https://doi.org/10.2352/ISSN.2470-1173.2018.16.COLOR-224

Distributed Fast Radio Burst Detection: Algorithm and Application

Stephen Itschner, Kevin Bandura and Xin Li - West Virginia University, Morgantown, West Virginia, USA

Abstract

Fast Radio Bursts (FRBs) are extra-galactic transient radio signals of great interest to astronomers. Due to their non-repeating random nature and short time duration (much less than one second), automatic and reliable detection of these events has been a significant challenge, with only 25 published detections since 2007. This research provides a toolset for simulation and distributed detection of FRBs based on well-known image processing techniques. Custom software was developed to simulate FRB events with unprecedented granularity based upon the currently known population of pulses, and represents them as color-mapped intensity images. These images are operated on directly by a Generalized Hough transform approach, followed by pattern recognition and machine learning steps, which yields a binary classifier that is successful in detecting Fast Radio Burst pulse profiles. When compared to the computationally expensive traditional process known as de-dispersion, our approach enjoys the advantage of no need for iterative data transformation.

Introduction

Fast Radio Bursts have potentially enormous importance to astrophysical scientific discovery. The theoretical foundation of their usefulness was first laid in [20], 1973, decades before such signals were anything more than theoretical. It wasn't until 2007 that theory became reality with the first ever identified FRB [10]. If FRBs are extragalactic in origin, their future study and efficient detection may be key to resolving longstanding mysteries of the universe, such as:

- Efficient Particle Acceleration
- Physics Beyond the Standard Model
- Nature of Strong Field Gravity
- The Nuclear Equation of State
- Cosmological Star Formation History
- Detection and Probing of Intervening Cosmological Media
- The Possibility of ET Civilizations

[6]. Given such high potential payoff, it is unfortunate (and motivating) that only 25 such signals have been discovered to date. This is because FRBs are notoriously hard to detect using traditional radio telescopes.

Physical Characteristics of FRBs

A salient identifying feature of Fast Radio Bursts is a dispersal curve evident when the signal is viewed in the time-frequency plane. Electromagnetic waves propagating through an ionized plasma experience a frequency dependant delay across the observed frequency band analogous to light moving through a prism. The difference in pulse arrival times, Δt, across an observed frequency band bounded by f_{high} and f_{low} for a given Dispersion Measure DM is given by the cold-plasma dispersion law:

$$\Delta t = 4.148808\text{ms}\left[\left(\frac{f_{low}}{\text{GHz}}\right)^{-2} - \left(\frac{f_{high}}{\text{GHz}}\right)^{-2}\right] \times \left(\frac{\text{DM}}{cm^{-3}\,pc}\right) \quad (1)$$

[10]

This effect becomes visually apparent for cases of high SNR when the radio telescope data is rendered into a time-frequency image

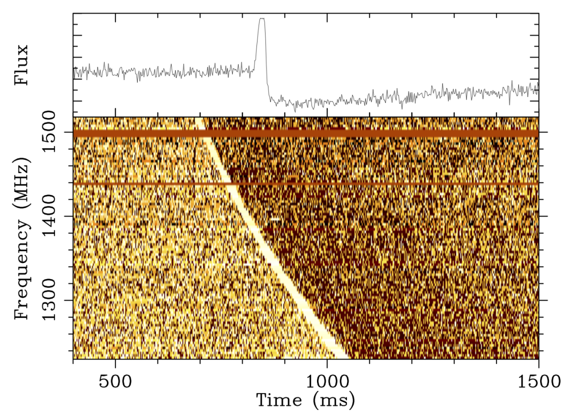

Figure 1. Plot of FRB010724, taken from [15]'s website, FRBCAT. The top 1-D line plot showing a single pulse represents the result of traditional de-dispersion, where all pulse energy has been aligned to a single time slice and then the image is integrated across the frequency axis (see "Traditional Detection Methods" section). The dark red horizontal striping near the top of the image indicates Radio Frequency Interference, and the color change after the pulse is the result of receiver saturation due to the pulse's high SNR [10].

