

Audio and Visual Processing (CMP-7016A)

Chandupraveen Gudi
Zhongye Chen

Overview:

- Image and Video
- Data collection and labeling
- Image cropping
- Binary image features
- DCT features
- Data interpolation
- Evaluation - Audio, Visual, Audio-Visual

Image:

- 2D grid or matrix, represented as a function $i(x,y)$
- Binary Image - (0 and 1)
- Grayscale Image - (0 to 255)
- RGB Image - $R(x,y)$, $G(x,y)$, $B(x,y)$

binary image



gray image



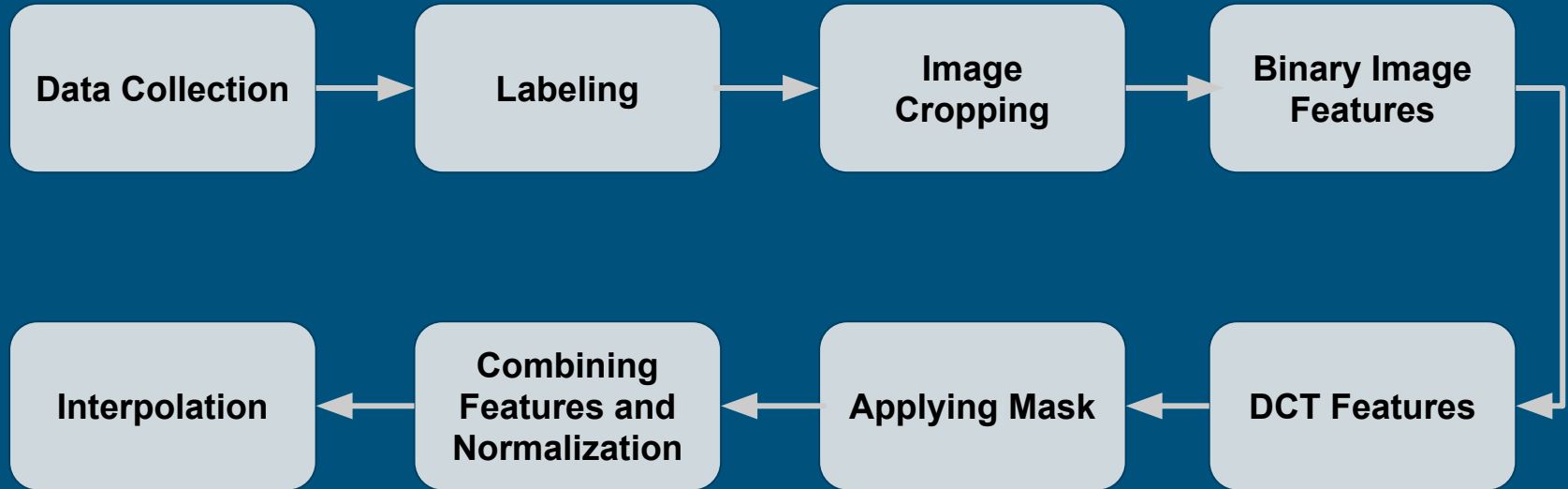
RGB image



Video and Video Processing:

- Sequence of Images (Frames)
- Image/Video enhancement.
- Compression for storing.
- Object classification in images/video.
- Collecting articulatory data (lips, teeth, tongue tip, mouth localization)

Feature Extraction Pipeline:



Data Collection : Video Recording

Collected data in two ways:

1. 20 different names - one time in one video

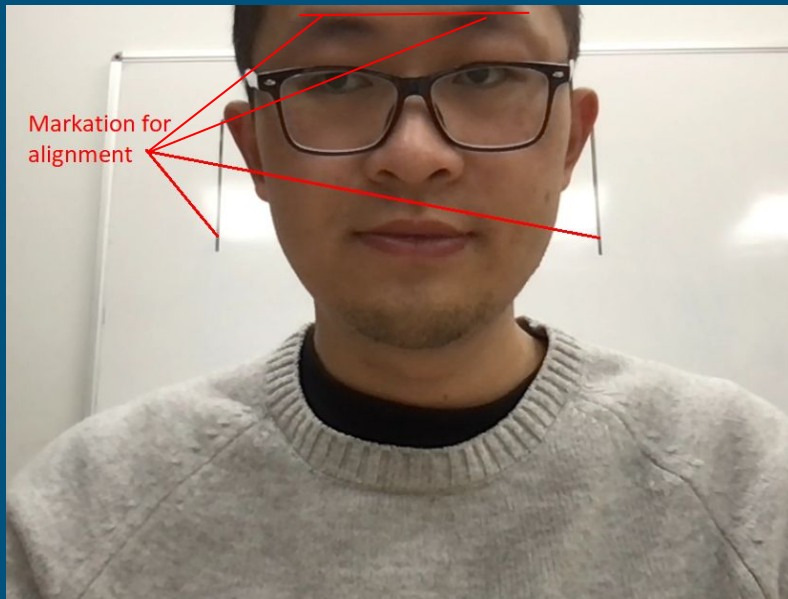
Eg: File1- Alex, Amelia, Anushka, ..., Will

2. One name - 20 times in one video

Eg: File1- Alex, Alex, Alex, ..., Alex

Data Collection : Video Recording

- Videos recorded in UEA library.
- Camera used : Macbook Pro 720p HD

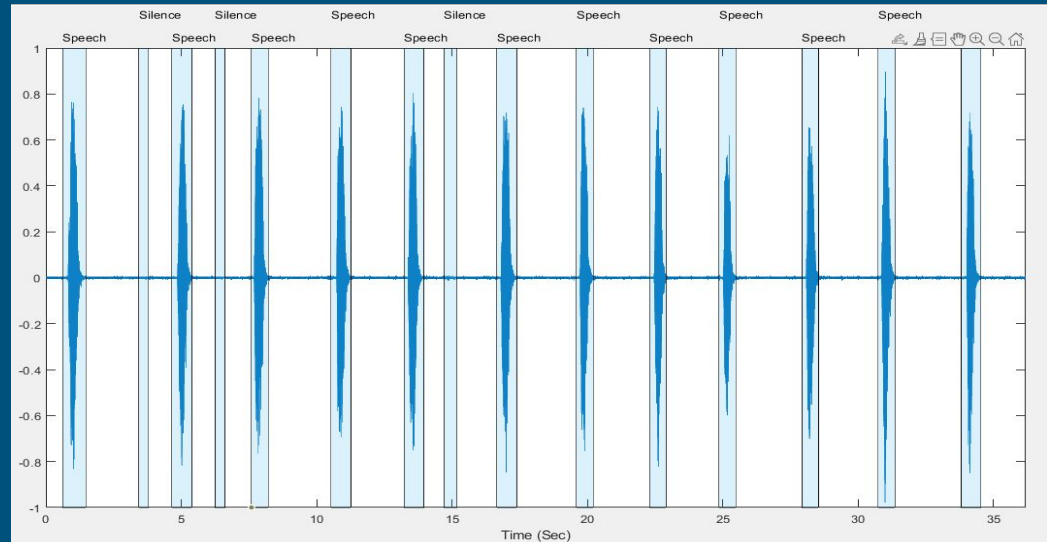


Data Collection : Audio

- Audio is extracted from video using **audioread** method available in Matlab.
- Normalized the audio signal to avoid clipping.
- Collected data with 40000 sampling frequency.
- Sampling frequency reduced to 16000 using **resample** method.
- Save the audio files using **audiowrite** method.

