International Journal of Computer Science Engineering & Technology
IJCSET
APPROVED BY
NATIONAL SCIENCE LIBRARY (NSL)
NATIONAL INSTITUTE OF SCIENCE COMMUNICATION AND INFORMATION RESOURCES (NISCAIR)
COUNCIL OF SCIENTIFIC AND INDUSTRIAL RESEARCH (CSIR)– NEW DELHI INDIA .
ISSN :2455-9091

IMAGE COMPRESSION PREPROCESSING USING A NOVEL DICTIONARY BASED LOSSLESS REVERSIBLE TRANSFORMATION ALGORITHM

[1] P. Jeyanthi,
[1] PG Department of computer science,
[1] Sree saraswathi thyagaraja college, pollachi.

ABSTRACT- Data compression is one of the empowering advances for each of these parts of the media transformation. It would not be functional to put images, not to mention sound and video, on sites in the event that it was not for Data compression algorithms. We proposed a novel word reference based lossless reversible transformation algorithm, Enhanced Transformation with Numerical Index. We have taken Length Index Preserving Transformation (LIPT) as our base work and killed some of its drawbacks. In this manner the overhead required in lexicon administering and upkeep of each of these word references is wiped out. In any case, this increase is acknowledged with the extra cost for looking the word in the lexicon. This mishap is additionally managed by actualizing word reference as table which is a swap for the double inquiry trees utilized by LIPT. In spite of the fact that the paired hunt tree used to look the words in the lexicon is space proficient, the inquiry expense is very high. Henceforth the pay in time and cost variables is accomplished with productivity by the table usage of the word reference. By the utilization of essential key, supreme disposal of duplication of words and swifter hunt of words is achieved.

Keywords: [Lossless reversible transformation algorithm, Image compression]

1. INTRODUCTION

At the point when the PC age came into being about in 1940's, storage room turned into a subject of issue. In the most recent decade we have been seeing a transformation of some call it an insurgency in the way we impart, and the procedure is still under way. This transformation incorporates the ever-present, perpetually developing Internet; the unstable advancement of versatile interchanges; and the always expanding significance of video correspondence. Information compression was the exit plan to that test. Also, therefore information compression has gotten to be universal inside the most recent decade. Compression is valuable since it decreases assets required to store and transmit information. Computational assets are devoured in the compression procedure and, as a rule, in the inversion of the procedure (decompression). Information compression is liable to a space–time intricacy exchange off. Case in point, a compression plan for video may require costly equipment for the video to be decompressed sufficiently quick to be seen as it is being decompressed, and the alternative to decompress the video in full before watching it might be badly designed or require extra stockpiling. The configuration of information compression plans includes exchange offs among different elements, including the level of compression, the measure of mutilation presented (when utilizing lossy information compression), and

the computational assets required to pack and decompress the information.

2. LOSSLESS

Lossless information compression algorithms typically misuse measurable excess to speak to information without losing any data, so that the procedure is reversible. Lossless compression is conceivable in light of the fact that most certifiable information shows measurable repetition. For instance, an image may have regions of shading that don't change more than a few pixels; rather than coding "red pixel, red pixel, ..." the information might be encoded as "279 red pixels". This is an essential case of run-length encoding; there are numerous plans to diminish document size by disposing of repetition.

The Lempel–Ziv (LZ) compression strategies are among the most mainstream algorithms for lossless storage.[7] DEFLATE is a minor departure from LZ advanced for decompression pace and compression proportion, however compression can be moderate. Flatten is utilized as a part of PKZIP, Gzip, and PNG. LZW (Lempel–Ziv–Welch) is utilized as a part of GIF images. Additionally vital is the LZR (Lempel-Ziv–Renau) algorithm, which serves as the premise for the Zip method.[citation needed] LZ techniques utilize a table-based compression model where table passages are substituted for rehashed strings of information. For most LZ strategies, this table is created progressively from before information in the information. The table itself is regularly Huffman encoded (e.g. SHRI, LZX). Current LZ-based coding plans that perform well are Brotli and LZX. LZX is utilized as a part of Microsoft's CAB position.

