The Rice Electronics Pattern Recognition Technology

Industry Relevance

And

Application Anecdotes

Rice Pattern Recognition Technology **– Affected Industries**

The Rice Pattern Recognition Intellectual Property (IP) qualifies as a "disruptive" technology. The IP has underlying capabilities which can be leveraged in numerous industries. Some areas are extremely significant due to their market scope, existing limitations or impact on society, and include:

1) **Human Interface to Machines and Internet.** This relates to recognition of speech, imagery, text and abstract symbols. Recognition capability is becoming an integral part of personal information devices (ranging from cell phones to computers). But many current recognition systems for personal devices require an active internet connection. This results in a cumbersome interaction between the device, and a remote computational resource which performs the recognition process. The Rice IP can enable at least part (and possibly all) of the recognition function to be migrated to the local personal device itself.

2) **Medical and BioInformatics Applications.** Recent advances in biological science are producing human genome information at an incredible rate. Additionally, such data is proliferating with respect to animals and even micro-organisms. It has become clear that exploitation of this data (and its timely benefit to mankind) requires new computing technology. Many of the most intense computational problems in this area relate to specialized pattern recognition. Advances in this area can be greatly accelerated by the subject IP.

3) **Wireless Communications.** The field of wireless communications may seem mature, including the area of networked wireless systems. However, certain derivatives of the Rice IP can affect the evolution of new wireless solutions. Shannon's laws dictate the capacity of wireless systems (as a function of bandwidth and SNR). Bandwidth is a valuable commodity and a definitive system limitation. However, in network architectures, the subject IP might optimize communication channels on the basis of significant improvements in SNR.

.

4) **Information Technology.** There are numerous applications of the IP in general areas of information technology. Opportunities exist wherever data must be recognized, routed, mined, secured or stored. Platforms would include servers, routers, terminals and mobile devices The Rice Electronics IP can be applied to the recognition of text, records, numerical patterns and abstract patterns including those of a 1-D and 2-D nature

Given below are examples of the IP utility.

Rice Pattern Recognition Technology **– Data Storage and Recognition**

Described below is a unique system for storage and retrieval of information based on the Rice Pattern Recognition IP. Aspects of the system bear similarity to "relational" memory or "content addressable" memory. However, the system under description transcends such entities in terms of both capacity and functionality.

There are 2 primary operational parameters for the system, M and N, where $M<N$. For purposes of this discussion, $M = \sqrt{N}$. (Although not a requirement, this may be typical, with N ranging from 2^{10} to 2^{16}).

Physically, the system consists of conventional memory circuitry supported by proprietary processing logic executing proprietary methods.

Functionally, the system stores data and is therefore characterized by certain "write" and "read" operations. However, the behavior of the system (and its capabilities) differ distinctly from any existing memory paradigm. This is discussed below.

The system has "M" storage channels. In each channel may be stored an N-point pattern, and some associated "data tag". (The data tag might be some identifier or other type of meta-data, typically consisting of $\log_2 N$ bits).

The detailed structure of a stored pattern cannot be explicitly retrieved. Instead, the system may be interrogated to determine whether or not an external pattern exists in a given system channel. This is done by presenting a "target" pattern to the system as a stimulus (where the target need not be an exact replica of the originally stored pattern). If the system has previously stored the target pattern, then it will respond affirmatively. In addition, the associated data-tag (in its exact, original form) will be retrieved by the system.

In this way, the system serves as both a powerful recognition and storage entity. Further, the effective "storage capacity" may be 30X to 50X times greater than the actual memory circuitry utilized.