Data Collection : Labeling

- Labeling process is automated.
- Used **ClassifySound** method available in Matlab.
 - <https://uk.mathworks.com/help/audio/ref/classifysound.html>
- Reduces manual effort.
- Gives more accuracy.

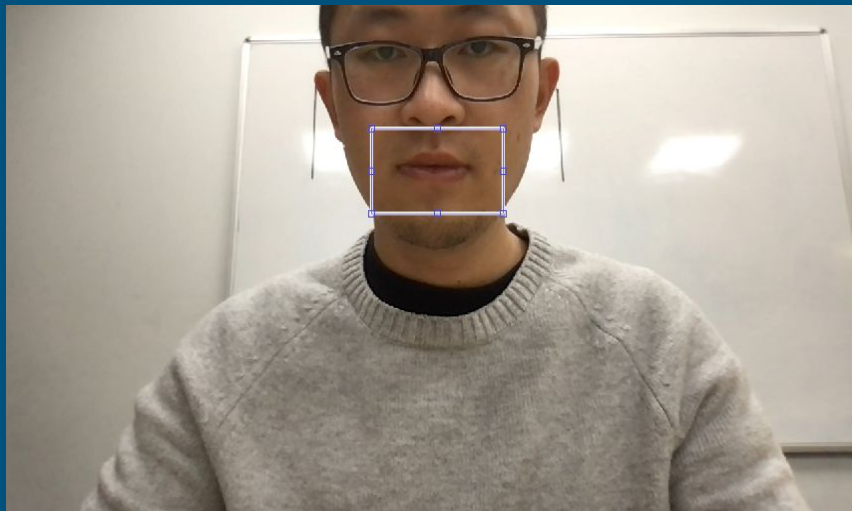


Video Processing : Image Cropping - Manual

- Used **imcrop** method to get the position and size of the crop rectangle.

Eg: [croppedImage,rect] = imcrop(Image);
croppedImage → Selected cropped image
rect → [xmin ymin width height]

- By taking the dimensions of first frame, applied the same to all the frames in video.

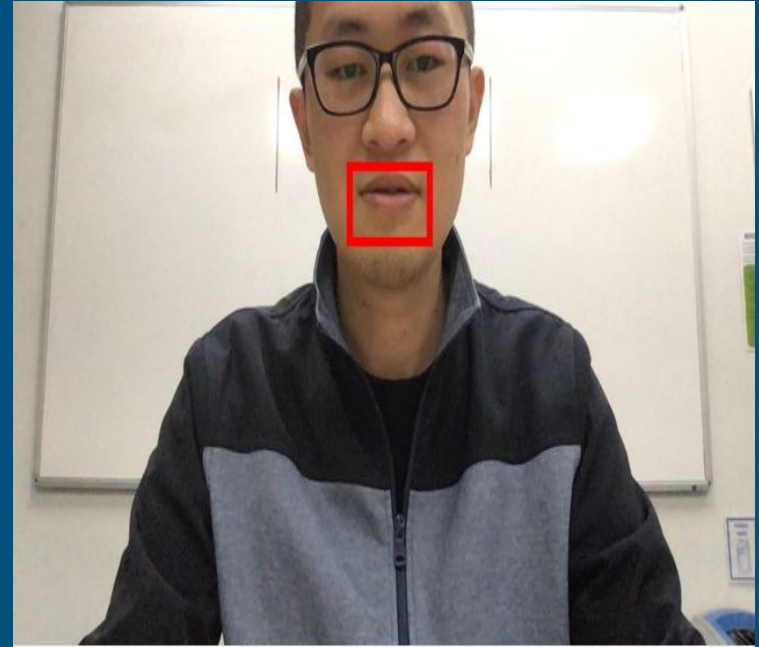


Video Processing : Image Cropping - Automated

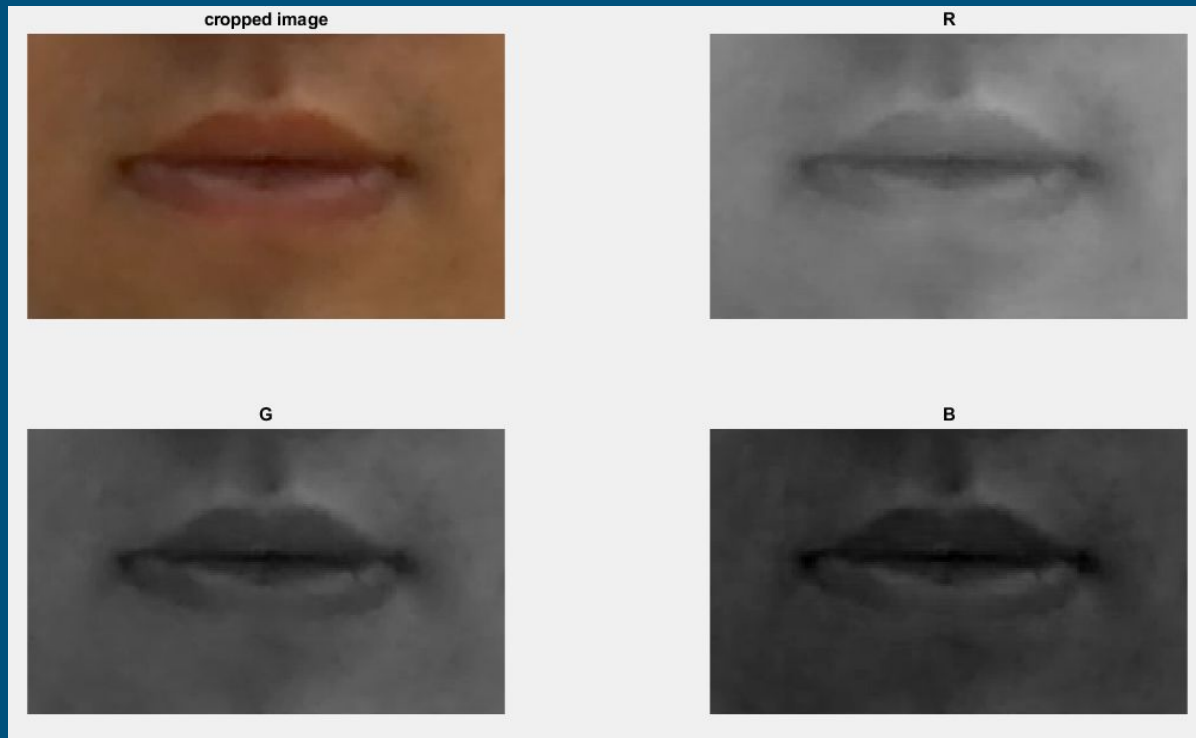
- Used **vision.CascadeObjectDetector** method for automatic detection.

```
img_frame = imresize(img_frame,0.5);  
detector = vision.CascadeObjectDetector('Mouth');  
detector.MergeThreshold = 100;  
bbox = step(detector, img_frame);  
cropped_image = imcrop(img_frame, bbox(1,:));
```

