



ESE150 Spring 2021



**ESE**



Lecture #12 – Psychoacoustic Compression Algorithm

**ESE 150 –  
DIGITAL AUDIO BASICS**

Based on slides © 2009–2021 DeHon, Koditschek  
Additional Material © 2014 Farmer

1

ESE150 Spring 2021

## LECTURE TOPICS

- × **Lab vs. Lecture**
- × **Big Picture reminder**
- × **Review and preclass**
- × **How do we take advantage of psychoacoustics in MP3**  
**Achieve this 6–12x reduction from CD Audio**
  - + Review Tricks
  - + Formulate Optimization
  - + **Algorithm for Adaptation**
- × **References**

2

ESE150 Spring 2021

## COMPARE TO LAB

- × **Lab 6**
  - + Capture spirit of reducing frequencies
  - + Simplified – only taking loudest fraction in each band
    - Rather than being rigorous about masking
    - Or trying to hit some fixed rate
- × **Lecture 11 & 12**
  - + Help understand more what real MP3 encoding looks like
  - + Thinking about fixed rate
    - And adaptation for variable rate encoding from Huffman
  - + Formulating masking more explicitly
    - But still simplistic
  - + Illustrating Optimization Approaches

3

ESE150 Spring 2021

## BIG PICTURE: WHY CARE

- × **MP3 shows**
  - + We can reduce bits needed too accurately represent music by 12x over PCM samples
- × **1998**
  - + Necessary to make MP3 player viable at all on 32MB Flash memory
    - Small, affordable by masses
  - + Also meant lower bandwidth to stream
- × **2006**
  - + Necessary to make Spotify streaming viable
- × **2021**
  - + With 64GB phones, maybe not essential for music on iPhone, but...

4

ESE150 Spring 2021

## BIG PICTURE: WHY CARE

- × **Networks / Internet bandwidth / Cell capacity**
  - + Allows us to reduce bandwidth needed for audio
  - + Alternative is
    - Building out more wires and capacity
    - Fewer cell calls supported simultaneously in a region
  - + Being efficient with bits
    - Reduces physical investment needed in wires
- × **Same ideas enable video**
  - + Viable to store (many) hours of video on phone/tablet
    - Of increasing resolution: HD, 4K...
  - + Viable to stream video into homes
  - + Zoom, Netflix, AppleTV, ... all enabled by compression

5

ESE150 Spring 2021

## REVIEW AND PRECLASS SETUP/REMINDER

6

## KNOBS WE CAN TURN

- × **Amplitude quantization**
  - + Per band
- × **Frequency quantization**
  - + Per band?
- × **Frequencies kept (per critical band)**
  - + Per band
- × **...and can perform lossless compression**

7

## PRECLASS 1

- × **44,300 samples/s**
- × **16b**
- × **26ms window**
- × **a) How many bits?**
- × 128Kb/s stereo → 64Kb/s per audio channel
- × **b) How many bits per 26ms window?**
- × **c) ratio?**

8

## OPTIMIZATION PROBLEM

- × **How fit in the resource constraints (128Kb/s) while maximizing goodness (sound quality)?**

- × **Quantify bits used:**  $\sum_{bands} \sum_{f \in freqs} Bits(f)$ 
  - + Cannot exceed 128Kb/s
  - + = 1,704 b / 26ms frame / channel

- × **Quantify goodness: minimize**  $\sum_{f \in freqs} Error(f) \times W(f)$

9

## PRECLASS 2

- × 576 frequencies
- × 16b amplitude
- × 1704b budget
- × (frequency, amplitude) pairs to represent
- × **How many frequencies can we keep?**
- × **Conclude:** cannot keep all frequencies and hit budget

10

## PRECLASS 3

- × **Bits if only have 3 non-masked, non-zero frequencies?**
- × **Conclude:** some frames won't use all their bits

11

## PRECLASS 4 AND 5

- × 60dB, 42dB, 30dB each are 25% of amplitudes
- × Other magnitudes remaining 25%
- × **How many bits to represent 80 non-zero/non-masked frequencies?**
- × **If all frequencies equally likely, how many bits to represent 80 non-zero/non-masked frequencies?**
- × **Conclude:** number of frequencies can keep depends on compressibility

