Journal of Theoretical and Applied Information Technology
20th October 2015. Vol.80. No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

DATA COMPRESSION METHODS

ALEKSANDR BORISOVICH VAVRENYUK, ARKADY PAVLOVICH KLARIN, VIKTOR VALENTINOVICH MAKAROV, VIKTOR ALEXANDROVICH SHURYGIN

National Research Nuclear University MEPhI
(Moscow Engineering Physics Institute)
Kashirskoe highway, 31, Moscow, 115409, Russia

ABSTRACT

Currently, there is a sharp worldwide increase in the volumes of the transmitted, stored, and processed information. Despite a certain share of superficiality, the term "information explosion" rather precisely describes the existing situation. According to the American research statistics, the volume of information created by the mankind by the year 2007 and stored on the artificial carriers amounted to 295 billion gigabytes (2.95×1020 bytes). These volumes determine the great importance of the "data compression" field of knowledge. The concept of "data compression" (DC) is very broad. Therefore, this article makes an attempt of the DC systematization (classification), which can facilitate the study of the problem as a whole as well as the choice of an appropriate method for solving a particular problem. The most known areas and specific examples of the DC application are concerned.

Keywords: *Data Compression, Application Areas, Information System, Physical Experiment, Collider, Space Researches, Efficient Data Presentation.*

1. INTRODUCTION

A characteristic feature of the post-industrial society is a constant growth of the transmitted and stored information in various fields of human activity. It is clear that this trend will continue as long as the modern civilization exists, and will only become stronger.

The sharp rise in the information volumes results in the appearance of problems concerning the overloading of the corresponding data transmission and storage systems. There are two methods for solving this problem:

The continuous improvement of technical characteristics of the data transmission and storage systems.

Efficient data presentation.

This article deals with the second method. It includes various techniques for transmission and storage of only the most significant data, i.e. of only that part of information, which is of the recipient's interest.

The fundamental difference of this trend from the first one is that such characteristics as the channel throughput, volume of storage devices, transmitter capacity, and computer speed are considered as specified. Within the framework of these characteristics, the data are required to be presented in such a way that the maximum possible unloading of the appropriate systems and reduction in the volume of the transmitted and stored data could be achieved.

This trend, that emerged in the early 1960s of the 20th century and known as the "data compression" (DC) (some sources also include such terms as the "information compression", "reduction of information redundancy", and "compression"), is based on the results of the information theory and in some cases takes into account human physiology [1], [2], [3].

Both trends feature simultaneous development and complementarity. The emergence of the new, high-performance hardware allows for implementation of the more sophisticated and effective methods of DC. On the other hand, no matter how good were the data storage and transmission devices, there is always a question of how to use them in the most efficient way. And this is where the DC is required.

Data compression is based on the entropy encoding, which in turn is based on averaging the frequency of occurrence of elements in the coded sequence. That is, only if the characters of a certain data sequence occur therein with different probability, then the data reduction can be carried out without loss of information. At the same time, the frequencies of the individual characters

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

occurrence are virtually identical after such a compression in the resulting sequence (entropy reaches its maximum value).

According to the converse of Shannon's coding theorem, the maximum degree of compression through coding without losses is limited to the source entropy. That is, in simple words, that the data cannot be compressed more than it is allowed by their entropy. At that, after any significant lossless data compression, the frequencies of character distribution in the resulting sequence tend to uniform distribution and, consequently, the degree of data recompression tends to zero.

It is not necessary to be familiar with the foundations of the information theory to verify the validity of this assertion. It suffices to run any data compression program, either zip or rar, and try to recompress the already compressed file. Either the degree of compression of such file will be negligibly low, or, most likely, the recompression of the file will cause increase in its size, as the data compression program will add to the existing volume of data some service information. The same will happen to other programs for data compression without loss of information, even to the most efficient ones.

Redundancy of information provides a wide field of activity in this area. Thus, the redundancy measurement of the natural languages (the ones we speak) demonstrates that almost 80 % of the information transmitted with the help of a language is redundant, i.e. excessive. The redundancy indices of various languages are very close. This figure roughly determines the theoretical limits of the text files compression. The availability of redundancy allows switching to a different coding system that would reduce it. Speaking on the transition to the codes that reduce the message size, the concept of compression codes is introduced [12]. Meanwhile, we differentiate between the concepts of lossless compression and lossy compression. It is obvious that when we deal with the information of a "phone number" type, the compression through the loss of a part of characters leads to no good.

