Brainlike Time Series Filtering and Data Reduction

Introduction

In recent years, Brainlike Inc., has patented and refined smart sensing methods that efficiently reduce cluttered data to useful information, in real time. With Brainlike sensing, real time data can be more readily understood, and transmitted data can be significantly reduced. Brainlike sensing offers special advantages for processing time series data. For example, in combined voice recognition and transmission applications, a person's voice may be hard to understand because of background clutter, and transmitting radio quality voice data may require transmission rates over 20,000 bytes per second. Brainlike sensing can continuously learn how to reduce voice data to a smaller number of feature values that are uniquely salient to a given individual. Once these feature values have been computed and transmitted, they can be transformed back to time domain values that reproduce the individual's same voice, but exclude clutter that was present in the original time series data. While many electronic filters are widely used to clarify time series data, Brainlike sensing adds a patented and patent pending process that continuously learns individuals' uniquely salient metrics.

Brainlike sensing has now been refined for real time use on small cell phone or remote sensor processors. Meanwhile, signal processing and computing advances have resulted in highly efficient feature extraction methods such as fast Fourier transforms (FFTs). FFTs are now readily available for low power, compact use on the latest generation of remote sensor and cell phone processors as well. These combined advances provide **enabling technology for the wireless revolution**.

In the human voice recognition case, established methods may be used to convert real time voice data to snippets, at the phoneme or word level. For example, a partitioning process on a caller's cell phone could first parse a person's voice into snippets. Assuming for simplicity that these snippets average one second in length, snippets measured in the time domain would contain an average of 20,000 amplitude values on a one byte gray scale. Established methods may be used to convert those values in the time domain to feature values. For example, an FFT could transform the 20,000 amplitude values to 20,000 frequency power values, which in turn could be reduced to 1,000 average power feature values. The first such feature value could be the average among frequency power levels between 1 and 20 Hz; the next feature could be the average among power levels between 21 and 40 Hz; and so on.

Brainlike sensing can first use an available FFT application that will reduce data to features in this way on a cell phone, during any given call. During each snippet's time span within the call, Brainlike sensing can continuously update learned baseline *salience* values for each such feature. Each salience value will show how much its corresponding feature contributes to accurate voice reproduction, for the person making the call. Brainlike sensing can then use an available FFT inverse transform application to convert only those salient features back to sounds like the sender's voice in the time domain. If the feature transformation function and inverse transformation function reside on the same cell phone, the output sound will be filtered so that the individual's learned voice will sound more prominent and background clutter will be reduced. If the transformation function resides on a sending cell phone, and the inverse transformation function resides on a receiving cell phone, then transmitted information will be reduced as well. In that case, as shown in Figure 1 below, only feature values, along with occasionally updated configuration values, will require transmission. The remainder of this report includes details on this and other Brainlike sensing alternatives.

