A New Paper/Computer Interface: Two-Dimensional Symbologies

Theo Pavlidis
Symbol Technologies
Holtsville, NY 11742
USA
t.pavlidis@ieee.org

Abstract

Two-dimensional symbologies have been around for a while but they did not reach the broad public until the last few years. Examples include the "bull's eye" code that appears in all United Parcel Service packages and the two-dimensional bar code PDF417 that appears in driver's licenses in several states, metered postage, and identity documents. The paper covers the basic principles of the design of such codes, problems associated with their reading, and examples of applications. It will also review how such symbologies related to OCR-based document reading and how the technologies can supplement each other.

1. Introduction

The proliferation of networks that connect one computer to another has decreased the need for scanning documents but it has not eliminated it, and it never will. There is always the need to interface between computers and physical objects. A major class of applications is found in the shipping industry. A package must contain both a human readable and a computer readable label. The latter need not be printed; it may be a Radio Frequency Identification (RFID) Tag. RFID tags have many advantages over printed labels but they also suffer from a major disadvantage: much higher cost! Even if the cost becomes as low as 50 cents per tag, it is not practical to use them to tag low cost products or label items when the shipping cost is under, say, $5. (Today RFID tags are used widely for tracking such big items as motor vehicles, especially in situations where scanning is not practical.) In contrast, the cost of adding a machine-readable code on a printed label is miniscule (less than a cent). Therefore the need for machine-readable printed labels is going to persist.

Early labeling systems were based on Optical Character Recognition using special fonts, such as OCR A. Such systems were replaced by bar codes in the early 1970's, mainly because it is much simpler to build a bar code reader than an OCR reader. (We will return to this point later in the paper.) While bar codes offer a solution to numerous applications, they suffer from a major limitation. They can only encode a limited number of digits and, therefore, they can function only as indices to a database. This limitation (and the relatively high cost of RFID) led to the development of two-dimensional codes where a label can hold a complete record of data and still be easily machine-readable.

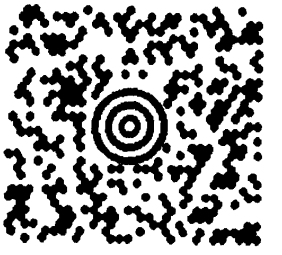

Figure 1: United Parcel Service MaxiCode

While everybody has seen barcodes, fewer people are familiar with the two-dimensional codes, so I provide two examples. Figure 1 shows the United Parcel Service (UPS) **MaxiCode** that is used for tracking packages. Such an image can be seen on the packages received at home as a result of Internet shopping.

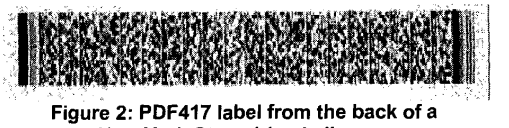

Figure 2: PDF417 label from the back of a New York State driver's license

Figure 2 shows the **PDF417** stacked bar code from the back of my driver's license. The code was developed at Symbol Technologies and, in contrast to the UPS MaxiCode which is limited to package tracking, PDF417 has been used in a range of applications, most recently in the "e-stamp."

Given the advances in OCR, one may ask why we persist in using special codes, thus requiring two labels, one machine-readable and the other human-

0-7695-0750-6/00 $10.00 © 2000 IEEE

Authorized licensed use limited to: Sterne Kessler Goldstein Fox. Downloaded on April 27,2026 at 14:50:35 UTC from IEEE Xplore. Restrictions apply.

Apple EX1015
IPR2026-00316
U.S. Patent No. 11,664,123

readable. The reason is that specially designed symbologies continue to provide better performance than OCR. The deeper reason for that is that human readers rely heavily on *high level* context, something that it is difficult to implement in a machine. My favorite example illustrates the limitations of spelling checkers (low level context). Quite often the word date is read incorrectly as dale. Since the latter is found in English dictionaries, no correction is made. However, if the document is a business letter the word date is far more likely than dale, while the opposite might be true in a poem. Bar codes and two-dimensional symbologies offer a substitute for high-level context by having a pre-defined structure that makes erroneous reading unlikely. As a result the algorithms needed to decode bar codes and two-dimensional symbologies for a given level of performance are much simpler than OCR algorithms. In particular linear bar code readers can be implemented in 8 bit micro-processors that cost about $2 a piece in large quantities.

