

生体情報を用いた User情動分析と解析

早稲田大学
基幹理工学部情報通信学科
亀山研究室
2014年5月

概要

- 従来のRecommendation Systemは、Userの視聴済みContentの傾向をMetadataを用いて分析し、Recommendを生成するApproachがほとんど。
- しかしながら、ある分析によれば、最も良いとされるRecommendation Systemでも20-30%の推薦精度しかない。
- これは、例えばGenre等はUser毎に受け止め方が違うため、言語で記述されたMetadataからの類推では限界があることを示している。
 - 例えば、ある人がSF映画と思うものを、他の人はAction映画と感じている等。
- 即ち、究極のRecommendation Systemを作るためには、Userが実際にAV Contentに対してどのように感じているか、その「情動」を知る必要がある。
 - しかも、User情動を客観的に取得しなければならない。
- 当研究室では、以上のことから、5年ほど前からUserの生体情報に着目し、それを用いたUser情動の分析と解析に注力している。

AV Systemと親和性の高い生体情報

- 視線・瞳孔径等
 - 興味度、注目点、集中等の反応を取得できる。
 - AV Contentを視聴するために目を用いるのは当たり前。
 - Eye-Trackerによる非拘束の計測が可能。
 - Eye-Trackerは年々価格が下落している。
- 脳波
 - 様々な情動反応を計測可能。
 - 簡易脳波計の出現により、手軽な測定ができるようになった。
 - 完全に非拘束に計測できないが、将来はメガネなどに組み込むことが期待される。
- 心拍・血圧等
 - 緊張・Relax等の反応を取得できる。
 - 装置を装着してもらう手間がかかる。
- 唾液Amylase
 - Stressの度合いを計測可能。
 - 試験用のChipを口中に入れる必要あり。
- 発汗
 - 主として手のひらの発汗を計測し、緊張度等を計測可能。
 - 非拘束に計測できない。

