

A Presentation of Project Progress on Speech Summarization

Under the guidance of
Dr. Shashidhar G. Koolagudi

Presented By
Yogendra Kumar
172CS026



Department of Computer Science and Engineering
National Institute of Technology Surathkal

Outline

1. Introduction
2. Applications
3. Motivation
4. Literature Survey
5. Research Gap
6. Problem Statement
7. Objectives
8. Progress
9. Results
10. Future Work

Introduction

1. Issues in processing the speech data available on the Internet.
2. Automatic identification of the relevant, significant and interesting parts of the speech.
- 3. Ways of the speech summarization**
 - 3.1. Extractive Summarization
 - 3.2. Abstractive Summarization

Applications

1. The summary of the recorded meetings.
2. For a television show that will have summary of the all broadcast news shows of the day.
3. Voicemail/Voice-call summary.

Motivation

1. Loss of information in text processing

1.1 In text summarization we may lose the information related to the emotions present in the speech.

2. Language Independent speech summarization.

Literature Survey

S.N.	Paper	Approach
1	From Text to Speech Summarization	The speech summarization using the techniques of text summarization.[3]
2	Summarizing Speech Without Text Hidden Markov Models	Summarizing speech documents without using any type of transcript/text in a Hidden Markov Using Model framework.[2]
3	Automatic Broadcast News Speech Summarization.	Speech summarization using lexical feature, acoustic/prosodic feature, structural feature, discourse feature .[1]

Research Gap

1. Loss of information in text summarization.
2. Unavailability of a proper evaluation measures in speech summarization.

Problem Statement

1. To generate an extractive summary of the speech using spectral and acoustic features.

Objectives

1. Extraction and analysis of the features efficient in speech summarization.
2. Automatic speech summarization using various combination of selected speech features.

Progress

1. Manual Work
2. Feature Extraction
3. Important Word Selection, Summary Generation and Validation.

Manual Work

1. Preprocessing
2. Sentence Segmentation
3. Word Segmentation

Feature Extraction

1. Average Pitch
2. Average Intensity
3. Glottal Parameters
 - a. Average Slope Ratio
 - b. Average Duration Ratio

Important Word Selection, Summary generation and Validation



Results

1. Dataset

The name of dataset is TED-LIUM . It contains about 118 hours of speech and the size is about 20 GB.

Extracted Features For Each Words

	A	B	C	D	E	F
1	Name	pitch	length	intensity	slope	duration
2	StewartBrand_2006_40	59.47	0.24	96.69	-1.5621	2.2104
3	StewartBrand_2006_20	60.52	0.45	198.5	-1.511	1.682
4	StewartBrand_2006_27	62.98	0.13	91.52	-1.5565	1.5589
5	StewartBrand_2006_143	64.1	0.41	280.54	-1.53	1.6164
6	StewartBrand_2006_38	64.26	0.74	89.94	-1.7914	2.2294
7	StewartBrand_2006_39	64.26	0.74	89.94	-1.7914	2.2294
8	StewartBrand_2006_47	64.71	0.15	84.39	-1.8808	1.2216
9	StewartBrand_2006_164	64.86	0.33	166.29	-1.6682	1.9257
10	StewartBrand_2006_72	65.01	0.68	82.83	-1.8193	2.4199
11	StewartBrand_2006_134	65.45	0.19	147.52	-1.6845	1.8445
12	StewartBrand_2006_171	65.7	0.61	95.14	-1.8582	2.3177
13	StewartBrand_2006_168	66.34	0.13	273.94	-1.1012	1.1356
14	StewartBrand_2006_69	66.77	0.56	171.19	-1.4393	1.623
15	StewartBrand_2006_203	67.24	0.21	96.21	-1.3487	1.9822