The best cutting edge lossless compressors use probabilistic models, for example, expectation by incomplete coordinating. The Burrows–Wheeler change can likewise be seen as a backhanded type of measurable modelling.[8]

The class of language structure based codes are picking up prevalence since they can pack very dull info amazingly adequately, for occurrence, a natural information accumulation of the same or firmly related species, a colossal formed report gathering, web documented, and so forth. The fundamental undertaking of sentence structure based codes is developing a connection free linguistic use inferring a solitary string. Sequitur and Re-Pair are commonsense punctuation compression algorithms for which programming is openly accessible.

In a further refinement of the immediate utilization of probabilistic displaying, measurable evaluations can be coupled to an algorithm called number juggling coding. Number juggling coding is a more current coding strategy that uses the numerical figurings of a limited state machine to deliver a string of encoded bits from a progression of info information images. It can accomplish better compression than different procedures, for example, the better-known Huffman algorithm. It utilizes an interior memory state to dodge the need to play out a coordinated mapping of individual info images to particular representations that utilization a whole number of bits, and it gets out the inner memory simply in the wake of encoding the whole string of information images. Number juggling coding applies particularly well to versatile information compression errands where the insights differ and are connection subordinate, as it can be effortlessly combined with a versatile model of the likelihood circulation of the info information. An early case of the utilization of number juggling coding was its utilization as a discretionary (yet not broadly utilized) element of the JPEG image coding standard. It has since been connected in different outlines including H.264/MPEG-4 AVC and HEVC for video coding.

Lossy

Lossy information pressure is the opposite of lossless information pressure. In these plans, some loss of data is satisfactory. Dropping superfluous point of interest from

the information source can spare storage room. Lossy information pressure plans are planned by exploration on how individuals see the information being referred to. For instance, the human eye is more touchy to unobtrusive varieties in luminance than it is to the varieties in shading. JPEG picture pressure works to some degree by adjusting off insignificant bits of information.[9] There is a comparing exchange off between saving data and lessening size. Various well known pressure groups misuse these perceptual contrasts, incorporating those utilized as a part of music documents, pictures, and video.

Lossy picture pressure can be utilized as a part of computerized cameras, to build stockpiling limits with insignificant debasement of picture quality. Additionally, DVDs utilize the lossy MPEG-2 video coding group for video pressure.

In lossy sound pressure, techniques for psychoacoustics are utilized to expel non-capable of being heard (or less discernable) parts of the sound sign. Pressure of human discourse is regularly performed with significantly more specific systems; discourse coding, or voice coding, is here and there recognized as a different order from sound pressure. Diverse sound and discourse pressure principles are recorded under sound coding positions. Voice pressure is utilized as a part of web communication, for instance, sound pressure is utilized for CD tearing and is decoded by the sound player.

3. METHODOLOGY

The compression procedure takes a unique information set and diminishes its size by taking out pointless information. The requirement for information compression expands step by step as the advancements requiring information move duplicate in extent. The outcome is the multiplication of more information compression algorithms. When LIPT was proposed, it was noticed that the 45% of aggregate Internet movement was involved by the content information. however, no lossless compression standard for content was proposed by then. Before that various refined algorithms have been proposed for lossless content compression of which Burrows Wheeler Transform (BWT) and Prediction by Partial Matching (PPM) beat the traditional algorithms like Huffman, number juggling and LZ groups of Gzip and Unix-pack. BWT's operation was in a manner that it first perform cyclic pivots of a piece of information and after that sorts lexicographically creating a rundown of each character and its subjectively long forward connection. The succession of characters at the last section of this piece is then compacted utilizing a plan called move to front (MTF) and an entropy coder as the backend compressor. By anticipating probabilities from going before images, PPM packs the succession of images. The first images considered in anticipating these probabilities make up the setting and the length of this connection is known as the request of PPM. A variation of PPM called PPMD processes escape likelihood as (u/2)/n, where u is the quantity of one of a kind token seen as such, and n is the quantity of token seen in this way. Among these algorithms, BWT has ended up being the most proficient and various endeavors have been made to enhance its effectiveness. In spite of the fact that PPM gives preferable compression proportion over all current compression algorithms, it is terribly moderate furthermore devours huge measure of memory to store connection in arrangement.