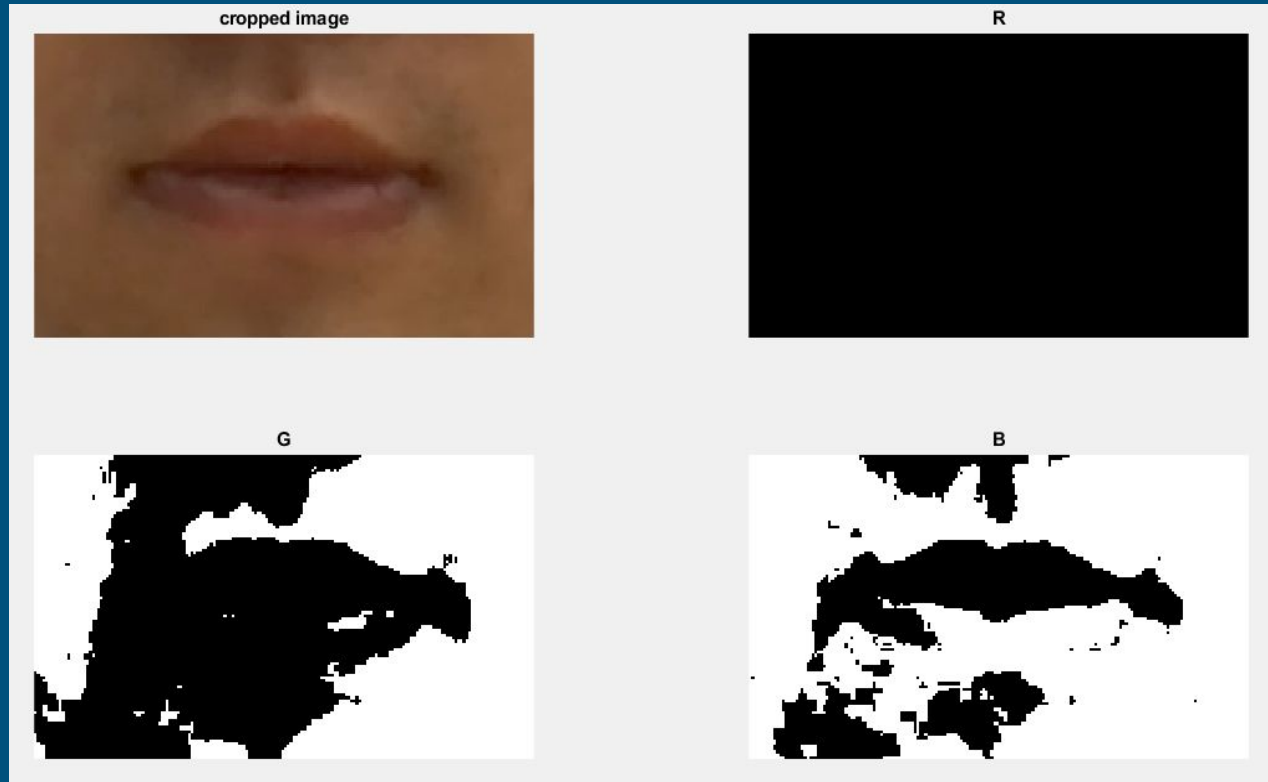
- Time consuming process.
- Gives promising results, even when the position of the face is changed.



