Analysis of MPEG-2 Video Streams

Damir Isović and Gerhard Fohler
Department of Computer Engineering
Mälardalen University, Sweden
{damir.isovic, gerhard.fohler}@mdh.se

Abstract

MPEG-2 is widely used as video coding standard for contents such as DVD or digital video broadcasting, DVB. It defines a layered structure, composing three different types of frames into groups for temporal and spatial compression of video information.

In this paper we present an exhaustive analysis av various MPEG streams, taken from original DVDs. The purpose is to get a more clear picture about what are valid assumptions about MPEG. The analysis showed that many common assumptions, in particular about relation of frame sizes, and equal importance of frames, do not hold in the general case.

1 Introduction

MPEG, the Moving Picture Experts Group standard for coded representation of digital audio and video, is used in a wide range of applications. In particular MPEG-2 has become the coding standard for digital video streams in consumer content and devices, such as DVD movies and digital television set top boxes for DVB, terrestrial TV broadcasts or via satellite. It should be noted that MPEG is a standard for the format, a syntax, not for the actual encoding: the same content, e.g., a movie, can be encoded in many ways while adhering to the same standard. In fact, MPEG encoding has to meet diverse demands, depending, e.g., on the medium of distribution, such as overall size in the case of DVD, maximum bitrate for DVB, or speed of encoding for live broadcasts.

In the case of DVD and DVB, sophisticated provisions to apply spatial and temporal compression are applied, while a very simple, but quickly coded stream will be used for the live broadcast. Consequently, video streams, and in particular their decoding demands will vary greatly between media, but also different types of contents or even different scenes within the same movie.

MPEG-2 video streams have a layered structure. The layer we are considering here is the picture layer, where the video data is organized in *Group of Pictures*(GOP), i.e., a sequence of pictures that consist of a number of frames. The three types of frames are I frames (*intra-coded pictures*), P frames (*predicted pictures*), and B frames *bi-directionally predicted pictures*. Simply speaking, I frames contain full pictures and are independent, P frames build a full picture using a previous I or P frame as reference, and B frames contain incremental changes to a full picture, based on both previous and later frames.

In this paper we present results of an analysis of realistic MPEG streams of DVD movies and match the analysis results against common assumptions. For example, an intuitive conclusion is that I will be the largest frames, followed by P and B frames, and frames have similar sizes within their respective frame type. While true on average, such assumptions do not hold for a considerable number of cases. The analysis of realistic streams presented in this paper shows, e.g., a case with 9% GOPs in which P have the largest size, and 1% of B frames, which corresponds to roughly to 8 and 1 minutes, resp., in a 90 minute feature film. Clearly, such deviations from average cannot be ignored.

The analysis showed that many common assumptions, in particular about relation of frame sizes, and equal importance of frames, do not hold in the general case.

2 MPEG Video Streams Properties

A complete description of the MPEG compression scheme is beyond the scope of this paper. For details on MPEG see e.g. [1, 4, 3]. Here we will focus on the MPEG video stream structure, and see how it can be analyzed and scheduled. In this work, we describe the most important characteristics of a MPEG-2 video stream. The text presented in this subsection is sumarized in figure 1.

2.1 Frame types

The MPEG-2 standard defines three types of frames, I, P and B.

I frames or *intra* frames are simply frames coded as still images. They contain absolute picture data and are self-contained, meaning that they require no additional information for decoding. I frames have only spatial redundancy providing the least compression among all frame types. Therefore they are not transmitted more frequently than necessary.

P frames The second kind of frames are P or *predicted* frames. They are forward predicted from the most recently reconstructed I or P frame, i.e., they contain a set of instructions to convert the previous picture into the current one. P frames are not self-contained, meaning that if the previous reference frame is lost, decoding is impossible. On average, P frames require roughly half the data of an I frame, but our analysis also showed that this is not the case for the significant number of cases.

B frames The third type is B or *bi-directionally* predicted frames. They use both forward and backward prediction, i.e., a B frame can be decoded from a previous I or P frame, and/or from a *later* I or P frame. They contain vectors describing where in an earlier or later pictures data should be taken from. They also contain transformation coefficients that provide the correction. B frames are never predicted from each other, only from I or P frames. As a consequence, no other frames depend on B frames. B frames require resource-intensive compression techniques such as Motion Compensation and Motion Estimation but they also exhibit the highest compression ratio, on average typically requiring one quarter the data of an I picture. Again, our analysis showed that this does not hold for a significant number of cases.

2.2 Group of Pictures

Predictive coding, i.e., the current frame is predicted from the previous one, cannot be used indefinitely, as it is prone to error propagation. A further problem is that it becomes impossible to decode the transmission if reception begins part-way through. In real video signals, cuts or edits can be present across which there is little redundancy. In the absence of redundancy over a cut, there is nothing to be done but to send from time to time a new picture information in absolute form, i.e., an I frame. As I decoding needs no previous frame, decoding can begin at I coded information, for example, allowing the viewer to switch channels. An I frame together with all of the frames before the next I frame form a *group of pictures (GOP)*. The GOP length is flexible, but 12 or 15 frames is a common value. Furthermore, it is common industrial practice to have a fixed pattern (e.g. $I\ BB\ P\ BB\ P\ BB\ P\ BB$). However, more advanced encoders will attempt to optimize the placement of the three frame types according to local sequence characteristics in the context of more global characteristics. Note that the last B frame in a GOP requires the I frame in the next GOP for decoding and so the GOPs are not truly independent. Independence can be obtained by creating a *closed* GOP which may contain B frames but ends with a P frame.