12

ESE150 Spring 2021

## SEEN

- × **May want to do something smarter than**
  - + Allocating fixed number of frequencies per band
  - + Allocating fixed quantization to a band
- × **Like to adapt our encoding to the data**
  - + If more Huffman compressible, we get more frequencies
  - + If fewer frequencies suffice for one band,
    - × Allow more frequencies for another
    - × ...or allocate less quantization

13

ESE150 Spring 2021

## OPTIMIZATION PROBLEM

- × **How fit in the resource constraints (128Kb/s) while maximizing goodness (sound quality)?**
- × **Optimization problems central to engineering**

14

ESE150 Spring 2021

## OPTIMIZATION PROBLEM

- × **How fit in the resource constraints (128Kb/s) while maximizing goodness (sound quality)?**
- × **Quantify bits used:**

$$\sum_{bands} \sum_{f \in freqs} Bits(f)$$
  - + Cannot exceed 128Kb/s
  - + = 1,704 b / 26ms frame / channel
- × **Quantify goodness: minimize**

$$\sum_{f \in freqs} Error(f) \times W(f)$$

15

ESE150 Spring 2021

## GOODNESS/SOUND QUALITY

- × **Error(Amp) = |Orig Amplitude – Encoded|**
  - + Whole OrigAmplitude if dropped
  - + |Orig Amplitude-Quantize(OrigAmplitude,bits)| if quantized
- × **W(freq)**
  - + 0 if below hearing threshold
  - + 0 if masked
  - + Value between 0 and 5 if partially masked in critical band
  - + Really depend on what already encoded
$$\sum_{f \in freqs} Error(f) \times W(f)$$

16

ESE150 Spring 2021

## EXAMPLE WEIGHT FUNCTION W(F)

- × **W(f)=CBWeight\*Mask**
- × **Mask = 0 if MaxAmp-FreqAmp<4**
- × **(MaxAmp-FreqAmp)/MaxAmap otherwise**
- + 1 otherwise
- + For example that follows, don't use second part

17

ESE150 Spring 2021

## APPROACH

$$\sum_{bands} \sum_{f \in freqs} Bits(f)$$

- × **Start with nothing**
- × **Start with bitbudget**

$$\sum_{f \in freqs} Error(f) \times W(f)$$

18

ESE150 Spring 2021

### APPROACH

$$\sum_{bands} \sum_{f \in freqs} Bits(f)$$

- Start with nothing
- Start with bitbudget

$$\sum_{f \in freqs} Error(f) \times W(f)$$

	Band 0	Band 1	Band 2	Band 3	
frequency	11 8 2 11	8 1 1 3	1 8 4	41 1	
sum	14	13	13	6	
weight	3	5	5	3	
weighted	42	65	65	18	

Bits = 0  
Error = 190

19

ESE150 Spring 2021

### APPROACH

$$\sum_{bands} \sum_{f \in freqs} Bits(f)$$

- Start with nothing
- Start with bitbudget
- While(bitbudget > 0)
  - Identify Largest Error component:  $Error(freq) \times W(freq)$

	Band 0	Band 1	Band 2	Band 3	
frequency	11 8 2 11	8 1 1 3	1 8 4	41 1	
sum	14	13	13	6	
weight	3	5	5	3	
weighted	42	65	65	18	

Bits = 0  
Error = 190

20

ESE150 Spring 2021

### APPROACH

$$\sum_{bands} \sum_{f \in freqs} Bits(f)$$

- Start with nothing
- Start with bitbudget

$$\sum_{f \in freqs} Error(f) \times W(f)$$

	Band 0	Band 1	Band 2	Band 3	
frequency	11 8 2 11	0 0 0 0	1 8 4	41 1	
sum	14	0	13	6	
weight	3	5	5	3	
weighted	42	0	65	18	