Binary Image Features: Normalization



Binary Image Features: Normalization

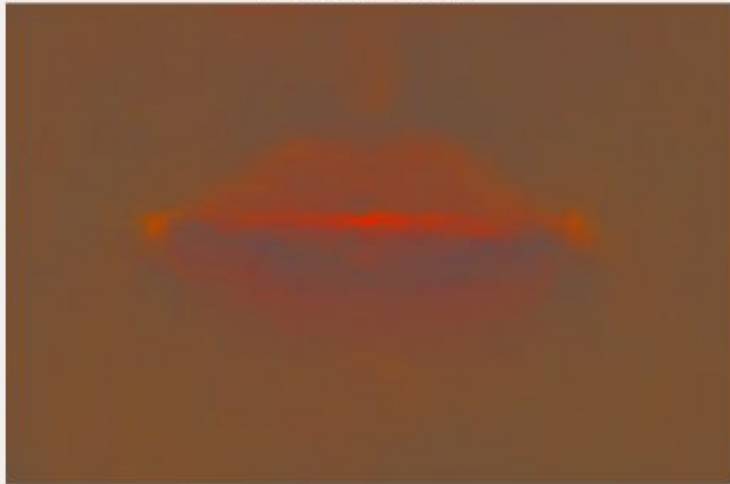


Binary Image Features: Normalization

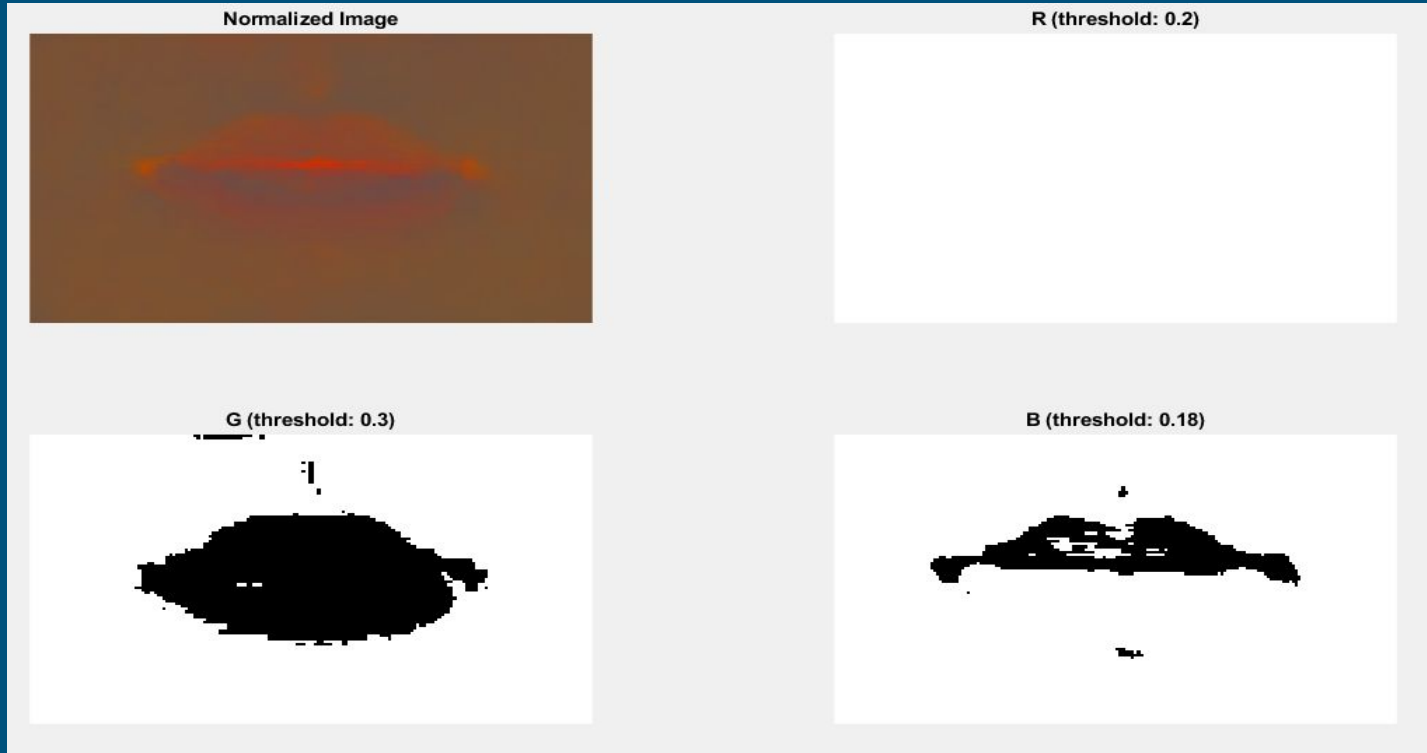
cropped image



normalized image

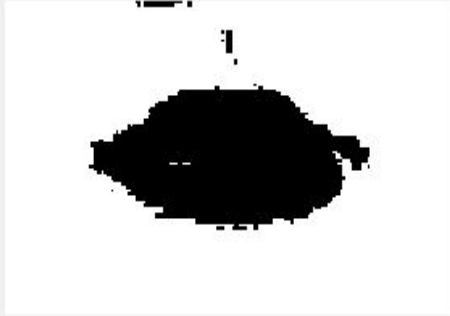


Binary Image Features: Normalization

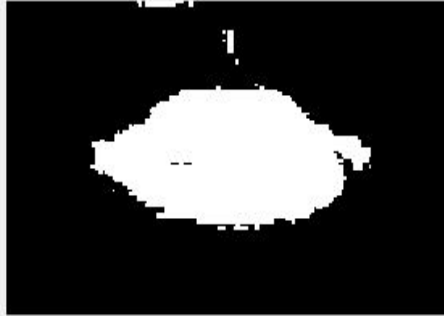


Binary Image Features: Morphological filters

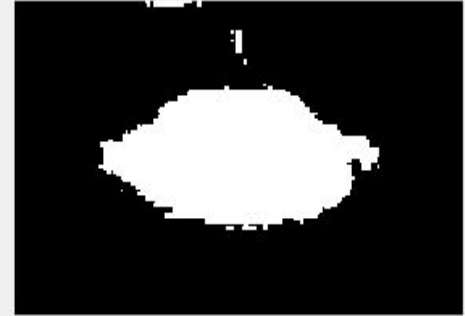
G (threshold: 0.3)



Imcomplement



filling holes



Binary Image Features: Regionprops

Centroid		BoundingBox				MajorAxisLength	MinorAxisLength
91.565	54.089	34.5	29.5	114	49	102.68	46.489
44	65	43.5	64.5	1	1	1.1547	1.1547
54.667	1.3333	53.5	0.5	2	2	2.582	1.7638
63.828	1.4483	56.5	0.5	16	2	17.273	2.2738
75.5	1.5	74.5	0.5	2	2	2.3094	2.3094
88	78	86.5	77.5	3	1	3.4641	1.1547
89	11	88.5	10.5	1	1	1.1547	1.1547
89	14	88.5	13.5	1	1	1.1547	1.1547
91.5	14.5	90.5	10.5	2	8	9.2376	2.3094
94	21.5	93.5	20.5	1	2	2.3094	1.1547
102	29	101.5	28.5	1	1	1.1547	1.1547
148	57	147.5	56.5	1	1	1.1547	1.1547

Video Processing : DCT features

- Convert the image to gray.
- Apply DCT to gray image.
- DCT processing only get real (Cosine) part of image.
- Cosine/real part of image has a strong "energy compaction" property.
- Most of the signal information tends to be concentrated in a few low-frequency components of the DCT.

Various Masks Used

- Square Shape of mask 10×10 in the top left corner of DCT Data
- Triangular Shape of mask in the top left corner of DCT Data
- New Mask(explain later)

Triangle Mask

Multiply the mask by the image you want to process. Keep in mind mask and image must be the same size.

