



NAIST Spring Seminar 2020/02/19



# Object detection for video summarization

NAIST, Ubi Lab  
Yohei Katayama

# もくじ

- 自己紹介
- 私の研究
- 2020SpringSeminarのテーマ
- スタート！





## 片山 洋平(かたやま ようへい)

研究分野：画像処理，動画要約

趣味：就活

好きなYouTuber：東海オンエア



@yo\_hey32

### 学部3年

Summer Seminar

Spring Seminar

### 経歴

大規模システム管理研究室 (IoT, ラズパイ)

ユビキタスコンピューティングシステム研究室 (スマートホーム)

### 学部4年

研究室配属

三重大学工学部情報工学科ヒューマンインタフェース研究室

卒研テーマ：動画像解析を用いた新生児の睡眠覚醒状態の自動判定

### 修士1年

研究室配属

7月～現在

ロボットビジョン研究室 (基幹研究室：ユビ研)

ユビキタスコンピューティングシステム研究室(ドラレコ要約)

# 自己紹介



## 片山 洋平(かたやま ようへい)

研究分野：画像処理，動画要約

趣味：就活

好きなYouTuber：東海オンエア



@yo\_hey32

### 学部3年

Summer Seminar

Spring Seminar

### 経歴

大規模システム管理研究室 (IoT, ラズパイ)

ユビキタスコンピューティングシステム研究室 (スマートホーム)

### 学部4年

研究室配属

三重大学工学部情報工学科ヒューマンインタフェース研究室

卒研テーマ：動画像解析を用いた新生児の睡眠覚醒状態の自動判定

### 修士1年

研究室配属

7月～現在

ロボットビジョン研究室 (基幹研究室：ユビ研)

ユビキタスコンピューティングシステム研究室(ドラレコ要約)

# 自己紹介



## 片山 洋平(かたやま ようへい)

研究分野：画像処理，動画要約

趣味：就活

好きなYouTuber：東海オンエア



@yo\_hey32

### 学部3年

Summer Seminar

Spring Seminar

### 経歴

大規模システム管理研究室 (IoT, ラズパイ)

ユビキタスコンピューティングシステム研究室 (スマートホーム)

### 学部4年

研究室配属 4月

三重大学工学部情報工学科ヒューマンインタフェース研究室

卒研テーマ：動画像解析を用いた新生児の睡眠覚醒状態の自動判定

### 修士1年

研究室配属 5月

7月～現在

ロボットビジョン研究室 (基幹研究室：ユビ研)

ユビキタスコンピューティングシステム研究室(ドラレコ要約)

# Video summarization



## What to summarize ?

- Target: **Drive Recorder Video**



## Why summarize ?

- In our algorithm, summarized video made for tourism.  
ex. Sightseeing planing, Memorial Movie
- Too long video (not real-time)
- It's difficult for viewer to know what is important part.

# Video summarization



## What to summarize ?

- Target: **Drive Recorder Video**



## Why summarize ?

- In our algorithm, summarized video made for tourism.  
ex. Sightseeing planing, Memorial Movie
- Too long video (not real-time)
- It's difficult for viewer to know what is important part.



We need to extract important scene !