In the later past, the M5 Data Compression Group, University of Central Florida has built up a group of reversible Star-transformations which connected to a source content alongside a backend compression algorithm, accomplishes better compression. The essential thought of the change module is to change the content into some middle of the road structure, which can be packed with better productivity. The changed content is given to a backend information compression module which packs the changed content. Be that as it may, execution time execution and

runtime memory consumption of these compression frameworks have stayed high contrasted and the backend compression algorithms, for example, bzip2 and gzip.

At a certain point of time the change made in the compression algorithms achieves an immersion stage. There shows up the need of Preprocessing algorithms which encourages information compression by changing over the first information into some middle of the road structure.

Again the preprocessing algorithms prospered in number among which Length Index Preserving Transform (LIPT) is one of the noticeable lossless content changing algorithms. LIPT accomplishes some compression at the preprocessing stage and also holds enough setting and excess for the compression algorithms to give better results. As uncovered by Fauzia. S Awan& Amar Mukherjee, BZip2 with LIPT accomplish 5.24% change in normal Bits per Character (BPC) over BZip2 without LIPT.

However, since the need for a lot of compression is more wanted, we climb with the proposition of this algorithm, Enhanced Transformation with Numerical Index (ETNI).

3.2. ENHANCED TRANSFOR MATION WITH NUMERICAL INDEX (ETNI)

The methodology proposed here, called Enhanced Transformation with Numerical Index (ETNI), is a word reference based reversible lossless content transformation which when put into operation on the source content upgrades the current information compression algorithms' capacity.

The crucialrecipe behind the achievement of information compression algorithms is excess of information ie., more the repetition, higher would be the compression proportion. What's more, henceforth the part to be played by the preprocessing algorithms is to change the content into some middle structure which can be compacted with better effectiveness which is really expanding the excess way of the information.

Officially existing algorithms assembled and used the word recurrence information as indicated by lengths for the Calgary, Canterbury and Gutenberg Corpus. These information are alluded in ETNI and our compression results depend on the content records got from Calgary Corpus. As the wellspring of our work is fundamentally taking into account LIPT, we display the ideas of ETNI in examination with LIPT.

We have utilized an English dialect lexicon that has 59,951 words and takes around 0.5 Mbytes in uncompressed structure. This word reference should be the same at both the compression and decompression closes. In LIPT this word reference D is apportioned into disjoint lexicons Di… Dn. furthermore, this algorithm utilizes letters of the letters in order to mean length of the words furthermore to indicate the balance inside a square of words having same length present in the English lexicon. Therefore the letters will be rehashed and again in the changed content bringing about better setting. The same standard is utilized in ETNI, with the exception of that rather than letters in order, numbers running from 0… 9999 are utilized to speak to the words as the utilization of numerals rather than letters in order accomplish better compression proportion. Additionally here we don't parcel the word reference D into disjoint lexicons.

Need for New Transformation

As quick change is likewise attractive, other than the compression proficiency, we need a transformation that thinks about both these angles. The key for accomplishment of the change effectiveness predominantly relies on upon the development of the information structure. Taking after are a few conceivable methodologies: the first is hash table, which is truly quick, however outlining a fine-skewed hash capacity is extremely troublesome. Also, unsuccessful ventures utilizing hash table are excessively moderate. Another alternative is advanced inquiry tries. Computerized seek tries are likewise quick, be that as it may, they

have over the top space prerequisites: assume every hub has 52-way fanning, then one hub will commonly involve 213 bytes. On the other hand we can utilize double inquiry tree (LIPT take this methodology), which is space productive. Sadly, the quest expense is entirely high for parallel hunt. Ternary inquiry tree to store the change lexicon is another methodology (connected in StarNT). This information structure gives a quick change encoding speed with a low stockpiling overhead. In any case, the structure is too intricate to ever be kept up.

