Astronomy and Computing 11 (2015) 55–63

Contents lists available at ScienceDirect

Astronomy and Computing

journal homepage: www.elsevier.com/locate/ascom

Full length article

Wavelet based recognition for pulsar signals

H. Shan [a,*], X. Wang [a], X. Chen [b], J. Yuan [a], J. Nie [a], H. Zhang [a], N. Liu [c], N. Wang [a]

[a] Xinjiang Astronomical Observatory of Chinese Academy of Sciences, Urumqi, Xinjiang, 830011, China

[b] School of Mechanical Engineering of Shandong University, Jinan, Shandong, 250100, China

[c] Nanchong Professional Technic College, Nanchong, Sichuan, 637000, China

ARTICLE INFO

Article history:
Received 3 July 2014
Accepted 8 March 2015
Available online 27 March 2015

Keywords:
Pulsar signals
Wavelets
Fuzzy C-means
Feature classification
Signal recognition

ABSTRACT

A signal from a pulsar can be decomposed into a set of features. This set is a unique signature for a given pulsar. It can be used to decide whether a pulsar is newly discovered or not. Features can be constructed from coefficients of a wavelet decomposition. Two types of wavelet based pulsar features are proposed. The energy based features reflect the multiscale distribution of the energy of coefficients. The singularity based features first classify the signals into a class with one peak and a class with two peaks by exploring the number of the straight wavelet modulus maxima lines perpendicular to the abscissa, and then implement further classification according to the features of skewness and kurtosis. Experimental results show that the wavelet based features can gain comparatively better performance over the shape parameter based features not only in the clustering and classification, but also in the error rates of the recognition tasks.

1. Introduction

Pulsars are neutron stars which are highly magnetized and rapidly-rotating. As the pulsars rotate, they radiate predominantly linearly polarized electromagnetic radiation. The radiating beam sweeps across the line of sight of the observer periodically, resulting in a train of narrow pulses of broad-band radiation. After the dispersion effect of the interstellar medium, and the degradation of the spatial and instrument noise, the pulsar signals are collected by the radio telescopes. These signals are processed by de-dispersion and denoising algorithms. Due to the radiation mechanism of pulsars, we can also receive signals with two to five peaks, and the number of signal peaks is also an important feature. The features of a received signal are considered as the unique signature of a given pulsar, and are due to its specific physical properties. This signature can distinguish itself from any of the others. These signals are very weak and the noise can affect the uniqueness of the signatures, leading to confusions during the recognition. Thus, the task of accurate recognition to determine whether a signal is newly discovered or has already been observed is crucial but arduous. Therefore, the analysis and recognition of pulsar signals need an efficient and expert approach.

Over the last decades, many algorithms and software have been proposed to do pulsar signal processing, mainly for de-dispersion, detection, recognition and classification. The Fast Fourier Transform (FFT) (Cooley and Tukey, 1965; Bracewell, 1965) has gained the most extensive applications in pulsar signal de-dispersion and phase restoration. The theories of multirate filter, filter banks and polyphase networks (Vaidyanathan, 1990) have been successfully applied to the analysis of pulsar signals. DSPSR (Van Straten and Bailes, 2011) is a mature digital signal processing software library in the field of radio pulsar astronomy. It is mainly applied to the analysis of phase-coherent dispersion removal algorithms. The wide bandwidth digital recording (WBDR) (Jenet et al., 1997) aims to overcome the limited capability of the detection system for accepting wide-band signals with much higher time resolution. It is also a combination of wide bandwidth and sustained data rate to implement a coherent de-dispersion. (Debosscher et al., 2007) proposed an automated method for the supervised classification of variable stars in terms of physical parameters from large databases. Such a classification can compare and assign new objects to a set of pre-defined variability training classes.