In the rest of the paper I review the encoding methodologies both for ordinary bar codes and for some of the two-dimensional codes and then I discuss their current and potential applications.

2. Linear Bar Codes

Linear bar codes encode information through the relative widths of bars and spaces [1]. Each code word consists of a fixed number of elements and the ratio of the widths to the narrowest element is computed during scanning. In order to provide noise resilience additional constraints are imposed. In the so-called *delta codes* the sum of the normalized widths is fixed to an integer value. In the so-called *binary codes* only two width values are allowed and the number of narrow and wide elements is fixed.

The use of length ratios ensures that bar codes are *self-clocking*, i.e. they can be decoded without knowledge of their absolute size. Their large height also allows scanning with significant tilt. The price for such advantages is low information density. For a given square area that can be divided reliably in n^2 cells we have storage capacity of n^2 bits. If the same area is used for a bar code the number of bits will be only a linear function of n. It can be shown [1] that the best one can do is:

$$B(n) = \log_2 [C(n-1, (n-1)/2)]$$

where $C(n,m)$ is the number of choices of m items out of n. This number can be significantly less than n. For example, for n equal to 7, the value is $\log_2 (20)$ or 4.32 (rather than 49 if the area was fully utilized).

Bar code decoding is trivial when they are of large size, they are printed carefully, and a high quality scanner (appropriate for OCR) is used. A simple decoder can be written in a day or two. The problem becomes far more difficult under realistic conditions. In many cases space on a package is at a premium and bar codes are printed with very small dimensions, as low as 5 mils for the narrowest element. Inexpensive printers produce poor quality symbols that may suffer further deterioration because of handling. Finally, cost pressures require the use of relatively low-resolution scanners. All these factors make bar code decoding a challenging problem. See [2] for more on modern techniques of bar code decoding. Additional information about bar code technology can be found in [3].

3. PDF417 Symbology

The low information density of linear bar codes led to the design of stacked bar codes with large number of bars and spaces per code word. Not only the vertical dimension would be utilized but the value of $B(n)$ would also be brought closer to n. The main challenge in the design of a stacked bar code is row separation, especially when a hand-held laser scanner scans the label (symbol).

PDF417 (Figures 2 and 3) is a stacked delta bar code designed to provide significantly higher density than liner bar codes while maintaining most of their advantages. Each code word has 4 bars and 4 spaces and the sum of the normalized width is constrained to 17 [4-8].

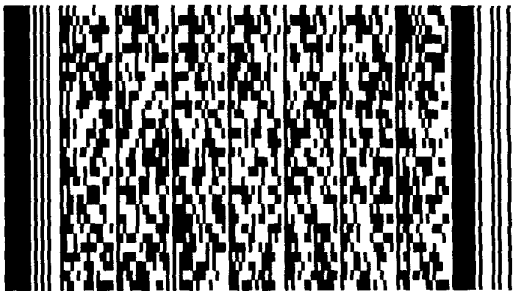

Figure 3: PDF 417 label magnified to show the detailed structure of the symbology

The solution to the row separation problem adopted by PDF417 is to use three different encoding schemes for each three adjacent rows. In particular, a *cluster number* is computed for each code word as following. Let x_i be the width of an element (x_1 is the first bar, x_2 the first space, etc). Then the cluster number is defined as

$$C(\underline{x}) = (x_1 - x_3 + x_5 - x_7)_{\ mod 9}$$

The symbology uses only code words with cluster numbers 0, 3, and 6. The words of the first row have

Authorized licensed use limited to: Sterne Kessler Goldstein Fox. Downloaded on April 27,2026 at 14:50:35 UTC from IEEE Xplore. Restrictions apply.

cluster number 0, those of the second 3, of the third 6, the fourth 0, and so forth. This provides both vertical self-clocking and error detection. As a result bar rows need be only three times as tall as the narrowest element.

In order to form a Galois field (and thus use a Reed Solomon error correction code) PDF417 selects 929 patterns from each cluster and the basic (low-level) encoding associates only a number in the 0 to 928 range with each code word. This provides considerable flexibility in the use of the PDF417 in applications. Shift words are used to switch between several modes, some of them predefined, others available for customization.

Each row contains a start and a stop word consisting of a wide black interval, followed by several narrow ones. Because these are the same in all rows the label is bounded by a black column on the left as shown in Figures 2 and 3.