The appearance of a bright dispersed pulse, however, does not necessarily indicate an FRB event. Many events, such as Radio Frequency Interference (RFI) can be recorded as a dispersed pulse, with a supposed class of signals known as "Perytons" being a key example - searches revealed the signals to emanate not from space but rather from the telescope compound's own microwave ovens [16]. For signals with high SNR, pulse shape can also be used to distinguish a signal which has truly traveled through an extraterrestrial cold plasma because the pulse exhibits an exponentially shaped pulse tail after dedispersion [19]. In some surveys the presence of the pulse in multiple beams of a phased array feed also indicates a terrestrial origin and those false positives can be discarded immediately [9].

There are also classes of legitimate astronomical signals which are subject to cold-plasma dispersion, such as Rotating Radio Transients (RRATs) and pulsars. Unlike (almost all) FRBs, RRATs and pulsars are periodically repeating in time [13]. Because they are repeating, pulsars can benefit from integration gain when the pulse period is known or derived by folding multiple repeating pulse profiles on top of one another. Currently known FRBs exhibit a higher DM than most pulsars and are non-repeating events (with the exception of FRB121102 [18]), and thus a high DM combined with context clues (such as telescope pointing, estimated interstellar plasma densities in that direction, etc.) allow astronomers to identify a FRB. Some of these key context clues are:

- Value of Dispersion Measure
- Pulse Shape/Profile
- Pulse Scintillation
- Pulse Scattering (width)

This underscores the need to implement a nuanced emulator which can capture and evaluate fine discerning pulse details when developing or evaluating any detection methodology, and to the author's best knowledge this paper debuts the most subtle pulse emulation tool to date.

Existing Detection Methods

Although dispersion is an important factor in identifying FRBs, it is sub-optimal with respect to pulse detection because the pulse's energy is then distributed across time throughout the image. The traditional approach to this is a process known as incoherent dedispersion, which is the intentional application of a time delay in the reverse dispersal direction. Because the pulse dispersal pattern is generally well behaved, dedispersion can align a pulse's energy into a narrow vertical time slice. Once de-dispersed, the signal is integrated across the measured frequency band and reveals a one dimensional data vector with a strong single pulse. However, the exact dispersion measure is never known a-priori, so an entire family of dispersion measures must be tried and evaluated. A family of boxcar filters for a set of candidate pulse widths (to account for pulse scattering) are applied as well as candidate DMs, so the solution space to be tried is computationally large [9],[13]. It is also possible to perform this operation coherently (i.e. apply a frequency-dependant phase shift to the Fourier Transform of the raw data used to generate the image).

This approach was initially designed for offline pulse searches, and in general incoherent is preferred over coherent due to computational concerns, despite the fact that coherent dedispersion is overall the more sensitive of the two [22]. Later, improvements were made for GPU parallelization of incoherent dedispersion during blind surveys [5],[1] and then more recently still algorithms based on data transform approaches harness the speed and scalability of the Fast Fourier Transform to facilitate real-time detection [22], [2]. It is entirely conceivable that those approaches would be low enough in computational complexity to facilitate citizen science with consumer-grade electronics, but the work presented here provides, in the worst case, an alternative tractable method and a novel set of FRB tools.

Motivation behind Our Approach

The number and localization of detected Fast Radio Bursts would improve immensely under a spatially distributed collection system. In the status-quo, a small number of specialized and expensive hardware installations observe the sky for FRB signals. While many new instruments are under construction with the goal of increasing FRB detection quantity and quality [14] [4], recent advances in RF and computing hardware make it possible for ordinary citizens to have significant scientific impact in this area at a relatively low cost. Systems made of many low-cost imagers dispersed across a wide geographical region would provide excellent source localization when an FRB event is detected, and localization is important in determining properties of the Inter-Galactic Medium (IGM) as well as to localize position and determine physical origin [12]. Additionally, distributed systems could easily distinguish RFI (Radio Frequency Interference) from true FRB detections by excluding sources that only influence a small regional portion of the sensor network. The potential benefits and drawbacks of such a citizen science project are theorized in [12]. We set out to provide a computationally efficient alternative method designed to run on an average PC, Raspberry Pi, Software Defined Radio program, or cell phone.

