Latency-Information Theory and Applications. Part I: On the Discovery of the Time Dual for Information Theory

Erlan H. Feria
The College of Staten Island (CSI) of the City University of New York (CUNY)
Department of Engineering Science and Physics
2800 Victory Blvd, Staten Island, New York, 10314
e-mail:feria@mail.csi.cuny.edu, website:http://feria.csi.cuny.edu

ABSTRACT

As part of research conducted on the design of an efficient clutter covariance processor for DARPA's knowledge aided sensor signal processing expert reasoning (KASSPER) program a *time-dual* for information theory was discovered and named *latency theory*, this theory is discussed in this first of a multi-paper series. While information theory addresses the design of communication systems, latency theory does the same for recognition systems. Recognition system is the name given to the time dual of a communication system. A recognition system uses prior-knowledge about a signal-processor's input to enable the sensing of its output by a processing-time limited sensor when the fastest possible signal-processor replacement cannot achieve this task. A processor-coder is the time dual of a source coder. While a source coder replaces a signal-source to yield a smaller sourced-space in *b*inary dig*its* (*bits*) units a processor coder replaces a signal-processor to yield a smaller processing-time in binary *o*perat*ors* (*bors*) units. A sensor coder is the time dual of a channel coder. While a channel coder identifies the necessary overhead-knowledge for accurate communications a sensor coder identifies the necessary prior-knowledge for accurate recognitions. In the second of this multi-paper series latency theory is successfully illustrated with real-world knowledge aided radar.

Keywords: space-time dual, time dual, latency, latency theory, information, latency-information theory, sourced-space, motion-time, processing-time, retention-space, bits, bors, processor ectropy, source entropy, sensor consciousness, channel capacity, knowledge aided, intelligent system, biology

1. INTRODUCTION

A real-world problem whose high performance is attributed to its use of an intelligent system (IS) is knowledge-aided (KA) airborne moving target indicator (AMTI) radar such as found in DARPA's knowledge aided sensory signal processing expert reasoning (KASSPER) [1]-[2] program. The IS's intelligence, or prior knowledge, is clutter synthetic aperture radar (SAR) imagery and its intelligence processor (IP), or on-line computer, is the associated clutter covariance processor. Unfortunately, however, the excellent signal to interference plus noise ratio (SINR) radar performance achieved directly depends on satisfying prohibitively expensive storage and computational requirements. The former 'storage' problem is easily addressed by using a highly efficient lossy source coding technique, e.g., a minimum mean squared error (MMSE) predictive-transform (PT) source coder [3] that compresses a typical SAR image by several orders of magnitude: a MMSE PT source coder is used rather than the JPEG2000 standard [4] since the PT coder outperforms JPEG by at least 5dB in this application [5]. Yet, a lossy source coder compressor seriously compromises the SINR

This work was supported in part by the Defense Advanced Research Projects Agency (DARPA) under the KASSPER Program Grant No. FA8750-04-1-004 and the PSC-CUNY Research Awards PSCREG-37-913, 38-247

Mobile Multimedia/Image Processing, Security, and Applications 2008
edited by Sos S. Agaian, Sabah A. Jassim, Proc. of SPIE Vol. 6982
698210, (2008) · 0277-786X/08/$18 · doi: 10.1117/12.784548

radar performance. Thus one seeks the replacement of the clutter covariance processor with a processor that is better matched to highly compressed SAR imagery.

In this paper, the discovery of a time dual for information theory, named latency theory, is reported. Latency-theory offers a time dual for all the concepts found in information theory. Among them is that of a recognition system as the time dual of a communication system where a recognition system uses prior-knowledge about a signal-processor's input to enable the sensing of its output by a processing-time limited sensor when the fastest possible signal-processor replacement cannot achieve this task. Other concepts are a sensor as the time dual of a channel, a processor coder as the time dual of a source coder, and a sensor coder as the time dual of a channel coder. While a source coder replaces a signal-source to yield a smaller sourced-space in *b*inary dig*its* (*bits*) units a processor coder replaces a signal-processor to yield a smaller processing-time in binary *o*perat*ors* (*bors*) units. Similarly to a source coder a processor coder can be lossy or lossless. While a channel coder identifies the necessary overhead-knowledge for accurate communications a sensor coder identifies the necessary prior-knowledge for accurate recognitions. Furthermore, there are time duals for the *source-space performance bounds* of *source-entropy* $\boldsymbol{H}$ in source-space bits per given discrete random variable source knowledge output to be communicated, and *channel-capacity* $\boldsymbol{C}$ in source-space bits per channel-space bits, a dimensionless quantity less than or equal to one, that are used in the design of *channel and source integrated* (CSI) coders. The time duals of $\boldsymbol{H}$ and $\boldsymbol{C}$ are the *processor-ectropy* $\boldsymbol{K}$ and *sensor-consciousness* $\boldsymbol{F}$, respectively, used in the design of *sensor and processor integrated* (SPI) coders. $\boldsymbol{K}$ and $\boldsymbol{F}$ are *processor-time performance bounds* in processor-time bors per given vector variable processor knowledge output to be recognized, and sensor-time bors per processor-time bors, a dimensionless quantity less than or equal to one, respectively.