The change word reference utilized as a part of this proposition is a table in which every one of the operations are done in an agreeable and easy way. The idea of essential key evades duplication of the information. Likewise seeking of a specific codeword amid deciphering stage is entirely straight forward as the pointer is quickly moved to the comparing record.

Dictionary

The change lexicon utilized as a part of the test is set up ahead of time, and shared by both the change encoding module and the change interpreting module. The words in the lexicon are sorted in diving request of recurrence of use. Contingent upon the position of the word in the sorted lexicon, the codeword is built which is a number went before by star(s). The confined codeword is then put away together with the relating word.

Transformation process

ETNI includes two stages,

Stage 1: Make a canny word reference

Stage 2: Encode the info content information

In stage 1 a canny lexicon DETNI is made in plummeting request of recurrence of words and in stage 2, the encoding process happens in which the word is changed to a codeword that will be transmitted. The codeword takes the numeric reach from 0 – 9999 went before by * to show that the accompanying is an encoded word. Accordingly the first word in the lexicon is spoken to as *0, second word as *1, third word as *2,… , 10000th word as *9999. In like manner the following after words are shown as **0, **1, **2, … , **9999, ***0, ***1, ***2, … , ***9999 et cetera. On the off chance that the word in the information content is not in the English lexicon, it will be passed to the changed content accordingly with no adjustment.

The change should likewise have the capacity to handle extraordinary characters, to indicate that the primary letter of the information content word is promoted, to signify that every one of the letters in order in the information word are promoted, and to mean that the letter sets in the word are in switch case.

In the interpreting process, the changed words with "*" shows that the relating word is in encoded structure. The quantity of stars and the number succeeding the star(s) show the position of the first word in the lexicon D from where it can be brought effectively. The words without "*" before them are non-changed words and subsequently are composed to decoded document as it is with no disentangling.

4.EXPERIMENTAL DISCUSSIONS AND RESULTS

The proposed algorithm Enhanced Transformation with Numerical Index (ETNI) is given an onset in view of Length Index Preserving Transform (LIPT). Be that as it may, the silly corrections made in ETNI brought an amount of refinements in the execution accomplished.

In LIPT the first lexicon D is divided into disjoint lexicons Di each containing expressions of length i where I = 1, 2, … , n and then these lexicons are sorted incompletely as per the recurrence of words. Be that as it may, in ETNI, the lexicon D is not apportioned but rather the entire arrangement of words are sorted in diving request of their recurrence. In this way the overhead required in lexicon dividing and upkeep of each of these word references is wiped out. In any case, this increase is

acknowledged with the extra cost for seeking the word in the lexicon. This misfortune is additionally managed by executing word reference as table which is a trade for the parallel inquiry trees utilized by LIPT. Despite the fact that the parallel hunt tree used to look the words in the lexicon is space effective, the pursuit expense is very high. Subsequently the remuneration in time and cost elements is accomplished with effectiveness by the table usage of the word reference. By the utilization of essential key, total disposal of duplication of words and swifter pursuit of words is achieved.

The following angle to be considered is the codeword structure. In LIPT the codeword is made of letters in order while, on account of ETNI, we utilize just numerals which diminish the bits expected to speak to a character. This thus diminishes the document size and consequently a littler encoded record. This littler record takes less transmission time. Additionally since the entire document is changed over n to numbers, more excess is accomplished, which thus again prepares for better compression proportion.

Various refined algorithms have been proposed for lossless content compression of which Burrows Wheeler Transform (BWT) and Prediction by Partial Matching (PPM) beat the established algorithms like Huffman, number juggling and LZ groups of Gzip and Unix-pack. BWT's operation was in a manner that it first perform cyclic revolutions of a piece of information and then sorts lexicographically creating a rundown of each character and its discretionarily long forward connection. The grouping of characters at the last section of this piece is then packed utilizing a plan called move to front (MTF) and an entropy coder as the backend compressor. With a handful of advantages in BWT, we chose to utilize Bzip2, a free and open source record compressor that utilizations BWT algorithm or compression.

Here is the table giving the BPC accomplished by Bzip2 alone for the 14 documents from Calgary corpus.