# Demo Movie



## Our System

# Demo Movie



奈良先端科学技術大学院大学 情報科学領域  
**ユビキタスコンピューティングシステム研究室**

Ubiquitous Computing System Laboratory, Nara Institute Science and Technology

<http://ubi-lab.naist.jp>



## Automatic Video Summarization System Using Drive Recorder Movie

**Summarized Video**

自分にあったルートを探してみましょう。



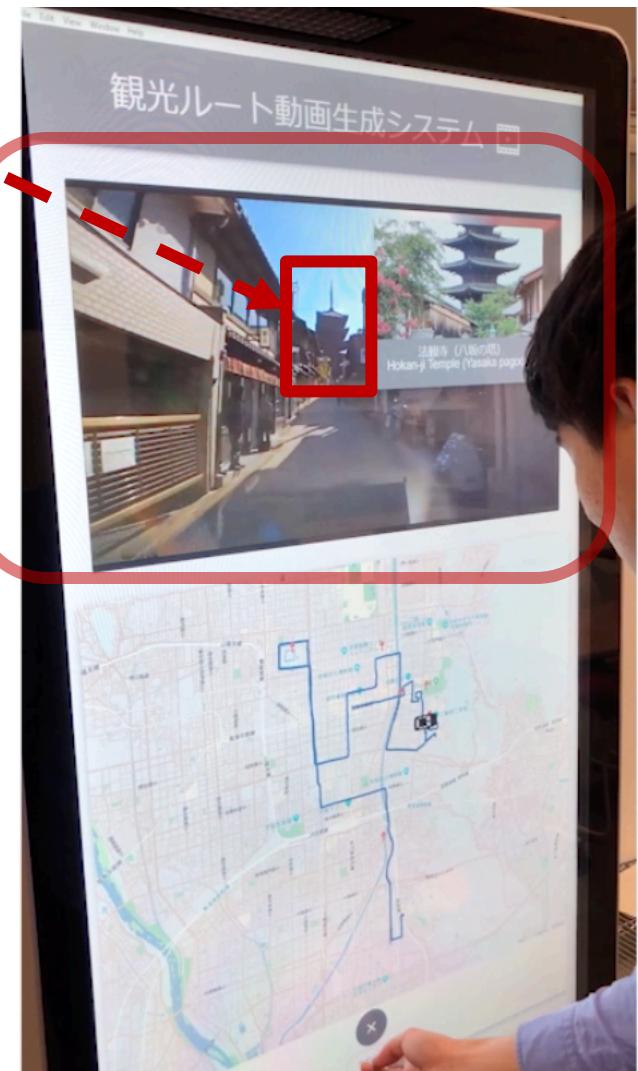


**Main target**  
(Japanese traditional object)

## Automatic Video Summarization System Using Drive Recorder Movie

**Summarized Video**

自分にあったルートを探してみましょう。



# Objet Detection



- There are a lot of Japanese traditional objects along the street.

Japanese Object



Pagoda



Torii



Latern

# Objet Detection



- There are a lot of Japanese traditional objects along the street.

