

Background

- Explosion of information
- No time to process everything

Thus, automatic summarization is required.

Not only text, but also speech

Newscast Speech Summarization via Sentence Shortening based on Prosodic Features

Kiyonori Ohtake¹, Kazuhide Yamamoto^{1,2}, Yuji Toma³, Shiro Sado³, Shigeru Masuyama³, and Seichi Nakagawa³

¹ ATR Spoken Language Translation Research Laboratories

² Nagaoka University of Technology

³ Toyohashi University of Technology

Newscast Speech Summarization via Sentence Shortening based on Prosodic Features - p.1/25

Newscast Speech Summarization via Sentence Shortening based on Prosodic Features - p.2/25

Speech Summarization

Our approach: direct speech summarizer based on prosodic features

Conventional one: speech recognizer + text summarizer

Why?

Speech conveys more information than its transcription.

Target: newscast speech

Newscast Speech Summarization via Sentence Shortening based on Prosodic Features - p.3/25

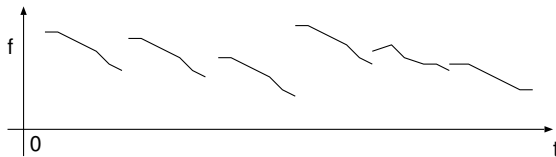
Newscast Speech Summarization

Five Steps:

1. Extract pitch pattern
2. Extract accent phrases
3. Detect summary units
4. Analyze dependency structures
5. Select important summary units

Newscast Speech Summarization via Sentence Shortening based on Prosodic Features - p.4/25

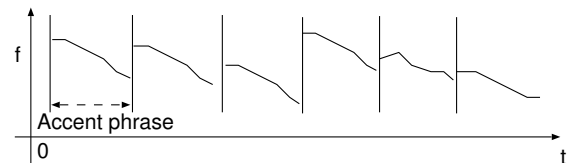
Extracting pitch pattern



By a method proposed by Hirose et al.

Newscast Speech Summarization via Sentence Shortening based on Prosodic Features - p.5/25

Extracting accent phrases

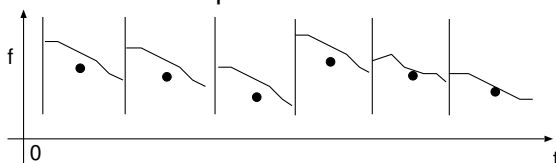


The beginning of each accent phrase is determined by picking up the rise in F0.

Newscast Speech Summarization via Sentence Shortening based on Prosodic Features - p.6/25

Detecting summary units (1/2)

Calculate base-pitch

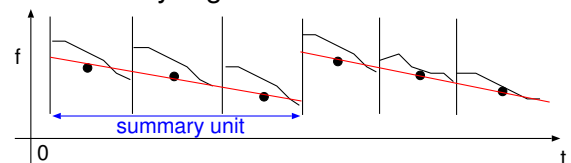


The base-pitch is a representative value of an accent phrase.

Newscast Speech Summarization via Sentence Shortening based on Prosodic Features - p.7/25

Detecting summary units (2/2)

Estimate by regression lines

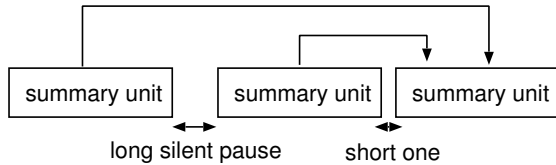


Heuristics: adjust the boundaries to correct mis-detection caused by exceptional pitch rises for emphasis or paraphrasing.

Newscast Speech Summarization via Sentence Shortening based on Prosodic Features - p.8/25

Dependency structure analysis

Approximate analysis



NewsCast Speech Summarization via Sentence Shortening based on Prosodic Features - p.9/25

Selecting summary units

Three methods:

method 1: based on common heuristics for news texts

method 2: based on heuristics of prosodic features with **absolute** parameters for each speaker

method 3: based on heuristics of prosodic features with **relative** parameters

NewsCast Speech Summarization via Sentence Shortening based on Prosodic Features - p.10/25

Estimating parameters

By trial and error

- for accent phrase detection
- for summary units selection

with NHK news speech database.

NewsCast Speech Summarization via Sentence Shortening based on Prosodic Features - p.11/25

Evaluate summary units detection

- Ten articles, 52 sentences
- Prepare correct summary unit boundaries by hand

heuristics	P	R	F
with	70.7%	77.7%	74.0%
without	47.9%	91.1%	62.8%

P: Precision R: Recall F: F-measure

NewsCast Speech Summarization via Sentence Shortening based on Prosodic Features - p.12/25

Speech summarization ratio

$$\text{SS-ratio} = \frac{\text{playback time of summary}}{\text{playback time of original}} \times 100(\%)$$

Ten articles from NHK database

	Avg	Max	Min	SD
Method 1	72.4%	98.7%	53.2%	8.25%
Method 2	70.1%	96.1%	53.3%	7.55%
Method 3	80.7%	91.2%	74.7%	6.63%

Method 2 produced shorter summaries.