File Name	Bzip2 (BPC)
paper5	3.24
paper4	3.12
paper6	2.58
Progc	2.53
paper3	2.72
Progp	1.74
paper1	2.49
Progl	1.74
paper2	2.44
Trans	1.53
Bib	1.97
News	2.52
Book2	2.06
Book1	2.42
Average BPC	**2.36**

Table 1: BPC achieved by Bzip2

Consequently as seen from the above table the normal BPC accomplished is 2.36 which when contrasted and the accompanying table demonstrates a striking distinction.

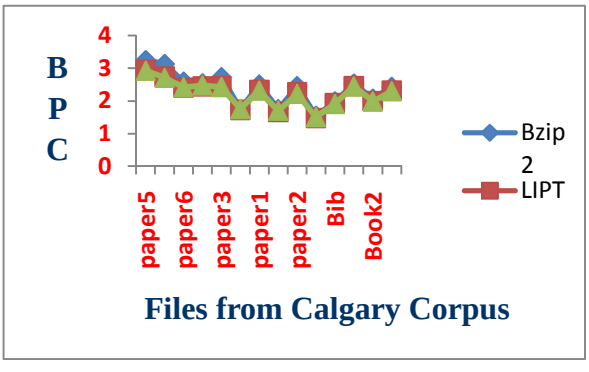
In this segment we center our regard for looking at the execution of LIPT utilizing Bzip2 with that of ETNI utilizing Bzip2. Our estimations have compression brings about terms of normal BPC (bits per character) as given in Table 5.1. Note postulations comes about incorporate some measure of pre-compression in light of the fact that the span of the ETNI content is littler than the extent of the first content document. By normal BPC we mean the un-weighted normal over the whole test corpus. The BPC figures are adjusted off to two decimal spots.

File	Original file size	Compressed file size	BZip2 with LIPT(BPC)	BZip2 with ETNI (BPC)
paper5	11954	4356	2.95	**2.92**
paper4	13286	4483	2.74	**2.70**
paper6	38105	11461	2.4	**2.41**
Progc	39611	12238	2.44	**2.47**

paper 3	46526	14067	2.45	**2.42**
Progp	49379	10756	1.72	**1.74**
paper 1	53161	15350	2.33	**2.31**
Progl	71646	15106	1.66	**1.69**
paper 2	82199	22716	2.26	**2.21**
Trans	93695	17536	1.47	**1.50**
Bib	111261	26464	1.93	**1.90**
News	377109	115057	2.45	**2.44**
Book 2	610856	151419	1.99	**1.98**
Book 1	768771	220062	2.31	**2.29**
Average BPC			**2.22**	**2.21**

Table 2: BPC Comparison between LIPT with BZip2 and ETNI withBzip2 for documents in Calgary Corpus

The documents in the table are recorded in climbing request of the document size. Note that for typical content documents the BPC diminishes as the record size increments. This can be obviously found in Table 5.2. From this table, it can be seen that ETNI demonstrates better BPC for the greater part of the documents and it has better normal BPC. From the table it is seen that the normal BPC utilizing Bzip2 with LIPT is 2.22 and utilizing Bzip2 with ETNI gives normal BPC of 2.21, a 1% change.



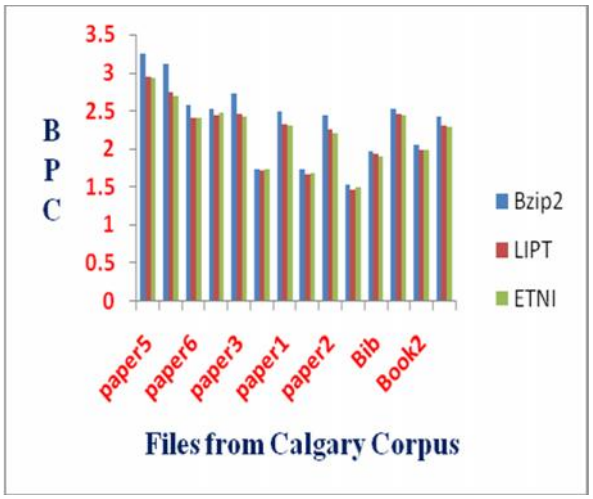


Figure 1: Comparison (1) of BPC achieved by Bzip2, B zip2with LIPT& Bzip2 with ETNI

The documents in the table are recorded in climbing request of the document size. Note that for typical content documents the BPC diminishes as the record size increments. This can be obviously found in Table 5.2. From this table, it can be seen that ETNI demonstrates better BPC for the greater part of the documents and it has better normal BPC. From the table it is seen that the normal BPC utilizing Bzip2 with LIPT is 2.22 and utilizing Bzip2 with ETNI gives normal BPC of 2.21, a 1% change.