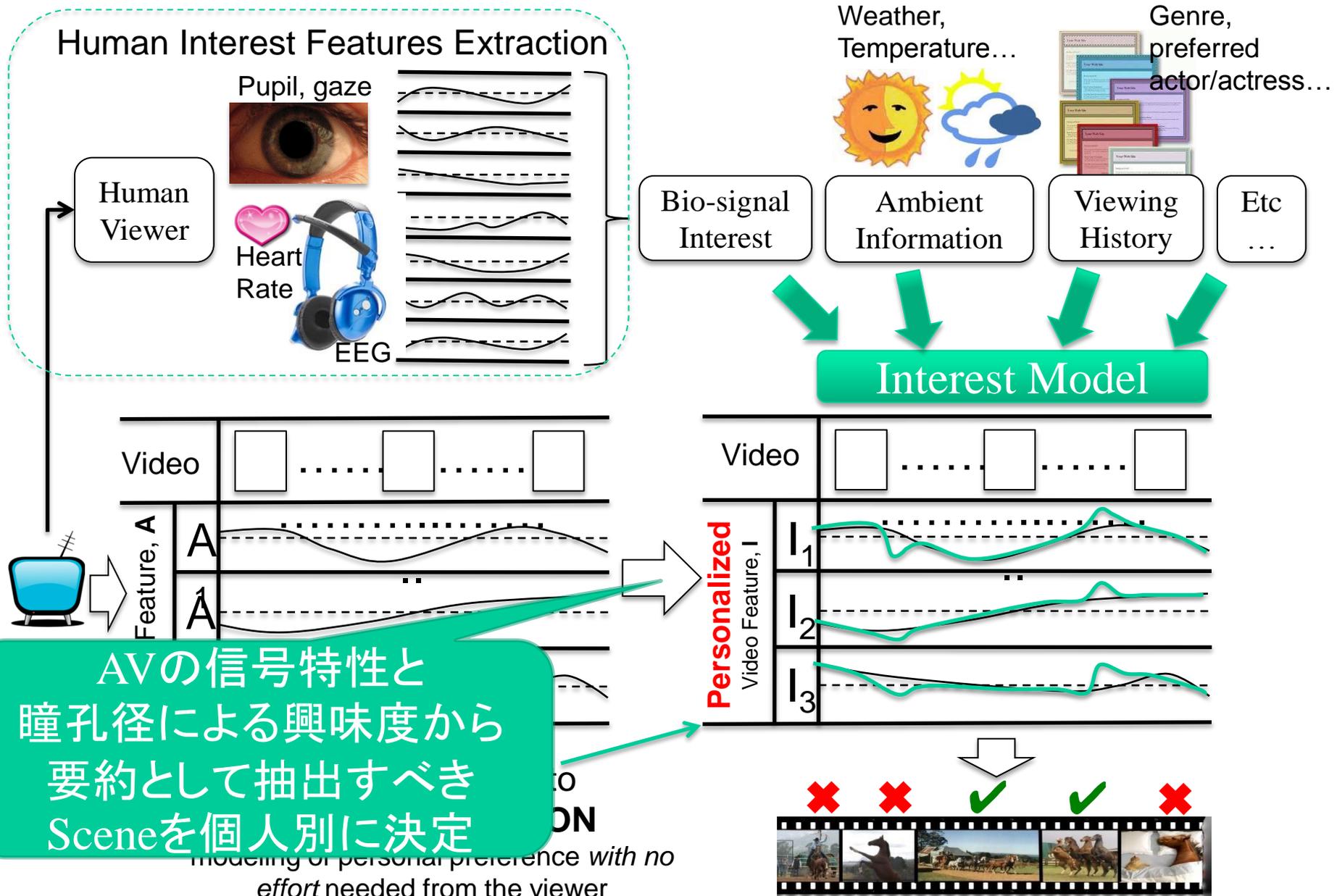


- 以上から、主に、Eye-Trackerによる**視線・瞳孔径計測**、及び、脳波計による**脳波計測**に着目して研究を進めている。

これまでの研究と進行中の研究

- 視線及び瞳孔径による映像要約のPersonalization
- 瞳孔径による映像Genreの自動分類
- 脳波と瞳孔径による映像に対する情動反応解析
- 視線、瞬目、瞳孔径を利用したComic読書時の情動反応解析
- 視線の類似度に着目したUser嗜好分析とUserのGroup化
- 脳波を利用した味覚と個人の味の好みの分析
- 視線情報を利用した個人の理解に合わせた画像分類
- 脳波を利用した音楽のGenre分類個人化
- 脳波による予期せぬEventに対する個人反応の分析

視線及び瞳孔径による映像要約のPersonalization (1/2)