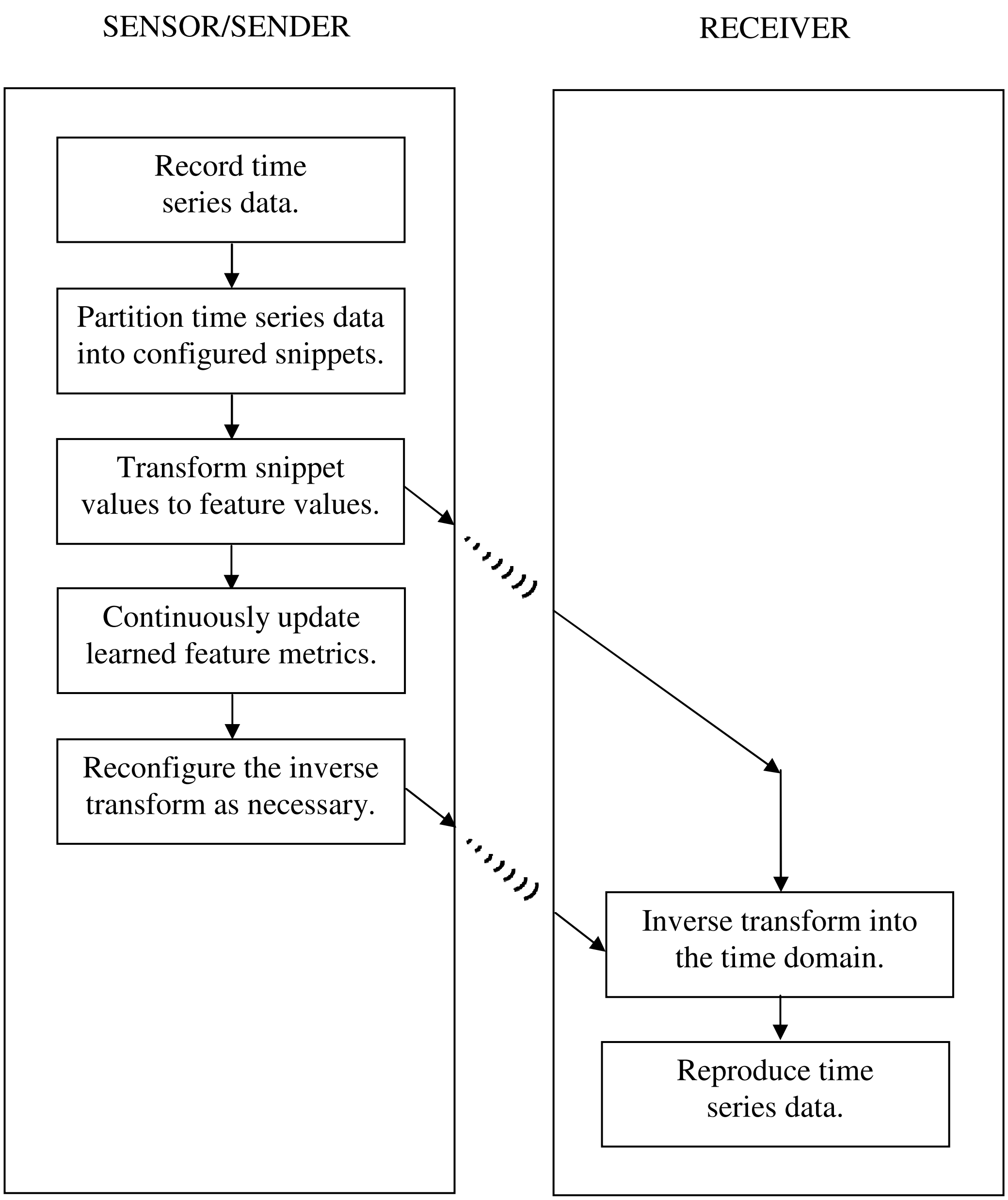

Figure 1. Time Series Reduction and Reproduction Operation

Details

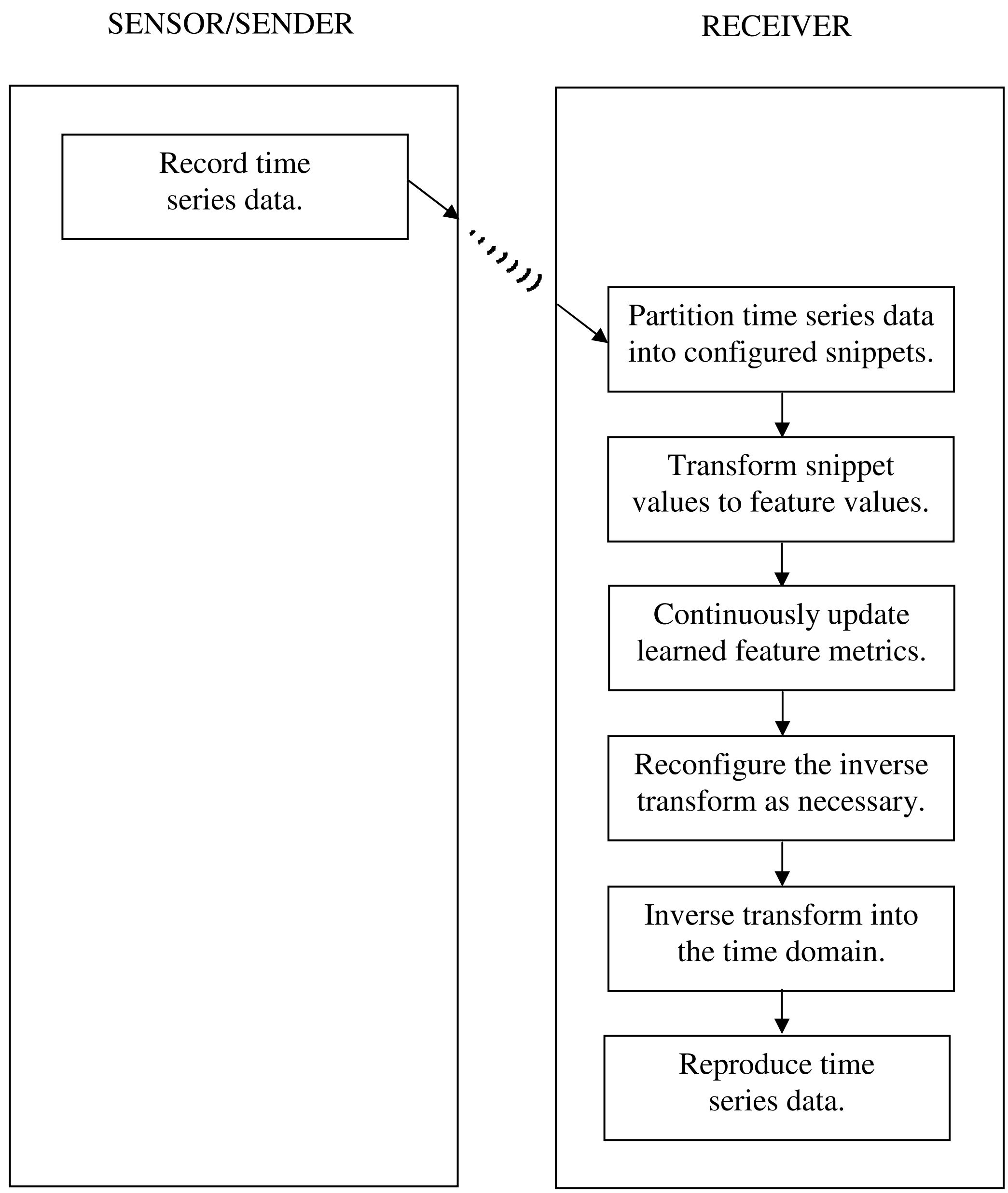

Figure 2. Receiver Filtering after Reception, without Data Reduction

To further explain data reduction for the above example, Brainlike sensing may continuously update average feature values for an individual and occasionally send a configuration packet, containing the corresponding most salient frequency ranges for that individual. Meanwhile, for each packet the sending phone would transmit only the power levels for those 1,000 frequency ranges on a one byte gray scale. Resulting data reduction would approach 20 to 1, depending on how often update configuration packets were sent. Update packets in this case could be 1,000 two byte words, pointing to the most salient features among as many as $2^{16} = 65{,}536$ possible features. In the worst case, the packet would be sent with every set of feature values, resulting in a data compression ratio of only 20 to 3. In practice, the packet would require transmission only rarely, resulting in a data compression ratio of nearly 20 to 1.

Brainlike sensing components may reside on a sensing and sending unit, a receiver unit, or both, as shown in Figures 1 through 3. The sensor/sender unit shown in all three figures records the data in the time domain, as shown in the top, left block of each figure. The sensor/sender may then perform other operations shown in blocks below it, or the receiver may perform any or all other operations, as shown in the three figures. Under one option, as shown in Figure 1, the sensor/sender and receiver may divide processing in a way that will both filter out clutter and reduce transmitted data. In this case, the sensor/sender may partition the data into snippets as shown, then reduce snippet values to feature values, and then transmit the feature values. The receiver may then inverse transform them into the time domain. The sensor/sender unit may also continuously update feature salience values, reconfigure data reduction as necessary, and transmit reconfigured values to the receiver in order to ensure proper time domain recovery, as shown. The sensor/sender may also occasionally send updated learned metrics, which the receiver would then receive and then reconfigure the inverse transformation function accordingly. The receiver may then play or display the reproduced time series values as appropriate.