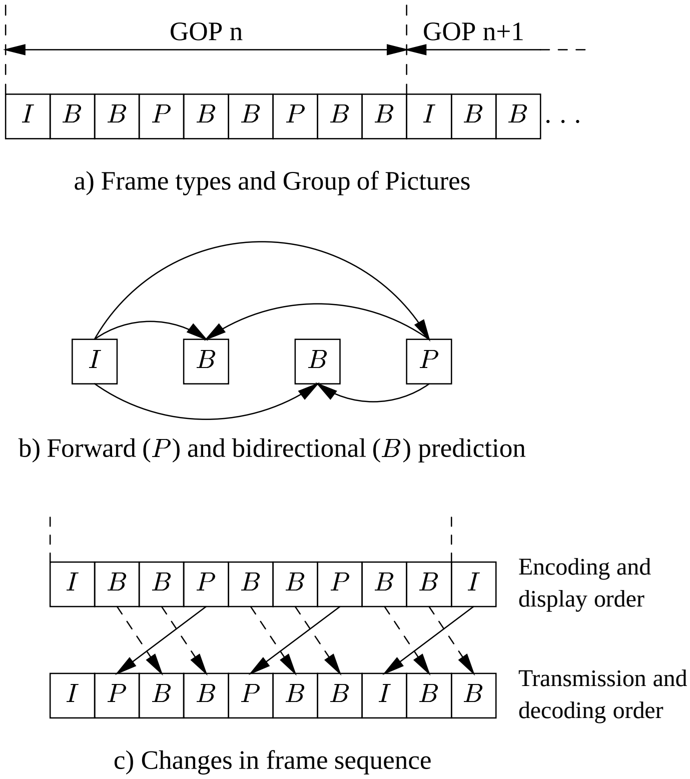


Figure 1: MPEG-2 video stream characteristics

2.3 Transmission vs display order

As we mentioned above, B frames are predicted from two I or P frames, one in the past and one in the future. Clearly information in the future has yet to be transmitted and so is not normally available to the decoder. MPEG gets around the problem by sending frames in the “wrong” order. The frames are sent out of sequence and temporarily stored. Figure 1-c shows that although the original frame sequence is $I\ BB\ P$ ..., this is transmitted as $I\ P\ BB$ ..., so that the future frame is already in the decoder before bi-directional decoding begins. Picture reordering requires additional memory at the encoder and decoder and delay in both of them to put the order right again. The number of bi-directionally coded frames between I and P frames must be restricted to reduce cost and minimize delay, if delay is an issue.

3 Analysis of Various MPEG streams

We have analyzed a number of realistic MPEG streams to get a more clear picture about which assumption about MPEG are valid. Some types of videos are more sensitive for frames dropping. For example, dropping 4 frames in an action video reduces half of the original video quality, 50%, while only 10% in a cartoon video [2]. Therefore we have analysed different types of movies such as action movie, drama, cartoons, etc.

3.1 Simulation environment

We have analysed the contents of original DVD movies. The movies were not encrypted or copy protected in any sense, which means that we managed to rip their context on a hard drive by using only legal ripping software, i.e., the one that will not try to break the CSS protection code on a DVD.

Ripped MPEG streams were analysed by an own-written piece of software (C-program). It takes approximately 10 minutes to analyse a 100 minutes long MPEG stream on a PC computer with the processor speed of 1,5 GHz.

3.2 Analysed DVD movies

An overview of the movies we analyzed is summarized in table 1. N and M refer to the GOP length and distance between reference frames respective, e.g. GOP(12,3) means I-to-I distance is 12, while I-to-P and P-to-P distance is 3.

3.3 Analysis results: Mission Impossible 2

Here is the data for the movie *Mission Impossible 2*. Table 2 sumarizes GOP and frame size properties for the movie. Minumum, maximum and average size is given in bits.

The size ration between average values for respective frame type is I:P:B=4:2:1, which means that on average I frames are twice as big as the P frames, and 4 times bigger than the B frames. However, this does not hold for a significant number of cases, which is depicted in table 3.

For example, in "Mission Impossible 2" we have a case with 10% GOPs in which P have the largest size, and 1% of B frames, which corresponds roughly to 13 and 1.5 minutes, resp, in a 90 minute feature film. Clearly, such deviations from average cannot be ignored. Furthermore we can see from table 3 that frames in a GOP are not sorted according to their bitsize, e.g., in 81% of the cases, the P frame that is closest to the I frame was not the largest among all P frames in the GOP.

We have also analysed the distribution of the frame sizes. We have divided the range between minimum and maximum frame size for respective frame type into ten size intervals, and identified the number of frames in respective interval. In that way we can e.g say that the majority of frames have bitsize between some X and Y. This is depicted in figure 2.

For example, from figure 2 we can see that 88% of the I frames has bitsize between 197737 and 790684 bits ($\approx$ 200 - 800 kB), which is a quite large interval. The assumptions about MPEG based on average frame size will not hold in this case, since the significant number of frames will have twice as large repective twice as small bitsize, compared to the average frame size (which is $\approx$ 500 kB).

