during the review of long unsupervised recordings: if the frequency parameters of the target species are known, detecting its presence on a CETPE would be much easier than on a standard TPE and, most importantly, would not require listening to a playback nor chasing visual signatures on a TFSI.

2.3 | User Parameters

To optimize informativeness, CETPE generation requires up to four user-input parameters that define a frequency window and a band pressure threshold. Parameters may be set by keyboard input (as in the proof-of-concept application described here) or by visual interactive controls such as sliders.

It may be convenient to restrict the CETPE generation to a range of overall volume: although not strictly necessary, one or two parameters may allow to discriminate samples, based on their overall Pressure Range (PR). Thanks to those optional parameters, it will be possible to restrict the colour rendering only to the loudest or to the feeblest samples, or to a desired mid-range of acoustic pressures.

If the interesting frequency should be searched regardless to the overall volume of the audio file at any given point, these two parameters:

- Overall Pressure Range Bottom (PRB) (dBFS);
- Overall Pressure Range Top (PRT) (dBFS).

Are unnecessary, and can be disregarded or respectively set at a very low negative value and at 0, respectively.

Back to the strictly necessary user input, two parameters will allow to define the frequency range FR relevant to the user purposes:

- Frequency Range Bottom (FRB) (Hz);
- Frequency Range Top (FRT) (Hz).

When the frequency range collapses to a single frequency, FRB and FRT coincide, and the possibility of mapping different frequencies to different colours ceases to exist. A single contrasting colour will mark the relevant CETPE portions.

The last indispensable parameter will allow to discriminate samples in the desired frequency range, based on the specific sound pressure of that range:

- Frequency Range Pressure Threshold (FRPT) (dBFS or sample values).

Two different kinds of CEPTE are proposed here, respectively a computationally simpler version named 2-colour CEPTE and a more information-rich version named Multicolour CEPTE.

In the 2-colour CEPTE, whenever a sample in the PR includes components in the FR above the FRPT, it is rendered on-screen in contrasting colour (in Fig. 2 and Fig. 3, red), otherwise is rendered in a standard colour (in Fig. 2, blue; in Fig. 3, black). This approach, as implemented in the proof-of-concept application, requires just two user-input parameters, a single interesting frequency and a FRPT. Fig. 3 is a cut-out of the output window of the proof-of-concept application, generated with the following input parameters: interesting frequency 42 kHz, FRPT sample value 10000, corresponding to a value of 9.2103 in the spectrogram log scale used for the TFSI. Image clarity was improved by excluding the outer frame and buttons not strictly required for the understanding of the example.

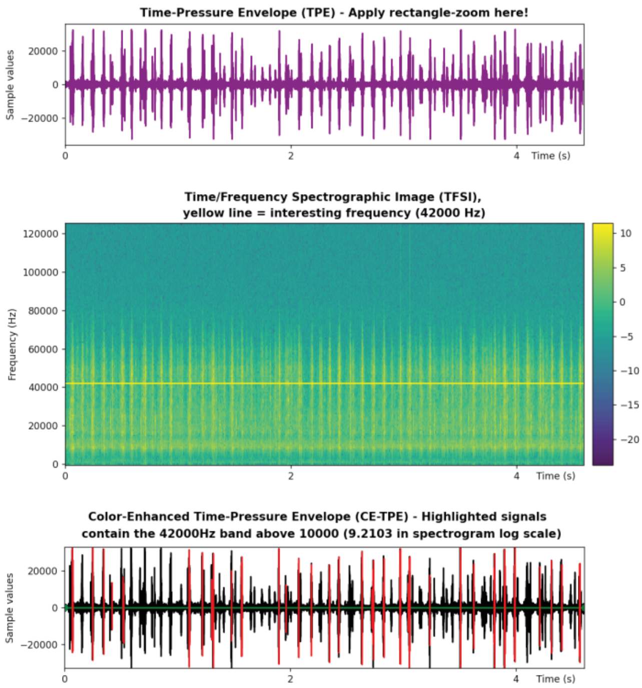

Fig. 3 – The output of the proof-of-concept Python application, after the analysis of the same recording appearing in Fig. 2. The outer frame and some control elements available in the output window were cut out to improve image readability