視線及び瞳孔径による映像要約のPersonalization (2/2)

Experiment Result for Short Film¹,
with high narrative nature

Method	Precision	Recall
Our Proposal	79.3% ↑	41.6% ↑
Arousal Model	65.3% ↑	33.6% ↑
Constant Interval	57.1%	29.3%

Experiment Result for
Documentary² video

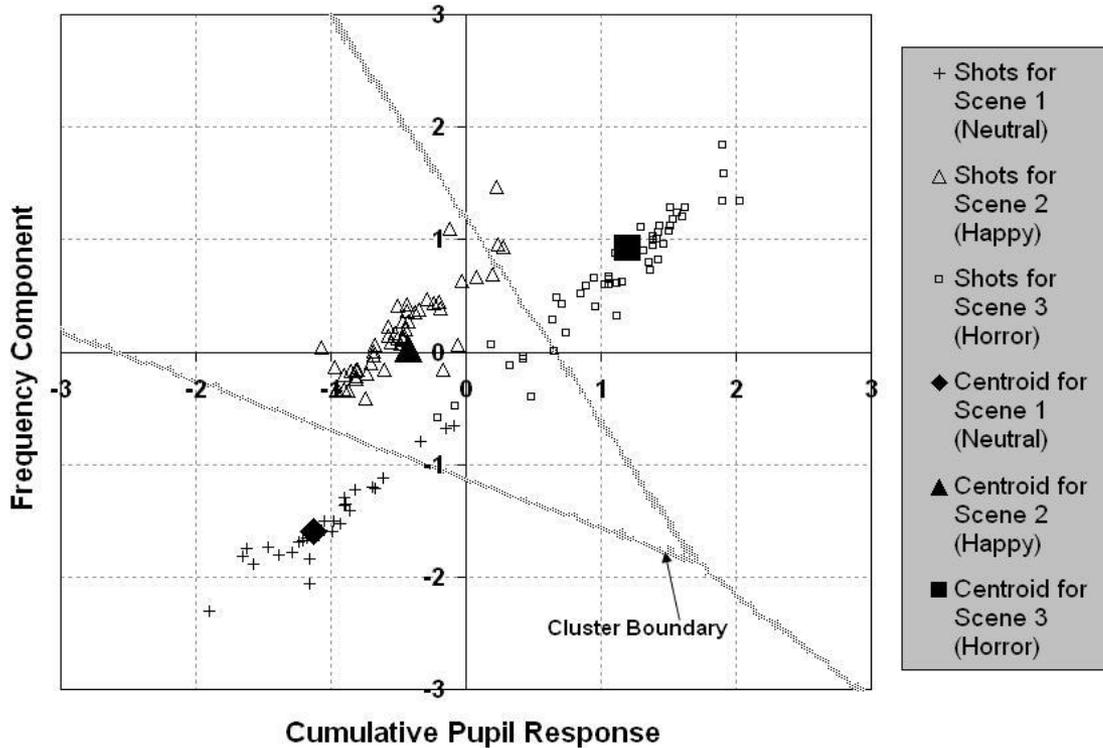
Method	Precision	Recall
Our Proposal	47.3% ↑	39.8% ↑
Arousal Model	39.3% ↑	29.5% ↑
Constant Interval	28.6%	17.5%