Since latency theory inherently encapsulates the idea of a *lossy processor* it directly addresses the aforementioned need for such a processor to replace the *lossless* clutter covariance processor of real world prior knowledge radar. In the second [6] of this multi-paper series a lossy clutter covariance processor is discussed that yields outstanding SINR radar performance while being several orders of magnitude faster than the original lossless clutter covariance processor. Due to its outstanding performance and low cost the advanced lossy processor architecture has been adopted by KASSPER. The discovery of latency theory occurred in an incremental fashion as part of contractual work conducted by the author for DARPA's KASSPER program [7] as well as part of proposed work for several PSC-CUNY Research Awards [8]. As the new latency theory concepts were found they were also disseminated to the general public at different stages of its development [9]-[16]. Thus in this paper the results of several years of research in the development of latency theory as the time dual of information theory are presented in a concise and straight forward manner for its general use in science, technology, humanities, and social sciences problems as is already found to be the case with information theory [17]. In this paper I will first review the basic elements of information theory and then present its time dual, i.e., latency theory.

2. INFORMATION THEORY

Information theory's communication system is presented in block diagram form in Fig. 1. It consists of three major subsystems. The first is a source-coder (encoder/decoder) which replaces an inefficient signal-processor to yield a smaller sourced-space. Source-coders can be lossy or lossless. For instance, in Fig. 2a a 4 Mbytes synthetic aperture radar (SAR) image is shown that is compressed/decompressed by a lossy MMSE PT source coder [5] by a factor of 8,172 as shown in Fig. 2b. Although this compressed image looks rather lossy, the actual application determines its usefulness, e.g., in the knowledge aided radar application of KASSPER this compressed image yields outstanding target detections when used in conjunction with a lossy processor [6]. The second subsystem is a channel coder (encoder/decoder) that identifies overhead-knowledge to transmit for a more accurate communication of the sourced-space compressed knowledge. The third subsystem is what I call a motion-coder (encoder/decoder) which is responsible for the motion of the channel encoder output knowledge for some prescribed motion-space interval (or space-dislocation). For instance, this motion-coder may be a modulation-antenna subsystem. However, the design of this subsystem, as opposed to that of a source coder and channel coder, is not directly addressed by information-theory. Instead for the design of a motion-coder we use the laws of motion in physics, e.g., as is done when Maxwell's equations and spectral analysis are used to design of an appropriate modulation-antenna system.

As mentioned earlier two performance bounds are used to guide us in the design of a channel and source integrated or CSI coder. The first bound is the *source-entropy* $\boldsymbol{H}$ which is defined as the average amount of information in a discrete random variable (or knowledge) $\mathbf{X} \in \{o_1, \ldots, o_U\}$, i.e.,

$$\boldsymbol{H} = \sum_{i=1}^{U} P[o_i] I_S[o_i] \quad \text{in bits per } \mathbf{X} \tag{2.1}$$

where $I_S[o_i]$ is the *source-information* advanced by the source outcome o_i and defined by

$$I_S(o_i) = \log_2(1/P[o_i]) \text{ in bits per } o_i \tag{2.2}$$

with the probability $P[o_i]$ being a measure of o_i's uncertainty, which is in turn driven by the *passing of time* [17]. A source-coder's design is then approached using $\boldsymbol{H}$ as a performance bound. In particular, the source encoder rate R_{SE} in bits per $\mathbf{X}$ reflects a *sourced-space penalty* incurred by the source coder. Thus when R_{SE} is greater than or equal to the source-entropy $\boldsymbol{H}$—and less than or equal to the source rate R_S, i.e.,

$$\boldsymbol{H} \le R_{SE} \le R_S \tag{2.3}$$

we say that the source coder is lossless and when R_{SE} is smaller than $\boldsymbol{H}$ we say that the source coder is lossy. Thus it can be said that a lossy source coder has a source-space advantage over a lossless source coder since its *sourced-space penalty* is lower. As an illustration see Fig. 2 where R_S= 16 bits/pixel > $\boldsymbol{H}$ > 4 bits/pixel > R_{SE}=1.24 x 10^{-6} bits/pixel. Note that the compression factor for this image is 512 bytes per 4Megabytes or 8,172. Thus the sourced-space penalty associated with this lossy source coder is at least four orders of magnitude less than that of the best lossless source coder since $\boldsymbol{H}$ > 4 bits/pixel.

The second bound used is *channel-capacity* $\boldsymbol{C}$ [17] which is defined here as the maximum achievable CSI coder-ratio R_{CSI}. R_{CSI} is defined as the ratio of communicated R_{SE}, R_{SE}^{com} (= k bits/outcome), to space-dislocated channel-encoder rate R_{CE} (= n bits/outcome) through a noisy-channel, i.e.,

$$0 \le R_{CSI} = \frac{R_{SE}^{com}}{R_{CE}} = \frac{k}{n} \le 1 \tag{2.4}$$

where R_{SE}^{com} is smaller than R_{CE}. R_{CSI} is achievable when R_{SE}^{com} is reconstructed by the channel-decoder with an arbitrarily small probability-of-error (a CSI-coder is said to be "lossless" if it satisfies this condition, otherwise it is said to be "lossy"). A simple example of a channel encoder is one that adds a parity-bit to each transmitted byte. This additional bit then permits the channel-decoder to determine if one of the 9 transmitted bits is in error. More than one bit can be concatenated with the byte to also produce error corrections. For a memoryless noisy-channel [17, 19], with its input E^n and output F^n denoting n-bits codewords for each communicated outcome, $\boldsymbol{C}$ is defined by