The maximum capacity of a PDF417 label is 1108 bytes, which can be expressed as 1850 ASCII characters or 2710 digits. The text/numerical capacity can be higher if data compression is used, but 1850 ASCII characters correspond to about 500 English words, which is a substantial part of a page. The size of such a PDF417 label depends on the resolution the decoder is capable of. If the minimum length that can be discerned by decoder is 10 mil, then one can store 250 characters per square inch. If the minimum length is 6.6 mil, then the capacity is 500 characters per square inch. Therefore the contents of a page of text (about 2500 characters) can be stored in 3 square inches, a small fraction of the page size.

Because the symbology can encode binary data it can be used to encode biometrics, including speech samples or the photograph of the face of a person. (The badges of employees of Symbol Technologies include a PDF417 label that contains the picture of the person. Temporary badges for visitors contain the picture of the building.)

The use of error correction allows the reading of labels with large parts missing as shown in Figure 4.

Figure 4: A damaged PDF417 label that can be read because of the Reed-Solomon error correcting code used by the symbology.

For a theoretical analysis of the code see [9]. Additional details about the code can also be found in [4] and in Appendix H of [3].

4. MaxiCode Symbology

MaxiCode [10, 11] was designed specifically for the United Parcel Service to allow conveyor belt reading at speeds of up to 150m per minute. It is based on a hexagonal grid, whose center is occupied by a bull's eye of three concentric dark circles (Figure 1). The bull's eye pattern has a distinct signature under an FFT, therefore it can be used both to locate the label reliably in a scanned image and for self-clocking.

Each label (symbol) has 33 rows with width alternating between 30 and 29 (because of the hexagonal structure). Six hexagonal cells in a row form a code word. Because of the space occupied by the bull's eye, there are only 144 code words. Each cell can be black or white, thus each code word has 6 bits. In order to encode all the 256 eight-bit patterns MaxiCode uses five character sets. Character set A encodes the upper case letters, the ten digits, and certain punctuation signs. Character set B encodes the lower case letters and the remaining punctuation signs, etc. Each set contains *shift* characters used to switch between sets.

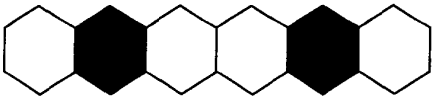

Figure 5: Example of a MaxiCode word

Figure 5 shows the image of a MaxiCode word (of course, the thin lines are not printed on the label). In this case the bit pattern is 010010, which corresponds to **R** in character set A and to **r** in character set B.

The specifications of the code insist on a fixed size label (about one square inch) as well as hexagon size. A Reed-Solomon error correction scheme is used at two selectable levels of security. Because of the fixed size and the error correction overhead the maximum number of data words per label is 93, or 558 bits.

Additional details about the code can be found in Appendix K of [3].

5. DataMatrix Symbology

A DataMatrix [12] label has a checkerboard type structure with each square denoting a 0 or 1 depending on its color. The frame of the square label as shown in the two examples of Figure 6 imposes structure. The figure on the right is a copy of an actual DataMatrix label, while the figure on the left is from a patent showing in detail the structure of the symbology. Two sides (marked by 12 in the left figure) are solid black and the other two contain alternating black (16) and

Authorized licensed use limited to: Sterne Kessler Goldstein Fox. Downloaded on April 27,2026 at 14:50:35 UTC from IEEE Xplore. Restrictions apply.

white (18) cells. The solid sides are used for localization and the other two for self-clocking.

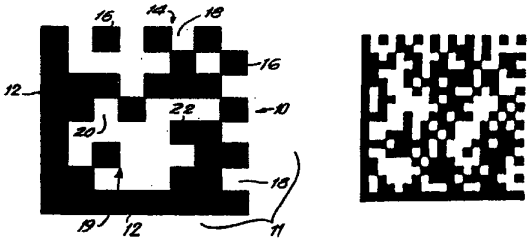

Figure 6: Left: Specification of the DataMatrix symbology as shown in one of the relevant patents [13]. Right: Example of an actual label.

The actual size of the DataMatrix label is quite small and is used in such applications as the marking of chips as well as other industrial parts. Corporate users include Alcatel and Mercedes-Benz. A convolution code is used for error correction. Usually it contains no more than 60 characters per label.

