Management of Information System Journal
Vol 2, No 3, Juli 2024, Hal. 10 - 17
ISSN: 2964-9455 (media online)
DOI: 10.47065/mis.v2i3.1916
https://journal.fkpt.org/index.php/MIS

Application of the Unary Codes Algorithm for Compressing Video Files in the Kinemaster Application

Johanes Anderson Situmorang
Faculty of Computer Science & Information Technology, Informatics Engineering Study Program,
Budi Darma University, Medan, Indonesia
Email: andersonsitumorang98@gmail.com,
Email Correspondent: andersonsitumorang98@gmail.com

Abstract- - Speed in sending information is something that is really needed in today's technology. Information is a message or a collection of messages or data that has been processed to produce something needed by those in need. Information can take the form of various data formats, including text, video or images. The problem of sending information speed is often a little hampered due to the large file size of the information. In addition, the security factor is also very important, because there are many irresponsible parties who will try to steal this information and can harm the party whose information has been stolen if the information is misused.
The data used has a very large size. For that we need a technique to change the size of the data to make it smaller. This technique is called compression or what we are more familiar with as compression.
Data compression is a process of converting a set of data into a form of code to save the need for data storage space. In this study, researchers used the Unary Codes algorithm in compressing video files.

Keywords: Video file compression, Algorithm unary codes

1. INTRODUCTION

Data compression is reducing/compressing file size. Data compression is a technique for reducing data so that a file with a size smaller than the original file size can be obtained. Compression is needed to save storage space and time for database transmission/exchange [1].
Video is an object that can move in digital form, making videos that suit your needs cannot be ignored at this time. Due to increasingly advanced technological developments, videos are not only used in the form of one file, but can be used as several files. Currently, video can be a tool to document daily activities or an event. Several types of video file formats are MP4, MPEG, and others. MP4 is a format that is often used for storing video files. Because MP4 is very easy to play in several video player applications [2].
Kinemaster is an application used to edit videos from Android. Kinemaster makes it easy for users to edit videos or upload them, so that almost everyone can use this application. All the tools needed are available in the display menu. With just a few touches and combining themes, animations, video effects you can produce works of imagination like a professional editor.
The current problem, which often occurs, is that a small number of compressed bits often result in duplication in a text, image or video file and the compression or compression that is carried out cannot produce maximum results in data compression. 2 in the context of computer science is the science and art of presenting information in short form.
In previous research conducted by Mhd. Rajani Pane, published in the scientific journal Vol. 5 No.1 of 2017 entitled "Designing Compression Applications Using the Shannon Fano Method and Unary Coding on Text Files" concludes that data file compression uses the Unary Coding method. In his work, he presented a simple example of using Unary Coding in his research. The string 'KING' will be encoded so that the probability for each character can be obtained. After the probability of each character is known, each character will be given a certain range whose value is between 0 and 1, according to the existing probability. In this case there is no provision for character ordering, as long as the encoder and decoder do the same thing [3]..

2. RESEARCH METHODOLOGY

2.1 Data Compression

Compression is a technique of reducing or compressing large files into smaller ones and reducing storage space requirements. The compression process is a process that leads to minimizing the number of bits for digital representations such as images, audio and video, which results in a smaller data size but still maintains the quantity of information in the data [1]. Video compression is the process of reducing or minimizing the number of each bit that represents a video with smaller data. Data compression is also defined as a technique for compressing data to obtain data with a size that is smaller than the original size [4].

2.1.1 Compression Technique

Data compression or data compression methods are divided into two types, namely lossless and lossy methods, namely [5]:

Management of Information System Journal
Vol 2, No 3, Juli 2024, Hal. 10 - 17
ISSN: 2964-9455 (media online)
DOI: 10.47065/mis.v2i3.1916
https://journal.fkpt.org/index.php/MIS

1. Lossless Compression
   In this type of compression, the information contained in the resulting image is the same as the information in the original image. The image resulting from the compression process can be perfectly restored to the original image, there is no loss of information, no information errors. Therefore, this method is also called error free compression. This type of compression is suitable for application to database files, spread sheets, word processing files, biomedical images, and so on.
2. Lossy Compression
   Lossy compression is a compression technique that produces data files
   compression results that cannot be restored to the previous data file. When the compressed data is decoded again, some parts of the data are lost and are not the same if the data is returned as a decoding result. And this method produces a higher compression ratio than the lossless method.