$$\begin{aligned} 0 \le \boldsymbol{C} &= \max_{\{P[e_i]\}} \Im(E^n, F^n) = \max_{\{P[e_i]\}} \frac{I(E^n, F^n)}{\boldsymbol{H}(E^n)} \\ &= \max_{\{P[e_i]\}} \frac{\boldsymbol{H}(E^n) - \boldsymbol{H}(E^n / F^n)}{\boldsymbol{H}(E^n)} = \frac{\boldsymbol{H}(\mathcal{E}^n) - \boldsymbol{H}(\mathcal{E}^n / \mathcal{F}^n)}{\boldsymbol{H}(\mathcal{E}^n)} = \frac{R_{SE}^{com}}{R_{CE}^{Min}} \le 1 \end{aligned} \tag{2.5}$$

where $I(E^n,F^n)$ is the mutual-information, $\boldsymbol{H}(E^n/F^n)$ is the decreased entropy of E^n after F^n is observed, and R_{CE}^{Min} is the minimum R_{CE} yielding the maximum achievable R_{CSI}. $\mathcal{E}^n$ is the random-code-word E^n whose probability-distribution $\{P[e_i]\}$ maximizes the mutual-information ratio $\mathcal{I}(E^n,F^n)=I(E^n,F^n)/\boldsymbol{H}(E^n)$. For instance, $\mathcal{E}^n$ has a uniform probability-distribution when the channel is binary-symmetric and is Gaussian for an additive-Gaussian channel [17,19].

The channel-capacity $\boldsymbol{C}$ guides the design of either lossless or lossy CSI-coders. A lossless CSI-coder is characterized by an achievable R_{CSI}, i.e.,

$$0 \le R_{CSI} \le \boldsymbol{C} \tag{2.6}$$

and is ideal when $R_{CSI} = \boldsymbol{C}$. A lossy CSI-coder has a R_{CSI} that is not achievable, i.e.,

$$\boldsymbol{C} < R_{CSI} \le 1 \tag{2.7}$$

The previously described *source-space performance bound methodology* is also known as channel-coding or "the mathematical theory of communication" of Shannon [20]. Information rate-distortion theory which defines a distortion channel-capacity (or rate-distortion function) [17, 19] is also available to guide the design of lossy CSI-coders. These two theories are embraced by information-theory which may include other information topics. However, as mentioned earlier motion-coders are not included as part of information-theory but rather their design is relegated to the use of the laws of motion in physics for their design.

3. LATENCY THEORY

In this section the time dual of information theory, i.e., latency theory, is advanced. Consider Fig. 3 which is the time dual of the communication system of Fig. 1 to which I have given the name recognition system. A recognition system uses prior knowledge about a signal-processor's input to facilitate the sensing of its output by a processing-time limited sensor when the fastest possible signal processor replacement cannot achieve this task. Similarly to a communication system a recognition system has three major subsystems.

The first subsystem is a processor-coder (the time dual of a source-coder) which replaces an inefficient signal-processor to yield a smaller processing-time. Processing-coders can be lossy or lossless. For instance, in Fig. 4a an inefficient full-adder signal processor is shown which requires at most six bor levels of computation-time to yield the vector output $\mathbf{y}=[s_i\ c_{i+1}]$ given the vector input $\mathbf{x}=[a_i\ b_i\ c_i]$. where a_i and b_i are the two added bits, c_i is the carry in bit, s_i is the sum bit, and c_{i+1} is the carry out bit. On the other hand, in Fig. 4b the fastest possible replacement for the inefficient full-adder signal processor is shown which requires at most 3 bor levels of computation-time to yield $\mathbf{y}$. Finally is Fig. 4b a lossy processor-coder is shown whose output $\hat{\mathbf{y}} = [0\ \ c_{i+1}] \neq \mathbf{y}$ does not require the evaluation of s_i and thus results in only 2 bor levels of computation-time with also a significantly simpler implementation. The second subsystem is a sensor coder (the time dual of a channel coder) that identifies the necessary prior-knowledge for a more accurate recognition of the processing-time compressed knowledge. The third subsystem is what I call a retention-coder (the space dual of a motion-coder) which is responsible for the retention of the sensor coder output knowledge for some prescribed retention-time interval (or time-dislocation). After latency theory is discussed we will return in the paper ending conclusion section to a discussion of the retention-coder in the context of a physics problem.

The sensor of a recognition system is the time dual of the channel of a communication system. While a channel is linked to a source-space space dislocation a sensor is linked to a processing-time time dislocation. A simple illustration of the necessary processing-time time dislocation of a recognition system is when a sequential adder that uses the fastest possible 1-bit full-adder of Fig. 4b to add two bytes is required by the sensor to yield its output in a maximum of 12 bors. Since this full adder requires at least 16 bors to add two bytes it is then necessary to perform a time dislocation of the processing-time to an earlier time by at least 16-12= 4 bors. Thus the only solution to this problem that satisfies the stated sequential processing constraint is to use prior knowledge about the sequential adder input to start the processor 4 bors earlier in time. The device that determines what this prior knowledge should be is called a *sensor coder* and is the time dual of a processor coder. Using our running example to illustrate this concept we can start the processing earlier in time by four bors if it is known that all the bits of the two added bytes always occur at the same time and it is assumed that the two least significant bits of the two bytes can be replaced with the pairs of bits 10 and 01, respectively, without resulting in a significant error on the resulting addition. The recognition system that results from this design approach is then called a sensor and processor integrated or SPI coder. In addition, the four bits 10 and 01 can be kept in storage in the retention-coder by the sensor coder where the design of the retention-coder is not addressed by the latency theory in its present form. In [22] the performance bound approach of information theory, developed next for its time dual, i.e., latency theory will also be applied to the design of motion-coders and retention-coders that are needed for general recognition/communication systems.