- Right bottom data is too small
- Left top corner present more information

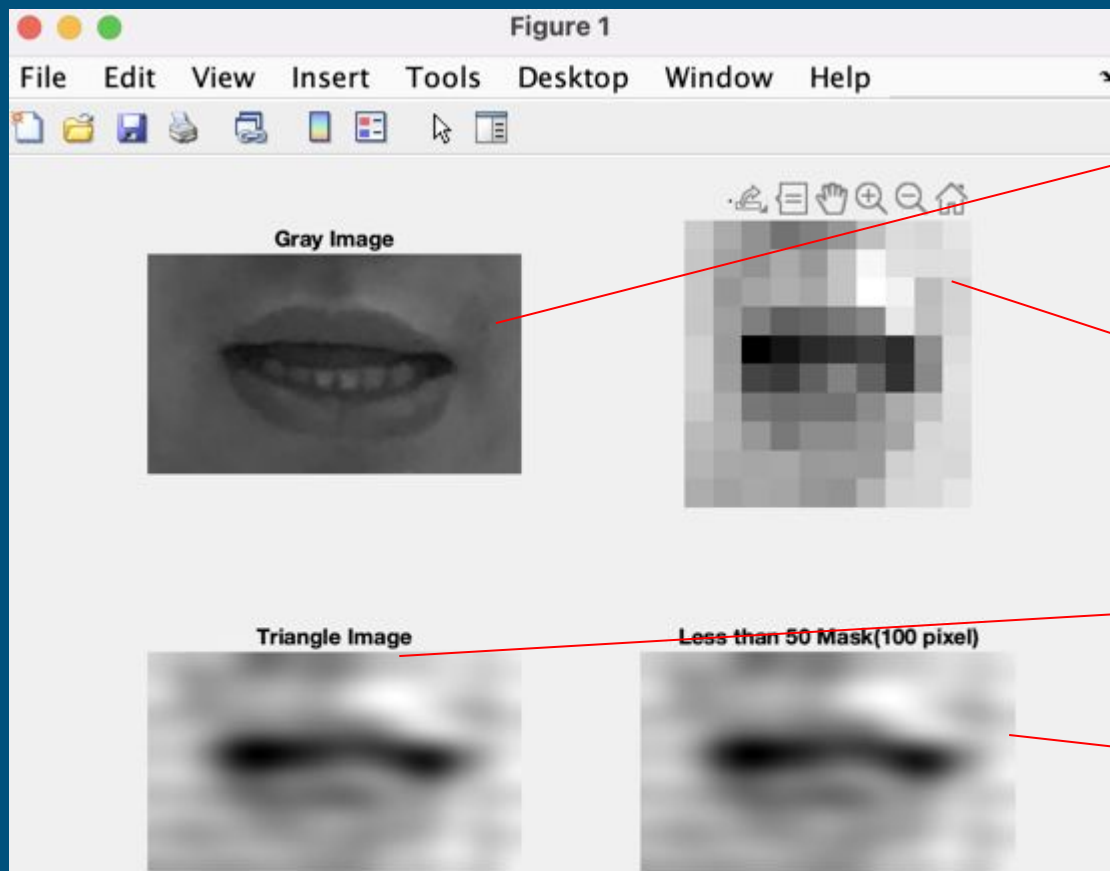
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	
2	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	
3	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	
4	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	
5	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	
6	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	
7	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	
8	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	
9	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	
10	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	
11	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
12	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
13	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	

613	566	190	0	0	0	19	9
101	65	0	0	0	47	40	14
0	0	12	42	38	11	0	0
3	10	12	0	0	0	0	2
1	0	0	0	13	22	14	3
18	10	0	0	5	2	0	0
0	0	0	6	14	11	1	0
2	5	5	0	0	0	0	2

Figure from Lecture 13

Original gray image size is 151*104
Remove Absolute Value Less Than 50

[illegible]



Original Image

Square Shape of Mask

Triangle Shape of Mask

New Mask

Frequency of Pixel Value Higher than 50

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL	AN	AO					
1	1897	1897	1897	1897	1353	916	1873	1814	959	116	1796	465	1081	258	1575	130	286	44	4	0	10	0	0	0	2	0	7	34	139	131	127	41	7	0	0	0	0	0	0	0	0	0	0		
2	1825	891	1749	1897	1797	1454	1440	545	342	145	33	231	1299	284	1148	22	14	1	23	1	0	0	0	0	0	0	8	5	24	40	22	0	7	1	0	0	0	0	0	0	0	0	0		
3	1897	1897	1897	1808	1897	1458	984	1497	1735	652	983	571	531	225	1356	43	162	29	0	4	10	0	10	1	13	7	24	71	196	138	86	18	31	4	0	0	0	0	0	0	0	0	0		
4	1609	1034	1628	390	1488	1588	1822	794	1048	698	1500	326	190	125	155	169	620	47	0	3	0	3	8	4	7	12	34	57	108	120	107	85	20	5	0	0	0	0	0	0	0	0	0		
5	1886	730	1839	524	1883	459	815	309	463	181	1112	120	241	13	150	57	142	9	39	3	2	0	15	19	7	0	34	20	81	28	48	30	16	1	4	0	0	0	0	0	0	0			
6	1883	1215	1897	1442	1519	767	728	889	1297	400	1301	50	130	3	230	458	687	62	16	0	44	4	5	3	0	8	12	15	80	97	73	7	13	0	4	0	0	0	0	0	0	0			
7	1619	586	1414	483	1418	1298	1545	884	1869	608	1116	126	456	22	321	91	62	4	19	1	73	0	5	0	1	2	22	19	38	34	31	63	9	2	0	0	0	0	0	0	0	0	0		
8	1800	605	1473	338	1465	508	1394	525	1559	326	518	79	423	324	721	148	88	1	28	10	45	1	10	1	0	1	19	2	41	15	28	4	12	0	0	0	0	0	0	0	0	0	0		
9	1672	138	1590	256	1196	333	1439	898	1234	210	544	277	848	163	96	16	59	28	4	8	2	4	6	0	7	7	10	2	9	3	8	7	7	0	0	0	0	0	0	0	0	0	0		
10	1598	400	1408	505	1019	340	1513	677	1318	235	557	266	449	72	54	61	93	63	27	15	26	10	8	3	10	5	1	0	7	0	7	4	0	0	0	0	0	0	0	0	0	0	0		
11	1583	687	1285	345	1279	294	1479	388	595	258	914	308	487	92	69	75	29	61	21	41	11	6	2	15	3	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
12	1398	117	730	647	1390	491	1286	506	715	297	1145	447	402	86	78	67	22	56	34	35	6	2	15	10	4	5	1	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
13	1224	28	877	667	1431	477	1124	787	729	310	1026	346	172	95	85	82	23	53	23	49	7	8	10	31	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	1488	153	1256	705	1300	419	567	478	744	286	932	217	116	56	25	46	10	9	14	20	10	4	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
15	1376	251	1323	301	1385	158	393	110	682	273	602	96	103	57	96	23	0	2	9	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
16	1384	17	1412	137	1216	55	227	75	691	138	564	78	187	74	71	16	3	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
17	1050	3	1029	74	884	25	70	22	444	32	228	52	263	20	14	10	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
18	852	2	733	36	558	28	70	15	319	13	34	49	143	6	13	12	0	2	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
19	808	0	527	21	344	22	38	2	175	5	22	14	39	9	0	2	1	6	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
20	342	2	226	13	255	9	21	8	70	6	3	10	25	7	5	1	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
21	192	0	85	5	123	9	15	2	14	8	0	0	2	3	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
22	159	1	84	7	144	4	18	4	31	2	0	0	1	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
23	139	0	73	3	172	3	9	0	8	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
24	43	0	106	1	186	1	20	1	4	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
25	31	0	109	2	90	0	7	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
26	32	0	61	0	82	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
27	19	0	57	1	42	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
28	6	0	31	0	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
29	1	0	13	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
30	3	0	11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
31	1	0	5	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
32	0	0	4	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
33	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
34	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	