Bits = 3+6=9  
Error = 125

21

ESE150 Spring 2021

### APPROACH

$$\sum_{bands} \sum_{f \in freqs} Bits(f)$$

- Start with nothing
- Start with bitbudget

$$\sum_{f \in freqs} Error(f) \times W(f)$$

	Band 0	Band 1	Band 2	Band 3	
frequency	11 8 2 11	0 0 0 0	0 0 0	41 1	
sum	14	0	0	6	
weight	3	5	5	3	
weighted	42	0	0	18	

Bits = 18  
Error = 60

22

ESE150 Spring 2021

### APPROACH

$$\sum_{bands} \sum_{f \in freqs} Bits(f)$$

- Start with nothing
- Start with bitbudget

$$\sum_{f \in freqs} Error(f) \times W(f)$$

	Band 0	Band 1	Band 2	Band 3	
frequency	0 0 0 0	0 0 0	41 1		
sum	0	0	6		
weight	3	5	5	3	
weighted	0	0	18		

Bits = 27  
Error = 18

23

ESE150 Spring 2021

### ALSO QUANTIZATION

$$\sum_{bands} \sum_{f \in freqs} Bits(f)$$

- To keep simple, assumed fixed quant.
- Incrementally assign bits
  - While(bitbudget > 0)
    - Identify Largest Error component:  $Error(freq) \times W(freq)$
    - Assign more bits to that frequency
      - Go from 0 bits to 1 bit
      - 1 bit to 2 bits

ESE150 Spring 2021

## APPROACH (GREEDY)

- × Start with nothing
- × Start with bitbudget
- × While(bitbudget>0)
  - + Identify Largest Error component
  - + Allocate some bits to reduce error
    - Add frequency or Add quantization bits to band
    - Pick one to most reduce the error

25

ESE150 Spring 2021

## ADAPTIVE REFINEMENT

- × Rediscovering where to allocate everything every time may be laborious
  - + Often same frequencies persist for more than 26ms
- × Maybe we can get close and adjust?
  - + Use critical band allocation from previous frames as a starting point guess
    - × bits, frequencies, quantization
  - + Try initial encoding with that

26

ESE150 Spring 2021

## APPROACH (ADAPTIVE)

- × Start with budget guess
  - + Quantization in bands
  - + Frequencies to keep in each band
- × Encode, compress
- × What do if haven't used up all bits?
- × What do if over budget?

27

ESE150 Spring 2021

Roundup

# PERCEPTUAL CODING & MP3

28

ESE150 Spring 2021

## MP3 FLOW CHART (ENCODING/DECODING)

*A type of DFT is performed*

**Basic MP3 Encoding Scheme**

PCM audio input → Time to frequency mapping filter bank → Bit/noise allocation, quantizer, and coding → Bitstream formatting → Encoded bitstream

Psychoacoustic model → Bitstream formatting

Ancillary data (optional) → Bitstream formatting

*Data is stored in frequency domain rep. Even Huffman coded too*

*We use psychoacoustic model to control quantization! (how we achieve compression)*

**MP3 Decoding Scheme**

Encoded bitstream → Bitstream unpacking → Frequency sample reconstruction → Frequency to time mapping → Decoded PCM audio

Ancillary data (if included) → Frequency sample reconstruction

*Notice: Psychoacoustics is embedded Your brain is involved in decoding!*

29

ESE150 Spring 2021

## MP3 ENCODING PROCESS

- × All MP3 files broken into "Frames"
  - + Each frame stores 1152 Audio Samples
  - + Lasts for 26 ms
  - + Frame also divided further into 2 "granuels"
    - × Each granuel contains 576 samples

30

ESE150 Spring 2021

## MP3 ENCODING PROCESS

acquire & transform "frame"