It follows that when choosing a particular method of DC, it is required not only to be guided by the quantitative characteristics of the method, but also to take into account the specific features of its application area.

The purpose of this paper is building the classification of the data compression methods as a tool to navigate in the broad area of science and technology entitled "data compression". This tool allows selecting the most appropriate data compression method for each particular case (analogue or digital, with known statistics or not, with loss of a part of information or lossless, etc.).

There are various options of the DC classification in the literature, but the proposed classification differs from the previous works by more fullness, and most importantly – by addressability of each block from top to bottom and vice versa.

2. APPLICATION AREAS AND EXAMPLES OF DATA COMPRESSION USAGE (IMPLEMENTATION TECHNIQUES)

Currently, a number of science and technology areas are characterized by the extremely large volumes of the transmitted and stored data, rather than it is caused by the need for DC in these areas. Listed below are the main areas of application:

1. Space researches.
2. Physical experiments.
3. Computers.
4. Communication means.
5. Telemetric systems.
6. Internet.
7. Meteorology.
8. Libraries.
9. Archives.
10. Video surveillance systems.
11. Television.

Figure 1 shows the structure typical of the majority of modern information systems.

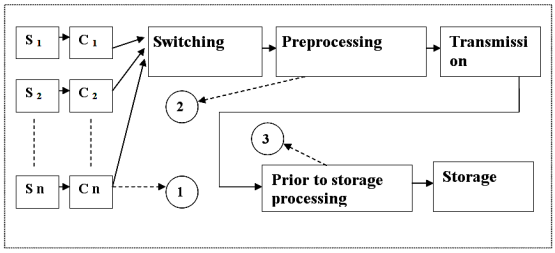

Figure 1. Typical structure of an information system

As a rule, there are several sources of information (S1.....Sn). These can be sensors, detectors, computer data, etc. Following the sources, there are always some primary converters (C1.....Cn), which convert signals from the sources to a certain unified presentation.

For example, an ADC in case of analogue sources and discrete processing. As there are

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

several sources, switching is required, i.e. polling of sources in accordance with the predetermined algorithm. This can be a cyclic polling, polling with priorities, etc. The preprocessing is usually understood as a set of data transformations before their entering into the transmission channel. This can be data screening, control, compression, conversion to the channel format, etc [4].

As a rule, processing before storage includes data control, processing, and presentation in the format that corresponds to the storage means. The blocks, in which the DC function can be implemented, are marked by the numbers 1, 2, and 3 in Figure 1.

Let us consider the Large Hadron Collider (LHC) as an example. The LHC is a ring accelerator with a circumference of 27 km located underground, in a monolithic rock not far from Geneva [5].

It should be noted that NRNU MEPhI is in close cooperation with the European Organization for Nuclear Research (CERN) in relation to the Collider subject (LHC sensors manufacturing, software development, data analysis, etc.). The MEPhI's teachers, postgraduates, and students are constantly working in the CERN, and the results of their cooperation are implemented in the MEPhI's training process.

This example has been chosen because it demonstrates what volumes of data the modern scientists have to deal with. It is obvious that this experiment is of great importance both for science and for education.

The LHC allows to operate with the smallest distances (down to a nanonanometer or 10–18 m) and the energies of the teraelectronvolt order (1 TeV = 1012 eV) [5].

At such energies, it is theoretically possible to fix the elusive Higgs particles that are supposed to be responsible for the existence of mass at other particles as well as of the particles forming the dark matter, which constitutes the major part of the matter in the cosmos.

Every two seconds, millions of data channels generate the total data volume of one petabyte order (one billion megabytes) from the sensors (blocks S1....Sn in Fig. 1) after the primary conversion (blocks C1.....Cn in Fig. 1). The raw data arrive at the processing system of the first level. This level is supported by hundreds of specialized computer boards with the circuit implementation of logic. At this stage, 100 thousand blocks of data per second (perspective events) are selected (preprocessing, block 2 in Fig. 1) for a more detailed analysis at the next stage of a higher level. In fact, the irreversible data compression is implemented since the total volume of information provided by the sensors is in the incommensurable excess of the processing capabilities.