L'output dell'applicazione Python proof-of-concept, dopo l'analisi della stessa registrazione che appare in Fig. 2. La cornice esterna e alcuni elementi di controllo della finestra di output sono stati esclusi per migliorare la leggibilità dell'immagine

To generate a Multicolour CETPE (for a simulation, see Fig. 4), the user-defined Frequency Window will be split in a fixed number of uniform, discrete Frequency Window Bands FWB's) that will be mapped to as many RGB bits. FWB number will determine the one-bit-per-band colour mapping method adopted. Alternatives, detailed in the Appendix, include:

- 64 colours (six-FWB's);
- 4096 colours (twelve-FWB's);
- 16M colours "full range" (24-FWB's).

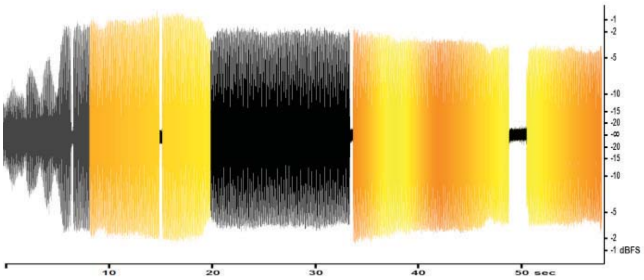

Fig. 4 – A simulated full range (24-FWB's) Multicolour CEPTE – the coloured portions correspond to samples whose overall volume falls in the PR above -5 dBFS (the lower volume portions are excluded from analysis and consequently are not coloured), and that contain components in the desired FR. Colour provides an approximate qualitative assessment of the frequency profile of each sample considered. Only frequency bands above a user-defined FRPT concur to colour definition.

Simulazione di un CETPE multicolore (24 bande di frequenza) – le porzioni colorate corrispondono a campioni il cui volume generale cade nel PR (ambito di pressione) superiore a -5 dBFS (le porzioni a volume inferiore sono escluse dall'analisi e di conseguenza non sono colorate), e che contengono componenti nel FR (ambito di frequenza) desiderato. Il colore fornisce un'approssimativa valutazione qualitativa del profilo di frequenza di ogni campione considerato. Solo le bande di frequenza sopra a una FRPT (soglia di pressione per l'ambito di frequenza) definita dall'utente concorrono alla definizione del colore

2.4 | Generation

Algorithmically, CETPE is an alternative way to display a time/frequency spectrogram: instead of the usual spectrographic view, delivered in a rectangular window on the display, with frequency as the vertical axis, a coloured time/pressure envelope (with pressure as the vertical axis) will be generated.

As usual for time/pressure envelopes, the sample-by-sample vertical envelope size will be proportional to overall sample pressure. The only variation occurs in the visual rendering of the vertical segment corresponding to the sound pressure (dBFS) of the current sample, or packet of samples depending on the current on-screen zoom level.

If an optional overall sample pressure range PR is defined by the PRB and PRT parameters, both mute samples (whose sound pressure is zero) and samples whose overall pressure is outside the PR shall be rendered in standard colour (such as black on white background or vice versa), and just the sample whose pressure value is inside the PR will be a candidate for colour assignment.

The steps of the colour rendering cycle are computationally trivial: for that reason, a flow chart illustration is deemed irrelevant. Descriptively:

- for each candidate sample, the TFSI arrays are scanned, to ascertain whether sound pressure in the desired frequency bins is equal to or higher than the pressure threshold FRPT. If not, the candidate sample shall be displayed in standard colour, otherwise the colour computation process is activated, in two possible fashions:
  - for 2-colour CEPTE, the above-FRPT sample is rendered in contrasting colour, and cycle passes to the next sample;
  - for Multicolour CEPTE, the discrete FWB's and the corresponding TFSI array rows come into play, depending on the FFT frequency bin size and on the colour mapping: 64, 4096 or 16M colours:
    - for each FWB, relevant rows (frequency bins) in the TFSI array are checked column by column for sound pressure;
    - the RGB colour byte triplet is composed from an empty three-byte binary mask (&H000000), by raising (setting to 1) each colour bit corresponding to any band whose pressure is above the user-defined FRPT: by cascading bitwise OR of the initial binary mask and each raised bit, the appropriate RGB colour is calculated;
    - the sample is rendered on screen in its appropriate size and colour[2];
    - cycle passes to the next column.