- Bring in 1152 PCM samples
- Use DFT/MDCT to transform to frequency-domain

Band	Center Freq	Bandwidth	Band	Center Freq	Bandwidth	Band	Center Freq	Bandwidth
1	100	100	11	1170	1270-1480	20	8000	7100-8000
2	150	100	12	1310	1480-1700	21	8500	8100-9000
3	200	100	13	1450	1700-1950	22	9000	9000-10000
4	250	100	14	1600	1950-2200	23	9500	10000-11000
5	300	100	15	1750	2200-2500	24	10000	11000-12000
6	350	100	16	1900	2500-2850	25	10500	12000-13000
7	400	100	17	2050	2850-3250	26	11000	13000-14000
8	450	100	18	2200	3250-3700	27	11500	14000-15000
9	500	100	19	2350	3700-4200	28	12000	15000-16000
10	550	100	20	2500	4200-4800	29	12500	16000-17000
11	600	100	21	2650	4800-5500	30	13000	17000-18000
12	650	100	22	2800	5500-6300	31	13500	18000-19000
13	700	100	23	2950	6300-7200	32	14000	19000-20000
14	750	100	24	3100	7200-8200	33	14500	20000-21000
15	800	100	25	3250	8200-9300	34	15000	21000-22000
16	850	100	26	3400	9300-10500	35	15500	22000-23000
17	900	100	27	3550	10500-11800	36	16000	23000-24000
18	950	100	28	3700	11800-13200	37	16500	24000-25000
19	1000	100	29	3850	13200-14800	38	17000	25000-26000
20	1050	100	30	4000	14800-16500	39	17500	26000-27000
21	1100	100	31	4150	16500-18500	40	18000	27000-28000
22	1150	100	32	4300	18500-20500	41	18500	28000-29000
23	1200	100	33	4450	20500-22800	42	19000	29000-30000
24	1250	100	34	4600	22800-25500	43	19500	30000-31000
25	1300	100	35	4750	25500-28500	44	20000	31000-32000
26	1350	100	36	4900	28500-31800	45	20500	32000-33000
27	1400	100	37	5050	31800-35500	46	21000	33000-34000
28	1450	100	38	5200	35500-39500	47	21500	34000-35000
29	1500	100	39	5350	39500-44000	48	22000	35000-36000
30	1550	100	40	5500	44000-49000	49	22500	36000-37000
31	1600	100	41	5650	49000-54500	50	23000	37000-38000
32	1650	100	42	5800	54500-60500	51	23500	38000-39000
33	1700	100	43	5950	60500-67500	52	24000	39000-40000
34	1750	100	44	6100	67500-75000	53	24500	40000-41000
35	1800	100	45	6250	75000-83000	54	25000	41000-42000
36	1850	100	46	6400	83000-91500	55	25500	42000-43000
37	1900	100	47	6550	91500-101000	56	26000	43000-44000
38	1950	100	48	6700	101000-111500	57	26500	44000-45000
39	2000	100	49	6850	111500-123000	58	27000	45000-46000
40	2050	100	50	7000	123000-135500	59	27500	46000-47000
41	2100	100	51	7150	135500-149000	60	28000	47000-48000
42	2150	100	52	7300	149000-163500	61	28500	48000-49000
43	2200	100	53	7450	163500-179000	62	29000	49000-50000
44	2250	100	54	7600	179000-195500	63	29500	50000-51000
45	2300	100	55	7750	195500-213000	64	30000	51000-52000
46	2350	100	56	7900	213000-231500	65	30500	52000-53000
47	2400	100	57	8050	231500-251000	66	31000	53000-54000
48	2450	100	58	8200	251000-271500	67	31500	54000-55000
49	2500	100	59	8350	271500-293000	68	32000	55000-56000
50	2550	100	60	8500	293000-315500	69	32500	56000-57000
51	2600	100	61	8650	315500-339000	70	33000	57000-58000
52	2650	100	62	8800	339000-363500	71	33500	58000-59000
53	2700	100	63	8950	363500-389000	72	34000	59000-60000
54	2750	100	64	9100	389000-415500	73	34500	60000-61000
55	2800	100	65	9250	415500-443000	74	35000	61000-62000
56	2850	100	66	9400	443000-471500	75	35500	62000-63000
57	2900	100	67	9550	471500-501000	76	36000	63000-64000
58	2950	100	68	9700	501000-531500	77	36500	64000-65000