A data processing system should reduce their flow to the controlled quantity, and has several levels. A starting system of higher level (block 3 in Fig. 1) transmits about 100 events per second to the concentrator of computing resources of the global LHC network – the distributed data network GRID. The network brings together the powers of computing centres and makes them available to users, including education institutions, providing invaluable "live" physical information. At an average time between the data blocks selected by the starting system of the first level (10 ms), the network software has enough time to reconstruct each event. The software binds the particle tracks to the common reference points. It further designs events, that is generates arrays of energies, pulses, and particle trajectories. Then the archives are made on the magnetic medium. At each of the processing stages, the additional data compression is carried out with the possibility of further data recovery.

In the field of space technologies, the data flows from the space vehicles, regardless of their purpose, have one common trend for increase. This is connected with the natural desire of researchers to expand the range and the amount.

It should be taken into account that on the onboard systems of data acquisition, processing, and transmission, the additional requirements are imposed relating to the restrictions on weight, dimensions, power consumption of the aerial vehicles' systems, which is a natural obstacle for meeting the increasing needs of the experimenters.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 1. Currently operating space research vehicles

Table 1 shows vehicles that are currently operating in space. The data are taken from the official website of the Federal Space Agency, and constitute only a small part of the launched satellites or those being planned for launching.

No.	Space vehicle	Object under study	Year of launching
1	The "Mars Odyssey" (NASA, USA) ("Hand" device, RF)	Mars. Scientific measurements on the orbit around Mars.	2001
2	The "Mars – Express" (ESA, EC) (6 devices, RF)	Mars. Scientific measurements on the orbit around Mars.	2003
3	The "Venus Express" (ESA, EC) (6 devices, RF)	Venus. Scientific measurements on the orbit around Mars.	2005
4	The "Resource – KD" (RF) ("Pamela" device)	Antimatter in the solar cosmic rays. Fluxes of antiprotons and positrons.	2006
5	The "Lunar Reconnaissance Orbital Complex" (NASA, USA) ("LAND" system, RF)	Search for water ice in the polar regions of the Moon.	2009
6	The "Coronas – Photon"(RF)	Solar flares. Monitoring of "space weather".	2009
7	The "Ultraviolet World Space Observatory" (RF)	Space objects and processes occurring inside them.	2011

The implementation of DC on the international space stations allowed reducing the volume of the transmitted to the Earth data by dozens of times, which in turn allowed increasing the operation time of devices due to the on-board memory saving, and transmitting more useful information under the same channel's carrying capacities.

In some systems, the main purpose of the DC usage is to save the information carriers, the areas for storing these carriers, etc. In this case, the DC is carried out at the stage of the data processing prior to storage. This problem is especially acute in the archives of the long-term storage based on the mainframe computers.

In the NRNU MEPhI, the similar issues are faced by the physicists receiving huge flows of information from the space objects, which work is assisted by the MEPhI (items 4, 6 in Table 1). Moreover, the efficient data storage is carried out in the educational computer labs and corporate networks of the MEPhI.

3. THE WORK RESULTS – CLASSIFICATION OF THE DATA COMPRESSION METHODS

Currently, the term "data compression" brings together a large number of different methods for reducing the volume of the data being transmitted and stored. These methods differ in a number of characteristics. They can be focused on the analogue or digital sources, be adaptive or nonadaptive, operate under conditions of the known or unknown statistical characteristics of the source of messages, etc.

Therefore, it makes sense to classify the existing compression methods that on the basis of this classification and among the plurality of compression methods, will allow us to select those ones that correspond to the assigned tasks is each specific case. This classification was developed at the Computer Systems and Technologies Department of the MEPhI and differs from the previous ones in that it takes into account considerably more features of compression techniques, and at the same time, for each method it is possible to determine a set of these features, and vice versa, for a given set of features to choose the appropriate methods. The proposed classification is shown in Figure 2.

The following features are used:

- Recoverability (with certain accuracy) of the original signal.
- Source type (analogue or digital).
- A priori knowledge of the source statistics.
- Adaptability.
- Kind of impact on the original signal.