Under another option, which is shown in Figure 2, the sensor/sender may transmit all recorded data in the time domain, in the usual way. The receiver may then partition the data into time series snippets, like the one second snippets in the above example, as shown. The receiver may then transform snippet values into feature values, like the 1,000 frequency domain feature values in the example. The receiver may then use Brainlike sensing to update learned feature metrics in real time. The receiver may then reconfigure the transform and inverse transform functions according to the most recently transmitted learned metrics. The receiver may then inverse transform the feature values for the snippet back into the time domain, so that the reproduced sound resembles the sender's voice. Finally, the receiver may play or display the reproduced time series values as appropriate. Under this option, the receiver would filter out clutter frequency components, but the overall system would not reduce transmitted data.

Under a third option, shown in Figure 3, filtering and clutter reduction will occur without data reduction, as with the Figure 1 case, but reduction will occur in the sensor/sender unit, instead of the receiver unit. Under yet another option (not shown), the sensor/sender unit will also inverse transform the feature values into the time domain, and then play or display the reproduced time series values as well.

Available voice recognition and synthesis technology may be coupled with Brainlike sensing to deliver affordable and valuable voice data reduction and filtering solutions, quickly. For example, currently available technology can efficiently convert voice data to text data, resulting in data reduction factors of about 1,000 from radio quality data (assuming that an individual says about 120, eight character words per minute). The text may then be transmitted, along with a feature configuration packet. The configuration packet would indicate features at the receiving should be used and how they should be combined to reproduce the caller's voice.