59	3000	100	69	9850	531500-563000	78	37000	65000-66000
60	3050	100	70	10000	563000-595500	79	37500	66000-67000
61	3100	100	71	10150	595500-629000	80	38000	67000-68000
62	3150	100	72	10300	629000-663500	81	38500	68000-69000
63	3200	100	73	10450	663500-700000	82	39000	69000-70000
64	3250	100	74	10600	700000-737500	83	39500	70000-71000
65	3300	100	75	10750	737500-776000	84	40000	71000-72000
66	3350	100	76	10900	776000-815500	85	40500	72000-73000
67	3400	100	77	11050	815500-856000	86	41000	73000-74000
68	3450	100	78	11200	856000-897500	87	41500	74000-75000
69	3500	100	79	11350	897500-940000	88	42000	75000-76000
70	3550	100	80	11500	940000-983500	89	42500	76000-77000
71	3600	100	81	11650	983500-1028000	90	43000	77000-78000
72	3650	100	82	11800	1028000-1073500	91	43500	78000-79000
73	3700	100	83	11950	1073500-1120000	92	44000	79000-80000
74	3750	100	84	12100	1120000-1167500	93	44500	80000-81000
75	3800	100	85	12250	1167500-1216000	94	45000	81000-82000
76	3850	100	86	12400	1216000-1265500	95	45500	82000-83000
77	3900	100	87	12550	1265500-1316000	96	46000	83000-84000
78	3950	100	88	12700	1316000-1367500	97	46500	84000-85000
79	4000	100	89	12850	1367500-1420000	98	47000	85000-86000
80	4050	100	90	13000	1420000-1473500	99	47500	86000-87000
81	4100	100	91	13150	1473500-1528000	100	48000	87000-88000
82	4150	100	92	13300	1528000-1583500	101	48500	88000-89000
83	4200	100	93	13450	1583500-1640000	102	49000	89000-90000
84	4250	100	94	13600	1640000-1697500	103	49500	90000-91000
85	4300	100	95	13750	1697500-1756000	104	50000	91000-92000
86	4350	100	96	13900	1756000-1815500	105	50500	92000-93000
87	4400	100	97	14050	1815500-1876000	106	51000	93000-94000
88	4450	100	98	14200	1876000-1937500	107	51500	94000-95000
89	4500	100	99	14350	1937500-1999000	108	52000	95000-96000
90	4550	100	100	14500	1999000-2061500	109	52500	96000-97000
91	4600	100	101	14650	2061500-2125000	110	53000	97000-98000
92	4650	100	102	14800	2125000-2189500	111	53500	98000-99000
93	4700	100	103	14950	2189500-2255000	112	54000	99000-100000
94	4750	100	104	15100	2255000-2321500	113	54500	100000-101000
95	4800	100	105	15250	2321500-2389000	114	55000	101000-102000
96	4850	100	106	15400	2389000-2457500	115	55500	102000-103000
97	4900	100	107	15550	2457500-2527000	116	56000	103000-104000
98	4950	100	108	15700	2527000-2597500	117	56500	104000-105000
99	5000	100	109	15850	2597500-2669000	118	57000	105000-106000
100	5050	100	110	16000	2669000-2741500	119	57500	106000-107000
101	5100	100	111	16150	2741500-2815000	120	58000	107000-108000
102	5150	100	112	16300	2815000-2889500	121	58500	108000-109000
103	5200	100	113	16450	2889500-2965000	122	59000	109000-110000
104	5250	100	114	16600	2965000-3041500	123	59500	110000-111000
105	5300	100	115	16750	3041500-3119000	124	60000	111000-112000
106	5350	100	116	16900	3119000-3197500	125	60500	112000-113000
107	5400	100	117	17050	3197500-3277000	126	61000	113000-114000
108	5450	100	118	17200	3277000-3357500	127	61500	114000-115000
109	5500	100	119	17350	3357500-3439000	128	62000	115000-116000
110	5550	100	120	17500	3439000-3521500	129	62500	116000-117000
111	5600	100	121	17650	3521500-3605000	130	63000	117000-118000
112	5650	100	122	17				