The wavelet transform (WT) (Chui, 1992; Daubechies, 1992; Vetterli and Kovacevic, 1995; Strang and Nguyen, 1996; Mallat, 1997; Mertins, 1999) was born in the mid-1980s as a new tool to analyze time–frequency information for temporal signals, especially the nonstationary signals (Yuan et al., 2011; Ordaz-Moreno et al., 2008). WT has gained extensive applications in the field of engineering and technology, for instance, signal and image process-

* Corresponding author. Tel.: +86 0991 3689057; fax: +86 0991 3838628.
E-mail address: shanhao@xao.ac.cn (H. Shan).

http://dx.doi.org/10.1016/j.ascom.2015.03.003
2213-1337/

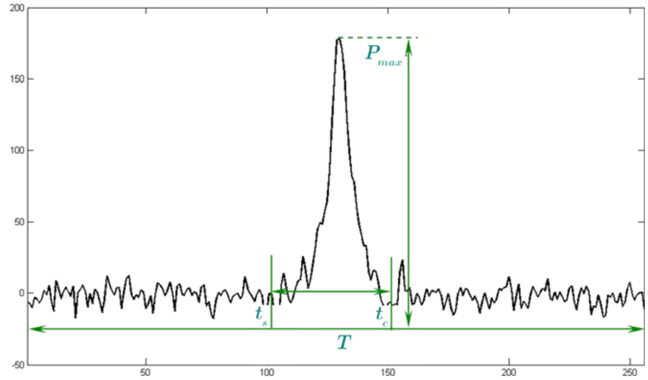

Fig. 1. The signal contour for Pulsar B0031-07, and the basic shape parameters can be derived from the contour.

ing, computer vision, seismic exploration, atmospheric and ocean wave analysis. Compared with the Fourier transform (FT), WT is a local transform in the time–frequency domain, which decides that it can efficiently extract information in a perspective of multiscale detailed analysis. Wavelet decomposition can cover the entire frequency domain, and this can provide a complete mathematical description for a signal. Different from Fourier analysis, WT uses short duration basis functions to measure the high-frequency contents and long duration basis functions for low-frequency contents (Constant-Q Quality). Properly selected wavelet filters can greatly reduce or eliminate the correlations between different signal features. WT also has fast transform algorithm, viz., the Mallat decomposition (Mallat, 1997). In the field of astronomy data analysis, the most successful applications include astronomical spectra denoising (Fligge and Solanki, 1997), weak lensing mass reconstruction (Starck et al., 2006), analysis of cosmic microwave background (CMB) radiation (Sanz et al., 1999; Starck et al., 2009), transversal coronal loop oscillations (Ireland and Moortel, 2002), 3D large scale structure maps (Lanusse et al., 2012), and cosmic web formation (Einasto et al., 2011). The outstanding advantages of wavelets for signal feature description should have broad application prospects in the field of pulsar signal processing.

Feature description stands as one of the most successful applications of wavelets in the field of signal analysis. Many methods about wavelet-based features have been proposed in the fields of machinery and equipment (Lin 2001, Seker and Ayaz 2003, Tse et al., 2004, Wu and Chen, 2006, Li et al., 2013), ultrasound analysis (Tsui and Basir, 2006), micro-electrical signals (Rosso et al., 2002), electrodynamic (Mao and Aggarwal, 2000) and electromagnetic (Mukhopadhyay and Srivastava, 2000) signal processing. Based on the concept of fourth-order normalized cumulant, (Akbarizadeh, 2012) proposed a statistic based kurtosis wavelet energy (KWE) as a feature for texture discrimination in SAR images.