Proposed Emulation and Detection Methods

Matlab-based FRB Emulation

Due to the small number of FRB detections in the wild, the design and analysis of any detection scheme is best facilitated using emulated data. This is an approach undertaken by many related papers [2], [22], and while they indeed allow the generation of dispersed radio data images, the distinguishing factors of pulse profile, scintillation, and scattering are largely unaddressed. Background noise is also modeled largely as a normal random process. This paper's emulation differentiates itself in that it incorporates many of the above nuances neglected by other approaches, but as the coherent form of de-dispersion is not addressed here, [22]'s emulator is more complete in the sense that it is able to generate dispersed data from a coherent phase process. As our detection algorithm is inspired by image processing techniques, however, the emulation engine works directly in the time-frequency domain and considers the data as a colormapped intensity image rather than a processed data stream.

In order to provide the most realistic emulation, data-driven processes were extensively used in image formation. Data-driven elements are derived from data made publicly available on FRB-CAT [15], an online repository maintained by Swinburne of all published FRB detections to date. Data (where available) was downloaded in the PSRCHIVE file format (*.ar) [7]. Transforming this data into time-frequency image data did not seem trivial for new users and requires a software learning curve. To facilitate this process for new researchers, the author is proud to provide a virtual machine image which includes the properly installed PSRCHIVE software and its related libraries, presented without guarantee on GITHUB. Running PSRCHIVE tools on *.ar files produces an integrated 2D time-frequency image representation of the data, which can be piped to an ASCII text file. The ASCII text file is then moved from the virtual machine and loaded into MATLAB, where the data is saved as a .mat file for further analysis.

As mentioned previously, dispersed burst events occurring in more than one beam of a phased array feed dish are considered RFI and are thus excluded from detection. For this reason, for each FRB event logged on FRBCAT, there exist multiple .ar files, and only 1 will produce an image with a detectable pulse when rendered in PSRCHIVE. The image data information for the remaining no-pulse files were used to evaluate and reproduce background noise in the emulator. A non-parametric statistical model of each frequency channel within the no-pulse data was generated (i.e. each image row is treated as an independent distribution), and then used as the kernel for random number generation within the emulator. This has the added benefit of faithfully representing channels which have been routinely excluded due to RFI, as can be seen by the rectangular region near the top of Figure 2 (rows 0 - 200) where values appear uniform.

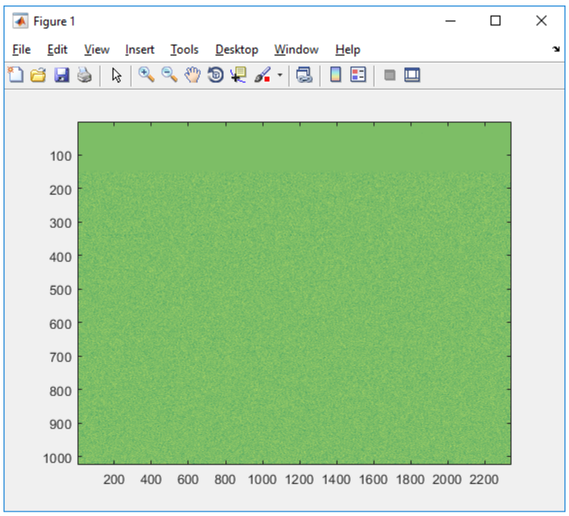

Figure 2. *Empirically Accurate Background Noise Generation, a frequency versus time artificial image.*

In addition to generating data-representative background noise, the emulator can specify scintillation, DM, signal strength, and pulse width/shape with individual frequency channel granularity. A sample of one such generated image can be seen in Figure 3, side-by-side with a real image generated in PSRCHIVE for FRB110220 (who's pulse properties it emulates). In order to generate the population of emulated FRBs with which to test and develop the classifier, FRBCAT [15] was an invaluable resource in cataloguing known distributions of these key parameters. Based on how the key parameters were reported in FRBCAT, the simulated population was distributed comparably, and this family of emulated images serves as input to the Hough Transform pulse detection step.