Details about the colour mapping strategy are illustrated in the Appendix.

3 | Discussion

While any exhaustive investigation of a digital audio file will still require a standard spectrographic image or frequency analysis, the partial information about frequencies provided by the CETPE may speed-up many routine activities, including the tedious process of identifying the relevant portions of long-duration, unsupervised recordings as those commonplaces in bioacoustical monitoring.

In that respect, the novel CETPE decreases the effort inherent in the on-screen visual analysis of TFSI's: while the latter require a time-consuming process of visual pattern recognition and expert interpretation to identify the relevant portions of an audio file, CETPE allows to discriminate the same portions relying on colour only: without the need to recognize visually a special spectral pattern, even an unskilled operator can identify colour.

Once criteria for relevancy are defined and user parameters are set consequently, thus suitably defining the FR and FRPT and, optionally, the overall PR, a CETPE will be generated, where only the interesting portions are coloured.

The definition of a frequency window may serve different qualitative purposes, e.g., locating the portions where:

- the target-signal occurs;
- the target-signal is not apparent;
- noise occurs.

[2] To generate, or to overwrite, an envelope line of adequate vertical length, one needs to check the maximum and the minimum value of the unitary samples, corresponding to (subsumed by) the current FFT array column. Knowing the sampling rate and the FFT size, also this passage is computationally trivial.

At the same time, the FRPT parameter, by excluding under-threshold occurrences of the desired frequency range, allows a qualitative discrimination, e.g., allows to locate the portions where:

- the target-signal is stronger;
- noise is higher (or lower).

Furthermore, the colour is not assigned randomly, but according to a rigid colour mapping in which reddish, greenish, and bluish hues are respectively related with higher, mid and lower frequency ranges. In that respect, the colour itself will be delivering useful information about frequency, the same way it delivers information about pressure in any standard TFSI.

Even though the idea was conceived in the context of bioacoustic investigations, any scientific field can benefit from CETPÈs, in so far as the on-screen (or on-print) analysis of digital audio is concerned. Examples abound in diverse fields including radio astronomy, other applications of signal processing, monitoring of industrial processes (performed in real time or on previously obtained recordings), etcetera. In as much as specific features of spoken language and music can be described in terms of special frequency patterns, CETPE may also fruitfully be applied to studies in linguistics and in the musical field.

At a higher level of abstraction, the concept of CETPE does not necessarily involve just the frequency domain: as an example, it would be relatively easy to associate colour to user-defined sound pressure ranges, regardless of frequency, or to time-based sound pressure patterns, with the same purpose of the frequency-based CETPE, that is to say, to highlight portions of a TPE that contain features uneasy or impossible to grasp in its single-color representation.

4 | Conclusions

At the price of a maximum time resolution equal to the TFSI on which it's based, necessarily lower than the maximum resolution of a standard TPE, the novel CETPE rendering contextually provides an on-screen visualization of three information items: time, sound pressure, and frequency – limited to the presence and above-threshold intensity of one or more user-defined frequency bands.

Besides its potential in speeding-up tedious activities, CETPE may allow to engage less skilled human resources in frequency-targeted spectral monitoring activities, such as the recognition of the portions that contain the song of a given species in a lengthy audio file generated by unmanned bioacoustical monitoring equipment.

As a matter of fact, contrary to the expert recognition and interpretation process of TFSI's, once an expert has set the user parameters according to the specific relevancy criteria, even an unskilled operator may identify the relevant, colour filled sections in the CETPE.

Likewise, the on-screen selection and deletion of noisy sections from a recording may be simplified without requiring a more time-consuming intervention on the TFSI screens.

Although the CETPE provides the same time resolution as the TFSI, in the visual presentation the frequency-related colour highlights can be overlaid on a previously generated ordinary TPE (as in the proof-of-concept Python application cited in this article), resulting in easier on-screen time measures of interesting frequency occurrences, than available via the TFSI alone.