Movie title	Genre	Length	Fps	Resolution	Mbit/s	GOP
Mission Impossible 2	Action	118 min	25	720x576	9800	(12,3)
Leaving Las Vegas	Drama	107 min	25	720x576	8700	(12,3)
Chicken Run	Cartoon	104 min	25	720x576	6000	(12,3)
The Usual Suspect	Thriller	106 min	30	720x480	9800	(12,3)
The Matrix	Action	122 min	30	720x480	7500	(12,3)
New Year's Concert	Music	120 min	25	720x576	7000	(12,3)
The Sea	Doc.	55 min	30	720x480	6500	(12,3)

Table 1: Analyzed MPEG streams

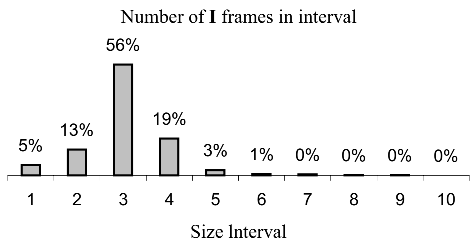
Item	Count	Minimum	Maximum	Average	Std deviation
I	16873	88	1976584	506109	187598
P	49679	16	1216000	234821	109889
B	112860	32	769048	148204	57615
GOP	16873	88	7541496	2222249	746767

Table 2: Mission Impossible 2 - Bitsizes for frames and GOPs

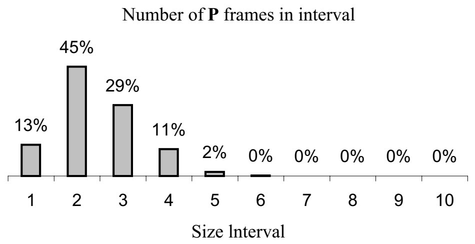
GOP property	Nr of GOPs	Percent
Open GOPs	12900	76%
Closed GOPs	3973	24%
GOPs with normal length (12)	12991	77%
Largest frame I	15061	89%
Largest frame P	1658	10%
Largest frame B	154	1%
GOPs where $P > I$	5256	31%
GOPs where $B > I$	4442	26%
GOPs where $B > P$	6545	39%
$P >$ some previous P in the GOP	13609	81%
$B >$ some previous B in the GOP	16326	97%

Table 3: Mission Impossible 2 - GOP properties

Interval	From	To	Nr of I	Percent
1	88	197737	876	5,2%
2	197737	395386	2190	13,0%
3	395386	593035	9410	55,8%
4	593035	790684	3137	18,6%
5	790684	988333	426	2,5%
6	988333	1185982	129	0,8%
7	1185982	1383631	79	0,5%
8	1383631	1581280	51	0,3%
9	1581280	1778929	23	0,1%
10	1778929	1976584	6	0,0%



Interval	From	To	Nr of P	Percent
1	16	121614	6377	12,8%
2	121614	243212	22355	45,0%
3	243212	364810	14460	29,1%
4	364810	486408	5496	11,1%
5	486408	608006	857	1,7%
6	608006	729604	102	0,2%
7	729604	851202	17	0,0%
8	851202	972800	11	0,0%
9	972800	1094398	3	0,0%
10	1094398	1216000	1	0,0%



Interval	From	To	Nr of B	Percent
1	32	76933	7365	6,5%
2	76933	153834	59938	53,1%
3	153834	230735	35827	31,7%
4	230735	307636	8370	7,4%
5	307636	384537	1195	1,1%
6	384537	461438	113	0,1%
7	461438	538339	32	0,0%
8	538339	615240	13	0,0%
9	615240	692141	3	0,0%
10	692141	769048	2	0,0%

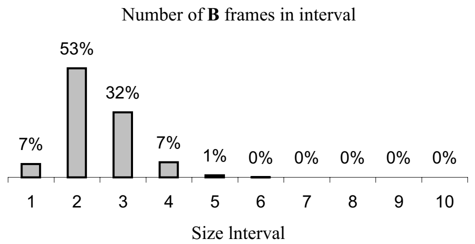


Figure 2: Mission Impossible 2 - Size distribution for I, P and B frames

3.4 Analysis results: Leaving Las Vegas

The GOP and frame sizes for the movie *Leaving Las Vegas* are presented in table 4. The GOP properties are described in table 5 and the size distribution is shown in figure 3.

3.5 Analysis results: Chicken Run

The size data and GOP properties for the cartoon *Chicken Run* is presented in tables 6 and 7. The size distribution is shown in figure 4.

3.6 Analysis results: The Usual Suspect

The size data and GOP properties for the movie *The Usual Suspect* can be found in in tables 8 and 9. The size distribution is depicted in figure 5.

3.7 Analysis results: The Matrix

The GOP and frame sizes for the movie *The Matrix* are presented in table 10. The GOP properties are described in table 11 and the size distribution is shown in figure 6.

3.8 Analysis results: New Year's Concert

The size data and GOP properties for the cartoon *New Year's Concert* is presented in tables 12 and 13. The size distribution is shown in figure 7.

3.9 Analysis results: The Sea

The GOP and frame sizes for the movie *The Sea* are sumarized in table 14. The GOP properties are described in table 15 and the size distribution is shown in figure 8.