Next the time duals for the two performance bounds of information theory are stated.

The first performance bound that is considered is the *processor-ectropy* $\boldsymbol{K}$—the time dual of the source-entropy $\boldsymbol{H}$—which is defined as the maximum amount of latency for a processor vector output $\mathbf{y} = [y_1,..., y_J]$, i.e.,

$$\boldsymbol{K} = \max(L_P(y_1),..,L_P(y_J)) \quad \text{in bor units per } \mathbf{y} \tag{3.1}$$

where $L_P[y_i]$ is the *processor-latency* of y_i which is defined as the minimum processing-time that is needed to obtain y_i after the original signal processor, say the 1 bit full adder of Fig. 4a, is redesigned subject to implementation constraints $\{C[y_i]\}$—for example, the constraints can be NAND gates that can have any number of inputs as is illustrated in Fig. 4b from which we derive $L_P(c_{i+1})$=2 bors/c_{i+1}, $L_P(s_i)$=3 bors/s_i and $\boldsymbol{K}$=max($L_P(c_{i+1})$, $L_P(s_i)$) =3 bors/**y**. Thus

$$L(y_i) = \mathrm{f}(C[y_i]) \quad \text{in bors units per } y_i \tag{3.2}$$

with f($C[y_i]$) indicating that $L(y_i)$ is a function of $C[y_i]$. The constraint $C[y_i]$ is the time dual of probability $P[o_i]$ and is driven by a *configuration of space certainty*, i.e., the ‘certain’ occupancy-space in m^3 occupied by the implemented NAND gates.

The processor-coder rate R_{PC} is the time-dual of R_{SC}, e.g., R_{PC}=3 bors/**y** for Fig. 4b**. A** lossless processor-coder is the time-dual of a lossless source-coder. It has a R_{PC} achievable, i.e.,

$$\boldsymbol{K} \le R_{PC} \le R_P \tag{3.3}$$

It is also ideal when R_{PC}=$\boldsymbol{K}$ and is equivalent to the signal processor when R_{PC}=R_P. In Fig. 4b an ideal processor-coder is displayed with $\boldsymbol{K}$=3 bors/y. On the other hand, a lossy processor-coder is the time-dual of a lossy source-coder. It has a R_{PC} that is not achievable, i.e.,

$$0 \le R_{PC} < \boldsymbol{K} \tag{3.4}$$

but is faster and simpler than a lossless one as illustrated in Fig. 4c where R_{PC} = 2 bors/y < $\boldsymbol{K}$ = 3 bors/**y** and s_i is not evaluated.

The second performance bound that must be used in the design of a recognition system is *sensor-consciousness* $\boldsymbol{\mathcal{F}}$—the time dual of channel capacity $\boldsymbol{\mathcal{C}}$. This performance bound is linked to the occurrence of a processing-time limited sensor (PTLS) condition. This condition is given by

$$\boldsymbol{\mathcal{K}} > W \tag{3.5}$$

where W is the maximum waiting time in bors of the sensor, e.g., W was assumed earlier in our example to be of 12 bors when the full output of a 1-bit full-adder of Fig. 4b was sensed. Since $\boldsymbol{\mathcal{K}}$ is at least 16 bors for this example it then follows that the PTLS condition is satisfied for this case. Notice that unless the PTLS condition is satisfied there is not need for the use of prior knowledge to advance or time dislocate the processing-time since all we need to do is build a lossless processor with $R_{PC} \le W$. Thus when the PTLS condition (3.5) is not satisfied we do not need a recognition system. On the other hand, when the PTLS condition is satisfied the positive difference

$$TD = \boldsymbol{\mathcal{K}} - W \tag{3.6}$$

tell us about the necessary time dislocation (TD) of computation-time that must be addressed by the sensor coder via the use of prior knowledge. The sensor consciousness $\boldsymbol{\mathcal{F}}$ can now be defined and is given by the expression

$$0 \le \boldsymbol{\mathcal{F}} = \tfrac{W}{\boldsymbol{\mathcal{K}}} \le 1 \tag{3.7}$$

where for our running example it is noted that $\boldsymbol{\mathcal{F}}$=12/16=0.75.