- Our proposed method achieved highest result for both test videos
- For short video which conform to film grammar, arousal model³ achieve precision at around 65%, while our proposal's precision is nearly 80%
- Documentary video has very low in narrative nature → the general precision and recall is low
- In general for both the cases, our proposed method achieves improvement around 20 points over constant interval sampling

提案方式の優位性を確認

1. "Signs", website, 2008. <http://patrickhughes.com.au/shorts/>
2. "ディープ・ブルー", Website, 2003, <http://www.deep-blue.jp/>
3. Alan Hanjalic and Li-Qun Xu, "Affective Video Content Representation and Modelling", IEEE Transactions on Multimedia, Vol. 7, No. 1, February 2005

瞳孔径による映像Genreの自動分類



Principal Component Analysis shows that 2 PC contributed to majority of the data variance

Subject	Variance Contribution of 2 PCs			features
	First PC	Second PC	Both	
1	69.58%	23.56%	93.14%	CPR, FC
2	66.03%	26.39%	92.42%	CPR, FC
3	66.52%	25.48%	92.00%	CPR, FC
4	66.68%	25.47%	92.15%	CPR, FC
5	64.17%	24.78%	88.95%	CPR, FC
6	67.26%	25.11%	92.37%	CPR, FC

Classification results for all subjects with the selected features

Subject	Accuracy
1	75.86%
2	66.96%
3	80.65%
4	70.97%
5	89.06%
6	47.86%
Average	71.89%

Clustering result for a test subject into 3 different scenes

Findings:

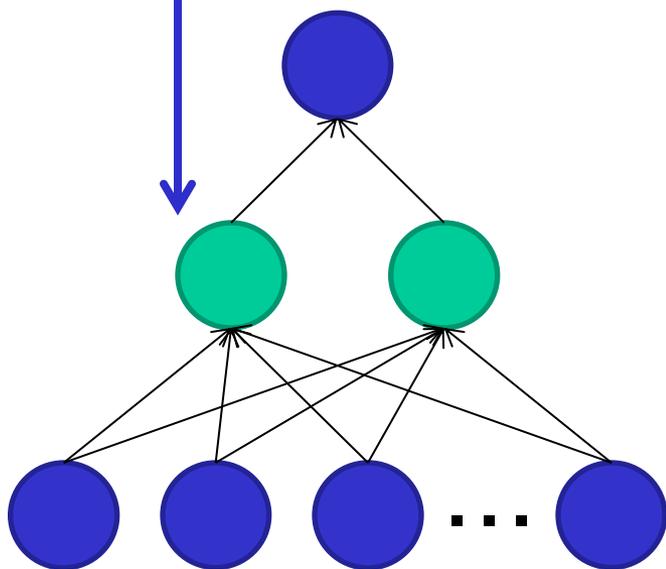
- An encouraging classification accuracy of 89.06% is achieved, with average accuracy 71.89%

瞳孔径と瞳孔径の周波数変化から、「中立」、「Happy」、「恐怖」の3 Sceneを自動分別に成功

脳波と瞳孔径による映像に対する情動反応解析 (1/2)

The intermediate layer : 2

The existence of functional modules in the brain responsible for the contraction and expansion of the pupil size[5]



Neural Network 3 layer model
(multi-layer perceptron)

The output layer : cumulating pupil size

- The temporal summation value of the cumulating pupil size in 1 second.
※We re-arrange the data to compensate the eye blinking by applying linear interpolation using the value before and after it.