The classification is designed so that all its blocks are addressable. There are lists of the methods names at the lower levels of the classification.

Suppose that we are interested in the name of "Run-Length Encoding" (block B6). Using the classification we can find out (following down-up) that this method relates to the efficient coding, suggests the statistics of the source of messages as the known one, is used for encoding the discrete sources, and is a quasi-reversible transformation. Conversely, suppose that it is required to choose a quasi-reversible method for the analogue source compression, which is independent of the statistics knowledge and does not require making complex adaptive systems. Following up-down, the group of the "Fixed Dc Procedures" (block E5) is reached.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

4. DISCUSSION OF THE DATA COMPRESSION METHODS

All DC methods can be divided into two main groups: the irreversible methods – lossy compression (block B2) and the quasi-reversible methods (block F2). The irreversible methods include such techniques, after application of which all the source data cannot be recovered. As a result of such methods, an experimenter obtains the targeted characteristics of the studied process, for example, probabilities of certain events, spectra, etc. With that, the on-board devices transmit not the source data, but the final results of the calculations, which naturally sharply reduces the volume of the data being transmitted.

The lossy compression is mainly applied for the graphics, sound, and video [20], i.e. for those areas where due to the enormous file sizes, the degree of compression is of great importance, and where it is possible to sacrifice those details that are unnecessary for the perception of information by a person. Specific capabilities for the compression of information are featured during the process of video compression. In some cases, the greater part of the image is being transmitted from frame to frame without changes, which allows building compression algorithms based on the selective tracing of just a part of the image. In a particular case, the image of a speaker, that does not change his position, can be updated only within the area of face or even mouth, i.e. in that part where the quickest changes from frame to frame take place. As a rule, the lossy compression of graphics provides for the high degrees of compression, though is almost imperceptible by a person.

The lossy compression methods have a number of drawbacks [11], [18], [13], [14]. The first drawback lies in the fact that the lossy compression is applicable not for all cases of the graphic information analysis. For example, if as a result of the image compression, a mole size will be changed on the face (but with that the person's face will remain completely recognizable), this photo would be quite suitable for sending it to friends by e-mail. However, if the sent photo is a lung image required to the medical examination for the purpose of shadow shape analysis, then in this case, the distortions are unacceptable. Furthermore, in the case of computer methods for the graphic information analysis, the lossy coding results (invisible for the naked eye) can be visible for the computer analyzer. The second drawback is that the repeated lossy compression and decompression lead to the effect of errors accumulation.

However, the irreversible compression is basically inapplicable if the researchers themselves cannot clearly formulate the targeted parameters of the initial process and require the compulsory transmission of all source data. This is a typical case for the practice of processing the results of a physical experiment.

The lossless coding can be applied for compression of any information as it allows for the absolutely accurate recovery of data after their encoding and decoding. The lossless compression is based on the simple principle of converting data from one group of characters to another more compact one. The best known are the two lossless compression algorithm [19]. These are the Huffman coding and the LZW coding (by the initial letters of its founders' names: Lempel, Ziv, and Welch), which are the main approaches to the information compression. The principle of the Huffman coding is to reduce the number of bits used to represent the frequently occurring characters, and thus, to increase the number of bits used for the characters of rare occurrence. The LZW method encodes strings of characters analyzing the input flow for building an extended alphabet, which allows the use of additional characters for presenting the strings of ordinary characters. For example, if to use the 9-bit ASCII codes instead of the 8-bit ones, 256 additional characters are obtained. The work of a compressor reduces to the construction of a table consisting of strings and corresponding to them codes. The compression algorithm is reduced as follows: the program reads the next character and adds it to the string. If the string is already in the table, the program proceeds with reading, otherwise, this string is added to the string table. The more duplicate strings it finds, the greater will be the degree of data compression. Returning to the example with a telephone and having drawn a simplified analogy, it is possible to say that if compressing the record 233 34 44 through the LZW method, we will come to adding the new strings – 333 and 444, and by expressing them with the additional characters, we will manage to reduce the record's length.

The quasi-reversible DC includes all methods that allow for the recovery of original representation of the studied process with the given accuracy. They can be divided into two main groups – efficient coding of a discrete source (block C3) and efficient analogue to digital conversion of signals (block F3).