2.1.2. Data Compression Performance Parameters

There are several techniques that are used as parameters to indicate the quality or performance of a compression method, namely [4]:

1. Compression Ratio
   Compression ratio is a compression parameter to compare the results of data before compression with data after compression. The compression ratio is an indicator to determine the results of compression. The following is the formula for writing the compression ratio.
   Rc=(Number of Bits before compression)/(Number of Bits after compression)
2. 2 Compression Ratio (RC)
   Compression ratio is a presentation of data that has been compressed which is obtained from the comparison between the size of data that has not been compressed and that which has been compressed.
   CR=(Data Size After Compression)/(Data Size Before Compression) 100%……(1)

2.2 Decompression

Decompression is a process of restoring data that has been compressed. Data that has been compressed must be able to be returned to its original form. To be able to change data that has been compressed, a different method is needed, such as when the compression process was carried out. So when decompressing there is a header note in the form of bytes containing notes about the contents of the file [5].

2.3 Unary Codes Algorithm

Unary Codes is a method inspired by Neuroscience, where existing data is usually encoded spatially, so that the location shows the location of the data. This is done by marking certain slots or marking them with different symbols to show the value or data in that slot [4]. The following are the steps in the Unary Codes compression algorithm:

1. low ← 0.0
2. high ← 1.0
3. while (input symbol still exists) do
4. take input symbols
5. CodeRange ← high – low
6. high ← low + CodeRange x high_range(s)
7. low ← low + CodeRange x low_range(s)
8. endwhile
9. low output

The following are the steps in the Unary Codes decompression algorithm:

1. take encoded symbol (ES)
2. repeat
3. Find the range of symbols that cover the encoded symbol (ES)
4. print symbols
5. CodeRange ← high_range – low_range
6. ES = ES – low_range
7. ES = ES/CodeRange
8. until symbols run out

Unary Codes are generally represented in a string of n 1 bits followed by one ending 0 bit which is defined as n-1 1 bits followed by one 0 bit or vice versa alternatively it can also equivalently start from n 0 bits followed by a ending 1 bit defined as n-1 bit 0 followed by one bit 1[4].

Table 1. Unary Codes

No	Kode	Kode Alternatif
1	0	1

Management of Information System Journal
Vol 2, No 3, Juli 2024, Hal. 10 - 17
ISSN: 2964-9455 (media online)
DOI: 10.47065/mis.v2i3.1916
https://journal.fkpt.org/index.php/MIS

No	Kode	Kode Alternatif
2	10	01
3	110	001
4	1110	0001
5	11110	00001
6	111110	000001
7	1111110	0000001
8	11111110	00000001
9	111111110	000000001
10	1111111110	0000000001
11	11111111110	00000000001
12	111111111110	000000000001
13	1111111111110	0000000000001
14	11111111111110	00000000000001
15	111111111111110	000000000000001
16	1111111111111110	0000000000000001
17	11111111111111110	00000000000000001

2.4 Video Files

Video is an electronic signal processing technology representing moving images. Video is a collection of moving images obtained from camera recordings or computer animation. Initially, this video information was stored analogue, as continuous changes in waveforms that represented changes in color and brightness of the recorded image. On the other hand, digital computers can only store and process binary data. For this reason, the computer industry defines color in 24-bit quantities which can be used to store around 16.7 million different possible colors. In this way, video data can be stored digitally as dots, each of which has a certain color and these dots, if arranged as a single unit, will form a complete image [7].