Hough Transform based FRB Detection

The Generalized Hough Transform is a well established method in the image processing community which allows the detection of arbitrary shapes in an image. Inspired by the concept of R-table laid out in [3], we sought to represent a family of emulated FRB events in the Hough transform space and use it as a feature to enable efficient classification decisions. An implementation of the Generalized Hough Transform operation is readily available in Matlab using command "hough", which was utilized in prototyping the operational mechanics of the method.

In general, the detection methodology proceeded based upon the ability of a binary classifier to learn boundaries in Hough space which distinguish an FRB with a believable DM (and other pulse features) from a burst detection with features that lie outside the scope of known detection parameters. It was found that the transformation operation was successful for cases where the FRB of interest exhibited an SNR large enough to be discerned visually on the color mapped intensity image. The figure below illustrates the changes a difference in DM makes when the image is converted to Hough space - the right image shows the Hough space transform of a pulse event with believable DM, while the left image shows the Hough space of a pulse event with DM beyond the realm of probability for an FRB.

An emulation set consisting of 100,000 images was generated. Of these, 5,000 images contain a "realistic" FRB pulse based on the parameters laid out in [15]. The 95,000 remaining images contain either no pulse at all (background noise only), or a pulse which is inconsistent with the set of currently known FRBs, or an artifically generated RFI signal. This uneven distribution was chosen to reflect the rarity of a true FRB event in a field of candidates, which is a condition commonly encountered by astronomers searching for FRBs. Next, a binary classifier step is used to interpret the resulting Hough transform after being trained in a supervised manner on the 100,000 emulated images. The real FRBs for which we have successfully recovered data (010724, 110220, 110626, 110703, 120127), along with several plausible and implausible emulated fakes are generated exclusively for validation, for a total validation set size of 500 images. A Convolutional Neural Net (CNN) binary classifier was trained using the MatConvNet library [21], using that package's boilerplate example network architecture, with slight modifications in order to report False Negative and False Positive rates as

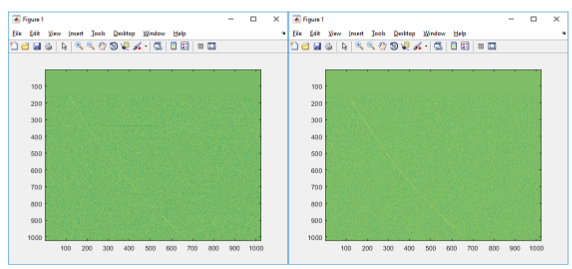

Figure 3. *Real FRB image 110220 (left) as generated in PSRCHIVE using telescope data, and emulated FRB data (right) generated by the emulator with specified burst parameters as reported in FRBCAT for FRB110220 [15]*

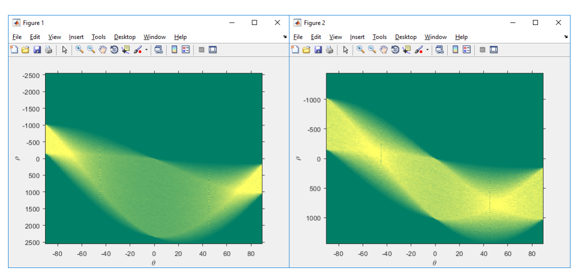

Figure 4. *Burst image transform with unrealistic DM (left) and transform with realistic FRB pulse profile parameters (right)*

the network trained. Its detection results are promising (Figure 5).

In the validation stage of training (red data points in the neural network training graphic), the binary classifier was tested on a set of images which it had never before seen. Included in this validation set were the actual FRB images generated for 010724, 110220, 110626, 110703, and 120127. Overall, the validation set's detection performance improved over time, to the point that it was exhibiting less than 1% error overall. Importantly, however, the detection performance when evaluated on just the published FRB images, it was able to successfully detect a pulse for FRBs 110220, 010724, and 110703. This is to be expected, because the basis of Hough transform begins with edge detection. For FRBs without SNR high enough to visually distinguish a pulse profile (110626, 120127), the transform is insufficient.