As an alternative to overlaying, a full-resolution TPE may be rendered de novo, by checking sample by sample whether the TFSI arrays' pressure data satisfy the interesting frequency/pressure requirements, and by rendering in a frequency-mapped colour all the related CETPE portions.

CETPE may serve different purposes in any field where the on-screen (or on-print) analysis of digital audio is concerned: the technique can be used to highlight the TPE portions where phenomena involving any numerical domain occur.

Frequency, sound pressure, and time – alone, or in combination – can provide opportunities to generate CETPÈs, in so far as the interesting digital quantities or the interesting patterns can be algorithmically recognised by a software implementation.

Conclusioni

Al costo di una risoluzione temporale massima uguale a quella dello spettrogramma tempo-frequenza (TFSI) su cui si basa, necessariamente inferiore della risoluzione massima di un inviluppo tempo/pressione (TPE) standard, la nuova rappresentazione CETPE fornisce contestualmente una visualizzazione a schermo di tre informazioni: tempo, pressione acustica e frequenza – limitatamente alla presenza ed intensità sopra soglia di una o più bande di frequenza definite dall'utente.

Al di là del suo potenziale nel velocizzare attività ripetitive, la CEPTE può consentire di impiegare risorse umane a minor competenza in attività di monitoraggio spettrale mirate alla frequenza, come il riconoscimento di porzioni che contengono il canto di una data specie in un lungo file audio generato da un apparato non presidiato di monitoraggio bioacustico.

Di fatto, contrariamente al processo esperto di riconoscimento e interpretazione dei TFSI, una volta che un esperto abbia definito i parametri utente secondo gli specifici criteri di rilevanza, anche un operatore incompetente può identificare le sezioni rilevanti, colorate, del CETPE.

Allo stesso modo, la selezione e cancellazione a schermo delle sezioni rumorose di una registrazione possono essere semplificate senza richiedere interventi più macchinosi sugli schermi delle TFSI.

Benché il CETPE fornisca la stessa risoluzione nel tempo della TFSI, nella presentazione visuale i colori correlati alle frequenze possono essere sovraimpressi su una normale TPE generata in precedenza (come avviene nell'applicazione Python proof-of-concept citata in questo articolo), fornendo a schermo misure del tempo delle occorrenze di frequenze interessanti, più agevoli di quelle ottenibili dalle sole TFSI.

Come alternativa alla sovraimpressione, una TPE a piena risoluzione può essere generata ex novo, controllando campione per campione se i dati di pressione contenuti nelle array della TFSI soddisfano i requisiti di frequenza/pressione interessanti, e restituendo in colore mappato sulla frequenza tutte le relative porzioni di CETPE.

Il CETPE può servire differenti scopi in qualsiasi ambito in cui sia richiesta l'analisi a schermo (o su stampa) dell'audio digitale: la tecnica può essere impiegata per evidenziare le porzioni di TPE in cui avvengono fenomeni coinvolgenti qualsiasi dominio numerico.

La frequenza, la pressione sonora e il tempo – da soli o in combinazione – possono fornire opportunità per generare CETPE, nella misura in cui le quantità digitali interessanti o i pattern interessanti possono essere riconosciuti algoritmicamente da un'implementazione software.

5 | Acknowledgements

The author thanks the anonymous reviewers for their insightful comments and suggestions.

6 | Supporting Material

The source code of the proof-of-concept Python software [12] [13], released under permissive Copyright Commons CC BY-SA 4.0 license [14] is freely available for download, along with the example audio file [15] based on the mix-down of the two songs analysed in Fig. 1, the same audio file used by the proof-of-concept application to generate Fig. 3.

7 | Appendix

7.1 | Colour Mapping

While 2-colour CETPE does not require any FWB-based colour mapping strategy, Multicolour CETPE can take advantage of a frequency-related colour rendition ("informative palette").

The concept of colour mapping as proposed here is strongly dependent on the 24-bit RGB colour model. By splitting the user-defined frequency range in 6, 12 or 24 uniform discrete FWB, it will be possible to map each FWB to a colour bit, thus generating a palette of 64, 4096 and 16M colours respectively.