NewsCast Speech Summarization via Sentence Shortening based on Prosodic Features - p.13/25

Preservation ratio of important information

Questionnaire with eight subjects
Scale: 1 to 5 (best: 5, worst: 1)

	Avg	Max	Min	SD
Original	4.9	5.0	4.5	0.17
Method 1	3.9	4.5	2.5	0.69
Method 2	4.3	4.8	3.0	0.60
Method 3	3.7	5.0	2.6	0.69

Method 2 showed the best and stable performance.

NewsCast Speech Summarization via Sentence Shortening based on Prosodic Features - p.14/25

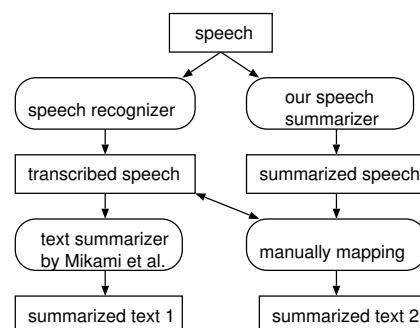
Evaluating phonological naturalness

	Avg	Max	Min	SD
Original	4.9	5.0	4.3	0.24
Method 1	4.0	4.5	3.5	0.50
Method 2	4.4	4.8	3.9	0.33
Method 3	3.6	4.3	2.4	0.58

Method 2 also showed the best and stable performance.

NewsCast Speech Summarization via Sentence Shortening based on Prosodic Features - p.15/25

Comparing with conventional method (1/2)



NewsCast Speech Summarization via Sentence Shortening based on Prosodic Features - p.16/25

Comparing with conventional method (2/2)

Standard: six subjects select important parts (S-ratio (summarization ratio): 56.8%)

	Ours	Mikami's
Avg S-ratio	73.7%	51.0%
SD of S-ratio	12.7%	20.8%
Avg Precision	69.1%	66.8%
Avg Recall	48.3%	36.5%

Mikami's produced shorter summaries.

Newscast Speech Summarization via Sentence Shortening based on Prosodic Features - p.17/25

Discussions (1/2) - speech summarization

- Fine summary unit detection whereas troublesome heuristics and parameter setting
- Method 2 is the best strategy thanks to the parameter settings

Newscast Speech Summarization via Sentence Shortening based on Prosodic Features - p.18/25

Discussions (2/2) - comparison

	Ours	Mikami's
Target S-ratio	70%	70-80%
Result	73.7%	51.0%

with manually transcribed speech

- Mikami's failed summary unit detection → longer parts were eliminated
- Ours seems robust

Quality evaluation is needed.

Newscast Speech Summarization via Sentence Shortening based on Prosodic Features - p.17/25

Conclusion

Feasibility of speech summarization is shown

Open problems:

- determine parameters dynamically
- consider other prosodic features

Newscast Speech Summarization via Sentence Shortening based on Prosodic Features - p.20/25

Thank you!

Appendix B

Example of detected summary unit
“ || ” indicates a boundary.

て 全国銀行協会連合会 のまとめによりますと || おとし以降事故が || 企業や個人に貸し出しが資金の総額がで先月九月末現在で || 二百九十五勝千三百一億円で || 去年の同じ月に比べて零点三パーセント増えました

Newscast Speech Summarization via Sentence Shortening based on Prosodic Features - p.23/25

Appendix A

Example of extracted accent phrase
“ | ” indicates a boundary.

て | 全国銀行協会連合会 のまとめによりますと | おとし以降事故が | 企業や個人に | 貸し出しが | 資金の総額が | 先月九月末現在で | 二百九十 | 五勝 | 千三百一億円で | 去年の同じ月に比べて | 零点三パーセント増えました

Newscast Speech Summarization via Sentence Shortening based on Prosodic Features - p.20/25

Appendix C

Underlines indicate eliminated parts.

て 全国銀行協会連合会 のまとめによりますと || おとし以降事故が || 企業や個人に 貸し出しが 資金の総額が で先月九月末現在で || 二百九十五勝千三百一億円で || 去年の同じ月に比べて零点三パーセント増えました

Newscast Speech Summarization via Sentence Shortening based on Prosodic Features - p.24/25

Appendix D

Manually transcribed:

え全国銀行協会連合会のまとめによりますと、え都市銀行十行が、企業や個人に貸し出した資金の総額は、え先月一う末現在で、二百う一十五兆千三百一億円で、え去年の同じ月に比べて〇・三%増えました。