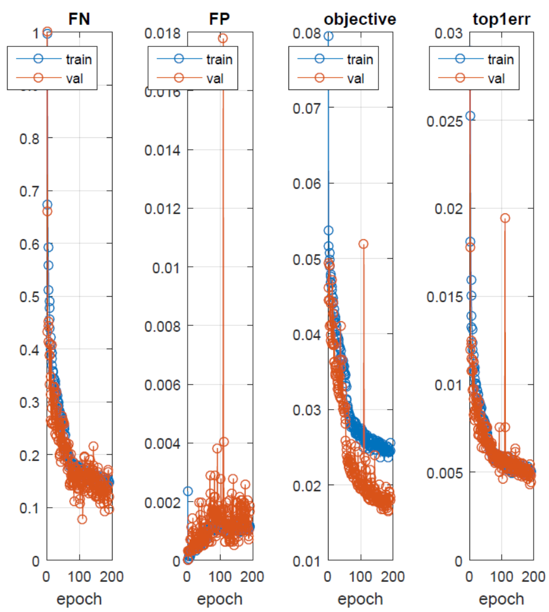

Figure 5. *Results of FRB detection training and validation run over 200 epochs. From left to right: false negative rate, false positive rate, training objective value (neural net, not FRB specific), and overall classification error rate ("top1err"). Results on training set (blue) as well as validation set (red) are shown. Overall, the plot demonstrates that the NN is capable of learning to differentiate realistic pulses from non-realistic ones by exploiting features in Hough space*

In terms of overall detection performance, the classic dedispersion approach is better because detection is possible for pulses which do not substantially exceed the noise floor. In terms of computation, however, a transform based approach such as this or [22], [2] excels because the scalability of the problem does not increase as the number of dispersion measures to be tested increases. In this case, the computational complexity of the evaluation of a neural network binary classifier is linear, and the Hough transform is O(Pn) or O(P + n), where P is the number of projections that the transform operates across (i.e. the size of Hough space generated) [8]. Traditional de-dispersion on the other hand has time complexity of roughly O(N^2)[11].

Conclusion

A classifier was trained in the subtleties of differentiating true FRBs from imposter signals, with an intervening step of signal transformation into Hough space. It is successful in detecting pulses with sufficiently high SNR to be seen by the naked eye in its dispersed form, as the nature of the transform algorithms are derived from image processing. This is a loss of detection performance with respect to conventional methods, and not all FRBs in the population are strong enough to appear within the Hough space. However, the most recent estimate puts the daily occurrence of FRBs on the order of 11,000 per day over the entire sky [17]. This means that some missed detections in favor of the strongest burst strength should be satisfactory. An additional useful product of the research is a nuanced data-driven FRB emulation tool which treats more factors than are currently considered in other developed detectors, and the performance of the classifier suggests that including these features in a machine learning process has potential to reduce false positive rates and consequently astronomer workload.

References

[1] Dalal Ait-Allal, Rodolphe Weber, Cédric Dumez-Viou, Ismael Cognard, and Gilles Theureau. Blind detection of giant pulses: GPU implementation. *Comptes Rendus Physique*, 13(1):80–87, 2012.

[2] Marwan Alkhweldi, Natalia A. Schmid, and Richard M. Prestage. A blind transform based approach for the detection of isolated astrophysical pulses. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, pages 3944–3948, 2017.

[3] D. H. Ballard. Generalizing the Hough transform to detect arbitrary shapes. *Pattern Recognition*, 13(2):111–122, 1981.

[4] Keith Bannister, Ryan Shannon, Jean-Pierre Macquart, Chris Flynn, Philip Edwards, Morgan O'Neill, Stefan Osłowski, Matthew Bailes, Barak Zackay, Nathan Clarke, Larry D'Addario, Richard Dodson, Peter Hall, Andrew Jameson, Dayton Jones, Robert Navarro, Joseph Trinh, James Allison, Craig Anderson, Martin Bell, Aaron Chippendale, Jordan Collier, George Heald, Ian Heywood, Aidan Hotan, Karen Lee-Waddell, Juan Madrid, Joshua Marvil, David McConnell, Attila Popping, Maxim Voronkov, Matthew Whiting, Graham Allen, Douglas Bock, David Brodrick, Francis Cooray, David DeBoer, Philip Diamond, Ron Ekers, Russell Gough, Grant Hampson, Lisa Harvey-Smith, Stuart Hay, Douglas Hayman, Carole Jackson, Simon Johnston, Baerbel Koribalski, Naomi McClure-Griffiths, Peter Mirtschin, Alan Ng, Ray Norris, Sarah Pearce, Chris Phillips, Daniel Roxby, Euan Robert Troup, and Tobias Westmeier. The detection of an extremely bright fast radio burst in a phased array feed survey. 2017.