7.1.1 | Matching the FFT frequency bin size, and the theoretically available number of bands in the RGB model

As clarified in the introduction, once a spectrogram array is generated, frequency-related information is not continuous, but is proposed in discrete units (bins) corresponding to array rows. So, rather than a natural frequency bracket (e.g., from 10000 Hz to 10500 Hz), it will be one or more frequency bin (e.g., in case of 100 Hz bins, array rows from 100 to 105) that will be mapped to a colour bit. The identification of the frequency bins included in a frequency band is computationally very trivial. The same frequency bin (e.g., the one including the 10000 Hz frequency) can be considered part of two adjacent frequency bands (e.g., the 9000 Hz-10000 Hz band and the 10000 Hz-11000 Hz band) with no relevant effect on the generation of a Multicolour CETPE. Hereinafter, for the sake of clarity, natural, continuous frequency bands will be cited even though the actual colour mapping computation will involve discrete frequency bins.

7.1.2 | One-bit-per-band, 24-bands full range colour mapping (white noise above threshold = &HFFFFFF)

In the full 24-bit range mapping, all possible colours are used, including lowest-brightness (darker) colours engaging the low-intensity bits 1-8 in each colour byte (R, G or B). This may allow for a more precise mapping, but – due to potentially reduced contrast – may be less effective visually. An "informative palette" is adopted, where the red channel maps the higher frequency range, the green channel maps the middle frequency range, and the blue channel maps the lower frequency range.

Table 1 includes an example: the definition of FWB boundaries is arithmetically trivial, based on user parameters.

Tab. 1 – Example of one-bit-per-band, 24-band full-range colour mapping for a frequency window spanning 12 kHz from 2 kHz to 14 kHz

Esempio di mappatura del colore a un bit per banda, su 24 bande (ambito completo) per una finestra di frequenza di 12 kHz da 2 kHz a 14 kHz

FWB	*From/To Hz*	*Corresponding RGB bit*	*Colour byte*	*Frequency window range*
Band 24	14000 13501	Red bit 128 (MSB)	Red	Higher Range
Band 23	13500 13001	Red bit 64		
Band 22	13000 12501	Red bit 32		
Band 21	12500 12001	Red bit 16		
Band 20	12000 11501	Red bit 8		
Band 19	11500 11001	Red bit 4		
Band 18	11000 10501	Red bit 2		
Band 17	10500 10001	Red bit 1 (LSB)		
Band 16	10000 9501	Green bit 128 (MSB)	Green	Mid-Range
Band 15	9500 9001	Green bit 64		
Band 14	9000 8501	Green bit 32		
Band 13	8500 8001	Green bit 16		
Band 12	8000 7501	Green bit 8		
Band 11	7500 7001	Green bit 4		
Band 10	7000 6501	Green bit 2		
Band 9	6500 6001	Green bit 1 (LSB)		

Tab. 1 – follows

FWB	From/To Hz	Corresponding RGB bit	Colour byte	Frequency window range
Band 8	6000 5501	Blue bit 128 (MSB)	Blue	Lower Range
Band 7	5500 5001	Blue bit 64		
Band 6	5000 4501	Blue bit 32		
Band 5	4500 4001	Blue bit 16		
Band 4	4000 3501	Blue bit 8		
Band 3	3500 3001	Blue bit 4		
Band 2	3000 2501	Blue bit 2		
Band 1	2500 2001	Blue bit 1 (LSB)		

7.1.3 | *One-bit-per-band, 12-bands partial range colour mapping (white noise above threshold = &HF0F0F0)*

In the 12-bit partial range mapping, only the most significant four bits in each RGB channel are engaged in the mapping. While the concept of "informative palette" is still supported, this implies the exclusion of the majority of dark colours, and increases brightness and contrast when the most significant bits in each colour byte are activated. This may improve display clarity at the expense of frequency resolution. Table 2 includes an example of 12-bit mapping for a frequency window spanning 12000 Hz from 2000 Hz to 14000 Hz.