Your Mission

**Object Detection**  
using machine learning

Japanese Object



Pagoda



Torii



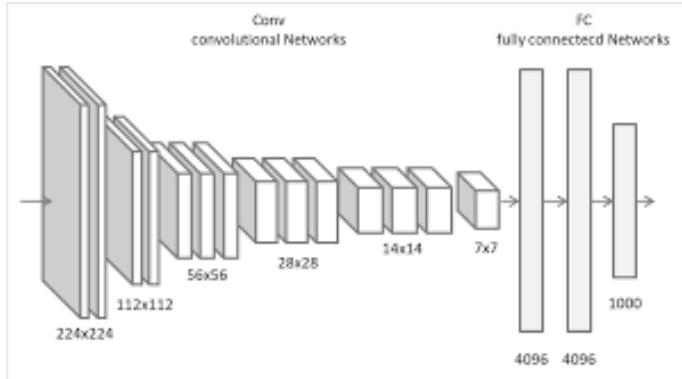
Latern



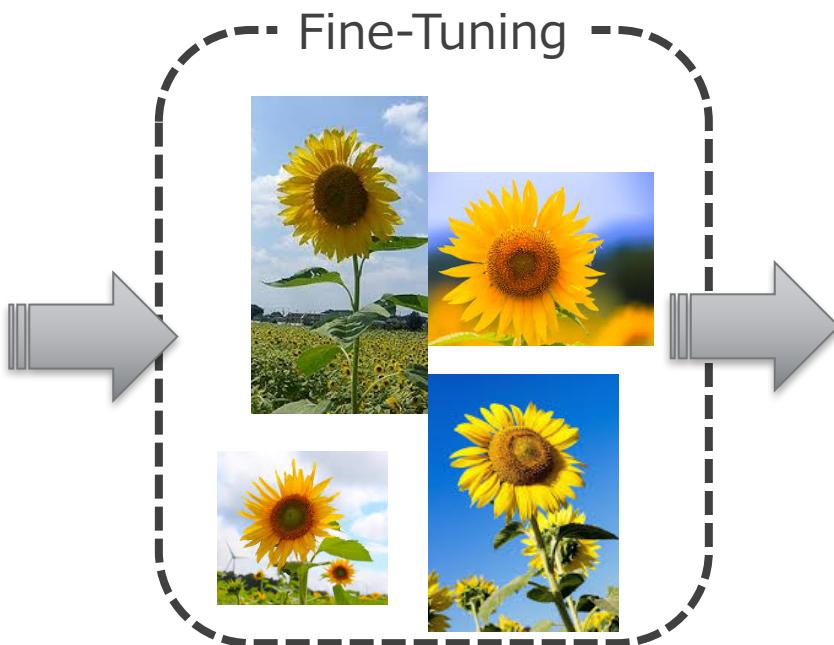
## Fine-Tuning (similar to Transfer Learning[1])

- One of the Machine learning methods.
- This method use trained models to adapt to different domains

ex) Pre trained model VGG16



1000 category  
(sunflower is not included)



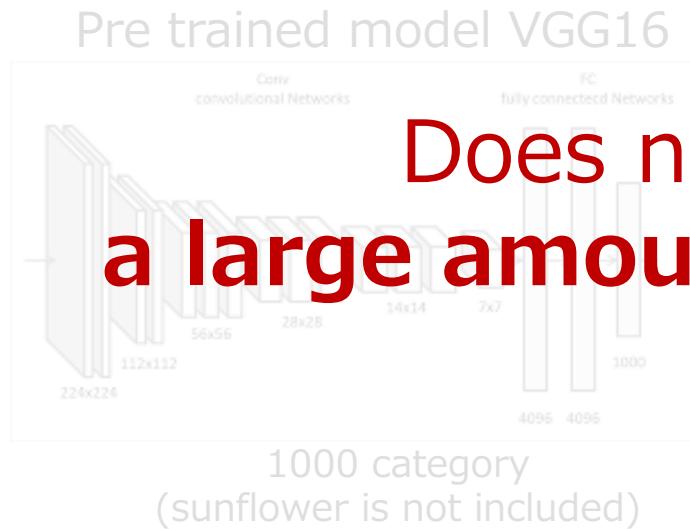
Sunflower detection Model



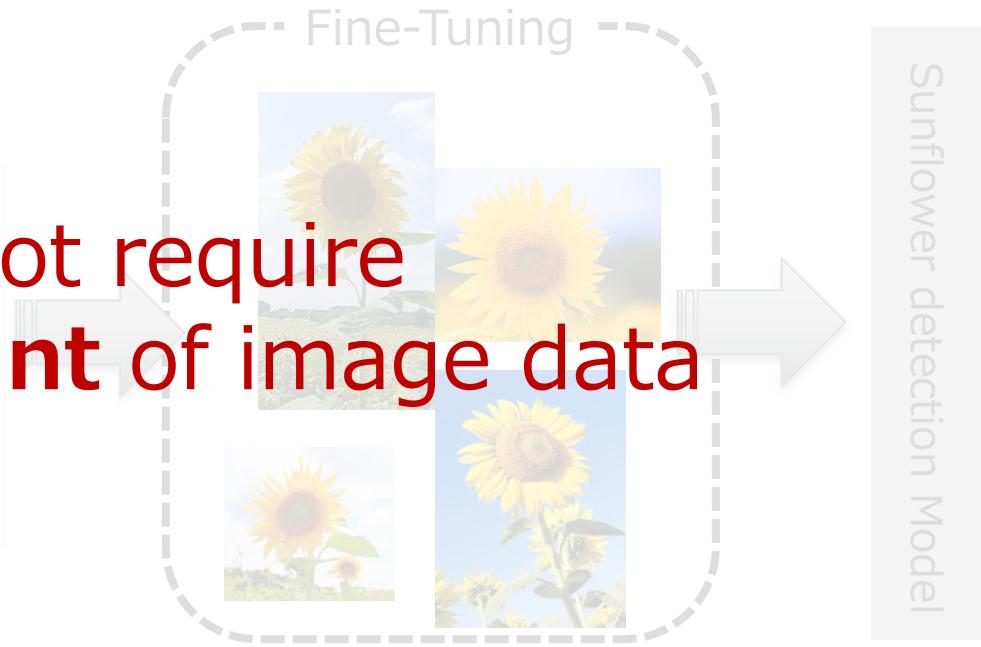
## Fine-Tuning (similar to Transfer Learning[1])