Combining DCT and Binary Data

- Combine 100 DCT data from gray image and 4 data from binary image

Visual Only Model in three Masks

	Square Shape Mask	Triangle Shape Mask	NewMask
Sentence Correctly	65%	60%	65%
word Correctly	96.06%	96.06%	100%
Word Accuracy	92.12%	93.10%	92%

Visual Only Using Square Shape of Mask

```
----- Overall Results -----
SENT: %Correct=65.00 [H=13, S=7, N=20]
WORD: %Corr=96.06, Acc=92.12 [H=195, D=5, S=3, I=8, N=203]
----- Confusion Matrix -----
      A  A  A  C  C  E  G  J  J  J  J  J  L  M  R  S  T  T  T  W
      l  m  n  h  h  r  e  a  a  o  o  o  u  a  o  h  h  i  o  i
      e  e  u  a  a  g  o  k  r  e  r  s  k  t  b  a  o  m  b  l
      x  l  s  n  r  u  r  e  d  d  e  a  t  u  m  y  l
      i  h  d  l  n  g  e  a  p  s  n  a
Alex  8   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
Amel  0  10   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   2
Anus  0   0  10   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
Chan  0   0   0  10   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
Char  0   0   0   0  11   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
Ergu  0   0   0   0   0  11   0   0   0   0   0   0   0   0   0   0   0   0   0   0
Geor  0   0   0   0   0   0   6   0   0   0   0   0   0   0   0   0   0   0   0   3
Jake  0   0   0   0   0   0   0  11   0   0   0   0   0   0   0   0   0   0   0   0
Jard  0   0   0   0   0   0   0   0   9   0   0   0   0   0   0   0   0   0   0   0
Joe   0   0   0   0   0   0   0   0   0   7   0   0   0   0   0   0   0   0   0   0
Jord  0   0   0   0   0   0   0   0   0   0   9   0   0   0   0   0   0   0   0   0
Jose  0   0   0   0   0   0   0   0   0   0   0   8   0   0   0   0   0   0   0   0
Luka  0   0   0   0   0   0   0   0   0   0   0   0  10   0   0   0   0   0   0   0
Matt  0   0   0   0   0   0   0   0   0   0   0   0   0  12   0   0   0   0   0   0
Rob   0   0   0   0   0   0   0   0   0   0   0   0   0   0  10   0   0   0   0   0
Shau  0   0   0   0   0   0   0   0   0   0   0   0   0   0   0  10   0   0   0   0
Thom  0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0  10   0   0   0
Tim   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0  14   0   0
Toby  0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0  10   0
Will  0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   3   9
Ins   0   0   0   1   1   0   0   1   0   0   0   0   0   0   0   0   3   0   2   0
Del [ %c / %e]
```

Visual Only Using Triangle Shape of Mask

```

----- Overall Results -----
SENT: %Correct=60.00 [H=12, S=8, N=20]
WORD: %Corr=96.06, Acc=93.10 [H=195, D=6, S=2, I=6, N=203]
----- Confusion Matrix -----

```

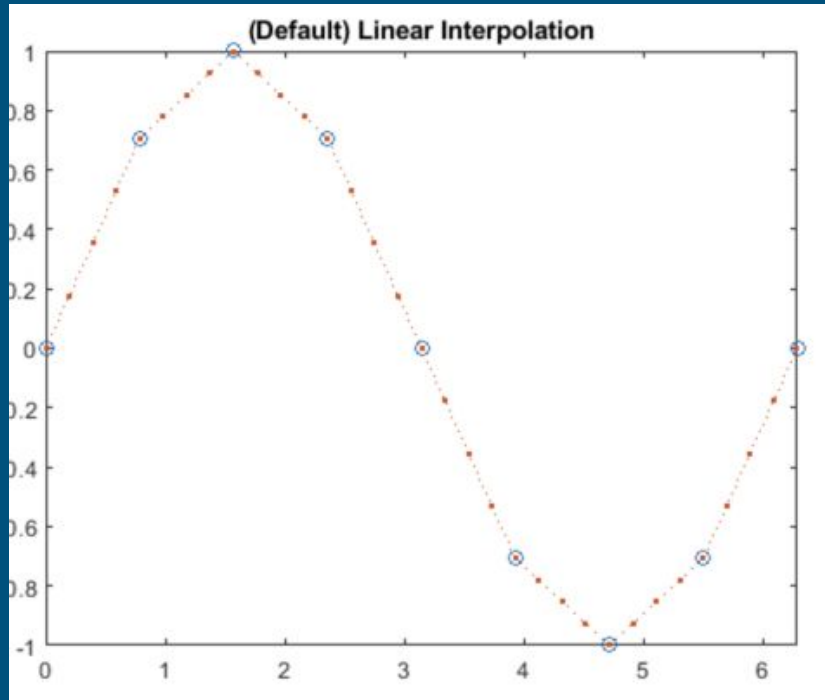
	A	A	A	C	C	E	G	J	J	J	J	J	L	M	R	S	T	T	T	W	
	l	m	n	h	h	r	e	a	a	o	o	o	u	a	o	h	h	i	o	i	
	e	e	u	a	a	g	o	k	r	e	r	s	k	t	b	a	o	m	b	l	
	x	l	s	n	r	r	e	d			d	e	a	t		u	m		y	l	
		i	h	d	l	n	g		e		a	p	s			n	a				Del [%c / %e]
Alex	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Amel	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
Anus	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Chan	0	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Char	0	0	0	0	11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ergu	0	0	0	0	0	11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Geor	0	0	0	0	0	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	3
Jake	0	0	0	0	0	0	0	11	0	0	0	0	0	0	0	0	0	0	0	0	0
Jard	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0
Joe	0	0	0	0	0	0	0	0	0	7	0	0	0	0	0	0	0	0	0	0	0
Jord	0	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0
Jose	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	0	0
Luka	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	1
Matt	0	0	0	0	0	0	0	0	0	0	0	0	0	12	0	0	0	0	0	0	0
Rob	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0	0	0
Shau	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0	0
Thom	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0
Tim	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	14	0	0	0
Toby	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0
Will	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	10	0 [83.3/1.0]
Ins	0	0	0	1	1	0	0	1	0	0	0	0	0	0	0	0	1	0	2	0	