The input layer : back ground activity of brain

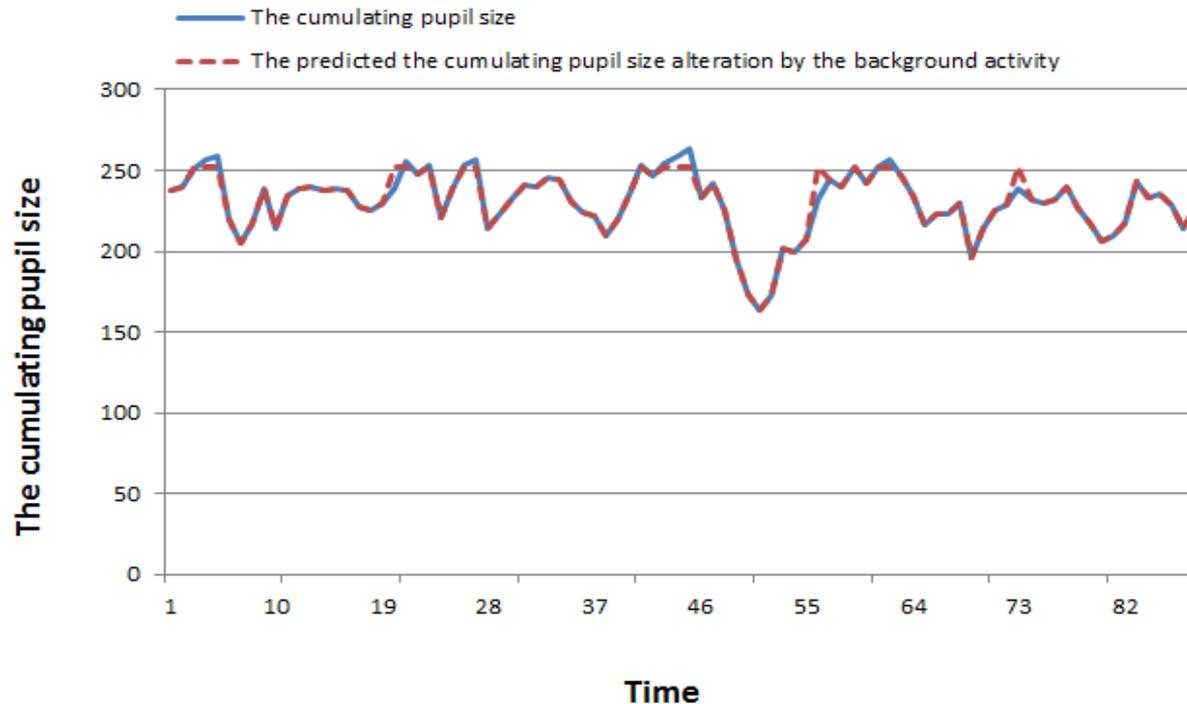
- The power ratio of the frequency component of the background activity of the past 5 seconds and the future 5 seconds from each time point.

α_{low}	8-9Hz	α_{high}	10-12Hz
β_{low}	13-17Hz	β_{high}	18-30Hz
γ_{low}	31-40Hz	γ_{high}	41Hz-

Neural Networkを用いて、瞳孔径と脳波の
間の関係性を分析

[5]Shigeyoshi ASANO, Ikki YASUIKE, Minoru NAKAYAMA, Yasutaka SHIMIZU, "A Pupil Size Estimation Model with Neural Network for Brightness Change", The Transactions of the Institute of Information and Communication Engineers, A Vol.J77-A No.5 (Japanese)

脳波と瞳孔径による映像に対する情動反応解析 (2/2)