CONCLUSION

We proposed a novel word reference based lossless reversible transformation algorithm, Enhanced Transformation with Numerical Index. We have taken Length Index Preserving Transformation (LIPT) as our base work and disposed of some of its drawbacks. In LIPT the first lexicon D is parceled into disjoint word references Di each containing expressions of length i where I = 1, 2, … , n and then these lexicons are sorted incompletely as indicated by the recurrence of words. Be that as it may, in ETNI, the lexicon D is not divided but rather the entire arrangement of words are sorted in sliding request of their recurrence. In this manner the overhead required in lexicon dividing and support of each of these word references is

wiped out. Be that as it may, this increase is acknowledged with the extra cost for looking the word in the lexicon. This misfortune is additionally managed by executing lexicon as table which is a swap for the paired hunt trees utilized by LIPT. Taking into account the new change the outcome got demonstrates a surprising change in the compression proportion by accomplishing lesser BPC. Bzip2 withLIPT gives 5.24% change in normal BPC over Bzip2 without LIPT. ETNI accomplishes a 1 % more change in normal BPC over Bzip2 withLIPT. In this manner an eminent update in BPC is figured it out.

REFERENCES

[1]. Wade, Graham (1994). Signal coding and processing (2 ed.). Cambridge University Press. p. 34. ISBN 978-0-521-42336-6. Retrieved 2011-12-22. The broad objective of source coding is to exploit or remove 'inefficient' redundancy in the PCM source and thereby achieve a reduction in the overall source rate R.

[2]. Mahdi, O.A.; Mohammed, M.A.; Mohamed, A.J. (November 2012). "Implementing a Novel Approach an Convert Audio Compression to Text Coding via Hybrid Technique" (PDF). International Journal of Computer Science Issues. **9** (6, No. 3): 53–59. Retrieved 6 March 2013.

[3]. Pujar, J.H.; Kadlaskar, L.M. (May 2010). "A New Lossless Method of Image Compression and Decompression Using Huffman Coding Techniques" (PDF). Journal of Theoretical and Applied Information Technology. **15** (1): 18–23.

[4]. Salomon, David (2008). A Concise Introduction to Data Compression. Berlin: Springer. ISBN 9781848000728.

[5]. S. Mittal and J. Vetter, "A Survey Of Architectural Approaches for Data Compression in Cache and Main Memory Systems", IEEE Transactions on Parallel and Distributed Systems, 2015.

[6]. Tank, M.K. (2011). Implementation of Limpel-Ziv algorithm for lossless compression using VHDL. Thinkquest 2010: Proceedings of the First International Conference on Contours of Computing Technology. Berlin: Springer. pp. 275–283.

[7]. Navqi, Saud; Naqvi, R.; Riaz, R.A.; Siddiqui, F. (April 2011). "Optimized RTL design and implementation of LZW algorithm for high bandwidth applications" (PDF). Electrical Review. **2011** (4): 279–285.

[8]. Mahmud, Salauddin (March 2012). "An Improved Data Compression Method for General Data" (PDF). International Journal of Scientific & Engineering Research. **3** (3): 2. Retrieved 6 March 2013.

[9]. Arcangel, Cory. "On Compression" (PDF). Retrieved 6 March 2013.

[10]. Marak, Laszlo. "On image compression" (PDF). University of Marne la Vallee. Retrieved 6 March 2013.

[11]. Mahoney, Matt. "Rationale for a Large Text Compression Benchmark". Florida Institute of Technology. Retrieved 5 March 2013.