Visual Only Using New Mask

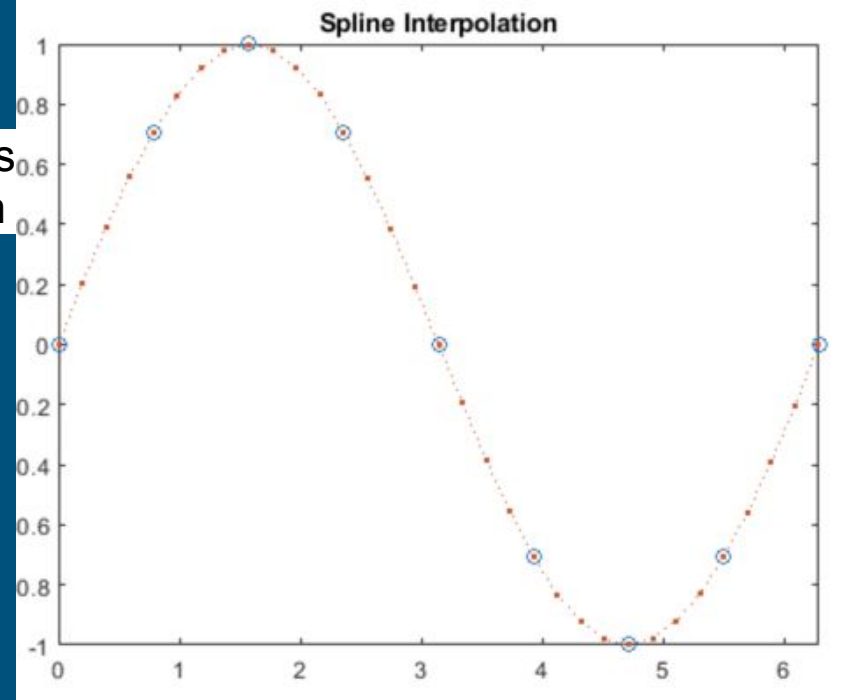
----- Overall Results -----																					
SENT: %Correct=65.00 [H=13, S=7, N=20]																					
WORD: %Corr=100.00, Acc=91.63 [H=203, D=0, S=0, I=17, N=203]																					
----- Confusion Matrix -----																					
	A	A	A	C	C	E	G	J	J	J	J	L	M	R	S	T	T	T	W		
	l	m	n	h	h	r	e	a	a	o	o	o	u	a	o	h	h	i	o	W	
	e	e	u	a	a	g	o	k	r	e	r	s	k	t	b	a	h	m	b	l	
	x	l	s	n	r	u	r	e	d		d	e	a	t		u	m	y	l		
		i	h	d	l	n	g		e		a	p	s			n	a			Del	[%c /
Alex	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Amel	0	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Anus	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Chan	0	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Char	0	0	0	0	11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ergu	0	0	0	0	0	11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Geor	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Jake	0	0	0	0	0	0	0	11	0	0	0	0	0	0	0	0	0	0	0	0	0
Jard	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0
Joe	0	0	0	0	0	0	0	0	0	7	0	0	0	0	0	0	0	0	0	0	0
Jord	0	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0
Jose	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	0	0
Luka	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0	0	0	0	0
Matt	0	0	0	0	0	0	0	0	0	0	0	0	0	12	0	0	0	0	0	0	0
Rob	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0	0	0
Shau	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0	0
Thom	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0
Tim	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	14	0	0	0
Toby	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0
Will	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	12	0
Ins	1	1	2	1	0	0	1	0	0	0	1	0	10	0	0	0	0	0	0	0	0

Combining audio and visual features

Data Interpolation



analysis
within
11).



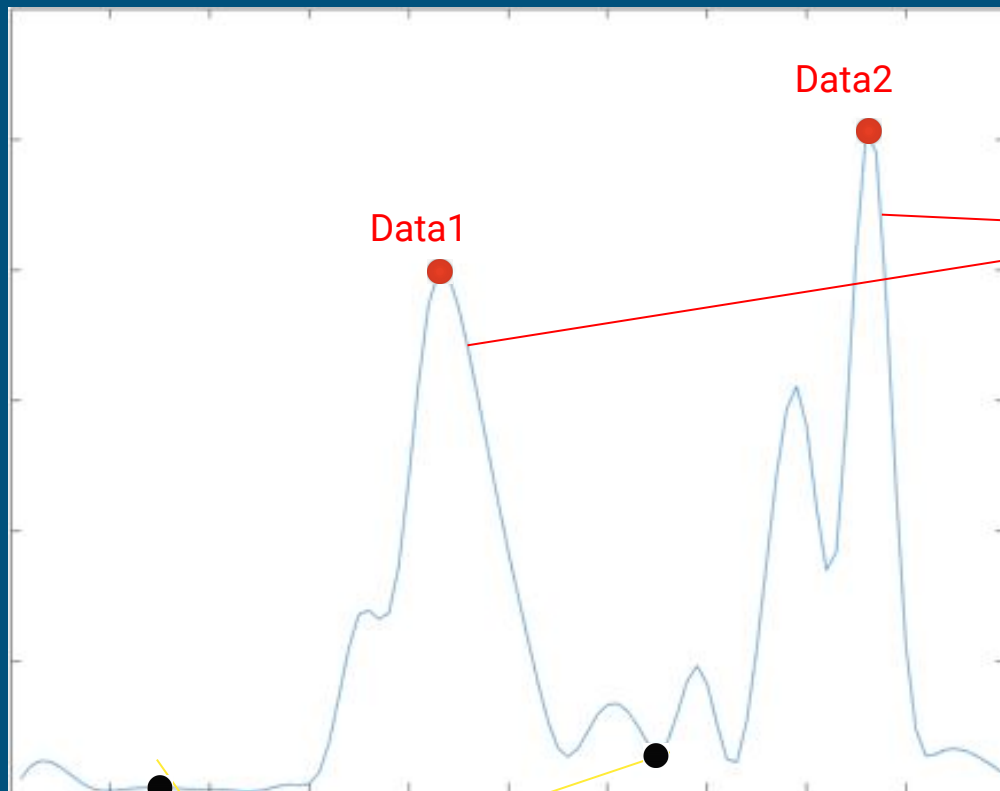
Data Interpolation Entire Video in Combine Model

```
----- Overall Results -----
SENT: %Correct=0.00 [H=0, S=20, N=20]
WORD: %Corr=100.00, Acc=29.06 [H=203, D=0, S=0, I=144, N=203]
----- Confusion Matrix -----
```