[5] B. R. Barsdell, M. Bailes, D. G. Barnes, and C. J. Fluke. Accelerating incoherent dedispersion. *Monthly Notices of the Royal Astronomical Society*, 422(1):379–392, 2012.

[6] James M. Cordes, T. Joseph W. Lazio, and M. A. McLaughlin. The dynamic radio sky. *New Astronomy Reviews*, 48(11-12):1459–1472, 2004.

[7] A. W. Hotan, W. van Straten, and R. N. Manchester. PSRCHIVE and PSRFITS:An OpenApproach to Radio Pulsar Data Storage and Analysis. *Publications of the Astronomical Society of Australia*, 21:302–309, 2004.

[8] J. Illingworth and J. Kittler. A survey of the hough transform. *Computer Vision, Graphics and Image Processing*, 44(1):87–116, 1988.

[9] E. F. Keane, D. A. Ludovici, R. P. Eatough, M. Kramer, A. G. Lyne, M. A. McLaughlin, and B. W. Stappers. Further searches for rotating radio transients in the parkes multi-beam pulsar survey. *Monthly Notices of the Royal Astronomical Society*, 401(2):1057–1068, 2010.

[10] D. R. Lorimer, M. Bailes, M. A. McLaughlin, D. J. Narkevic, and F. Crawford. A Bright Millisecond Radio Burst of Extragalactic Origin. *Science*, 318(5851):777–780, 2007.

[11] A. Magro, A. Karastergiou, S. Salvini, B. Mort, F. Dulwich, and K. Zarb Adami. Real-time, fast radio transient searches with GPU de-dispersion. *Monthly Notices of the Royal Astronomical Society*, 417(4):2642–2650, 2011.

[12] Dan Maoz and Abraham Loeb. Searching for giga-Jansky fast radio bursts from the Milky Way with a global array of low-cost radio receivers. 000(February), 2017.

[13] M. A. McLaughlin, A. G. Lyne, D. R. Lorimer, M. Kramer, A. J. Faulkner, R. N. Manchester, J. M. Cordes, F. Camilo, A. Possenti, I. H. Stairs, G. Hobbs, N. D'Amico, M. Burgay, and J. T. O'Brien. Transient radio bursts from rotating neutron stars. *Nature*, 439(7078):817–820, 2006.

[14] Cherry Ng. CHIME FRB: An application of FFT beamforming for a radio telescope. *32nd URSI GASS*, (232214567787831183), 2017.

[15] E. Petroff, E. D. Barr, A. Jameson, E. F. Keane, M. Bailes, M. Kramer, V. Morello, D. Tabbara, and W. van Straten. FRBCAT: The Fast Radio Burst Catalogue. *Publications of the Astronomical Society of Australia*, page 15, 2016.

[16] E. Petroff, E. F. Keane, E. D. Barr, J. E. Reynolds, J. Sarkissian, P. G. Edwards, J. Stevens, C. Brem, A. Jameson, S. Burke-Spolaor, S. Johnston, N. D.R. Bhat, P. Chandra S. Kudale, and S. Bhandari. Identifying the source of perytons at the Parkes radio telescope. *Monthly Notices of the Royal Astronomical Society*, 451(4):3933–3940, 2015.