- One of the Machine learning methods.
- This method use trained models to adapt to different domains

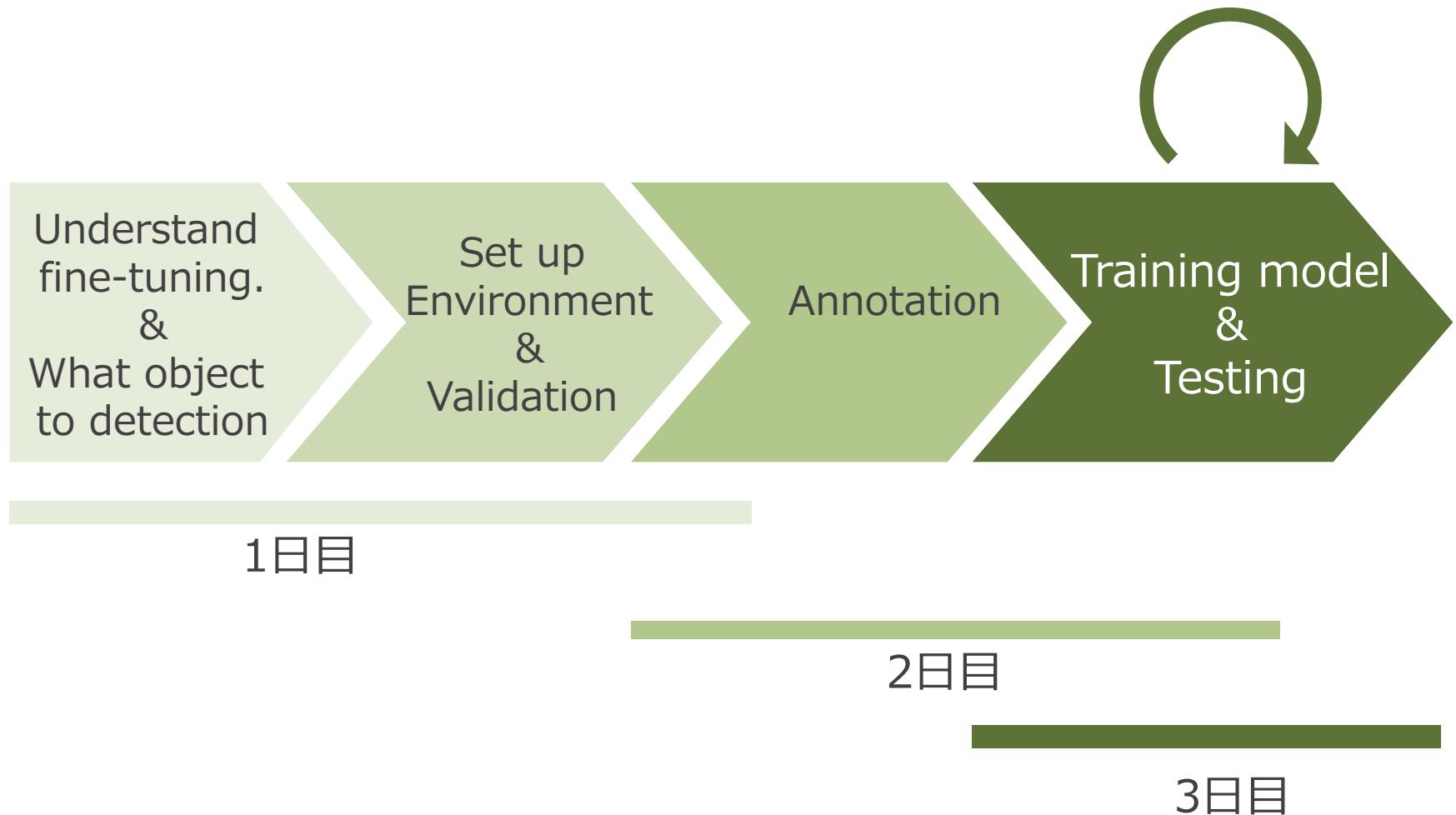
ex)