Result After Column Standardization

	A	B	C	D	E	F	G	H
1		Name	pitch	length	intensity	slope	duration	sum
2	0	StewartBrand_2006_40	-3.2011086515531	0.24	-0.773792658441288	-0.214901374301503	0.71445702592561	-3.47534565837028
3	1	StewartBrand_2006_20	-2.97126433575789	0.45	2.14441934559698	-0.057338359708044	-0.224607897023477	-1.10879124689243
4	2	StewartBrand_2006_27	-2.43277193875197	0.13	-0.921981992110651	-0.197634194620028	-0.443379456476604	-3.99576758195926
5	3	StewartBrand_2006_143	-2.18760466857042	0.41	4.49595763074873	-0.115923433627334	-0.341191278827905	1.85123824972307
6	4	StewartBrand_2006_38	-2.15258077283019	0.74	-0.967270028319083	-0.921930713759037	0.748223554192137	-3.29355796071618
7	5	StewartBrand_2006_39	-2.15258077283019	0.74	-0.967270028319083	-0.921930713759037	0.748223554192137	-3.29355796071618
8	6	StewartBrand_2006_47	-2.05407606606082	0.15	-1.12635142132971	-1.19758890367401	-1.04282419249236	-5.42084058355691
9	7	StewartBrand_2006_164	-2.02124116380436	0.33	1.2211739998542	-0.542052760766589	0.208492257637183	-1.13362766707957
10	8	StewartBrand_2006_72	-1.9884062615479	0.68	-1.17106619125703	-1.0079582696721	1.08677742970652	-3.0806532927705
11	9	StewartBrand_2006_134	-1.89209054826229	0.19	0.683163595023649	-0.592312587339454	0.064184778940238	-1.73705476163786
12	10	StewartBrand_2006_171	-1.83736571116819	0.61	-0.818220795228041	-1.12790349995949	0.905149051346574	-2.87834095500915
13	11	StewartBrand_2006_168	-1.6972701282073	0.13	4.30677975797933	1.20624918198274	-1.19566216254085	2.62009664921392
14	12	StewartBrand_2006_69	-1.60314340840546	0.56	1.36162423872845	0.163743208713696	-0.329461853219533	-0.407237814182844
15	13	StewartBrand_2006_203	-1.50026071466855	0.21	-0.787551049188154	0.443101508560415	0.308903249587644	-1.53580700570865
16	14	StewartBrand_2006_53	-1.46961480589586	0.17	-0.845737576721772	-0.412240570661216	0.237104947378819	-2.49048800590003
17	15	StewartBrand_2006_217	-1.41051198183423	0.74	-0.759461001413303	0.242678887257582	-0.055775255690844	-1.9830693516808
18	16	StewartBrand_2006_181	-1.33608620338626	0.54	-1.07131785834225	-1.01381677706403	0.709480905970543	-2.711739932822
19	17	StewartBrand_2006_142	-1.15221075075009	0.74	-0.874974157058861	0.658016227095915	-0.706580658385688	-2.07574933909873
20	18	StewartBrand_2006_34	-1.03838308959437	0.31	-0.805322303902855	-0.352422126764678	0.631107016678237	-1.56502050358367

Opinion Score

1	Shiva Sir	7
2	Pradyoth Sir	9
3	Nagaratna Mam	8
4	Kapil	7
5	Kunal	6
6	Kemanth	4.5
7	Debu	6
8	Shubham	6
9	Preeti	8
10	Nehal	7.5
11	Arnab	7
12	Prateek	6.5
13	Amit	6
14	Praveen Sir	6
15	Manjunath Sir	8
16	Alok	7
17	Aniket	7
18	Dhawal	6.5
19	Smriti	8
20	Apoorva	9
21	Swapnil	8.5
22	Average	7.071428571

Future work

1. Try to implement the paper [2] and compare the result for same set of inputs.

References

- [1] Sameer Raj Maskey. "Automatic Broadcast News Speech Summarization" Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Graduate School of Arts and Sciences COLUMBIA UNIVERSITY, 2008.
- [2] Sameer Maskey, Julia Hirschberg "Summarizing Speech Without Text Using Hidden Markov Models". Proceedings of the Human Language Technology Conference of the North American Chapter of the ACL, pages 89–92, New York, June 2006.
- [3] Kathleen McKeown, Julia Hirschberg, Michel Galley and Sameer Maskey "FROM TEXT TO SPEECH SUMMARIZATION". Conference Paper in Acoustics, Speech, and Signal Processing, 1988. ICASSP-88., 1988 International Conference on , April 2005.

Thank You