Additional details about the code can be found in Appendix J of [3].

6. Why Reading Symbologies is easier than Reading Characters

Reading a printed character requires the analysis of a two-dimensional shape. Before that process can be carried out the system must also decide whether a pixel belongs to foreground (usually black) or background (usually white). The latter is a non-trivial problem because of the point-spread function of any scanning system. Narrow areas appear as gray because the sensors include the color of their neighboring areas.

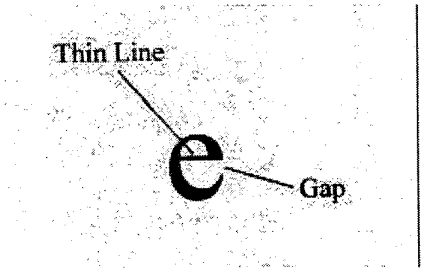

Figure 7: Examples of parts of a character that are difficult to binarize.

For example, in a lower case "e" (as shown in Figure 7) both the thin crossbar and the gap on the right may have the same gray level; therefore there is no threshold value that will separate them. In the worst case this may cause the "e" to be classified as "o," another common vowel that is also likely to satisfy any grammatical constraints imposed.

There is a large literature on threshold selection that illustrates the difficulty of the binarization problem. In contrast, binarization of two-dimensional symbologies is much easier. An overview of the symbol can provide some information about the scale. This can be obtained from the start and end bars in a PDF417 label, the bull's eye in a MaxiCode label, or the frame in a DataMatrix label. A gray area that is smaller than the minimum interval expected by the scale can be safely assigned the color of the surrounding areas. A larger gray area may be assigned the opposite color of the surrounding black or white areas. It also possible to map a gray area into a combination of black and white areas on the basis of a deeper analysis of the surrounding areas [15]. In linear bar codes binarization may even be avoided altogether [2, 14-17]. Not only binarization is easier for symbologies than for OCR, but also the next step in the recognition process does not require two-dimensional shape analysis. For PDF417 the start/stop bars and for DataMatrix the frame provides information about the orientation of the label, so that the direction of the rows is readily identifiable. For MaxiCode there are three possible edge directions only and from them it is possible to infer the direction of the rows as well. In either case the problem of reading the label is reduce to finding length ratios (for PDF417) or black and white cells (for MaxiCode and DataMatrix). Because of the redundancy imposed by the coding schemes used, any misreading will be identified immediately.

Therefore two-dimensional symbologies offer far more reliable means than OCR for interfacing printed material to the electronic world. It is probably true that label decoding is a solved problem and the main research issues are how the technology can be used in novel applications.

7. General Applications

The major area of applications is the labeling of objects in case where a database may not be available, or where there is a need to eliminate the possibility of errors because of look-up problems. Therefore the shipping industry is one of the larger users of two-dimensional symbologies.

While the United Parcel Service is committed to their own MaxiCode, most other shippers are selecting PDF417. Figure 8 shows the label used by a Brazilian carrier (Part of VASP Brazilian Airlines).

Authorized licensed use limited to: Sterne Kessler Goldstein Fox. Downloaded on April 27,2026 at 14:50:35 UTC from IEEE Xplore. Restrictions apply.

Figure 8: Shipping label from a Brazilian carrier illustrating the use of PDF417.

Besides the shipping industry, PDF417 is used in the Healthcare industry (for example, the labeling of blood samples with patient information when they are shipped from the collection site to the laboratory that performs the analysis). A growing domain of applications is in Government documents such as drivers license (New York, Arizona, North Carolina, Peru, Philippines, etc) and identity cards (Egypt, etc). A highly visible application was the use of PDF417 in the 1996 Olympics identity cards.

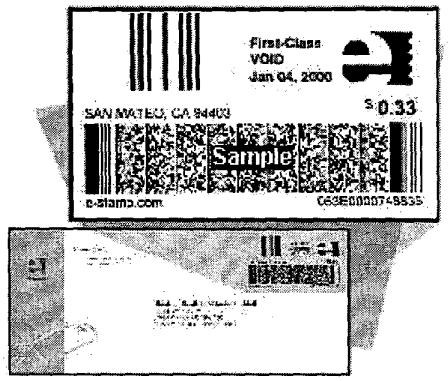

Figure 9: Example of electronic stamp based on the PDF-417 symbology.

