THE HARTLEY PHASE CEPSTRUM AS A TOOL FOR SIGNAL ANALYSIS

I. Paraskevas, E. Chilton,
Centre for Vision Speech and Signal Processing (CVSSP)
School of Electronics and Physical Sciences,
University of Surrey
Guildford, GU2 7XH, Surrey, UK

e-mail: paraskevas@env.aegean.gr, E.Chilton@ee.surrey.ac.uk

M. Rangoussi
Department of Electronics
Technological Education Institute of Piraeus
250, Thivon str., Aigaleo-Athens,
GR-12244, GREECE
Phone/Fax: +302105381222,6
e-mail: mariar@teipir.gr

1. Introduction

This paper proposes the use of the Hartley Phase Cepstrum as a tool for signal analysis. The phase of a signal conveys critical information, which is exploited in a variety of applications. The role of phase is particularly important for the case of speech or audio signals. Accurate phase information extraction is a prerequisite for speech applications such as coding, synchronization, synthesis or recognition. However, signal phase extraction is not a straightforward procedure, mainly due to the discontinuities appearing in it ('wrapping' effect). A variety of phase 'unwrapping' algorithms have been proposed to overcome this point, when the extraction of the accurate phase values is required. In order to extract the phase content of a signal for subsequent utilization, it is necessary to choose a function that can encapsulate it. In this paper we propose the use of the Hartley Phase Cepstrum (HPC).

2. The Hatrley Phase Cepstrum

In general, computation of the cepstrum of a signal belongs to a class of methods known as homomorphic deconvolution processes, [1]. A homomorphic process describes the invertible procedure in which a signal is transformed into another domain via an orthogonal transform Ξ, a non-linear process is applied to the transformed signal in the new domain, and the result is transformed back to the original domain, via the inverse transform, Ξ^{-1}:

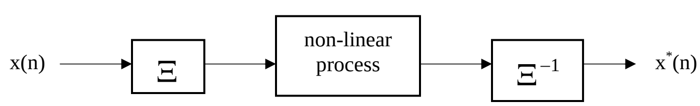

Figure 1: Summary of the homomorphic deconvolution process

In the special case where Ξ and Ξ^{-1} represent the *DTFT* (Discrete-Time Fourier Transform) and the *IDTFT* (Inverse Discrete-Time Fourier Transform), respectively, while the non-linear process is the evaluation of the Fourier phase spectrum,

$$\varphi(\omega) = \arctan\left(\frac{\Im(S(\omega))}{\Re(S(\omega))}\right) \quad (1)$$

where $\Re(S(\omega))$ and $\Im(S(\omega))$ are the real and imaginary components of the Fourier transform $S(\omega)$ of the signal $s(t)$, respectively, we obtain the so-called Fourier Phase Cepstrum, $c_F(\tau)$:

$$c_F(\tau) = IDTFT(\varphi(\omega)) \qquad (2)$$

The Fourier phase spectrum (2nd stage of figure 1) experiences two categories of discontinuities. The first category of the discontinuities ('extrinsic') is related to the use of the *arctan* function and is overcome using the 'unwrapping' algorithm, [2]. The second category of discontinuities ('intrinsic') originates from the properties of the signal itself and is overcome with their compensation, [3]. Hence, for the Fourier case, the non-linear process (2nd stage of figure 1) can be divided into the three stages:

Figure 2: Stages of the non-linear part of the homomorphic deconvolution process applied to the Fourier case

For the Hartley Phase Cepstrum case, the first, the second and the third stages of figure 1 are the *DTHT* (Discrete-Time Hartley Transform), the evaluation of the Hartley phase spectrum [4], i.e.

$$Y(\omega) = \cos(\varphi(\omega)) + \sin(\varphi(\omega)) \quad (3)$$

and the *IDTHT* (Inverse Discrete-Time Hartley Transform), respectively. Hence, the Harley Phase Cepstrum is defined as:

$$c_H(\tau) = IDTHT(Y(\omega)) \qquad (4)$$

The Hartley phase spectrum (equation (3)), unlike its Fourier counterpart (equation (1)), does not have 'wrapping' ambiguities. Hence, it experiences only the 'intrinsic' category of discontinuities, which can be detected and compensated, [4].