Item	Count	Minimum	Maximum	Average	Std deviation
I	13716	136	1469848	471886	140329
P	52860	32	1009832	231145	76835
B	106478	32	636416	152435	45746
GOP	13716	136	7185768	2543520	627856

Table 4: Leaving Las Vegas - Bitsizes for frames and GOPs

GOP property	Number of GOPs	Percent
Open GOPs	13381	98%
Closed GOPs	335	2%
GOPs with normal length (12)	12573	92%
Largest frame I	12904	94%
Largest frame P	758	6%
Largest frame B	54	0,4%
GOPs where $P > I$	786	6%
GOPs where $B > I$	230	2%
GOPs where $B > P$	5072	37%
$P >$ some previous P in the GOP	11481	84%
$B >$ some previous B in the GOP	13715	100%

Table 5: Leaving Las Vegas - GOP properties

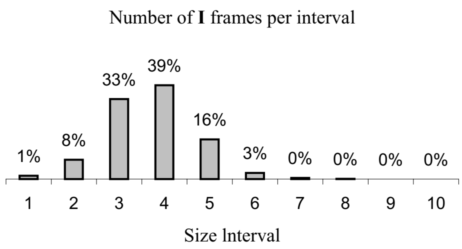
Frame type	Nr of frames	Min	Max	Avg	Std dev
I	10139	57424	1121216	674549	216068
P	30406	1272	1097336	255551	133372
B	80861	1264	891240	115185	53982
GOP	10139	69712	4680200	2360795	622665

Table 6: Chicken Run - Bitsizes for frames and GOPs

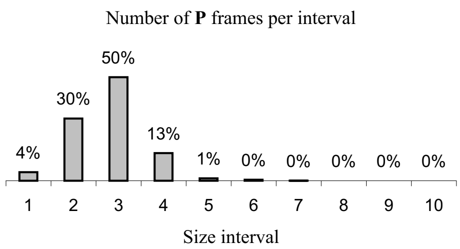
GOP property	Number of GOPs	Percent
Open GOPs	10123	100%
Closed GOPs	16	0,2%
GOPs with normal length (12)	10056	99%
Largest frame I	9291	92%
Largest frame P	842	8%
Largest frame B	14	0,1%
GOPs where $P > I$	841	8%
GOPs where $B > I$	79	1%
GOPs where $B > P$	1180	12%
$P >$ some previous P in the GOP	8260	81%
$B >$ some previous B in the GOP	10138	100%

Table 7: Chicken Run - GOP properties

Interval	From	To	Nr of I	Percent
1	136	146985	188	1,4%
2	146985	293970	1089	7,9%
3	293970	440955	4503	32,8%
4	440955	587940	5281	38,5%
5	587940	734925	2231	16,3%
6	734925	881910	344	2,5%
7	881910	1028895	63	0,5%
8	1028895	1175880	14	0,1%
9	1175880	1322865	1	0,0%
10	1322865	1469850	2	0,0%



Interval	From	To	Nr of P	Percent
1	32	100984	2206	4,2%
2	100984	201968	15999	30,3%
3	201968	302952	26630	50,4%
4	302952	403936	7098	13,4%
5	403936	504920	618	1,2%
6	504920	605904	252	0,5%
7	605904	706888	35	0,1%
8	706888	807872	3	0,0%
9	807872	908856	0	0,0%
10	908856	1009840	2	0,0%



Interval	From	To	Nr of B	Percent
1	32	63642	914	0,9%
2	63642	127284	14630	13,7%
3	127284	190926	25275	23,7%
4	190926	254568	8385	7,9%
5	254568	318210	693	0,7%
6	318210	381852	54	0,1%
7	381852	445494	12	0,0%
8	445494	509136	2	0,0%
9	509136	572778	0	0,0%
10	572778	636420	0	0,0%

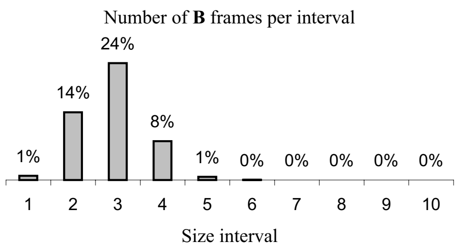
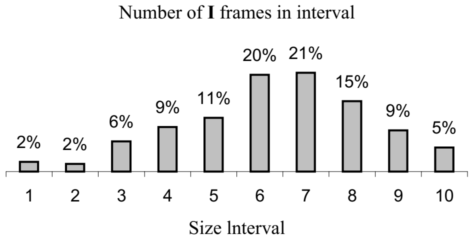


Figure 3: Leaving Las Vegas - Size distribution for I, P and B frames

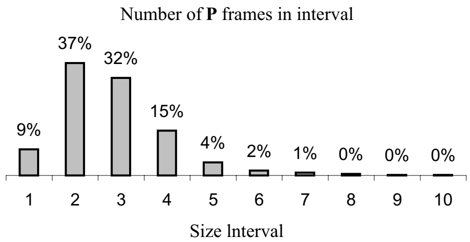
Item	Count	Minimum	Maximum	Average	Std deviation
I	13404	2856	1282720	514744	174307
P	40088	32	1204808	281867	92779
B	98868	32	762048	129537	48890
GOP	13404	13312	5896728	2324673	583043