Does not require  
a large amount of image data



# Provisional Schedule





- ・アカウントの有無(GitHub, Googleアカウント)
- ・PC環境確認(OS, Python)
- ・理解度確認(コマンド操作, 使用言語)
- ・スタート！

# 研究背景と目的



## 研究背景

観光動画の需要拡大

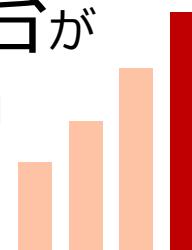
40%の観光客が  
観光計画に  
動画を利用<sup>[1]</sup>



× 動画の収集が困難

ドライブレコーダの普及

139万台が  
年間に販売<sup>[2]</sup>  
(2018年)



× 活用が限定的

+



ドライブレコーダ動画に観光という新たな**価値**を与える。

[1] Ipsos MediaCT, et al. The 2014 traveler's road to decision\_Google TravelStudy, 2014.

[2] GfKジャパン調べ、2018年ドライブレコーダの販売動向



## 研究目的

- ・ドライブレコーダ動画を自動要約するアルゴリズムの検討
- ・要約における重要シーン検出

## 研究課題

Step 1. ユーザが理解しやすい動画に自動要約

→ 動画の要約アルゴリズムの検討



Step 2. 各ユーザの特性に合わせた要約

→ ユーザに必要な情報の取捨選択



# 本システムの全体像



## 解析 (研究対象)



ドライブレコーダ動画



GPSデータ



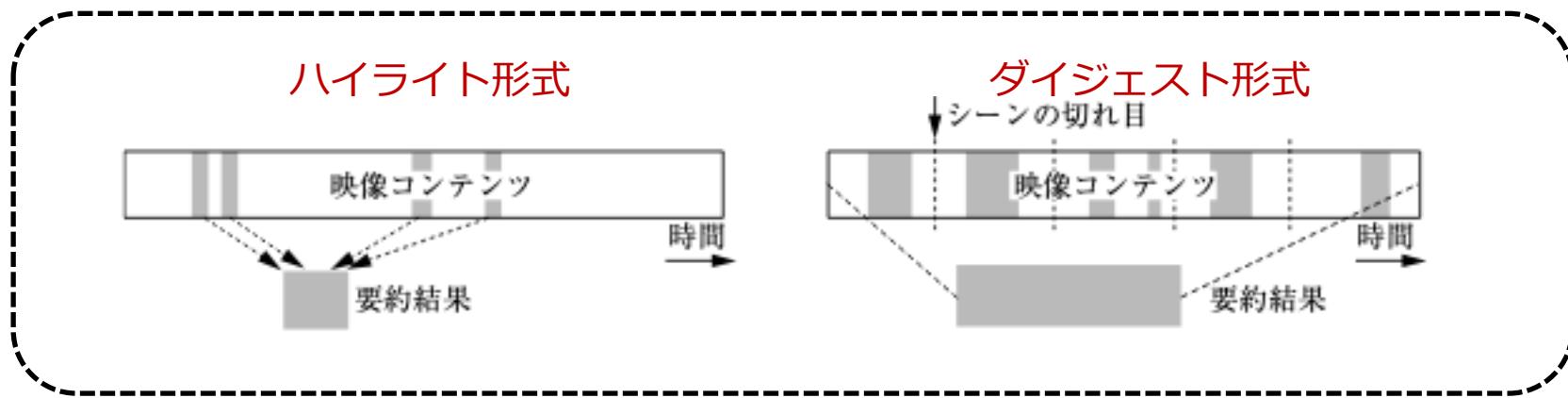
キュレーション動画





## 動画要約技術について<sup>[3]</sup>

- ・ スポーツ：得点に関わる，盛り上がったシーンを抽出