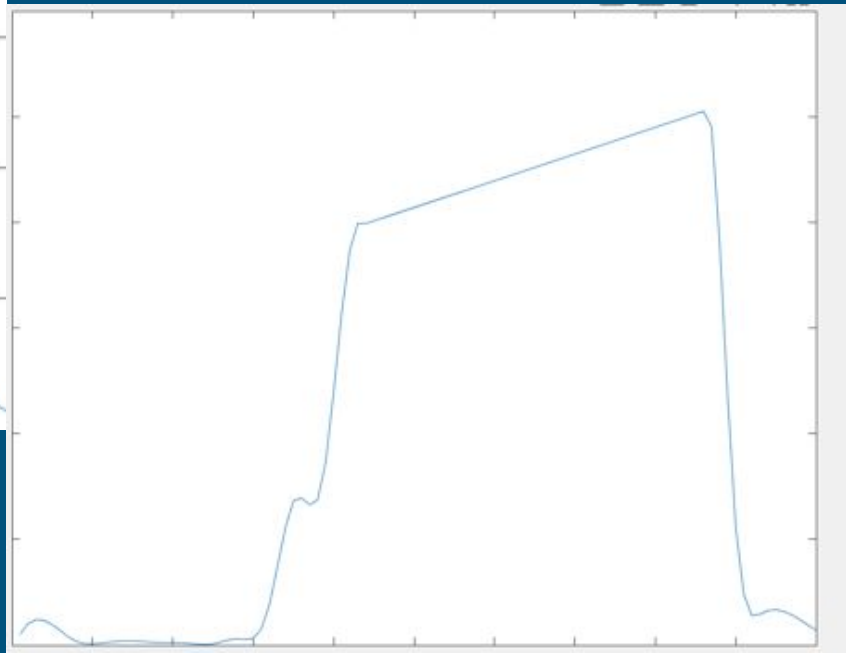
	A	A	A	C	C	E	G	J	J	J	J	J	L	M	R	S	T	T	T	W	
	l	m	n	h	h	r	e	a	a	o	o	o	u	a	o	h	h	i	o	i	
	e	e	u	a	a	g	o	k	r	e	r	s	k	t	b	a	o	m	b	l	
	x	l	s	n	r	u	r	e	d		d	e	a	t		u	m	y	l		
		i	h	d	l	n	g		e		a	p	s			n	a				Del [%c / %e]
Alex	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Amel	0	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Anus	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Chan	0	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Char	0	0	0	0	11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ergu	0	0	0	0	0	11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Geor	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Jake	0	0	0	0	0	0	0	11	0	0	0	0	0	0	0	0	0	0	0	0	0
Jard	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0
Joe	0	0	0	0	0	0	0	0	0	7	0	0	0	0	0	0	0	0	0	0	0
Jord	0	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0
Jose	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	0	0
Luka	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0	0	0	0	0
Matt	0	0	0	0	0	0	0	0	0	0	0	0	0	12	0	0	0	0	0	0	0
Rob	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0	0	0
Shau	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0	0
Thom	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0
Tim	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	14	0	0	0
Toby	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0
Will	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	12	0
Ins	1	4	11	15	2	1	2	1	0	0	1	11	2	0	0	23	0	21	6	43	

Data Interpolation Every Second in Combine Model

Overall Results																				
SENT: %Correct=40.00 [H=8, S=12, N=20]																				
WORD: %Corr=100.00, Acc=79.31 [H=203, D=0, S=0, I=42, N=203]																				
Confusion Matrix																				
	A	A	A	C	C	E	G	J	J	J	J	J	L	M	R	S	T	T	T	W
	l	m	n	h	h	r	e	a	a	o	o	o	u	a	o	h	h	i	o	i
	x	l	s	n	r	u	r	e	d	e	d	e	a	t	b	u	a	m	b	l
		i	h	d	l	n	g	e		a	p	s				n	a		y	l
	Del [%c / %e]																			
Alex	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Amel	0	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Anus	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Chan	0	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Char	0	0	0	0	11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ergu	0	0	0	0	0	11	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Geor	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0
Jake	0	0	0	0	0	0	0	11	0	0	0	0	0	0	0	0	0	0	0	0
Jard	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0
Joe	0	0	0	0	0	0	0	0	0	7	0	0	0	0	0	0	0	0	0	0
Jord	0	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0
Jose	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	0
Luka	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0	0	0	0
Matt	0	0	0	0	0	0	0	0	0	0	0	0	0	12	0	0	0	0	0	0
Rob	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0	0
Shau	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0
Thom	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0
Tim	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	14	0	0
Toby	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0
Will	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	12
Ins	0	0	5	1	0	<u>10</u>	0	1	1	0	1	1	6	0	0	1	3	9	0	3



High Energy or Speech or Noise



Data1

Data2

Silence Or Low Energy

Without Data Interpolation In Combine Model

```

----- Overall Results -----
SENT: %Correct=60.00 [H=12, S=8, N=20]
WORD: %Corr=100.00, Acc=95.07 [H=203, D=0, S=0, I=10, N=203]
----- Confusion Matrix -----

```

	A	A	A	C	C	E	G	J	J	J	J	J	L	M	R	S	T	T	T	W	
	l	m	n	h	h	r	e	a	a	o	o	o	u	a	o	h	h	i	o	i	
	e	e	u	a	a	g	o	k	r	e	r	s	k	t	b	a	o	m	b	l	
	x	l	s	n	r	u	r	e	d			d	e	a	t		u	m		y	l
		i	h	d	l	n	g		e		a	p	s			n	a				Del [%c / %e]
Alex	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Amel	0	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Anus	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Chan	0	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Char	0	0	0	0	11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ergu	0	0	0	0	0	11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Geor	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Jake	0	0	0	0	0	0	0	11	0	0	0	0	0	0	0	0	0	0	0	0	0
Jard	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0
Joe	0	0	0	0	0	0	0	0	0	7	0	0	0	0	0	0	0	0	0	0	0
Jord	0	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0
Jose	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	0	0
Luka	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0	0	0	0	0
Matt	0	0	0	0	0	0	0	0	0	0	0	0	0	12	0	0	0	0	0	0	0
Rob	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0	0	0
Shau	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0	0
Thom	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0
Tim	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	14	0	0	0
Toby	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0
Will	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	12	0
Ins	1	0	0	1	1	0	0	1	0	0	1	1	1	0	0	0	3	0	0	0	0

References:

- Ben Milner (2020) - Audio Processing lecture notes, University of East Anglia.
- David Greenwood (2020) - Visual Processing lecture notes, University of East Anglia.
- Sheppard, William Fleetwood (1911). "Interpolation" . In Chisholm, Hugh (ed.). Encyclopædia Britannica. 14 (11th ed.). Cambridge University Press. pp. 706–710.

Thank you for your listening
Question?