Table 8: The Usual Suspect - Bitsizes for frames and GOPs

Interval	From	To	Nr of I	Percent
1	57424	163803	207	2,0%
2	163803	270182	165	1,6%
3	270182	376561	643	6,3%
4	376561	482940	948	9,4%
5	482940	589319	1140	11,2%
6	589319	695698	2056	20,3%
7	695698	802077	2098	20,7%
8	802077	908456	1494	14,7%
9	908456	1014835	878	8,7%
10	1014835	1121216	510	5,0%



Interval	From	To	Nr of P	Percent
1	1272	110878	2616	8,6%
2	110878	220484	11245	37,0%
3	220484	330090	9768	32,1%
4	330090	439696	4473	14,7%
5	439696	549302	1298	4,3%
6	549302	658908	478	1,6%
7	658908	768514	279	0,9%
8	768514	878120	141	0,5%
9	878120	987726	63	0,2%
10	987726	1097336	45	0,1%



Interval	From	To	Nr of B	Percent
1	1264	90261	29767	36,8%
2	90261	179258	41449	51,3%
3	179258	268255	8605	10,6%
4	268255	357252	941	1,2%
5	357252	446249	83	0,1%
6	446249	535246	9	0,0%
7	535246	624243	5	0,0%
8	624243	713240	0	0,0%
9	713240	802237	1	0,0%
10	802237	891240	1	0,0%

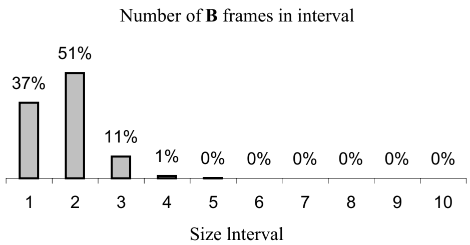
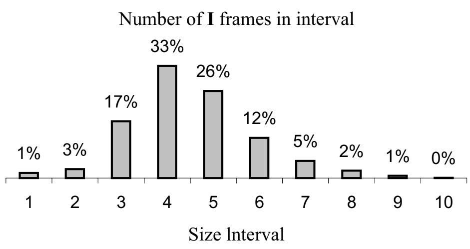


Figure 4: Chicken Run - Size distribution for I, P and B frames

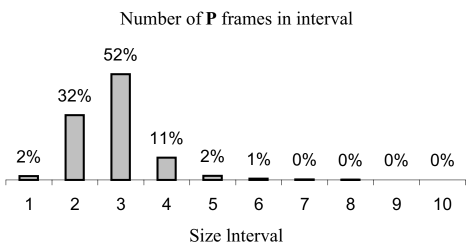
GOP property	Number of GOPs	Percent
Open GOPs	11443	85%
Closed GOPs	1961	15%
GOPs with normal length (12)	11005	82%
Largest frame I	11874	89%
Largest frame P	1477	11%
Largest frame B	53	0%
GOPs where $P > I$	4253	32%
GOPs where $B > I$	2035	15%
GOPs where $B > P$	1112	8%
$P >$ some previous P in the GOP	9587	72%
$B >$ some previous B in the GOP	13264	99%

Table 9: The Usual Suspect - GOP properties

Interval	From	To	Nr of I	Percent
1	2856	130842	198	1,5%
2	130842	258828	356	2,7%
3	258828	386814	2238	16,7%
4	386814	514800	4408	32,9%
5	514800	642786	3422	25,5%
6	642786	770772	1579	11,8%
7	770772	898758	675	5,0%
8	898758	1026744	289	2,2%
9	1026744	1154730	88	0,7%
10	1154730	1282720	12	0,1%



Interval	From	To	Nr of P	Percent
1	32	120509	752	1,9%
2	120509	240986	12809	32,0%
3	240986	361463	20900	52,1%
4	361463	481940	4428	11,0%
5	481940	602417	793	2,0%
6	602417	722894	253	0,6%
7	722894	843371	103	0,3%
8	843371	963848	33	0,1%
9	963848	1084325	13	0,0%
10	1084325	1204808	4	0,0%



Interval	From	To	Nr of B	Percent
1	32	76233	8255	8,3%
2	76233	152434	66744	67,5%
3	152434	228635	21015	21,3%
4	228635	304836	2016	2,0%
5	304836	381037	439	0,4%
6	381037	457238	161	0,2%
7	457238	533439	202	0,2%
8	533439	609640	25	0,0%
9	609640	685841	7	0,0%
10	685841	762048	3	0,0%

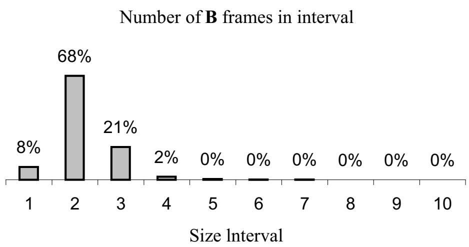


Figure 5: The Usual Suspect - Size distribution for I, P and B frames

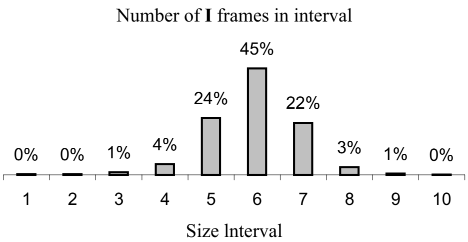
Item	Count	Minimum	Maximum	Average	Std deviation
I	14663	41104	760000	430088	70920
P	43920	1272	809016	249576	65226
B	117090	3184	664968	136725	41336
GOP	14667	76088	4322648	2269353	408220