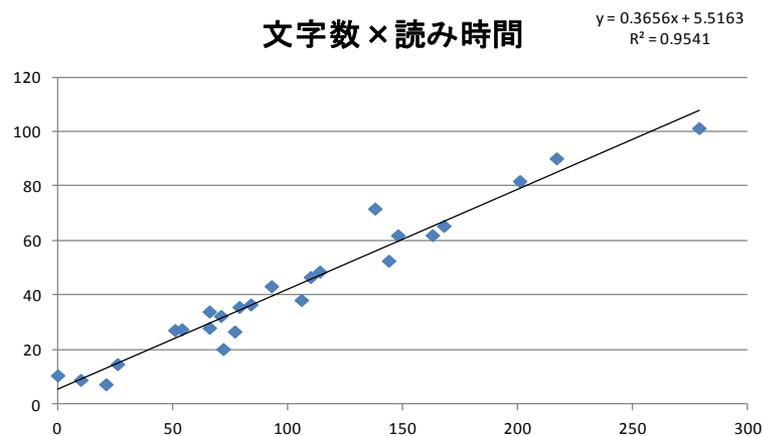
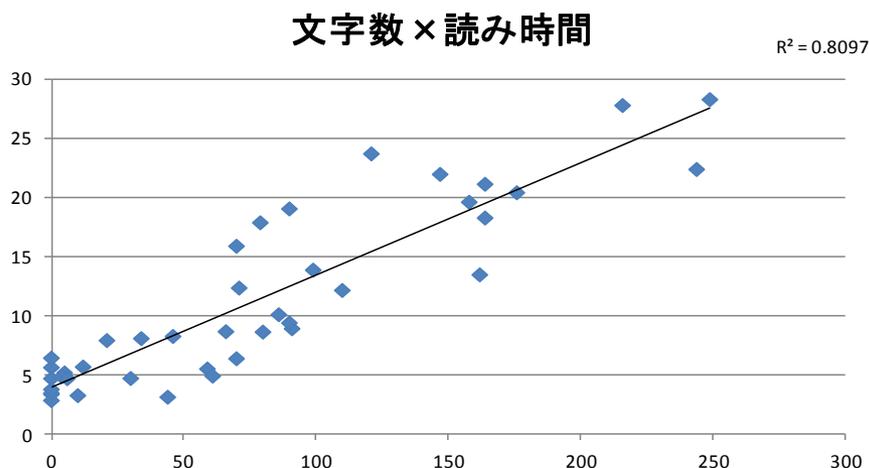
R-square value :
0.967368

Fig.1 Interpolation Result of Subject 1 (The cumulating pupil size)

It is possible to predict the pupil size from the ratio of the power of the background activity.

非常に強い相関関係があることを確認。
瞳孔径だけでは測定できない複雑な情動反応を分析できる可能性が示唆される。

視線、瞬目、瞳孔径を利用した Comic読書時の情動反応解析 (1/2)



横軸: 1 Page当たりの吹き出し中の文字数

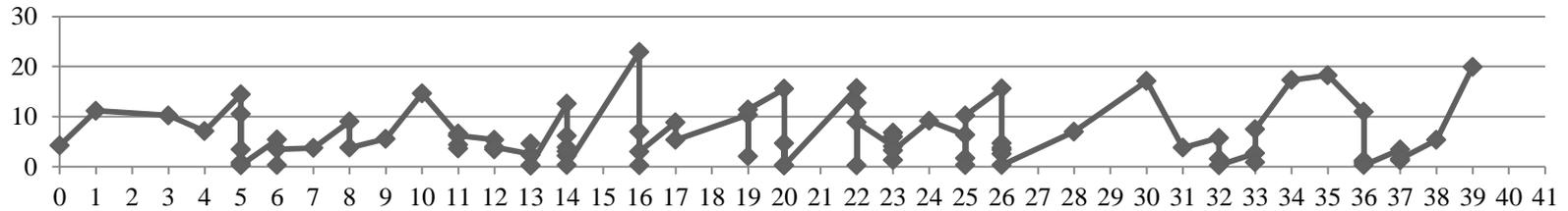
縦軸: 1 Pageを読むのに費やした時間(秒)

Test Comic: 左「ゴルゴ13」さいとうたかお, 右「ミスター・ホー」鳥山明

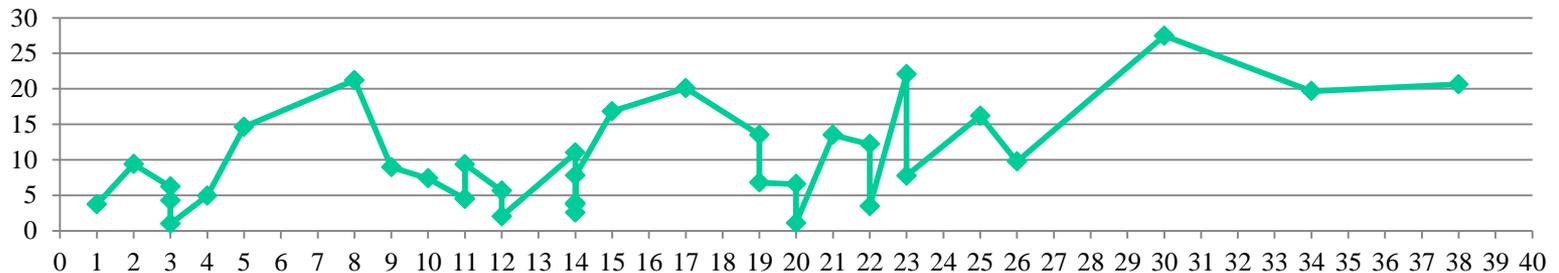
吹き出し中の文字数と読む速度に強い相関関係が存在
決定係数による直線から大きく外れているものは、そのPageに対する興味度等を表していることが示唆される。