- ・ ニュース：司会者の出演するシーンを抽出
- ・ ドラマ：会話やアクションシーンを抽出



→ 観光の要約動画ではシーンの重要度に合わせた**再生速度の変更**が望ましい。

[3] 滝嶋康弘, 映像の自動要約技術.映像情報メディア学会誌, Vol. 62, No. 5, pp.714-716, 2008.



## 金谷の研究

*Automatic Tour Video Summarization Focusing on Scene Change for Advance Touristic Experience [4]*

- 観光動画のカラーヒストグラムを利用して**自動要約**.
- 観光動画の提示において**4倍から8倍速**が適切であるとの知見.
- ユーザ評価において70%が**観光計画に有効**との回答.

### 類似点

- ・観光動画の自動要約技術の研究

### 相違点

- ・自動車と歩行者での移動速度の違い

### 問題点

- ・定量的な評価が為されていない



# 先行研究との比較



	目的	移動方法	最適な 再生速度	要約 アルゴリズム
金谷の研究	観光計画	徒歩	4~8倍速	カラー ヒストグラム
Ours	観光計画 + 振り返り	自動車  	?  	後述  

**○** 

信号による  
一時停車あり

移動速度が  
速い

自動車の移動に  
合わせて**最適化**

# 提案手法



金谷手法

- ・カラーヒストグラム
- ・観光地の位置情報
- ・ユーザ嗜好(PoI)

提案手法

- ・速度
- ・オブジェクト検出
- ・綺麗な景色の検出



ドライブレコーダ動画



GPSデータ



キュレーション動画



# 提案手法

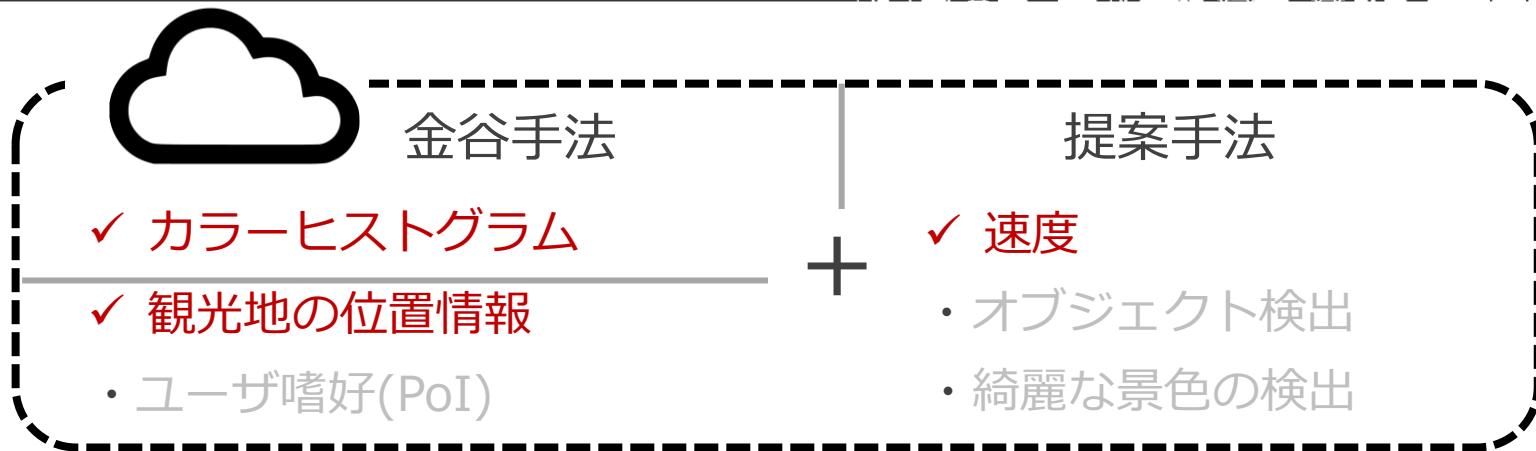
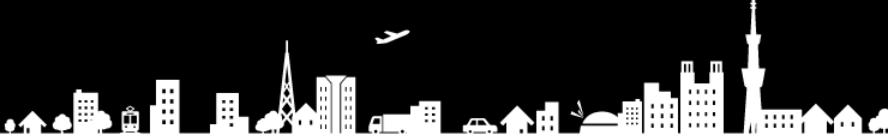