So, for the Hartley case, the non-linear process (2nd stage of figure 1) can be divided in the following two stages:

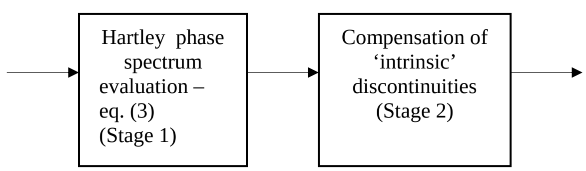

Figure 3: Stages of the non-linear part of the homomorphic deconvolution process applied to the Hartley case.

The proposed HPC is a signal feature that bears certain advantages over its Fourier counterpart, especially useful for practical applications in speech. These advantages are based on the properties of the respective spectra, which carry over to the cepstral domain thanks to the analytic relations that hold between the two domains. Localization ability and robustness to noise are two such advantages. As a simplified example of a signal, let us consider a pulse signal in the time domain. The Fourier Phase Cepstrum can identify only a single pulse, due to the ambiguities introduced by the use of the 'unwrapping' algorithm, whereas the Hartley Phase Cepstrum can indicate the location of a sequence of pulses, even for the case where noise is present [5]. Moreover, the HPC, unlike its Fourier counterpart, is more tolerant to noise – a property justified via the shape of the probability density function of the HPC, in the case where the time domain signal is pure Gaussian noise (section 3). Another property of the HPC, of interest in speech synthesis, is its invertibility: Unlike the Fourier case (figure 2), both stages of the non-linear process of the evaluation of the HPC (figure 3) are invertible, because the 'unwrapping' algorithm is not used.

3. Noise robustness of the HPC

The aim of this section is to show why the Hartley phase spectrum is more immune to noise as compared to the Fourier phase spectrum. To this end are employed the Probability Density Functions (PDFs) of the Hartley and of the Fourier phase spectra, in the special case of a pure Gaussian noise signal. The PDF of the Hartley phase spectrum is given by:

$$p_H(\beta) = \frac{1}{\pi\sqrt{2}\sqrt{1-\left(\frac{\beta}{\sqrt{2}}\right)^2}}, \; -\sqrt{2} < \beta < \sqrt{2}, \qquad (5)$$

where β denotes the Hartley phase function values, $Y(\omega)$. (See [5] for the proof of eq. (5) – proof not shown here because of lack of space).

The shape of $p_H(\beta)$ is shown in figure 4 (left). It can be observed from figure 4 (left) that the peaks of this PDF are in its upper and lower range of the horizontal axis (i.e. $\sqrt{2}$ and $-\sqrt{2}$). However, the information content of the signal, in the Hartley phase spectrum, is encapsulated in the zero crossings with respect to the frequency axis rather than in the minimum / maximum values of the cosinusoidal signal (i.e. $\pm\sqrt{2}$), [5]. Consequently, noise mainly affects the higher and the lower domain of the Hartley phase spectrum values and hence, its information content (encapsulated in the zero-crossings, middle part of its domain) is less affected.

For the Fourier phase spectrum case though, assuming again a Gaussian noise signal, and if the *arctan* function is omitted from eq. (1), then the $p_F(\beta)$ is a Cauchy distribution, assuming that the real and the imaginary parts of the Fourier spectrum are independent, [6]. In this case:

$$p_F(\beta) = \frac{1}{\pi\sqrt{1+\beta^2}}, \; -\infty < \beta < \infty \qquad (6)$$

where now β denotes the Fourier phase spectrum, $\varphi(\omega)$. The Cauchy distribution $p_F(\beta)$ (shown for $-5 < \varphi(\omega) < 5$ in figure 4 (right)), similarly to the Gaussian distribution, is symmetrical about $\beta = 0$ with its maximum value at $\beta = 0$. However, the Cauchy distribution falls more rapidly as $|\beta|$ increases and also its tails are heavier, compared to the Gaussian. Audio signals (e.g. speech, mechanical sounds etc.), convey a heavy noise additive component and hence, the PDFs of their phase spectra are similar to the Cauchy distribution for the Fourier case and to the distribution in eq. (5) for the Hartley case, respectively.