Another type of applications involves cases where there is a need for secure encryption. The original text is transformed into binary data (for example, by the method of [18]) and these data are then printed using a two-dimensional symbology. This approach is used in the **e-stamp®** service. Figure 9 shows how "electronic stamps" will be used and has been copied from the web site of the service (www.e-stamp.com).

PDF417 has been accepted as a standard by China and it is already been used in banking operations. General Motors has also accepted it as a standard.

All two-dimensional symbologies can be read with a CCD scanner, but only a few can also be read by laser scanners. PDF417 was designed to be readable by laser scanners while MaxiCode and DataMatrix are readable only by CCD scanner. Because laser scanners use coherent light they have greater scanning range and are also less sensitive to ambient illumination conditions.

8. Document Applications

All digital communications use various error detection and/or error correction schemes and thus operate reliably even in noisy channels. For example, a "raw" recognition rate of, say 90%, can be converted to 100% through the error correction mechanism. Printed text does not contain explicit error recovery mechanisms; therefore the "raw" recognition rate is usually the final recognition rate. While linguistic context (spelling or grammar checking) can be used to perform some corrections there are limits to the approach because usually it cannot handle names or strings of numbers.

I believe that Document Analysis has been asked to meet unrealistically high performance standards, such as recognition rates well in excess of 99.5%. The problem is that even human readers have trouble achieving this rate without high level semantic context. A simple experiment can confirm this point. Just try to read a phone book in less than ideal conditions. You will have no trouble reading the names but not the numbers. You will probably be confusing "6" and "8" or "4" and "9" the way OCR systems do.

Lopresti and Sandberg [19] have suggested a method where a certificate is computed for a page of text and printed at one of the margins using a two-dimensional symbology. Because the latter can be read quite reliably, its contents can be used to correct OCR errors.

An alternative is the inclusion of a two-dimensional symbology label containing either a summary or the most important information of a larger document. Tax forms in Spain and custom forms in Mexico contain PDF417 labels.

Attempts to print a page with PDF417 labels containing the full text have not been successful because it is difficult to compress the text so that the PDF417 is no larger than a footnote.

Authorized licensed use limited to: Sterne Kessler Goldstein Fox. Downloaded on April 27,2026 at 14:50:35 UTC from IEEE Xplore. Restrictions apply.

I believe that paper input to computers will be greatly facilitated by the inclusion of bar codes or related symbologies to help the conversion. Note that an additional benefit that such labels may provide is protection against unauthorized altering of the document.

Another potential application area is transmittal of graphical information. While electronic mail is used routinely for text, the absence of universal standards for drawings encourages the use of FAX transmission. If the FAX image is going to be converted back to some graphics format, a two-dimensional symbology label could contain information useful for the conversion.

9. Conclusions

A challenge facing pattern recognition researchers is the blending of the new technology with the old. How can we use bar codes or two-dimensional symbologies to facilitate image recognition tasks? The use of special labels to encapsulate information related to a larger document is only one such example.

I have mentioned only three symbologies, but there are many others that are described in such sources as [3]. These include Code 49 (the first stacked bar code), Code One (a matrix symbology), and several others. This paper is by no means a comprehensive survey of such codes and anyone interested in the topic should consult the sources listed for additional references. The web site of the Auto ID Manufacturers (AIM) provides considerable information for all symbologies that AIM has standards for. The URL of the site is `www.aimi.org/aimstore` and contains both brief online descriptions and means for ordering the (rather voluminous) standards for each symbology. Unless you plan to do serious development work, you do not need the complete standards, and the other sources cited in the paper should suffice.

Finally, I should point out that all symbologies discussed in this paper (as well as all that can be found in the AIM web page) are in the public domain. In other words anyone is free to use them without paying a fee and anyone is free to develop novel scanning and decoding algorithm, provided they do not infringe on earlier patents. Originally, many of these symbologies were proprietary but they met with resistance in the market place. Users did not want to be beholden to a single source supplier and, as a result, the symbologies were placed in the public domain. While the originators of symbologies still have an advantage when it comes to the development of decoders and scanners, that advantage is not absolute.

10. Acknowledgements

I want to thank the following people from Symbol Technologies: Alan Melling and Michael O'Haire for providing me with samples of symbols and examples of PDF417 applications and helpful discussions; Joseph Katz and Daniel McGlynn for helpful comments on the paper; and Katy Horn for help with locating the patents cited.