The definition for sensor-consciousness can also be stated as the time-dual of that for channel-capacity (2.4)-(2.7). $\boldsymbol{\mathcal{F}}$ is then the maximum achievable SPI-coder ratio R_{SPI} where R_{SPI} is given by the ratio of R_{PC}^{rec} to the sensor-coder rate R_{SC}. R_{PC}^{rec} is the recognized R_{PC} and thus is the same as the maximum waiting time W of the sensor, i.e., $R_{PC}^{rec}=W$, while R_{SC} is equal to R_{PC}^{rec} plus the amount of time-dislocation, i.e., $R_{PC}- R_{PC}^{rec}$, that it can provide via the use of prior knowledge. Thus $R_{SC}= R_{PC}^{rec} + R_{PC}- R_{PC}^{rec} =R_{PC}$ where R_{PC} is equal to T, the processing-time of the processor-coder in bors, per **y**. Thus

$$0 \le R_{SPI} = \frac{R_{PC}^{rec}}{R_{SC}} = \frac{W}{T} \le 1. \tag{3.8}$$

R_{SPI} is achievable when the processor-coder is both lossless and has an output arbitrarily close to the signal-processor's output (MSE can be used as a measure). The sensor-consciousness then follows from

$$\begin{aligned} 0 \le \boldsymbol{\mathcal{F}} &= \max_{\{C[y_i]\}} \mathscr{L}(\mathbf{y},\mathbf{z}) = \max_{\{C[y_i]\}} \frac{L(\mathbf{y},\mathbf{z})}{\mathcal{K}(\mathbf{y})} \\ &= \max_{\{C[y_i]\}} \frac{\mathcal{K}(\mathbf{y}) - \mathcal{K}(\mathbf{y}/\mathbf{z}(t_i+W)=\mathbf{y}(t_i+T))}{\mathcal{K}(\mathbf{y})} = \frac{\boldsymbol{\mathcal{K}}(\mathbf{y}) - \boldsymbol{\mathcal{K}}(\mathbf{y}/\mathbf{z}(t_i+W)=\mathbf{y}(t_i+T))}{\boldsymbol{\mathcal{K}}(\mathbf{y})} = \frac{R_{PC}^{rec}}{\boldsymbol{\mathcal{K}}(\mathbf{y})} = \frac{R_{PC}^{rec}}{R_{SC}^{Min}} \le 1 \end{aligned} \tag{3.9}$$

where: a) $\{C[y_i]\}$ is the signal-processor redesign constraint; b) t_i is the original signal-processor starting-time; c) $\mathcal{K}$(**y**) is **y**'s ectropy under the constraint $\{C[y_i]\}$; d) $\mathcal{K}(\mathbf{y}/\mathbf{z}(t_i+W)=\mathbf{y}(t_i+T))=\mathcal{K}(\mathbf{y})- R_{PC}^{rec}$ is the decreased $\mathcal{K}$(**y**) after **y** is time-dislocated from t_i+T to t_i+W and sensed. $\mathbf{z}(t_i+W)$ is the sensor output at time t_i+W; e) $L(\mathbf{y},\mathbf{z})=\mathcal{K}(\mathbf{y})-\mathcal{K}(\mathbf{y}/\mathbf{z}(t_i+W)=\mathbf{y}(t_i+T))= R_{PC}^{rec}$ is **y**'s recognized ectropy after **y** is time-dislocated and sensed. This quantity is a mutual-latency because a time-dislocation reversal yields the same recognized ectropy, i.e., $L(\mathbf{y},\mathbf{z})=L(\mathbf{z},\mathbf{y})$. Note that $L(\mathbf{z},\mathbf{y})=\mathcal{K}(\mathbf{z})-\mathcal{K}(\mathbf{z}/\mathbf{y}(t_i+T)=\mathbf{z}(t_i+W))= R_{PC}^{rec}$ since $\mathcal{K}(\mathbf{z})= R_{PC}^{rec}$ is **z**'s ectropy and $\mathcal{K}(\mathbf{z}/\mathbf{y}(t_i+T)=\mathbf{z}(t_i+W))=0$ is **z**'s remaining ectropy after **z** is time-dislocated to t_i+T and sensed; f) $\mathscr{L}(\mathbf{y},\mathbf{z})=L(\mathbf{y},\mathbf{z})/\mathcal{K}(\mathbf{y})$ is the mutual-latency ratio; g) $\boldsymbol{\mathcal{K}}$(**y**) is the smallest among all possible $\mathcal{K}$(**y**) cases. For instance, $\boldsymbol{\mathcal{K}}$(**y**)=min(6,3)=3-bors/**y** for the example of Fig. 4 since $\mathcal{K}$(**y**)=6-bors/**y** when $\{C[y_i]\}$ is limited to 2-

input NAND-gates (Fig. 4a) and $\mathcal{K}(\mathbf{y})$=3-bors/$\mathbf{y}$ when limited to 2 or more input NAND-gates (Fig. 4b); and h) $R_{SC}^{Min}=\boldsymbol{K}(\mathbf{y})$ is the minimum R_{SC} yielding the maximum achievable R_{SPI}.

The sensor-consciousness $\boldsymbol{F}$ guides the design of either lossless or lossy SPI-coders. A lossless SPI-coder is characterized by an achievable R_{SPI}, i.e.,

$$0 \leq R_{SPI} \leq \boldsymbol{F} \tag{3.10}$$

and is ideal when $R_{SPI} = \boldsymbol{F}$. A lossy SPI-coder has a R_{SPI} that is not achievable, i.e.,

$$\boldsymbol{F} < R_{SPI} \leq 1 \tag{3.11}$$

The previously described sensor-consciousness performance-bound viewpoint is named sensor-coding and guides the design of recognition-systems. Thus, sensor-coding is the same as "the mathematical theory of recognition" just like channel-coding is the same as "the mathematical theory of communication" [21]. Moreover while information is the central-theme of communication, latency is of recognition.