3. RESULT AND DISCUSSION

3.1 Discussion

Based on this research, analysis and design of software for compressing video files in mp4 format will be carried out using the Unary Codes algorithm. The analysis stage of a system is carried out before the design stage is carried out. The purpose of analyzing a system is to find out the reasons why the system is needed, namely by formulating the needs of the system to minimize excess resources and help plan the scheduling of the system. The problem discussed is how to design a video file compression application so that the application that has been designed can compress large video files, so that the data size becomes smaller, and is good and safe in the data storage and transfer process without reducing the content of the data. The compression and decompression diagram for MP4 files can be seen in the image below:

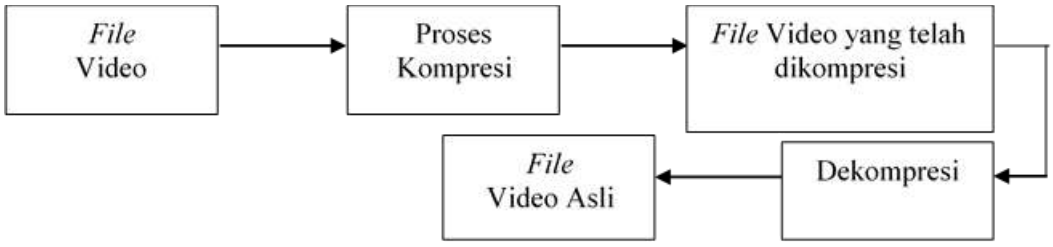

Figure 1. How Unary Codes Work

Based on the image above, the first input is an MP4 file that has not gone through the compression process. The output from the compression is an mp4 file whose size has changed from the initial capacity. The decompression process is the process of returning the original capacity of the mp4 file.

3.1.1 Case Example

In analyzing this video file, what you have to do is take a sample of the MP4 file by taking the original video file as an edit from the Kinemaster application and saving it in the system used in MP4 format. An example of this problem is a video file in MP4 format. When compressing an MP4 file, you must first analyze the MP4 file that will be compressed. MP4 files are files that are very well known and the most popular among other files. Reading MP4 files. Reading mp4 files is done to get the value of the data in an mp4 file in the form of hexadecimal numbers. Then the MP4 hexadecimal value data is compressed using the Unary Codes algorithm.

Management of Information System Journal
Vol 2, No 3, Juli 2024, Hal. 10 - 17
ISSN: 2964-9455 (media online)
DOI: 10.47065/mis.v2i3.1916
https://journal.fkpt.org/index.php/MIS

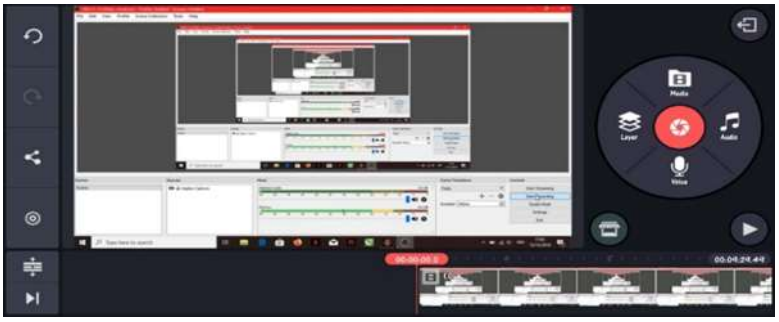

Figure 2. Kinemaster Video File

Based on the image above, there is information about the video file that will be compressed, this information can be seen in the table below.

Table 1. Sample MP4 File Information

Information	
MP4 File Type	MP4
Video Title Screen Recording on Laptop	Video Rekaman Layar Pada Laptop
Size 181 Mb (189,982,105 bytes)	181 Mb (189.982.**105** byte)

At this stage, the character set will be formed, namely the characters contained in the contents of the file. To get the binary value of the video file to be compressed, use the Binary Viewer application.