Table 10: The Matrix - Bitsizes for frames and GOPs

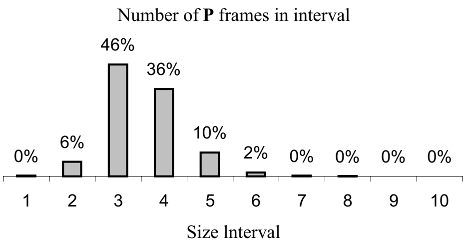
GOP property	Number of GOPs	Percent
Open GOPs	14664	100%
Closed GOPs	23	0%
GOPs with normal length (12)	14595	100%
Largest frame I	13665	93%
Largest frame P	954	7%
Largest frame B	48	0%
GOPs where $P > I$	2453	17%
GOPs where $B > I$	449	3%
GOPs where $B > P$	1491	10%
$P >$ some previous P in the GOP	7424	51%
$B >$ some previous B in the GOP	14662	100%

Table 11: The Matrix - GOP properties

Interval	From	To	Nr of I	Percent
1	41104	112993	53	0,4%
2	112993	184882	42	0,3%
3	184882	256771	156	1,1%
4	256771	328660	653	4,5%
5	328660	400549	3471	23,7%
6	400549	472438	6529	44,5%
7	472438	544327	3197	21,8%
8	544327	616216	474	3,2%
9	616216	688105	80	0,5%
10	688105	760000	8	0,1%



Interval	From	To	Nr of P	Percent
1	1272	82046	159	0,4%
2	82046	162820	2618	6,0%
3	162820	243594	20183	46,0%
4	243594	324368	15747	35,9%
5	324368	405142	4294	9,8%
6	405142	485916	707	1,6%
7	485916	566690	150	0,3%
8	566690	647464	43	0,1%
9	647464	728238	14	0,0%
10	728238	809016	5	0,0%



Interval	From	To	Nr of B	Percent
1	3184	69362	2967	2,5%
2	69362	135540	65311	55,8%
3	135540	201718	39890	34,1%
4	201718	267896	7782	6,6%
5	267896	334074	966	0,8%
6	334074	400252	127	0,1%
7	400252	466430	28	0,0%
8	466430	532608	9	0,0%
9	532608	598786	5	0,0%
10	598786	664968	3	0,0%

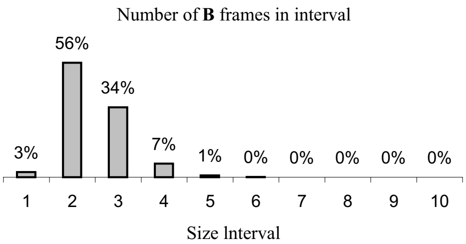


Figure 6: The Matrix - Size distribution for I, P and B frames

Item	Count	Minimum	Maximum	Average	Std deviation
I	14541	3432	1895088	1019897	363358
P	55248	32	1459952	396579	98782
B	110396	24	1565960	184918	51664
GOP	14541	8912	10635840	4000410	806103

Table 12: New Year's Concert - Bitsizes for frames and GOPs

GOP property	Number of GOPs	Percent
Open GOPs	14322	98%
Closed GOPs	219	2%
GOPs with normal length (12)	12292	85%
Largest frame I	13402	92%
Largest frame P	1079	7%
Largest frame B	60	0%
GOPs where $P > I$	3897	27%
GOPs where $B > I$	2999	21%
GOPs where $B > P$	2206	15%
$P >$ some previous P in the GOP	13566	93%
$B >$ some previous B in the GOP	13996	96%

Table 13: New Year's Concert - GOP properties

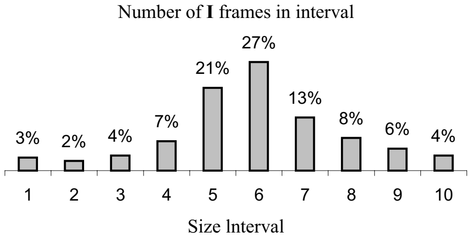
Item	Count	Minimum	Maximum	Average	Std deviation
I	8036	14392	819736	568826	116259
P	23929	32	764696	414148	48855
B	63747	32	423880	199928	26295
GOP	8036	80672	6412152	3396570	291106

Table 14: The Sea - Bitsizes for frames and GOPs

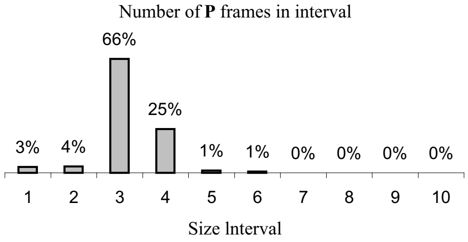
GOP property	Number of GOPs	Percent
Open GOPs	8020	100%
Closed GOPs	16	0%
GOPs with normal length (12)	7674	95%
Largest frame I	7672	95%
Largest frame P	357	4%
Largest frame B	7	0%
GOPs where $P > I$	1317	16%
GOPs where $B > I$	1532	19%
GOPs where $B > P$	333	4%
$P >$ some previous P in the GOP	5872	73%
$B >$ some previous B in the GOP	7997	100%