A Note on the Selection of the Terms Latency, Ectropy and Consciousness. 'Latency' in latency theory was selected as the time dual of the term 'information' in information theory because it describes, for both artificial and living systems, the time interval "when something is initiated and the moment one of its effects begins" [24], which is the fundamental 'time interval issue' addressed by latency theory. On the other hand, 'ectropy' in processor-ectropy has been selected as time dual of 'entropy' in source-entropy because while entropy refers 'to tropy [25]' (or to evaluate) the 'en' (or inner) space of a source by finding its sourced-space information, ectropy refers 'to tropy' (or to evaluate) the 'ec' (or outer) time interval of a processor by finding its processing-time latency. Finally, the term 'consciousness' in sensor consciousness has been selected as the time dual of 'capacity' in channel capacity because it refers to the waiting time interval over which the sensor is conscious.

4. CONCLUSIONS

In this paper, the time dual of information theory has been advanced and named latency theory. Since latency theory is one of two pillars of a space-time duality that is firmly anchored on information theory [17] it promises to have general applicability in all science, technology, humanities and social sciences problems that require the replacement of complex signal-processors with simpler and faster processors that are better matched to processor inputs that are highly compressed in a lossy fashion due to severely taxing memory constraints. In the second of this multi-paper series [6] the latency theory philosophy is successfully applied to DARPA's KASSPER program. The success of our latency theory in this real-world problem should also brings to light a better understanding of how biological systems achieve outstanding decisions while apparently only using highly compressed knowledge. This is the case, for instance, when a human after a new viewing expertly detects a human face seen only once before even though that face cannot be accurately described prior to such new viewing [23]. In the third paper of this multi-paper series [22], the recognition system of latency theory (Fig. [3]) is found to inherently lead to the discovery of a seminal space-time duality for physics. This space-time duality surfaces when it is noted that the design of a write/read device for prior knowledge retention (a recognition system problem) is the 'space dual' of the design of a transmitter/receiver device for knowledge motion (a communication system problem). Thus it is concluded that laws of motion in physics used in the engineering of a transmitter/receiver motion device must have a space dual that can be used in the engineering of a write/read retention device. It is further hoped that this advanced latency-information theory (LIT) perspective for physics should be useful in a unified guidance of both *engineering science and physics* problems. Such an approach may conceivably address in a rather straight forward manner significant theoretical questions in physics such as the development of a satisfactory quantum gravity theory as well as the production of more reliable predictions regarding future technological advancements. Finally it is hoped that LIT can serve in the near future as a significant pedagogical tool for a superior understanding and guidance of the design and implementation of complex systems.

REFERENCES

1. *Proceedings of DARPA Kassper,* 2002-2005
2. Guerci, J.R. and Baranoski, E, " Knowledge-Aided Adaptive Radar at DARPA", *IEEE Signal Processing Magazine*, vol. 23, no. 1, pp. 41-50, January 2006
3. Feria, E.H., Agaian, S**.,** *Accelerated Predictive-Transform,* USA Patent, May 2006
4. Taubman, D. S. and Marcellin, M., *JPEG2000: Image Compression Fundamentals, Standards and Practice*, Kluwer Academic Publishers, MA, 2002
5. Feria, E. H. and Licul, D., "A Bit Planes Predictive-Transform Source Coder Illustrated with SAR Imagery", *Proceedings of SPIE Defense and Security Symposium,* 17-21 April 2006
6. Feria, E.H., "Latency-Information Theory and Applications. Part II: On Real World Knowledge Aided Radar", *Proceedings of SPIE Defense and Security Symposium,* 17-21, March 2008
7. ______, "A Predictive-Transform Compression Architecture and Methodology for KASSPER"*,* Tech. Rep. *DARPA Grant FA8750-04-1-0047*, May 2006
8. ______."Latency-Information Theory", *PSC-CUNY Research Awards*, PSCREG-37-913, 38-247, 39-834, 2006-2009
9. ______, "Compression-Designs for Knowledge-Aided AMTI Radar Systems", *Proceedings of SPIE Defense and Security Symposium*, 17-21 April 2006
10. ______, "The Mathematical Theory of Signal Processing and Compression-Designs", *Proceedings of SPIE Defense and Security Symposium,* 17-21 April 2006
11. ______, **"**The Mathematical Theory of Signal Processing and Compression-Designs", *Dean's Sponsored Talk*, College of Staten Island, Wed., May 10th, 2006
12. ______, **"**Latency-Information Theory Illustrated With Real-World Knowledge-Aided Radar", *Dean's Sponsored Talk*, College of Staten Island, Wed., Sept. 27th, 2006
13. ______, "Compression-Designs in Artificial and Living Systems", *Proceedings of Annual IEEE System, Man & Cybernetics Conference*, Oct. 8-11, Taipei, Taiwan, 2006
14. ______, "On a Nascent Latency-Information Theory", *Abstract in Proceedings of SPIE,* April 2007
15. ______, "Universal Range Bin Centroids from SAR Imagery for Knowledge-Aided Radar*", Abstract in Proceedings of SPIE*, April 2007
16. ______, **"**A Predictive-Transform Sidelobe Canceller for Adaptive Radar Systems*", Abstract in Proceedings of SPIE*, April 2007
17. Lucky, R. "Reflections on the Rise of Information Technology", *Coordinated Science Laboratory (CSL),* Urbana, Illinois, October, 2001
18. *The Communications Handbook*, IEEE Press, 1997
19. Peebles, D.Z., *Probability, Random Variables and Random Processors*, McGraw-Hill, 1991
20. Haykin, S., *Communication Systems*, Wiley, New York, ed. 3, 1994
21. Shannon, C.E., W. Weaver, *The Mathematical Theory of Communication*, University of Illinois Press, Chicago, 1949
22. Feria, E.H., "Latency-Information Theory and Applications. Part III: On the Discovery of the Space Dual for the Laws of Motion in Physics ", *Proceedings of SPIE Defense and Security Symposium,* 17-21, Mar. 2008
23. *Handbook of Face Recognition*, Springer, 2005
24. *Compact Edition of the Oxford English Dictionary* (Oxford, 1991)
25. *Greek Roots and Their Modern English Spelling,* (Leonardo Press, 1989)
26. Prof. Alfred M. Levine is gratefully acknowledged for reviewing earlier drafts of this paper.