[17] P. Scholz, L. G. Spitler, J. W. T. Hessels, S. Chatterjee, J. M. Cordes, V. M. Kaspi, R. S. Wharton, C. G. Bassa, S. Bogdanov, F. Camilo, F. Crawford, J. Deneva, J. van Leeuwen, R. Lynch, E. C. Madsen, M. A. McLaughlin, M. Mickaliger, E. Parent, C. Patel, S. M. Ransom, A. Seymour, I. H. Stairs, B. W. Stappers, and S. P. Tendulkar. The repeating Fast Radio Burst FRB 121102: Multi-wavelength observations and additional bursts. *ArXiv e-prints*, page 18, 2016.

[18] L. G. Spitler, P. Scholz, J. W.T. Hessels, S. Bogdanov, A. Brazier, F. Camilo, S. Chatterjee, J. M. Cordes, F. Crawford, J. Deneva, R. D. Ferdman, P. C.C. Freire, V. M. Kaspi, P. Lazarus, R. Lynch, E. C. Madsen, M. A. Mclaughlin, C. Patel, S. M. Ransom, A. Seymour, I. H. Stairs, B. W. Stappers, J. Van Leeuwen, and W. W. Zhu. A repeating fast radio burst. *Nature*, 531(7593):202–205, 2016.

[19] D. Thornton, B. Stappers, M. Bailes, B. Barsdell, S. Bates, N. D. R. Bhat, M. Burgay, S. Burke-Spolaor, D. J. Champion, P. Coster, N. D'Amico, A. Jameson, S. Johnston, M. Keith, M. Kramer, L. Levin, S. Milia, C. Ng, A. Possenti, and W. van Straten. A Population of Fast Radio Bursts at Cosmological Distances. *Science*, 341(6141):53–56, 2013.

[20] V. L. Ginzburg. Possibility of Determining Intergalactic Gas Density by Radio Observations of Flares of Remote Sources, 1973.

[21] Andrea Vedaldi and Karel Lenc. MatConvNet - Convolutional Neural Networks for MATLAB. 2016.

[22] Barak Zackay and Eran O. Ofek. An accurate and efficient algorithm for detection of radio bursts with an unknown dispersion measure, for single dish telescopes and interferometers. *The Astrophysical Journal*, 835(1):1–13, 2017.

Author Biography

Stephen Itschner received the B.S. degree in Electrical Engineering and Biometric Systems in 2013 from West Virginia University. He has continued as a PhD student at WVU with a research focus in signal processing for radio astronomy and radar.

Kevin Bandura received the B.S. degree and the Ph.D. degree in physics from Carnegie Mellon University in 2003 and 2011, respectively. From 2011 to 2015, he was a Postdoctoral Research Associate at McGill University. He is currently an Assistant Professor at West Virginia University in the Department of Computer Science and Electrical Engineering, and Adjunct Assistant Professor in Department of Physics and Astronomy. His current research interests are signal processing for radio astronomy and cosmology.

Xin Li received the B.S. degree with highest honors in electronic engineering and information science from University of Science and Technology of China, Hefei, in 1996, and the Ph.D. degree in electrical engineering from Princeton University, Princeton, NJ, in 2000. He was a Member of Technical Staff with Sharp Laboratories of America, Camas, WA from Aug. 2000 to Dec. 2002. Since Jan. 2003, he has been a faculty member in Lane Department of Computer Science and Electrical Engineering. His research interests include image/video processing, biometrics and cybersecurity, social media data mining and biomedical informatics. Dr. Li received a Best Student Paper Award at the Conference of Visual Communications and Image Processing as the junior author in 2001, a Runner-up prize of Best Student Paper Award at IEEE Asilomar conf. on Signals, Systems and Computers as the senior author in 2006, a Best Paper Award at the Conference of Visual Communications and Image Processing as the single author in 2010 and three Top 10% Paper Awards in 2013-2014. Dr. Li has published dozens of journal articles on premier IEEE journals and conference papers on top CS conferences including CVPR, ICCV and NIPS. Dr. Li is currently serving as a Senior Area Editor for IEEE Signal Processing Letters, an Associate Editor for IEEE Transactions on Image Processing and an Associate Editor for IEEE Transactions on Circuits and Systems for Video Technology. Dr. Li was elected to an Fellow of IEEE in 2017 for his contributions to image coding, restoration, and interpolation.