Surprisingly, despite of these advantages, the applications of wavelets in the astronomical field of pulsar signal analysis and processing remain almost a vacancy. Motivated by the eximious advantages of multi-resolution analysis and localization of WT, we introduce WT into a special field-pulsar signal analysis and recognition. The contribution is that a framework of wavelet based features for the recognition and classification of pulsar signals is constructed. This is a reference method for online observation and post-processing. To our knowledge, it is a field in which few people are involved. We propose two types of wavelet features for pulsars signals. (1) Coefficient energy based features. These features are derived mainly from wavelet coefficients. The multiscale meticulous analysis of WT makes it accurate and complete for the defined features to depict the pulsar signals through a scale series. For instance, the mean of the absolute value of the coefficients in each subband can provide information about the frequency distribution of pulsar signals. The standard deviation of the coefficients provides information about the amount of change of the frequency distribution over time. These quantity can be used to derive features. (2) Singularity based features. The localizations of singularities (or signal jumps) by wavelet are not functions of detecting resolution, which change with energy (Sullivan et al., 2006). Singularities of a pulsar signal can be detected through exploring the slope of the scalogram along a modulus maxima line, i.e., the so-called Lipschitz exponent, which is a measure of the differentiability of a signal in a region. The signals are separated into a class of single peak and a class with double peaks according to the singularities based feature-number of the straight wavelet modulus maxima lines perpendicular to the abscissa. The high order cumulants are selected as features of singularities to do further classification. The fuzzy C-means (FCM) clustering is used as the classification method for the pulsar database. The Euclidean and Fisher distance criteria, and the minimum risk Bayesian decision are adopted as criteria to do discriminations for a newly collected pulsar signal.

Table 1
The basic shape parameters for pulsar signals.

Shape parameters for pulsar signals	Formulae of parameters
Height	P_{max}
Width	$Tl_p/256$
Area for the external rectangle	$P_{max}TL_p/256$
Perimeter for the external rectangle	$2(P_{max} + TL_p/256)$
Conical degree for the external rectangle	$P_{max}/(Tl_p/256)$
Concentration	$(perimeter)^2/4\pi\,(area)$
Energy	$\Sigma_{t=t_s}^{t_e} x(t)$
Skewness	$\|e'_{max-} - e'_{max+}\|$

2. The basic shape parameters for pulsar signals

A pulsar signal at a certain frequency (e.g., we choose 1.642 GHz for all signal contours in this paper) can be decomposed in a set of features. This set is a unique signature for a given pulsar. It can be used to decide whether a pulsar is newly discovered or not. Sometimes, feature extraction depends on the basic shape parameters. Intuitively, the observable shape parameters include the pulse period, the pulse width, pulse energy, the slenderness, and their spatial and temporal variations, etc. What is reflected in the signals about the properties of the interstellar medium through which the pulsar signals propagate will be ignored in this paper. Consider the case of a signal of one peak, suppose a pulsar signal (e.g., Pulsar B0031-07 in Fig. 1) in one period T with random selected phase is denoted as $x(t),\ t \in [1, n]$, and the number of the sampling points is $n = 256$ per period. Through the pulse detection (Savran et al., 2010), suppose the beginning and ending points of the pulse portion are t_s and t_e, respectively, and the signal values satisfy $x_s = x(t_s)$ and $x_e = x(t_e)$. Let $l_p = t_e - t_s$ be the sampling length of the pulse portion, then the time duration is $Tl_p/256$. Suppose the width of the external rectangle of the pulse is $Tl_p/256$, and the length of the rectangle is the peak value $P_{max} = x(t_{max})$, which appears at the sampling point t_{max}. Then the conical degree (or slenderness) is $P_{max}/(Tl_p/256)$. We assume that the pulse portion contains sharp ends, and the signal envelope is $e(t) \in [t_s, t_e]$. Suppose $e(t_{max}) = P_{max}$, the left and right derivatives of $e(t)$ are $e'_{max-} = e'_-(t_{max})$ and $e'_{max+} = e'_+(t_{max})$, respectively. The skewness of the signal can be defined as the absolute value of the difference of the left and right derivatives in the signal envelope, i.e., $k = |e'_{max-} - e'_{max+}|$, where $|\cdot|$ is the absolute value. In addition, there are also many other shape parameters, and we list them and the foresaid ones in Table 1. The pulsar contours in the database we construct in this paper are downloaded from the ATNF Pulsar Catalogue by Australia Telescope National Facility Pulsar Group 2004.