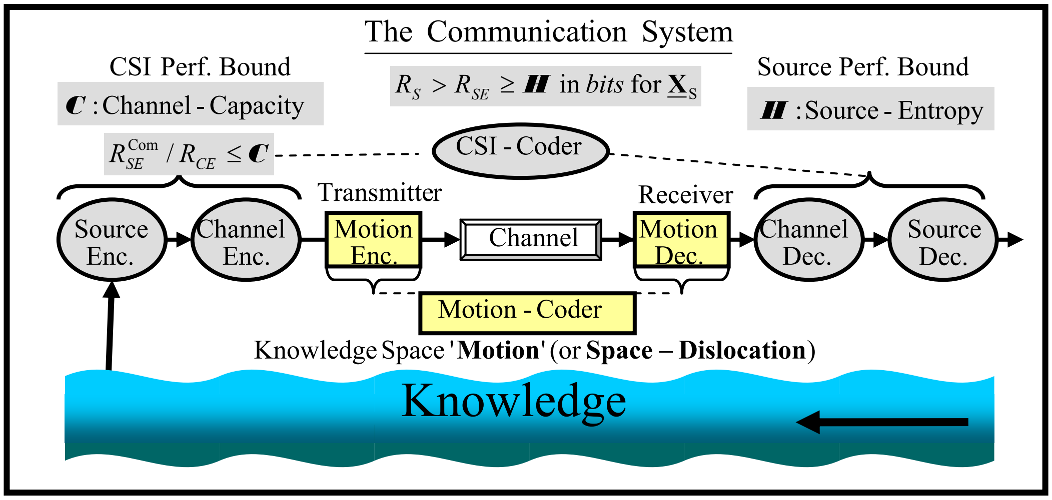

Fig. 1 Information Theory's Communication System

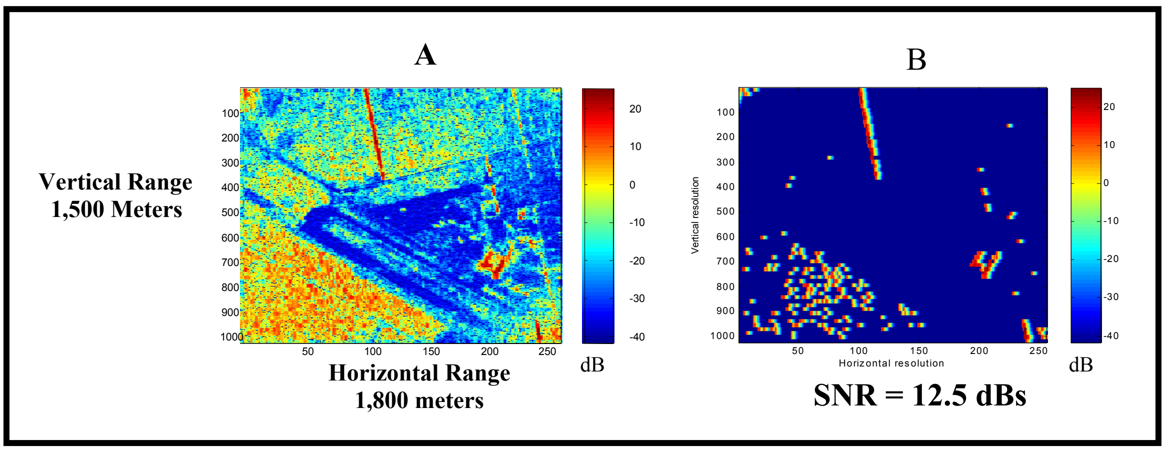

Fig. 2 A) 4 Megabytes SAR Image of Mojave Airport in California
B) 512 Bytes SAR Image Encoded/Decoded With a MMSE PT Source Coder