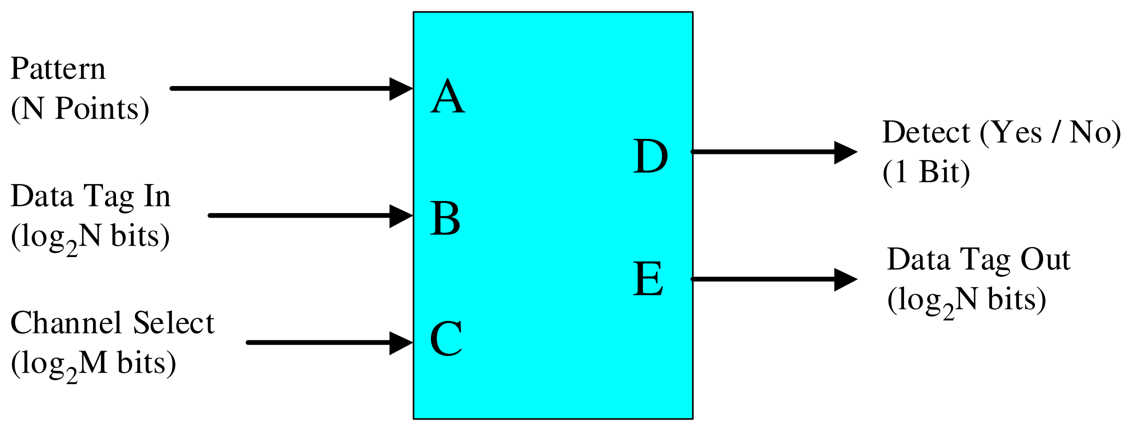
For a hypothetical set of parameters (e.g., N=4096, M=64), the system is characterized as follows:

- 64 Channels
- A single 4096 point pattern may be stored in each channel
- A 12 bit "data tag" may be stored for each pattern

Figure 1 illustrates the system inputs and outputs. Inputs A, B and C are necessary for a store (write) operation. (These are the pattern, the data tag and the channel number, respectively). Inputs A and C are also used for a read (i.e. interrogate, or “retrieval”) operation.

As described above, retrieval of information from the system resembles pattern recognition. In this case, a “stimulus pattern” is presented at input A of Figure 1, while the channel of interest is designated to input C. If the system has a matching pattern stored at the designated channel, then this is indicated by the “detect output” D. Additionally, the “data-tag” associated with the pattern will be produced by the system at output E.

There are no restrictions on the stochastic features of the subject data patterns. They may consist of random numbers. Further, the input stimulus may be corrupted in various ways (e.g., by noise, missing fragments, etc). Recognition can still take place, given sufficient integrity of the input stimulus.



Pattern Storage System – Inputs / Outputs

Figure 1

As referenced above, the implementation of the system relies on conventional memory technology, and proprietary processing. Integral to the Rice IP technology base, are optimized processing structures. These minimize the cost, complexity and power consumption associated with the recognition system.

Rice Pattern Recognition Technology
– A Communications Example

The Rice Pattern Recognition IP can encode data simultaneously into multiple domains (time, frequency and code domains). This facilitates the storage and transmission of data patterns with unique levels of energy efficiency.

The IP may process large data patterns for applications such as speech or image recognition. However, smaller patterns are sometimes more suitable for communication systems. The application below involves data patterns (sequences) with length suited for a particular communications application (i.e., 128 points in length).

The Rice IP can enhance the efficiency of communications networks by its multi-domain encoding approach. Shannon's laws dictate the capacity of a communications channel (as a function of bandwidth and SNR) Bandwidth is a valuable commodity and generally a constraining factor in systems design. However, the subject IP can enhance wireless network performance through significant improvement in SNR levels. This is illustrated below by its postulated use for an "enhanced" system. This enhance system would have elements of both CDMA and OFDM, but with distinct advantages over either.

A conventional CDMA system imparts a sequence of phase transitions on a carrier signal. A sequence of 128 phase transitions might be typical. In the basic system, only 2 phase states might be employed (e.g., 0 degrees or 180 degrees). Thus, a sequence might be expressed as follows;

0° 0° 180° 0° 180° 0° 0° 180° 0° 180° 180° 180° 0° 0° - - -

A grouping of 128 such transitions would traditionally be termed a "symbol".