Table 15: The Sea - GOP properties

Interval	From	To	Nr of I	Percent
1	3432	192597	474	3,3%
2	192597	381762	358	2,5%
3	381762	570927	550	3,8%
4	570927	760092	1073	7,4%
5	760092	949257	3036	20,9%
6	949257	1138422	3978	27,4%
7	1138422	1327587	1945	13,4%
8	1327587	1516752	1200	8,3%
9	1516752	1705917	802	5,5%
10	1705917	1895088	551	3,8%



Interval	From	To	Nr of P	Percent
1	32	146024	1866	3,4%
2	146024	292016	2059	3,7%
3	292016	438008	36194	65,5%
4	438008	584000	13921	25,2%
5	584000	729992	762	1,4%
6	729992	875984	415	0,8%
7	875984	1021976	27	0,0%
8	1021976	1167968	2	0,0%
9	1167968	1313960	0	0,0%
10	1313960	1459952	2	0,0%



Interval	From	To	Nr of B	Percent
1	24	156617	24255	22,0%
2	156617	313210	85485	77,4%
3	313210	469803	561	0,5%
4	469803	626396	73	0,1%
5	626396	782989	5	0,0%
6	782989	939582	4	0,0%
7	939582	1096175	3	0,0%
8	1096175	1252768	6	0,0%
9	1252768	1409361	0	0,0%
10	1409361	1565960	3	0,0%

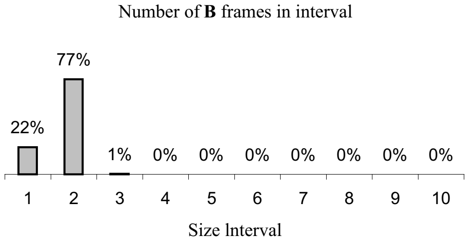
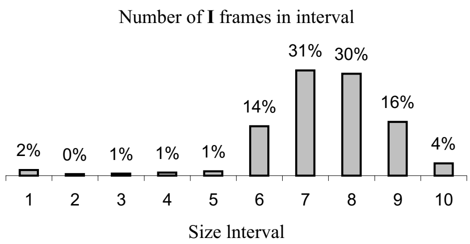
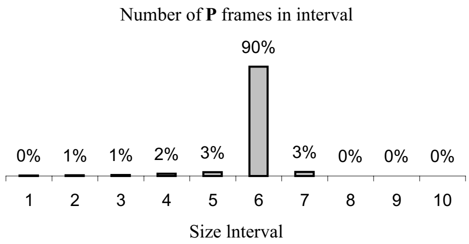


Figure 7: New Year’s Concert - Size distribution for I, P and B frames

Interval	From	To	Nr of I	Percent
1	14392	94926	134	1,7%
2	94926	175460	36	0,4%
3	175460	255994	51	0,6%
4	255994	336528	71	0,9%
5	336528	417062	102	1,3%
6	417062	497596	1165	14,5%
7	497596	578130	2480	30,9%
8	578130	658664	2403	29,9%
9	658664	739198	1267	15,8%
10	739198	819736	289	3,6%



Interval	From	To	Nr of P	Percent
1	32	76498	104	0,4%
2	76498	152964	127	0,5%
3	152964	229430	180	0,8%
4	229430	305896	446	1,9%
5	305896	382362	759	3,2%
6	382362	458828	21526	90,0%
7	458828	535294	781	3,3%
8	535294	611760	1	0,0%
9	611760	688226	3	0,0%
10	688226	764696	2	0,0%



Interval	From	To	Nr of B	Percent
1	32	42416	345	0,5%
2	42416	84800	262	0,4%
3	84800	127184	707	1,1%
4	127184	169568	1924	3,0%
5	169568	211952	42148	66,1%
6	211952	254336	18096	28,4%
7	254336	296720	200	0,3%
8	296720	339104	57	0,1%
9	339104	381488	4	0,0%
10	381488	423880	3	0,0%

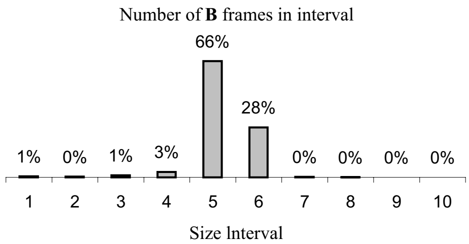


Figure 8: The Sea - Size distribution for I, P and B frames

Movie title	Avg size ratio	I frames		P frames		B frames	
	I:P:B	average	std dev	average	std dev	average	std dev
Mission Impossible 2	4:2:1	506109	187598	234821	109889	148204	57615
Leaving Las Vegas	6:3:2	471886	140329	231145	76835	152435	45746
Chicken Run	6:2:1	674549	216068	255551	133372	115185	53982
The Usual Suspect	4:2:1	514744	174307	281867	92779	129537	48890
The Matrix	3:2:1	430088	70920	249576	65226	136725	41336
New Year's Concert	6:2:1	1019897	363358	396579	98782	184918	51664
The Sea	3:2:1	568826	116259	414148	48855	199928	26295

Table 16: Comparrison of bitsize properties for all analysed movies

Movie title	Number of GOPs where					
	I largest	P largest	B largest	$P > I$	$B > I$	$B > P$
Mission Impossible 2	89%	10%	1%	31%	26%	39%
Leaving Las Vegas	94%	5%	1%	6%	2%	37%
Chicken Run	91%	8%	1%	8%	1%	12%
The Usual Suspect	88%	11%	1%	32%	15%	8%
The Matrix	93%	7%	0%	17%	3%	10%
New Year's Concert	92%	7%	0%	27%	21%	15%
The Sea	95%	4%	0%	16%	19%	4%

Table 17: Comparisson of GOP properties for all analysed movies

4 Comments on analysis results

An overview of the movies we analyzed is summarized in table 16 and 17. Here we mach the most common assumptions about MPE video streams with our analysis results.