視線、瞬目、瞳孔径を利用した Comic読書時の情動反応解析 (2/2)

被験者1の瞬目 Pace (ゴルゴ13)



被験者2の瞬目 Pace (ゴルゴ13)

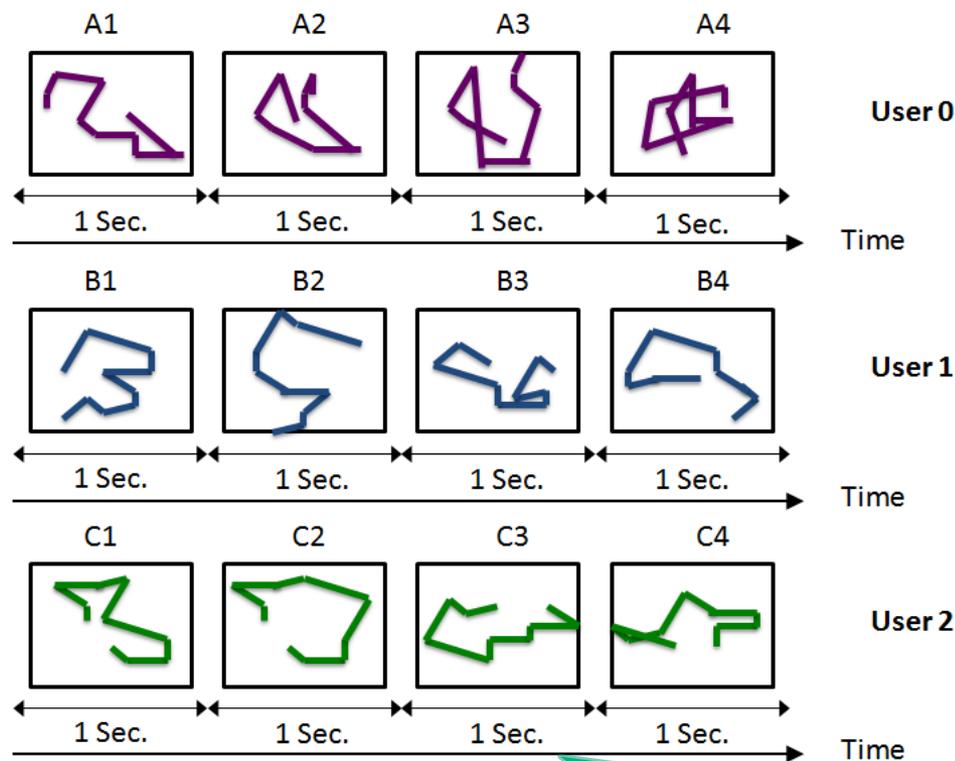


横軸: Page

縦軸: 前の瞬目との間の時間間隔

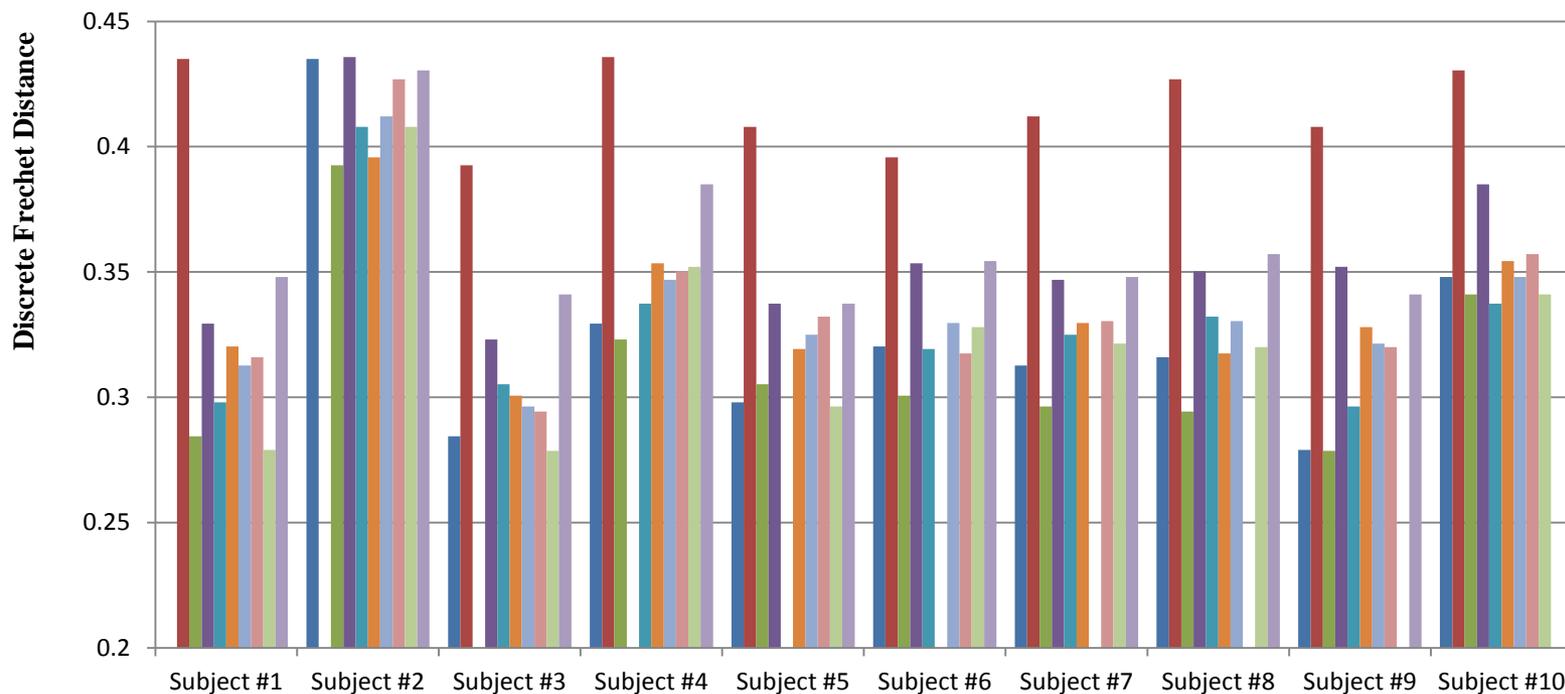
瞬目の回数から、Page毎の興味度が抽出できる可能性が示唆される

視線の類似度に着目した User嗜好分析とUserのGroup化 (1/2)