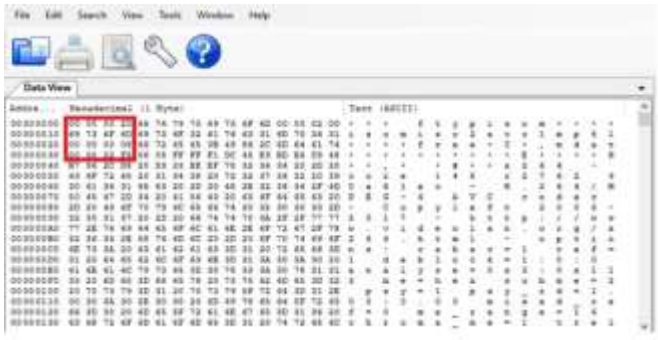

Figure 3. Hexadecimal and Binary Values of MP4 Files

Based on the image above, the hexadecimal value of the mp4 file is obtained. For manual calculation purposes, only a sample value of 16 characters of the hexadecimal value of the MP4 file is taken. The hexadecimal numbers for the sample MP4 file are: 00 00 00 20 69 73 6F 6D 00 00 00 08 00 00 02 F5.

Table 2. String to be compressed

Nilai Hexadesimal	Nilai Biner	Bit	Frekuensi	Bit x Frekuensi
00	00000000	8	8	64
20	00100000	8	1	8
69	01101001	8	1	8
73	01110011	8	1	8
6F	01101111	8	1	8
6D	01101101	8	1	8
08	00001000	8	1	8
02	00000010	8	1	8

Management of Information System Journal
Vol 2, No 3, Juli 2024, Hal. 10 - 17
ISSN: 2964-9455 (media online)
DOI: 10.47065/mis.v2i3.1916
https://journal.fkpt.org/index.php/MIS

F5	11110101	8	1	8
	Jumlah Bit x Frekuensi			128

After calculating the string data before compression, from the table above, compression is then carried out using the Unary Codes Algorithm, below is the code from the Unary Codes which can be seen in the table below:

Table 3. Unary Codes algorithm codewords

N	Kode	Kode Alternatif
1	0	1
2	10	01
3	110	001
4	1110	0001
5	11110	00001
6	111110	000001
7	1111110	0000001
8	11111110	00000001
9	111111110	000000001
10	1111111110	0000000001
11	11111111110	00000000001
12	111111111110	000000000001
13	1111111111110	0000000000001
14	11111111111110	00000000000001
15	111111111111110	000000000000001
16	1111111111111110	0000000000000001
17	11111111111111110	00000000000000001

Based on the table above, sorting can be done based on Unary Codes, which come from hexadecimal values as a sample obtained from table 3. The compression process for sample MP4 files is as follows:

Table 4. Sample MP4 file compression process

Nilai hexadesimal	Frekuensi	Unary Codes	Jumlah Bit	Total Bit
00	8	0	1	8
20	1	10	2	2
69	1	110	3	3
73	1	1110	4	4
6F	1	11110	5	5
6D	1	111110	6	6
08	1	1111110	7	7
02	1	11111110	8	8
F5	1	111111110	9	9
Total	16			52 bit

Based on table 4, it can be seen that the bit string form of the hexadecimal value "00 00 00 20 69 73 6F 6D 00 00 00 08 00 00 02 F5," after compression is:
“00010110111011110111110000111111000111111101111111110”.
So a file with 16 characters has a size of 52 bits after compression. Because the number of bit values of 46 is not divisible by 8 and has a remainder of 4, padding can be added in the amount of the remainder and made in the form of a bit, namely "1" so that it is divisible by 8. Thus, the flag bits which are the binary value of the length of the padding bit are 7 - 4 + ” 1 ” = 0001 final bit = 9 – n >> 9 – 4 = 5 in binary becomes “00000101”. So the total previous bits after adding padding and flag bits is 52 + 4 + 8 = 64.
After adding padding bits and flag bits, the bit string after compression is:
“00010110111011110111110000111111000111111101111111100000100000101”
So the total number of bit strings after adding padding and flags is 52+4+8 = 64 bits. To decode the bit string into ASCII characters, each bit string is divided into 8 bits, then read in ASCII characters.