Assumption 1: - *I frames are the largest and B frames are the smallest*. This assumption holds on average. In all the movies that we analysed, the average sizes of the I frames were larger than the average sizes of the P frames, and P frames were larger than B frames on average, with frame size ratio I:P:B = 4:2:1. Of course, the ratio depends also on the movie content, i.e., the ratio for the New Year's Concer movie that we analyzed was 6:2:1, reflecting the fact that the we have a quite static background which is not cahnged often, so the difference between current frame and the next one gets smaller. In other words, we need less bits for predicted frames.

However, our analysis showed that this assumption is not valid for a significant number of cases. For example, in "The Usual Suspect" we have a case with 11% GOPs in which P have the largest size, and 1% of B frames, which corresponds roughly to 14 and 2 minutes of the movie. Clearly, such deviations from average cannot be ignored.

Assumption 2: - *I frame is always the largest one in a GOP*. This is not true. For example in the movie "Mission Impossible 2" the P frame was larger than the I frame in 31% of the GOPs. The I frame might be the most important one in the GOP from the reconstruction point of view, but it does not necessarily has to be the largest one.

Assumption 3: - *B frames are always the smallest ones in a GOP*. Neither this assumption is true. For example, in "Mission Impossible 2" a B frame was largest in 1% of the cases. And in 39% of the cases, a B frame was larger than all P frames in the same GOP. This implies that even the assumption that P frames are always larger than B frames is also not valid. Another example is GOP nr 393 in "Mission Impossible 2" where the B frame is almost 100 times larger than the I frame ($B \approx 1MB, I \approx 12kB$).

Assumption 4: - *The sequence structure in a GOP is fixed to a specific I,P,B frame pattern*. Not true. In 23% of the GOPs in "Mission Impossible 2" the GOP length was not 12 frames. Not all GOPs consist of the same fixed number of P and B frames following the I frame in a fixed pattern. That is because more advanced encoders will attempt to optimize the placement of the three picture types according to local sequence characteristics in the context of more global characteristics. For instance scene changes or large changes in video content do not occur regularly, and hence the need for I frames in most video sequences is not at regular intervals.

Assumption 5: - *Frame properties for all movies are the same*. Neither this is true. Our analysis showed big variations between frame sizes, GOP pattern and the impact on the overall output video quality depending on the number of dropped frames. Different kinds of video will also effect the perceived quality of the video. For instance, the viewer will perceive jerky motion much easier if we drop frames in an action movie than in a cartoon.

Assumption 6: - *B and P frames are sorted in a GOP according to their sizes in descending order*. This is not true. There is no such an ordering within a GOP. As a matter of fact, our analysis showed that the largest B frames are placed towards the end of the GOP. So, the "best-effort" algorithms will perform badly when skipping the last B frames in the GOP.

Assumption 7: - *All B frames are equally important*. Not true. B sizes vary a lot. In our analysis we could see that e.g. in "Leaving Las Vegas" almost 90% of the B frames is in a pretty large interval between 6000 and 300000 bits. So, if we drop a large B frame, the entire GOP could be ruined. On the other hand, more bits does not necessarily mean better quality. That is because motion vectors give the highest compression ratio, but are smallest. So, a B frame with a lot of motion vectors would have less data than some annother frame with more row picture information, but still give better output quality when decoded All this implies that selection of B frames to be dropped should be performed carefully.

Assumption 8: - *Frame sizes vary with minor deviations from the average value*. Not true. For example, from figure 2 we can see that 88% of the I frames has bitsize between 197737 and 790684 bits ($\approx$ 200 - 800 kB), which is a quite large interval. The assumptions about MPEG based on average frame size will not hold in this case, since the significant number of frames will have twice as large repective twice as small bitsize, compared to the average frame size (which is $\approx$ 500 kB).

References

[1] ISO/IEC 13818-2: Information technology - Generic coding of moving pictures and associated audio information, Part2: Video. 1996.

[2] J. K. Ng, K. R. Leung, W. Wong, V. C. Lee, and C. K. Hui. Quality of Service for MPEG Video in Human Perspective. In *Proceedings of the 8th International Conference on Real-Time Computing Systems and Applications (RTCSA 2002)*, Tokyo, Japan, March 2002.

[3] L. Teixera and M. Martins. Video compression: The MPEG standards. In *Proceedings of the 1st European Conference on Multimedia Applications Services and Techniques (ECMAST 1996)*, Louvian-la-Neuve, Belgium, May 1996.

[4] J. Watkinson. *The MPEG handbook*. ISBN 0 240 51656 7, Focal Press, 2001.