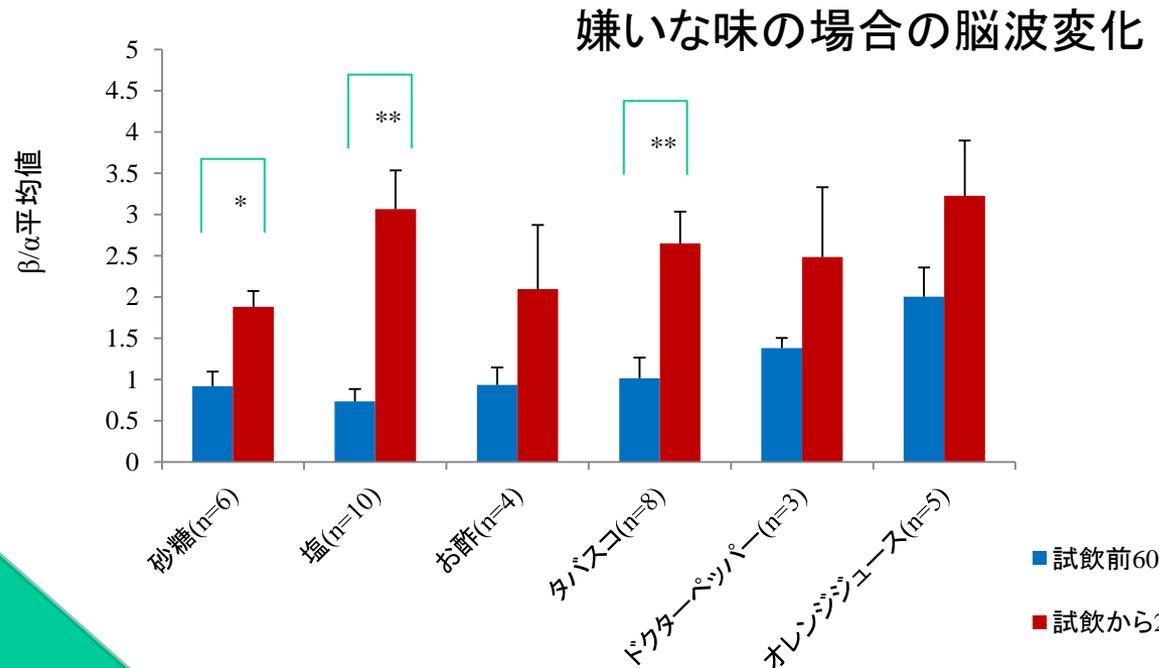
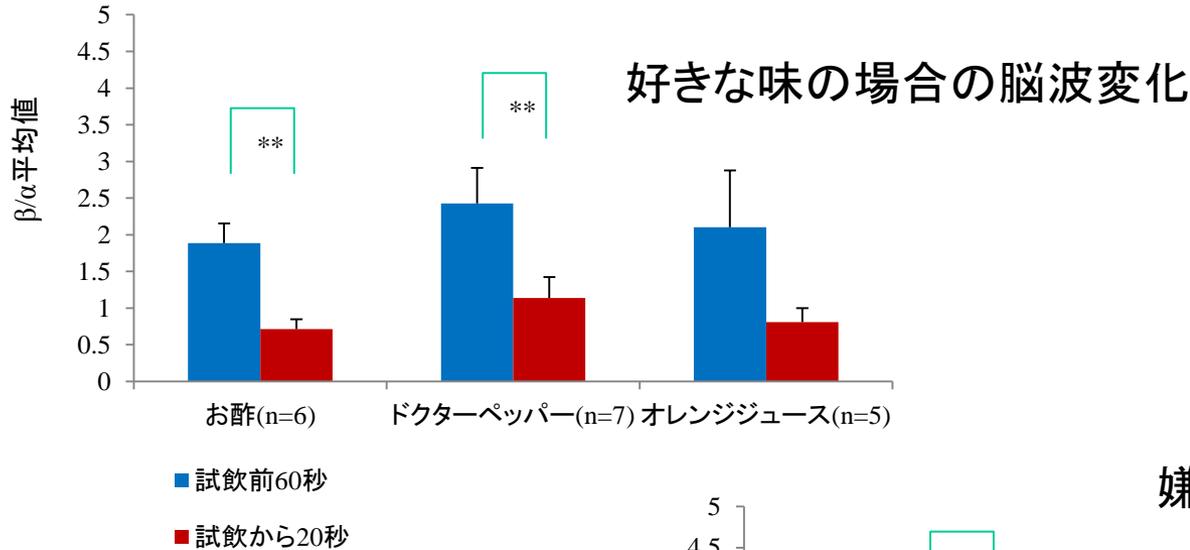
視線の軌跡が類似している場合、それらのUserの嗜好が似ているのではないかと推測される。Fretchet Distanceを用いて、視線の類似度を検証

視線の類似度に着目した User嗜好分析とUserのGroup化 (2/2)



User間で視線の類似度が観測され、嗜好の違いによるUser Groupに分類できる可能性が示唆された。

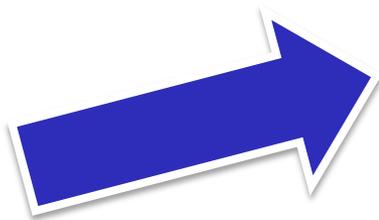
脳波を利用した味覚と個人の味の好みの分析



脳波(α波とβ波の比)によって味の好みを分析できる

視線情報を利用した個人の理解に合わせた画像分類 (1/2)

Target: "Ocean"



分類



類似クラス



比類似クラス



仮説:

Target画像をUserに提示して類似の画像を選んでもらう場合、類似画像と比類似画像では画像に対する視線情報が異なる

→Personalizeされた画像分類

視線情報を利用した個人の理解に合わせた画像分類 (2/2)

(1) 11次元

		予測	
		p	n
正解	p*	115	141
	n*	146	270

precision=0.441

recall=0.479

f値=0.445

(2) 13次元

		予測	
		p	n
正解	p*	129	127
	n*	123	293

precision=0.508

recall=0.504

f値=0.508

視線の特徴量から、個人毎に、類似画像を弁別可能

適用Application

- Recommendation System
 - TabletやSmart PhoneをTarget
- リモコンなしTV
- 「食べログ」評価の自動入力
- Life Log
- 映画Genreの個人化
- 音楽Genreの個人化
- 正確なUser反応をMarketing情報へ反映
- その他、様々な応用が可能