The second group includes the methods of discretization and quantization of a continuous

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

signal that take into account its natural redundancy. If the statistics is known, the most applicable procedures are the prediction, interpolation, transmission of the expansion coefficients instead of their original continuous process to any series (blocks D6, D7, and D8). If the statistics is unknown, the aperture algorithms (block E6) are usually applied in practice, since they do not require the knowledge of the source statistics, can be easily implemented in practice, and ensure high compression coefficients.

The DC methods accounting for the physical properties of a receiver have set themselves apart (blocks G5, G6, G7, and G9). They focus on the specific features of the human organs of senses. For example, it is possible to permit such distortions in the voice timbre that allow comprehending the message's meaning, and thereby reducing the volume of thee data transmitted. When transmitting images, the decrease in contrast, brightness, etc is usually used (Richardson, 2005).

The efficient coding is used for compressing the data from the discrete sources. In particular, it can be regarded as the second stage of the continuous source data compression after discretization and quantization of the original signal.

As a rule, in case of the known source statistics, various variants of the series length coding and address-positional coding (blocks B6, B7) are used, since these methods can be easily implemented and do not require the bulky correspondence tables. In those cases where there is a change in time of any value, the predictive coding (block B8) is used. In order to code the discrete sources with small alphabets, the best methods for statistical coding are used (Shannon-Fano and Huffman codes) (block B5), which allow to completely eliminate the redundancy at the known statistics.

The problem of the discrete source data compression becomes much more complicated if the statistical characteristics of source are unknown, partially known, or change in time (blocks C5-C9). The basic theoretical studies on the solution of this problem are presented in the works of the following authors [2], [3], [6].

For a long time, it was impossible to find any information on the practical application of the results of these studies in literature. This is mainly due to the following reasons:

- The high labor intensity of implementation (adaptive coding, universal coding – blocks C8, C9).
- The insufficient efficiency of the methods simple for implementation (for example, iterations of simple substitutions, differential coding – blocks C5, C6).
- The narrow classes of sources, on which the methods are focused. For example, the matrix coding (block C7) is used in case of the discrete data monotony.
- At the present time, in connection with the development of flash-technologies [7], [8] the methods of universal coding, which labor intensity previously seemed unthinkable, have become quite feasible in practice [9].

It should be noted that each block of the classification is exactly the DC method rather than a means or algorithm. As a rule, a DC means or algorithm is understood as a combination of several methods, for example of those used in the archivers [10], [16], [17], [20].

(See figure 2 in the appendix)

Figure 2. Classification of the data compression methods

For clarity, the data compression methods with indication of the blocks in the classification considered as well as the numbers of the previously listed areas of their application are summarized in Table 2.

Table 2. Data compression methods

Compression method	Block	Area No.
Improved processing	A4	1-11
Preliminary rejection of events	A5	1, 2, 3, 5
Spectral analysis	A6	2, 4, 5
Transmission of histograms, quantiles	A7	1, 2, 5, 7
Measurement of probabilities	A8	1, 2, 5
Optimal coding (the Shannon-Fano and Huffman codes)	B5	3, 6, 8, 9
Series length coding	B6	1, 2, 4, 5, 10, 11
Address-positional coding	B7	1, 2, 4, 5, 10, 11
Predictive coding	B8	4, 5, 7
Iterations of simple substitutions	C5	3, 4, 9
Differential coding	C6	4, 5
Matrix coding	C7	3
Adaptive coding	C8	1-6
Universal coding	C9	1-9

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Transformation into the Karhunen–Loève series	D6	4, 5
Optimal prediction	D7	1, 2, 5
Discretization by Kotelnikov	D8	1, 2, 4, 5
Aperture algorithms	E6	1, 2, 4, 5
Fourier transformation	E7	1, 2, 4, 5, 7
Differential methods	E8	2-5
Adaptive quantization	E9	2, 4, 5
Variable polling rate	E10	1, 2, 5
Accumulation of statistics	F6	2, 3, 6, 7
	F7	1, 4, 5
Interactive probability procedures	F10	3-6
Analysis and synthesis of speech	G6	1, 3, 4, 6, 11
DC at transmission of images and films	G7	1, 5, 6, 7
DC at transmission of a TV signal	G9	10, 11