In practice, the conventional CDMA system would employ an encoding technique based upon some orthogonal system of basis functions (e.g. Walsh functions), such that the phase sequence associates the "symbol" with a specific channel. That is, the phase sequence within the symbol represents a channel number. With such a system, it might be possible to represent 128 distinct channels.

Therefore, in the conventional system, a symbol would belong to one of 128 channels. However, the actual "data" transmitted (carried) by a symbol might only be a single bit. In the basic CDMA system, this is encoded by simply manipulating the sign of the symbol. The system would then transmit information by means of a stream of contiguous symbols. This might resemble the following, (where +1 or –1 is indicative of a symbol, and hence a single data bit);

–1 +1 +1 –1 +1 +1 +1 –1 +1 –1 –1 +1 –1 +1 +1 –1 –1 - - -

It should be noted that in the conventional CDMA, the only “multi-bit” encoding performed on a symbol is in the “code-domain” as described above, via phase sequences.

In contrast, the Rice IP can encode information by three means. These are;

1) manipulation of the phase sequencing within the symbol,

2) modulation of spectral (frequency domain) characteristics within the symbol

3) control of temporal (time domain) spacing among the symbols

As described above, the conventional CDMA employs only the first of these means (i.e., manipulation of the phase sequence). Also, the conventional approach uses a contiguous stream of symbols (as has been described). In contrast, the Rice IP may produce a “low energy” symbol stream, exemplified as follows;

+1 0 0 0 0 0 0 –1 0 0 0 0 0 –1 0 0 0 0 0 0 +1 0 0 0 0 0 0 0 +1 0 0 - - -

Time →

In the “low energy” representation, a “0” indicates the absence of a symbol from a specific channel at a specific time. However, this representation may convey more information than the conventional CDMA symbol stream, but with nearly 90% reduction in energy of the sequence.

At the same time, the alternative “low energy” signal may serve as a multi-user access waveform in network systems. This is because “multi-bit” data is encoded in each domain, with the opportunity to use either the spectral or code domain to discriminate among multiple channels. Ad hoc communications is possible, obviating the need for explicit synchronization among channels. (It is in non-synchronous networks where the IP approach may render the most distinct advantages relative to traditional CDMA).

The use of a low energy signal as described above may seem inefficient (as a channel may appear to be under-utilized). Therefore, the symbols might be further “spread” along the time axis, in order to make better use of channel resources. This can be done by convolving the “low energy” symbol stream with a pseudo-random code (such as a Barker code). In this case, a short but (illustrative) Barker code might consist of following 5 digits:

+1 +1 +1 –1 +1

Also, use of a spreading code can enable reduction of the symbol amplitudes at the transmitting device, further conserving transmitter power. Accordingly, after convolution with the Barker code and suitable amplitude scaling, a resulting “low-energy sequence” might resemble that below:

+K +K +K –K +K 0 0 –K –K –K +K –K 0 –K –K –K +K –K - - -

→ Time

where K < 1.

As seen above, the channel is now more fully utilized with lower amplitude symbols. This serves to minimize the peak power of any transmitted symbol in the sequence. In theory, the K parameter can be related to the inverse of the length of the spreading code (in this case, the 5 digit Barker code). Longer spreading codes are possible.

Additionally, the nature of the Rice IP is to mitigate requirements for highly linear generation of the symbols. This represents another level of power reduction of the transmitted signal. (Such linearity requirements are an impediment to power conservation in the use of modern multi-user access waveforms, such as those based on OFDM).

There are distinct advantages to these "low-energy sequences" in certain wireless network communications. Among these are reduction in mutual interference among channels, and lower overall ambient power levels. Further, the signaling technique may be used to reduce power consumption of small mobile devices.

As with the data storage IP, the Rice technical base for communications includes special structures for high speed processing. These greatly reduce the footprint (size, power and cost) of the IP, which is particularly important for use in small and mobile devices. The functional utility of these structures, includes multi-domain pattern encoding and recognition for the communications application.