

インターネット上の大量の事例に基づく プレゼンテーションの解析・分類と自己学習応用

Toshihiko Yamasaki

Associate Professor,
Department of Information and Communication Engineering,
Graduate School of Information Science and Technology,
The University of Tokyo



プレゼンテーション



スティーブ・ジョブズ 驚異のプレゼン
(カーマイン・ガロ 著、日系BP社)

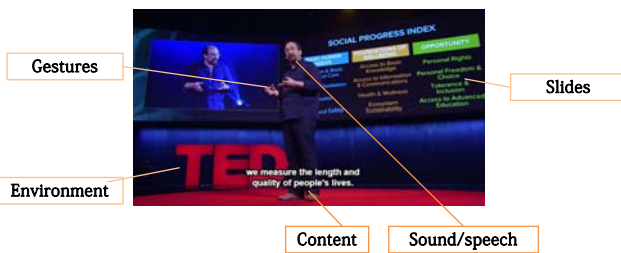


世界最高のプレゼン術
(ウィリアム・リード 著、角川書店)

- ✓ プレゼンテーションは現代人にとって必須スキルの1つ
- ✓ にもかかわらず、プレゼンは体系的な学習支援がない

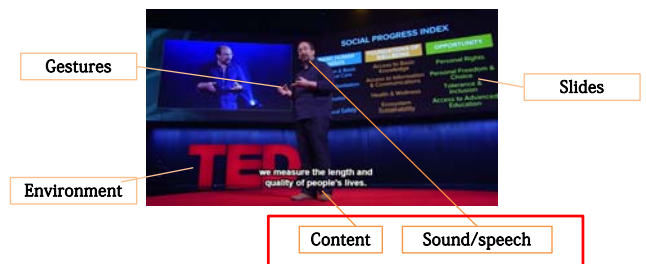
Presentation and Online Lecture Analysis

- What makes the presentation more attractive/impressive?
- What impression do the listeners have?



Presentation and Online Lecture Analysis

- What makes the presentation more attractive/impressive?
- What impression do the listeners have?



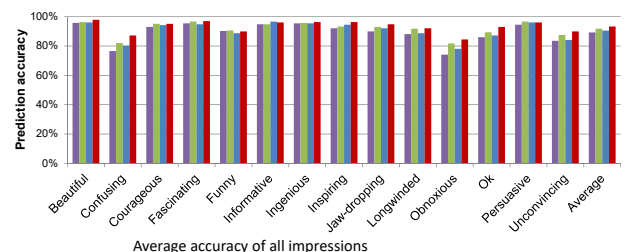
Features

- The six features were extracted from each video (65,825-dim)
 - Content features
 1. Bag-of-words (BoW): 43,408 dim.
 2. Latent semantic indexing (LSI) [Deerwester, JASOS90]: 3,000 dim.
 3. Latent dirichlet allocation (LDA) [Blei, JMLR03]: 3,000 dim.
 4. Word2vec [Socher, EMNLP11] + spherical k-means : 10,000 dim.
 - Surface-level linguistic features
 5. [Yamasaki, MVE15]: 44 dim.
 - Acoustic feature
 6. OpenSMILE [Eyben, ACMMM13]: 6,373 dim.
- Support vector machine (SVM) was employed as a classifier
- The accuracy was calculated by the leave-one-out method.

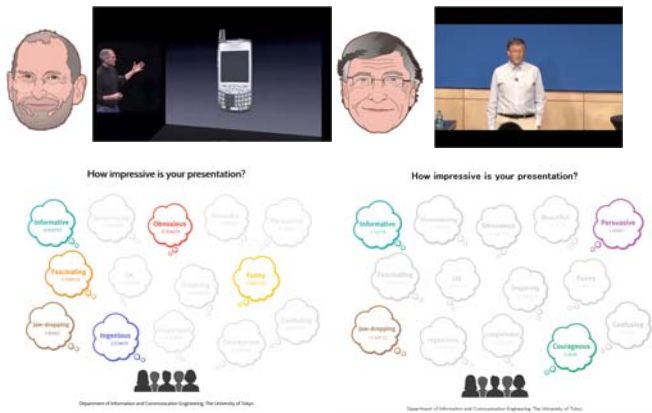
[Yamasaki, ACMMMW16] 5

Experiments with 1600+ TED Talks

- early fusion (concatenating all features)
- early fusion + label correlation
- our late fusion (feature correlation)
- our late fusion + label correlation