Management of Information System Journal
Vol 2, No 3, Juli 2024, Hal. 10 - 17
ISSN: 2964-9455 (media online)
DOI: 10.47065/mis.v2i3.1916
https://journal.fkpt.org/index.php/MIS

00010110 01111100 00111111 00011111 11011111 11100001 00000101
| ? ▼ ß á ♣

Video files that have been compressed with the Unary Codes algorithm, the file type will be changed to a new file type with a specified extension as an identifier that the file is a file that has been compressed with a compression application that applies the Unary Codes algorithm. The specified file extension is (*.mp4).

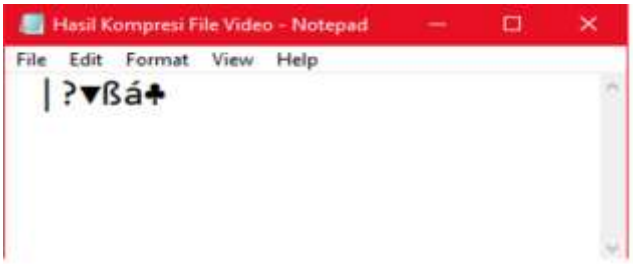

Figure 4. Video File Compression Results

Based on the image above, if the compression results of the video file are opened with notepad, unique characters will be displayed.

To determine the performance level of a compression algorithm, we must measure the compression results of the video file. The smaller the size of the compressed file, the higher the quality of the compression performance and vice versa. From the compression results with the Unary Codes algorithm above, the level of compression performance can be calculated according to the parameters that have been determined, namely:

Compression Ratio (CR)

CR=(Data Size After Compression)/(Data Size Before Compression) □ 100%…….(2.1)

CR=64/128 x 100 %

CR = 50 %

Equation 2.1 Compression Ratio (CR) results using the Unary Codes algorithm are 50%.

Table 4. Video File Compression Results Using the Unary Codes Algorithm

Information	
MP4 File Type	MP4
Video Title Screen Recording on Laptop	Video Rekaman Layar Pada Laptop
Size 181 Mb (189,982.97 bytes)	181 Mb (189.982.97 byte)

3.1.2 Analysis of the Decompression Process Using the Unary Codes Algorithm

In this decompression process, the results of the compressed video file are included in the previous discussion below. You can see the results of the compressed image that will be decompressed.

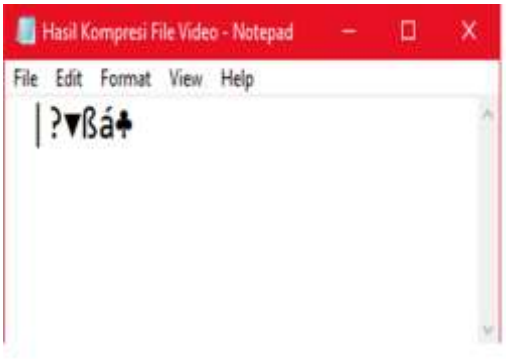

Figure 5. Video File Compression Results

Based on the image above, the steps above can be decomposed based on the characters being changed based on the binary value of the existing bit string.