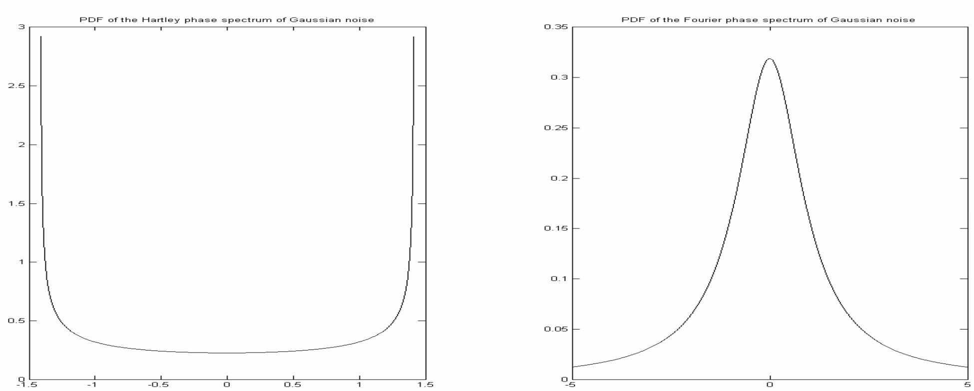

Figure 4: PDFs (a) of the Hartley phase spectrum, $Y(\omega)$, (left) and (b) of the Fourier phase spectrum, $\varphi(\omega)$, (right) for a pure Gaussian noise signal.

It should be noted here that, if the definition of the conventional Fourier phase spectrum is used (i.e. if the *arctan* function is not omitted in eq. (1)) then the PDF of the phase spectrum is no more a Cauchy, (see, eg., [6], chapter V.5); rather, it becomes uniform in $-\pi < \varphi(\omega) < \pi$.

Nevertheless, in either of the above two choices for the definition of the Fourier phase spectrum (i.e., either including the *arctan* function or not), the information content is distributed across the whole range of $\varphi(\omega)$ values and hence there does not exist a specific region of the PDF horizontal axis where the information content is mainly encapsulated. This constitutes a major difference to the case of the Hartley phase spectrum, where, as pointed out earlier, information lies mainly towards the two endpoints of the PDF range. This difference in the shapes of the respective PDFs justifies the relative noise immunity of the proposed HPC.

4. Conclusions

The phase of a signal as a function of frequency conveys meaningful information that is particularly useful for speech or audio signals. Accurate phase extraction is crucial in various speech processing applications, such as localization, synchronization, coding, etc. The major disadvantage of the computation of the phase spectrum via the Fourier transform is the heuristics employed of the compensation of the 'extrinsic' discontinuities ('wrapping' ambiguities). The effect of the 'wrapping' ambiguities is more severe in the case where noise is present. The Hartley phase spectrum, on the other hand, is advantageous as (a) it does not convey 'extrinsic' discontinuities and (b) due to its structure, it is less affected by the presence of noise, as justified through comparison of the shapes of the respective PDFs. As signal localization applications show, the phase content of a signal is encapsulated in a more efficient and easy to identify manner in the Hartley Phase rather than in the Fourier Phase Cepstral function. Hence, the Hartley Phase Cepstrum is proposed here as a promising and viable substitute to its Fourier counterpart.

References

[1] 'Theory and Applications of Digital Signal Processing', Rabiner, L.R., Gold, B., ch. 12, Prentice-Hall, Englewood Cliffs, New Jersey, 1975.

[2] 'A new phase unwrapping algorithm', Tribolet, J., Acoustics, Speech, and Signal Processing [see also IEEE Transactions on Signal Processing], IEEE Transactions on, vol. 25, no. 2, Apr 1977, pp. 170 – 177.

[3] 'Combination of Magnitude and Phase Statistical Features for Audio Classification', Paraskevas, I., Chilton, E., Acoustics Research Letters Online (ARLO), Acoustical Society of America, July 2004.

[4] 'Audio Classification Using Features Derived From The Hartley Transform', Paraskevas, I., Chilton, E., Rangoussi, M., 13th Int. Conference on Systems, Signals and Image Processing (IWSSIP'2006), Budapest, Hungary, September 2006.

[5] 'Phase as a Feature Extraction Tool for Audio Classification and Signal Localisation', Paraskevas I., Ph.D. thesis, University of Surrey, 2005.

[6] 'Introduction to the theory of statistics', Mood, A.M., Graybill F.A. and Boes, D.C., MacGraw-Hill International, 1974.

Acknowledgement

Authors gratefully acknowledge Dr. Alex Kadyrov of the CVSSP, University of Surrey, for his contribution in the derivation of eq. (5).