Tab. 2 – Example of one-bit-per-band, 12-bands colour mapping for a frequency window spanning 12 kHz from 2 kHz to 14 kHz
Esempio di mappatura del colore a un bit per banda, su 12 bande per una finestra di frequenza di 12 kHz da 2 kHz a 14 kHz

FWB	From/To Hz	Corresponding RGB bit	Colour byte	Frequency window range
Band 12	14000 13001	Red bit 128 (MSB)	Red	Higher Range
Band 11	13000 12001	Red bit 64		
Band 10	12000 11001	Red bit 32		
Band 9	11000 10001	Red bit 16		
UNUSED		Red bit 8 Red bit 4 Red bit 2 Red bit 1 (LSB)	UNUSED	

Tab. 2 – follows

FWB	From/To Hz	Corresponding RGB bit	Colour byte	Frequency window range
Band 8	10000 9001	Green bit 128 (MSB)	Green	Mid-Range
Band 7	9000 8001	Green bit 64		
Band 6	8000 7001	Green bit 32		
Band 5	7000 6001	Green bit 16		
UNUSED		Green bit 8 Green bit 4 Green bit 2 Green bit 1 (LSB)	UNUSED	
Band 4	6000 5001	Blue bit 128 (MSB)	Blue	Lower Range
Band 3	5000 4001	Blue bit 64		
Band 2	4000 3001	Blue bit 32		
Band 1	3000 2001	Blue bit 16		
UNUSED		Blue bit 8 Blue bit 4 Blue bit 2 Blue bit 1 (LSB)	UNUSED	

7.1.4 | *One-bit-per-band, 6-bands (2-MSB) mapping*

This six-FWB mapping will engage only the two most significant (and higher-contrast) bits in each colour byte, as illustrated in Table 3. The resulting 64-colours, 2-MSB per channel RGB palette, appearing in Fig. 5, has the desirable

Tab. 3 – Example of one-bit-per-band, 6-band colour mapping for a frequency window 12 kHz from 2 kHz to 14 kHz
Esempio di mappatura del colore a un bit per banda, su 6 bande per una finestra di frequenza di 12 kHz da 2 kHz a 14 kHz

FWB	From/To Hz	Corresponding RGB bit	Colour byte	Frequency window range
Band 6	14000 12001	Red bit 128 (MSB)	Red	Higher Range
Band 5	12000 10001	Red bit 64		
UNUSED		Red bit 32 Red bit 16 Red bit 8 Red bit 4 Red bit 2 Red bit 1 (LSB)	UNUSED	

Tab. 3 – follows

FWB	*From/To Hz*	*Corresponding RGB bit*	*Colour byte*	*Frequency window range*
Band 4	10000 8001	Green bit 128 (MSB)	Green	Mid-Range
Band 3	8000 6001	Green bit 64		
UNUSED		Green bit 32	UNUSED	
		Green bit 16		
		Green bit 8		
		Green bit 4		
		Green bit 2		
		Green bit 1 (LSB)		
Band 2	6000 4001	Blue bit 128 (MSB)	Blue	Lower Range
Band 1	4000 2001	Blue bit 64		
UNUSED		Blue bit 32	UNUSED	
		Blue bit 16		
		Blue bit 8		
		Blue bit 4		
		Blue bit 2		
		Blue bit 1 (LSB)		