[12]. Shmilovici A.; Kahiri Y.; Ben-Gal I.; Hauser S. "Measuring the Efficiency of the Intraday Forex Market with a Universal Data Compression Algorithm" (PDF). Computational Economics, Vol. 33 (2), 131-154., 2009.

[13]. I. Ben-Gal. "On the Use of Data Compression Measures to Analyze Robust Designs" (PDF). IEEE Trans. on Reliability, Vol. 54, no. 3, 381-388, 2008.

[14]. Korn, D.; et al. "RFC 3284: The VCDIFF Generic Differencing and Compression Data Format". Internet Engineering Task Force. Retrieved 5 March 2013.

[15]. Korn, D.G.; Vo, K.P. (1995), B. Krishnamurthy, ed., Vdelta: Differencing and Compression, Practical Reusable Unix Software, New York: John Wiley & Sons, Inc.

[16]. The Olympus WS-120 digital speech recorder, according to its manual, can store about 178 hours of speech-quality audio in .WMA format in 500MB of flash memory.

[17]. Coalson, Josh. "FLAC Comparison". Retrieved 6 March 2013.

[18]. Jaiswal, R.C. (2009). Audio-Video Engineering. Pune, Maharashtra: Nirali Prakashan. p. 3.41. ISBN 9788190639675.
[19]. Faxin Yu; Hao Luo; Zheming Lu (2010). Three-Dimensional Model Analysis and Processing. Berlin: Springer. p. 47. ISBN 9783642126512.
[20]. "File Compression Possibilities". A Brief guide to compress a file in 6 different ways.
[21]. "Summary of some of Solidyne's contributions to Broadcast Engineering". Brief History of Solidyne. Buenos Aires: Solidyne. Retrieved 6 March 2013.
[22]. Zwicker, Eberhard; et al. (1967). The Ear As A Communication Receiver. Melville, NY: Acoustical Society of America.
[23]. "Video Coding". Center for Signal and Information Processing Research. Georgia Institute of Technology. Retrieved 6 March 2013.
[24]. Graphics & Media Lab Video Group (2007). Lossless Video Codecs Comparison (PDF). Moscow State University.
[25]. Lane, Tom. "JPEG Image Compression FAQ, Part 1". Internet FAQ Archives. Independent JPEG Group. Retrieved 6 March 2013.
[26]. Faxin Yu; Hao Luo; Zheming Lu (2010). Three-Dimensional Model Analysis and Processing. Berlin: Springer. p. 47. ISBN 9783642126512.
[27]. Bhojani, D.R. "4.1 Video Compression" (PDF). Hypothesis. Retrieved 6 March 2013.
[28]. Ahmed, N.; Natarajan, T.; Rao, K.R. (January 1974). "Discrete Cosine Transform". IEEE Transactions on Computers. **C–23** (1): 90–93. doi:10.1109/T-C.1974.223784.
[29]. Chanda P, Bader JS, Elhaik E; Elhaik; Bader (27 Jul 2012). "HapZipper: sharing HapMap populations just got easier" (PDF). Nucleic Acids Research. **40** (20): e159. doi:10.1093/nar/gks709. PMC 3488212 . PMID 22844100.
[30]. Christley S, Lu Y, Li C, Xie X; Lu; Li; Xie (Jan 15, 2009). "Human genomes as email attachments". Bioinformatics. **25** (2): 274–5. doi:10.1093/bioinformatics/btn582.
PMID 18996942.
[31]. Pavlichin DS, Weissman T, Yona G; Weissman; Yona (September 2013). "The human genome contracts again". Bioinformatics. **29** (17): 2199–202. doi:10.1093/bioinformatics/btt362.
PMID 23793748.
[32]. PCSX2 team (January 8, 2016). "1.4.0 released! Year end report". Retrieved April 27, 2016.
[33]. Hilbert, Martin; López, Priscila (1 April 2011). "The World's Technological Capacity to Store, Communicate, and Compute Information". Science. **332** (6025): 60–65. Bibcode:2011Sci...332...60H.
doi:10.1126/science.1200970.
PMID 21310967. Retrieved 6 March 2013.