4. CONCLUSION

The constant volume growth of the transmitted and stored information leads to the overloading of the storage devices and data transfer channels, which necessitates the use of data compression (DC). Currently, there are a lot of such methods of data compression. They differ by a number of characteristics. In this article, we have examined the approaches to the systematization and arrangement of these characteristics. The results of the work done are presented by the following materials:

- A generalized block diagram of an information system.
- Specific examples of the DC methods implementation.
- Detailed classification of the DC methods.
- The table of the DC methods and application areas conformity.

The use of these materials can help to facilitate the orientation within the wide field of science and technology under the name of "Data Compression".

In particular, the results of this work were used to choose the data compression methods in experiments in the field of gamma–ray astronomy ("Natalia – 1" device, MEPhI) and the “Landysh” e-capsule for medical research of the gastrointestinal tract (MEPhI). The need for the classification of the data compression methods is due to the fact that in each particular case the choice of the compression method is a very important and difficult task, which can be solved with the help of the approach, suggested by the authors.

Over time, with the replenishment of the list of DC application areas and new methods, the developed classification scheme will be expanded and detailed, but according to the authors, will not lose its value for the practical use.

REFERENCES:

[1] Fano, R. (1965). *Information Transmission. Statistical Theory of Communication* (pp. 438). Moscow: Mir. ISBN: 978-5-458-39451-2.

[2] Kolmogorov, A.N. (1965). Three Approaches to the Definition of "Information Amount". *Problems of Information Transmission, 1*, 3-11.

[3] Fitingof, B.M. (1967). Compression of Discrete Information. *Problems of Information Transmission, 3*, 28-36.

[4] Solovyev, V.F. (1970). *Efficient Coding at Transmission of Messages* (pp. 64). Moscow: Energy.

[5] Collins, G. (2008). Factory of Discoveries. *In the World of Science, 5*, 3-12.

[6] Klarin, A.P., & Shurygin, V.A. (1984). Study of the Universal Coding Efficiency Depending on the Block Length. *Problems of Information Transmission, 2*, 105-110. ISSN: 0555-2923 (print).

[7] Shurygin, V.A. (2011). *Flesh Memory* (pp. 116). Moscow: MEPhI.

[8] Vasilev, N.P., Makarov, V.V., & Shurygin, V.A. (2011). Flesh Memory – Formation, Development, Perspectives. *Automation in Industry, 3*, 55-59.

[9] Shurygin, V.A. et al. (2011). Method of Universal Coding of Binary Data. *Issues of Radio Electronics, 2*, 94-115.

[10] Vatolin, D, Ratushniak, A, Smirnov, M, & Yukin, V. (2002). *Data compression Methods. Design of Archivers, Compression of Images and Video* (pp. 384). Dialog-MEPhI. ISBN 5-86404-170-X.

[11] Salomon, D. (2006). *Compression of Data, Image, and Sound* (pp. 368). Moscow: Technosfera. ISBN 5-94836-027-X, 0-387-95260-8.

[12] Welstead, S. (2003). *Fractals and Wavelets for Image Compression in Action (with CD-ROM)* (pp. 320). Moscow: Triumph. ISBN: 5-89392-079-1.

[13] Kovalgin, Yu.A., & Vologdin, E.I. (2004). *Digital Coding of Audio Signals* (pp.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

240). Moscow: Korona-Print. ISBN: 978-5-7931-0290-2.

[14] Artiushenko, V.M., Shelukhin, O.I., & Afonin, M.Yu. (2003). *Digital Compression of Video Information and Sound* (pp. 428). Moscow: Dashkov & Co. ISBN 5-74978-258-7.

[15] Richardson, Y. (2005). *Video Coding. H.264 and MPEG-4 – the New Generation Standards* (pp. 368). Moscow: Technosfera. ISBN: 5-94836-041-5.

[16] Pearlman, W.A., & Said, A. (2011). *Digital Signal Compression: Principles and Practice* (pp. 41-76). Cambridge University Press. ISBN: 10- 0521805031, 0521899826.