	Label correlation	
	Without	With
Early fusion (concatenating all features)	89.2%	91.7%
Our late fusion (feature correlation)	90.5%	93.3%



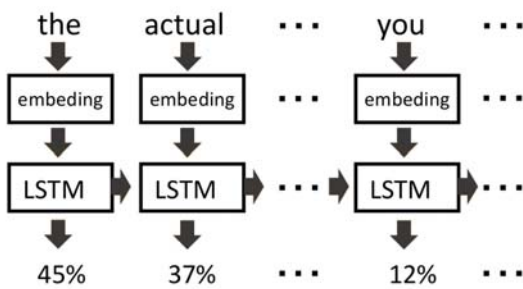
TED Talksに対しては93%の精度で14種類の印象の有無を予測 7

- Hand-craftedの特徴量
 - 14種類の印象を93.3%の精度で予測
 - 判定には5分程度の入力が必要

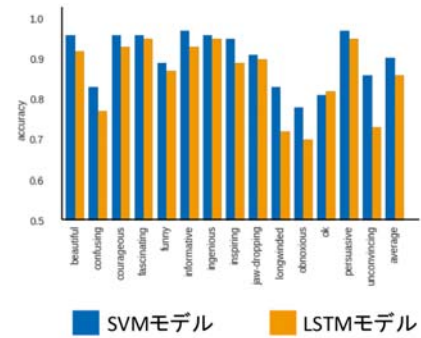


- DNNを用いた予測へ
 - Hand-craftedよりも高精度を実現できるか？
 - 時間的にもっと粒度の高い予測ができるか？

LSTMを用いたプレゼンの印象予測

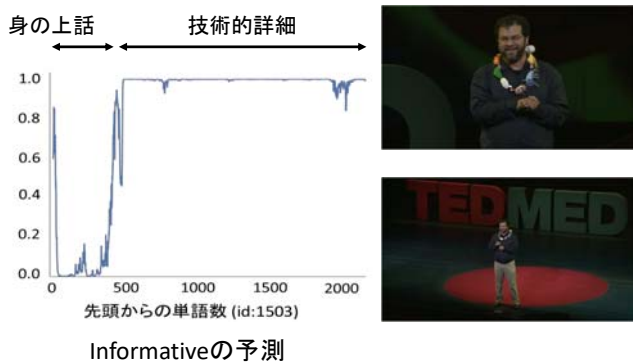


実験結果



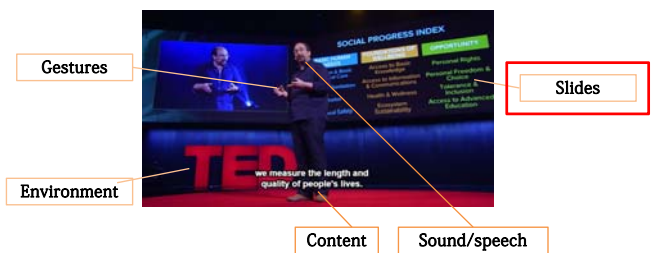
Hand-craftedのほうがDNNよりも性能高い

LSTMを用いることの利点



Presentation and Online Lecture Analysis

- What makes the presentation more attractive/impressive?
- What impression do the listeners have?



Contributions

1. Dataset creation



+ Clarity evaluation

2. Clarity prediction (Classification)



3. Support improvement of slides



Score: 67.6/100



13

Prepare 10 topics for slide collection

10 topics

- Introduction (Albert Einstein)
- Grand Slam (tennis)
- Eyeglasses vs Contact Lenses
- Rock-paper-scissors
- Windows vs Macintosh
- Redox
- Wage-price spiral
- Food chain
- Dogs & Cats
- Seasons