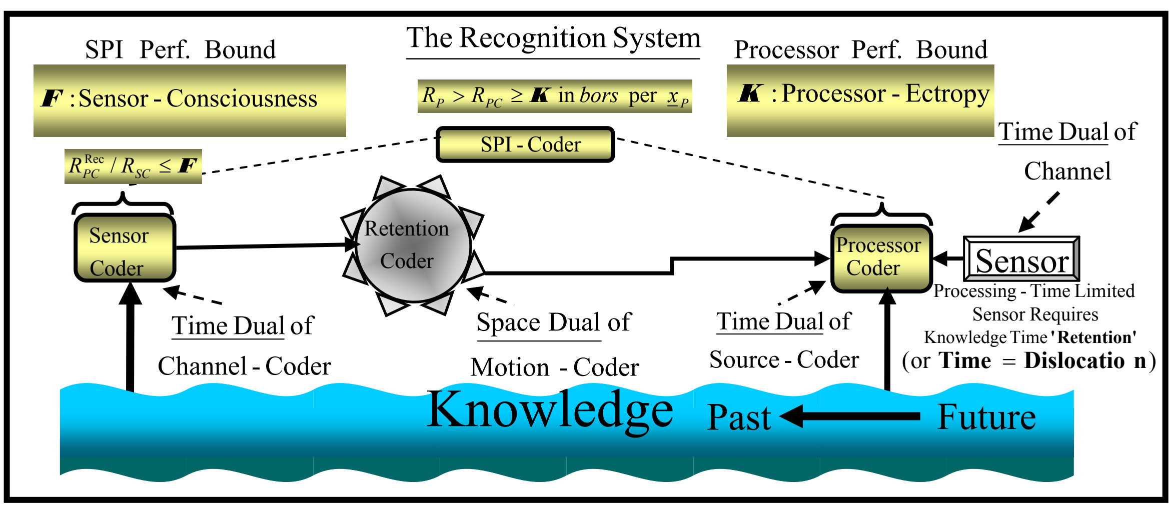

Fig. 3 Latency Theory's Recognition System

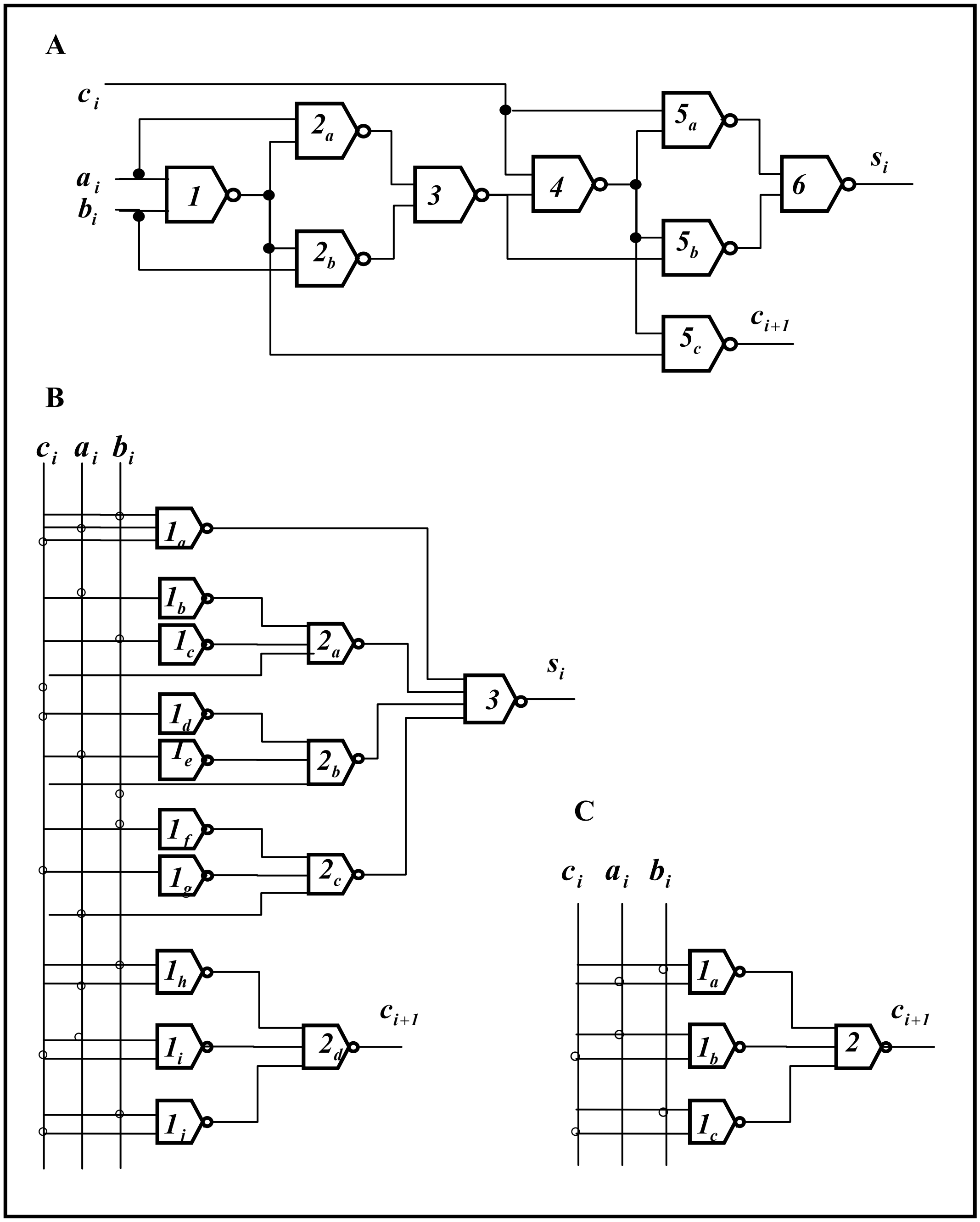

Fig. 4 A) Original Signal Processor; B) Lossless Processor Coder C) Lossy Processor Coder