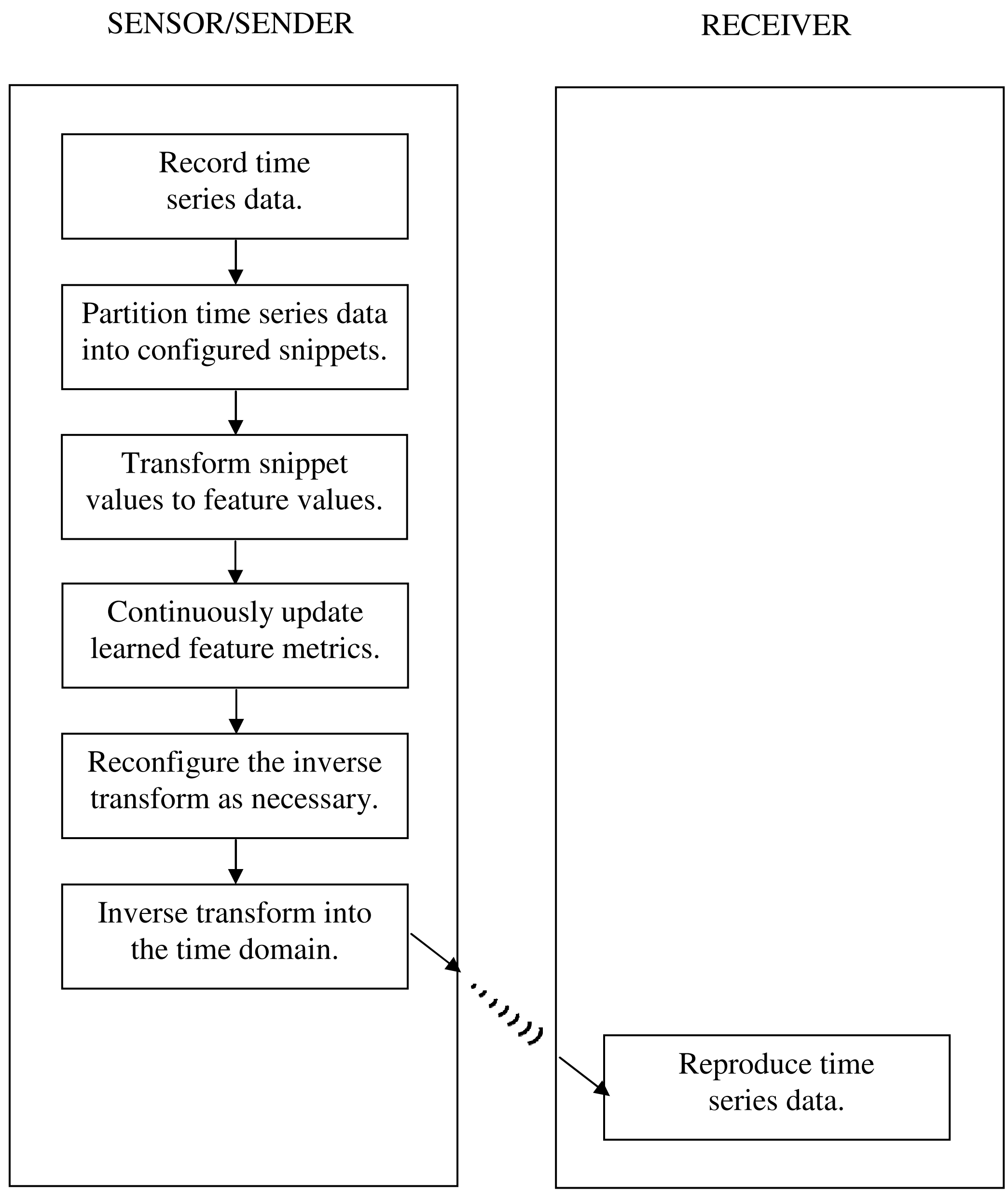

Figure 3. Sender Filtering before Transmission, without Data Reduction

The features in this case would not be FFTs, but state-of the art features for reproducing a person's voice from text. Any variety of features can be used as well for greater efficiency, such as orthogonal counterparts to FFTs that can be transformed and inverse transformed linearly. Closely held features may be used as well, allowing time series to be encrypted before transmission and then decrypted after transmission. In addition, straightforward extensions of Brainlike sensing usage in the univariate time series case can produce similar clutter and bandwidth reduction in bivariate time series such as video images, as well as higher dimensional time series. Thus, Brainlike sensing may also be applied in many ways, where "individuals" may be any variety of sensors, generating any variety of time series data in real time.

Technology for converting voice to text and for converting text to a person's voice is not new. Manual voice conversion technology is as old as stenography and manual text conversion is as old as voice impersonation. Automatic voice recognition, transformation, and synthesis have also been studied and developed for decades, resulting in their effective use in many available products today. Brainlike sensing adds the key elements of being able learn an individual's metrics in real time and then using the learned metrics to reproduce that individual's time series.

Brainlike sensing has been designed to update learned metrics, including feature means, covariances, and estimation weights quickly and compactly. With Brainlike learned metrics for a person's voice readily available, they can be used to identify and suppress noisy snippets that do not contain the voice, and enhance the person's voice while suppressing noise in snippets that contain both voice and noise. Learned weights may also be used to impute voice features that may not have been transmitted. In the Figure 1 case, for example, available bandwidth may allow time series to be transmitted usually at 20 KHz, but sometimes only 10 KHz may be available. In that case, learned weights for imputing higher frequency components for a person from that person's lower frequency components may be used to enhance his or her voice, even though the higher frequency components could not be reproduced from the arriving signal.

The numbers presented in the above example were intended to illustrate how Brainlike sensing can reduce data. Automated feature refinement methods, available in the **Brainlike Studio™ Toolkit** may reduce the number of features substantially, improving data reduction to factors above 20 to 1. In practice, transmission bandwidth, along with voice reproduction accuracy, will decrease with the number of salient features being transmitted. Brainlike sensing applications may easily be configured to make the number of salient features configurable, so that cell phone users and network providers can adjust bandwidth as well as filtering accordingly.

In summary, patented and patent pending Brainlike sensing technology is capable of uniquely filtering and reducing time series data, in a general purpose form that can be efficiently deployed on cell phone or remote sensor processors. Voice clutter and data reduction is one key application, but Brainlike sensing can add similar value in many other time series applications, ranging from **video surveillance** to **health care monitoring**.