All product names mentioned (including names of symbologies) are trademarks or registered trademarks of their respective companies.

This paper was written as part of the author's consulting with Symbol Technologies. However the contents of the paper reflect only the author's personal views and not those of the corporation.

Author's main affiliation: Department of Computer Science, SUNY, Stony Brook, NY, 11794-4400.

11. References

[1] T. Pavlidis, J. Swartz, and Y. P. Wang "Fundamentals of Bar Code Information Theory," *IEEE Computer Magazine*, **23** (April 1990), pp. 74-86.

[2] S. J. Shellhammer, D. P. Goren, and T. Pavlidis "Novel Signal-Processing Techniques in Barcode Scanning," *IEEE Robotics and Automation Magazine,* **6** (March 1999), pp. 57-65.

[3] R. C. Palmer *The Bar Code Book,* third edition, Helmers Publishing, 1995.

[4] T. Pavlidis, J. Swartz, and Y. P. Wang "Information Encoding with Two-Dimensional Bar Codes," *IEEE Computer Magazine*, **25** (June 1992), pp. 18-28.

[5] T. Pavlidis, Y. P. Wang, and J. Swartz "High Density Two-Dimensional Bar Code Symbol," *U.S. Patent 5,304,786*, April 19, 1994.

[6] T. Pavlidis, Y. P. Wang, and J. Swartz "Systems Utilizing a High Density Two Dimensional Bar Code Symbol," *U.S. Patent 5,399,846*, March 21, 1995.

[7] T. Pavlidis, Y. P. Wang, and J. Swartz "High Density Two Dimensional Bar Code Symbology," *U.S. Patent 5,504,322*, April 2, 1996.

[8] T. Pavlidis, Y. P. Wang, and J. Swartz "Microfilm Reader for High Density Two-Dimensional Bar Code Symbology," *U.S. Patent 5,796,090*, August 18, 1998.

Authorized licensed use limited to: Sterne Kessler Goldstein Fox. Downloaded on April 27,2026 at 14:50:35 UTC from IEEE Xplore. Restrictions apply.

[9] Y. P. Wang "Spatial Information and Coding Theory," *Ph.D. Dissertation*, Dept. of Computer Science, SUNY, Stony Brook, Dec. 1989.

[10] D. G. Chandler, E. P. Batterman, and G. Shah "Hexagonal, Information Encoding Article, Process and System," *U.S. Patent 4,874,936*, October 17, 1989.

[11] D. G. Chandler, E. P. Batterman, and G. Shah "Polygonal, Information Encoding Article, Process and System," *U.S. Patent 4,896,029*, January 23, 1990.

[12] D.G. Priddy and R.S. Cymbalski "Dynamically Variable Machine Readable Binary Code and Method for Reading and Producing Thereof," *U.S. Patent 4,939,354*, July 3, 1990.

[13] J. He, *personal communication.*

[14] E. Joseph and T. Pavlidis "Deblurring of Bilevel Waveforms," *IEEE Trans. on Image processing,* **2** (April 1993), pp. 223-235.

[15] E. Joseph and T. Pavlidis "Bar Code Waveforms Recognition Using Peak Locations," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, **16** (June 1994), pp. 630-640.

[16] E. B. Joseph and T. Pavlidis "Analog Waveform Decoder utilizing Histogram of Edge Sizes," *U.S. Patent 5,311,001*, May 10, 1994.

[17] E. B. Joseph and T. Pavlidis "Analog Waveform Decoder using Peak Locations," *U.S. Patent 5,504,318*, Apr. 2, 1996.

[18] J. Pastor "CRYPTOPOST™: A Universal Information Based System for Automated Mail Processing," *Proceedings of U.S. Postal Service Advanced Technology Conference*, Vol. 1, Nov. 5-7, 1990, pp. 429-442.

[19] D.P. Lopresti and J. S. Sandberg "Certifiable Optical Character Recognition," *Proceedings of the Second ICDAR*, Tsukuba, Japan, October 20-22, 1993, pp. 432-435.

Authorized licensed use limited to: Sterne Kessler Goldstein Fox. Downloaded on April 27,2026 at 14:50:35 UTC from IEEE Xplore. Restrictions apply.