Management of Information System Journal
Vol 2, No 3, Juli 2024, Hal. 10 - 17
ISSN: 2964-9455 (media online)
DOI: 10.47065/mis.v2i3.1916
https://journal.fkpt.org/index.php/MIS

00010110 01111100 00111111 00011111 11011111
| ? ▼ß
11100001 00000101
á ♣

The decompression process is carried out by reading the bit string from index 1 to index n of the length of the bit string. Each bit string reading index is checked in the Unary Codes table to see whether it matches or is the same. If it matches, the character is written in the character table with the same index in the Unary Codes code table. The following is an example of the decompression process of a compressed file in Table 5

Table 5. Hexadecimal and Binary Values After Compression

Nilai Hexa	*Codeword*
00	0
00	0
00	0
20	10
69	110
73	1110
6F	11110
6D	111110
00	0
00	0
00	0
08	1111110
00	0
00	0
02	1111111110
F5	11111111110

Based on the table above, all binary values are taken and combined as follows:
00010110111011110111110000111111000111111101111111110

1. First Stage

The first decompression process is to read the flag bits value from the entire bit value by changing the last 8 bit values into hexadecimal values as below:
00010110011111000011111100011111110111111110000100000101 Obtained the value of the last 8 bits as follows, 00000101 in the binary value is the value 5. The value 5 indicates that the previous compression result is divisible by 8 so there is additional padding and flag bits. Next, remove the padding and flag bits from the overall bit value so that it looks like the one below:
"00010110111011110111110000111111000111111101111111110"
Based on the decompression results above, the initial hexadecimal value before compression is found as follows: 00 00 00 20 69 73 6F 6D 00 00 00 08 00 00 02 F5

Table 6. Hexadecimal Decompression Results

00	00	00	20
69	73	6F	6D
00	00	00	08
00	00	02	F5

Based on the table above, it can be seen in the previous figure 6. The decompression results above show that the hexadecimal number value has returned to its original value. With this, there are also results where the video file returns to its original state and information such as about the video file returning to its original state.

Management of Information System Journal
Vol 2, No 3, Juli 2024, Hal. 10 - 17
ISSN: 2964-9455 (media online)
DOI: 10.47065/mis.v2i3.1916
https://journal.fkpt.org/index.php/MIS

Table 7. Information on the Results of Decompressing Video Files

Information	
MP4 File Type	MP4
Video Title	Video Rekaman Layar Laptop
Laptop Screen Recording	
Size 181 Mb (189,982,105 bytes)	181 Mb (189.982.105 byte)

4. CONCLUSION

1. The application of the Unary Codes algorithm results in compression of video files from large sizes to small sizes so as to save storage capacity.
2. In system design using the Microsoft Visual Studio 2008 application where the application is able to apply the Unary Codes algorithm to compress video files.
3. In this research the author has taken sample video files from the Kinemaster video editing application so that he can produce samples that will be compressed. This makes it easier for research to compress video files.

REFERENCES

[1] DARMA.Putra, *Pengolahan Citra Digital*. yogyakarta: C.V Andi Offset, 2010.
[2] R. Syahputra, "Kompresi File Video Mp4 dengan Menggunakan Metode Discrete Cosine Transform," *JURIKOM (Jurnal Ris. Komputer)*, 2016.
[3] D. V. S. O.D.N.M.W.M.K.T.S.M.kom, E.M.Kom, *Teori Pengolahan Citra Digital*. yogyakarta: C.V Andi Offset, 2009.
[4] G. M. D. Salomon, *Handbook Of Data Compression*. London Dordrecht Heidelberg New York, 2010.
[5] C. D. T. Sukiman, "Aplikasi Kompresi File dengan algoritma Elias Gamma," 2013.
[6] A. S. dan S. M. Rosa, *Pemograman Berorientasi Objek. Bandung*. Bandung: Modula, 2010.
[7] I. Binanto, *Multimedia Digital Dasar Teori dan Pengembangannya*. yogyakarta: C.V Andi Offset, 2010.
[8] J. Enterprise, *Rahasia Menjadi Jago Download*. Jakarta: PT. Elex Media Komputindo, 2010.
[9] Hendrayudi, *Dasar-Dasar Pemrograman Microsoft Visual Basic 2008*. Bandung: PT. Sarana TutorialL NuraniI Sejahtera, 2011.
[10] R. H. S, *Visual Basic 2008 Pemula Banget*. Jakarta: PT. Elex Media Komputindo, 2009