[17] Miano, J. (2003). *Formats and Algorithms of Image Compression in Action* (pp. 236). Moscow: Triumf. ISBN 5-89392-078-3, 0-201-60443-4.

[18] Salomon, D. (2006). *Data Compression: The Complete Reference* (pp. 1118). Springer. ISBN: 1-84628-602-6 (978-1-84628-602-5).

[19] Sayood, K. (2012). *Introduction to Data Compression, 4th Edition* (pp. 768). Printbook. ISBN: 1849690308.

[20] Wayner, P. (1999). *Compression Algorithms for Real Programmers* (pp. 252). Morgan Kaufman. ISBN: 9780127887746.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

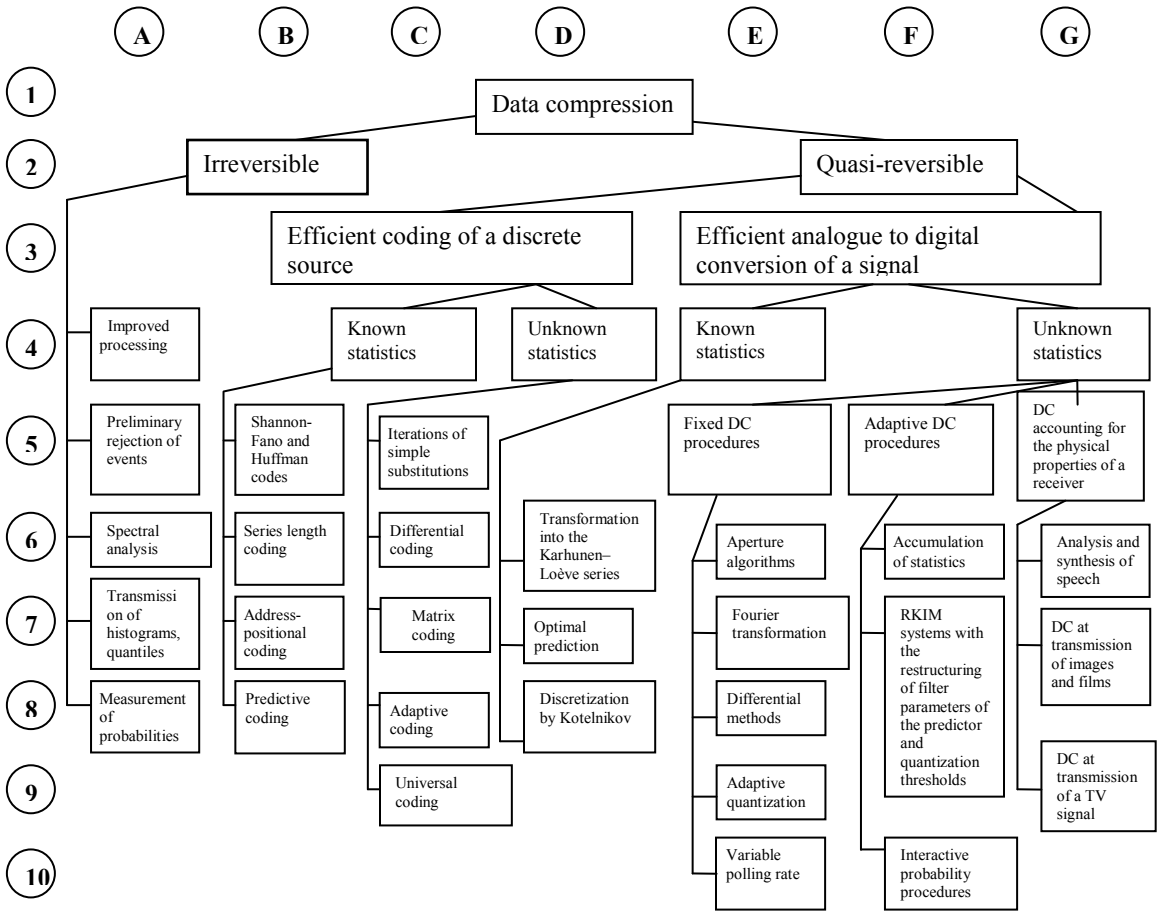

Figure 2. Classification Of The Data Compression Methods