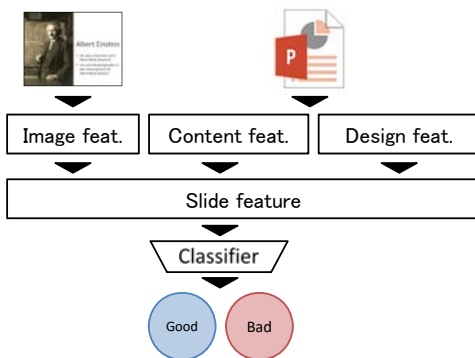
(1) Introduction (Albert Einstein)	(2) Grand Slam (tennis)	(3) Eyeglasses vs Contact Lenses	(4) Rock-paper-scissors	(5) Food chain
(6) Redox	(7) Wage-Price Spiral	(8) Windows vs Macintosh	(9) Dogs vs Cats	(10) Seasons

14

Example of top/bottom 10 slides (1)

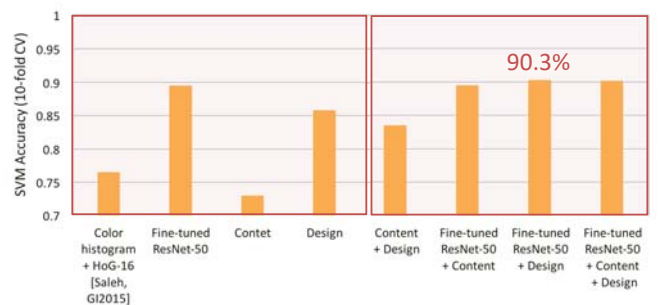
Example of top/bottom 10 slides (2)

Proposed prediction model



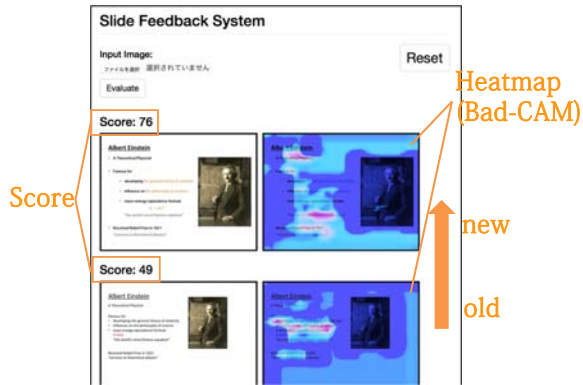
17

Result of visual clarity prediction



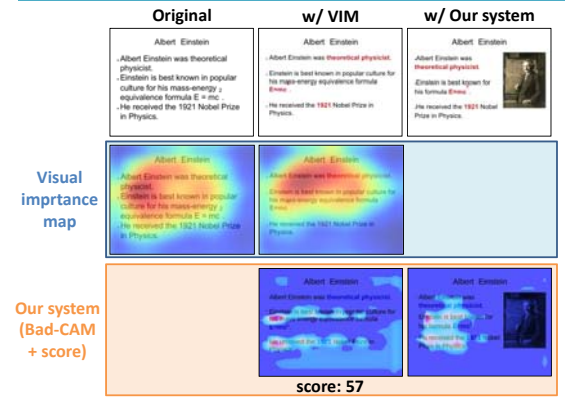
18

Proposed feedback system



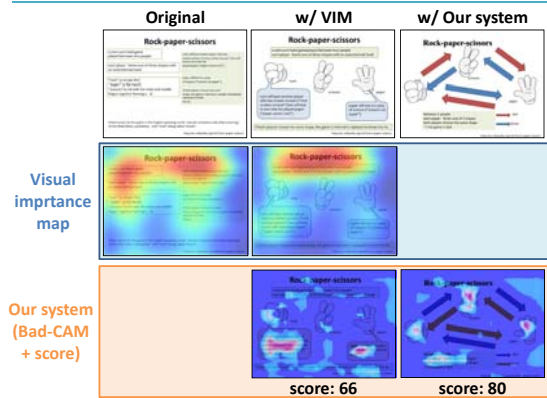
19

User study | results of improvement



20

User study | results of improvement



21

Conclusions

AIを用いたプレゼンテーションスキルの評価・向上

- オーラルプレゼンテーションスキルの評価・向上
93.3%の精度で14種類の印象を予測
- プレゼンテーションスライドの評価・向上
初心者でも大きく視覚的明瞭性を向上



WEB・スマホアプリの開発
別領域への応用・展開

22