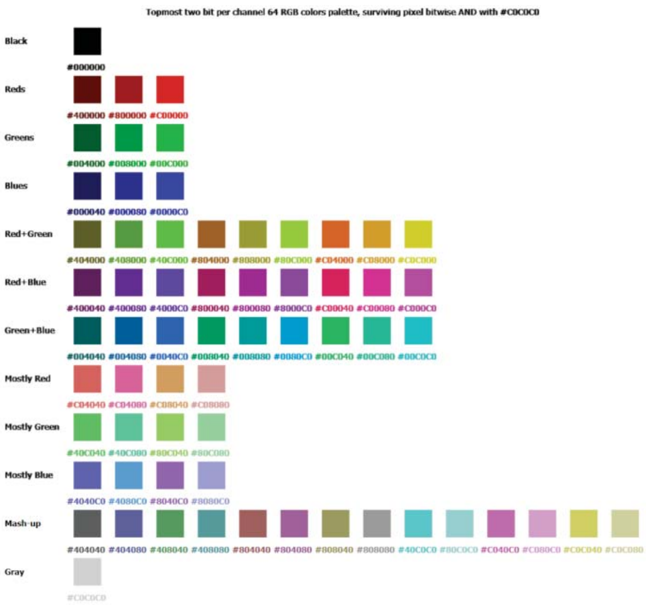

Fig. 5 – 64-colours, 2-MSB per channel, 24-bit RGB high-contrast informative palette. Under each square, its respective 24-bit hexadecimal colour code. In each channel, only the two MSB are considered, allowing only the values &H00, &H40, &H80 and &HC0

Tavolozza informativa ad alto contrasto a 64 colori RGB 24-bit, con 2 MSB (bit più significativi) per canale. Sotto ogni quadrato, il relativo codice colore esadecimale a 24 bit. In ogni canale, solo i due MSB sono considerati, permettendo solamente i valori &H00, &H40, &H80 and &HC0

property to provide a sharp increase of brightness as soon as the most significant bits in each RGB channel are activated. Considering the limited amount of display pixels engaged by each sample in the CETPE, the reduced palette coupled with the fewer bands considered may improve clarity, at the expense of a limited number of FWB's (six) enabled by this reductive mapping.

References

[1] H. Nyquist, Certain topics in telegraph transmission theory, Trans. AIEE. 47 (2) (1928) 617-644.

[2] C.E. Shannon, Communication in the presence of noise, P. IRE 37 (1) (1949) 10-21. https://doi.org/10.1109/jrproc.1949.232969

[3] R. Boulanger, V. Lazzarini (Eds.), The Audio Programming Book, The MIT Press, Cambridge, Massachusetts (2011) 209-212.

[4] M.T. Heideman, D.H. Johnson, C.S. Burrus, Gauss and the history of the fast Fourier transform, IEEE ASSP Magazine 1 (4) (1984) 14-21.

[5] J. Smith, Pulse Code Modulation (PCM) in Mathematics of the Discrete Fourier Transform (DFT) with Audio Applications, 2nd ed., online book (2007). https://ccrma.stanford.edu/~jos/mdft/mdft.html (accessed December 29, 2022).

[6] T.B. Welch, H.G. Wright Cameron, M.G. Morrow, Real-Time Digital Signal Processing from MATLAB to C with the TMS320C6x DSPs, 2nd ed., CRC Press, Boca Raton, USA, 2012.

[7] National Instruments Corporation, White Paper 4278–The Fundamentals of FFT-Based Signal Analysis and Measurement in LabVIEW and LabWindows/CVI, 2018. www.ni.com/white-paper/4278/en/ (accessed June 10, 2022).

[8] R.B. Blackman, J.W. Tukey, The measurement of power spectra from the point of view of communications engineering, Bell Syst. Tech. J. 37 (1958) 185-282.

[9] F.J. Harris, On the use of windows for harmonic analysis with the discrete Fourier transform, Proc. IEEE 66 (1978) 51-83.

[10] A.H. Nuttall, Some windows with very good sidelobe behavior, IEEE Trans. Acoust. Speech Signal Process 29 (1981) 84-91.

[11] C. Brizio, F.M. Buzzetti, G. Pavan, Beyond the audible: wide band (0-125 kHz) field investigation on Italian Orthoptera (Insecta) songs, Biodiversity Journal, 11 (2) (2020) 443-496. https://doi.org/10.31396/Biodiv.Jour.2020.11.2.443.496

[12] Proof-of-concept, Python source code. www.cesarebrizio.it/CE-TPE_Python_V4.0.py (accessed November 15, 2023).

[13] Proof-of-concept, Python source code, HTML color-highlighted version. www.cesarebrizio.it/CE-TPE%20HTML%20source.html (accessed November 15, 2023).

[14] Copyright Commons – CC BY-SA 4.0 License deeds https://creativecommons.org/licenses/by-sa/4.0/deed.en (accessed November 15, 2023).

[15] Proof-of-concept, example WAV file. www.cesarebrizio.it/CETPE_Audio_Sample.wav (accessed November 15, 2023).