	Input	Output	技術
観光地の位置情報	GPS	観光地近くの検出	閾値で判定可能
速度	速度計	停車の検出	
オブジェクト検出	画像	オブジェクトの検出 e.g. 門, 鳥居	YOLOv3 <sup>[5]</sup> CNNベースの物体検出モデル
綺麗な景色の検出	画像	非オブジェクトの検出 e.g. 海, 空	NIMA <sup>[6]</sup> CNNベースの景観判定モデル
ユーザ嗜好	検討中 (ソーシャルメディア)	ユーザ特性から 必要・不要シーンの判定	多数論文あり

[5] Redmon, Joseph, and Ali Farhadi. "Yolov3: An incremental improvement." *arXiv preprint arXiv:1804.02767* , 2018

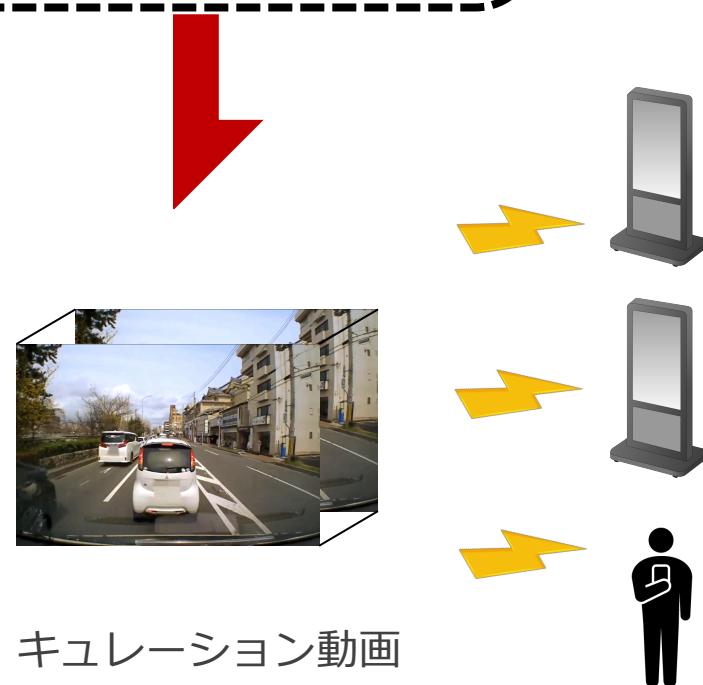
[6] Talebi, Hossein, and Peyman Milanfar. "Nima: Neural image assessment." *IEEE Transactions on Image Processing* 27.8 , 2018

# 現在の進捗



ドライブレコーダ動画

GPSデータ



# 現在の進捗



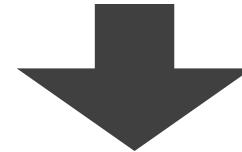
14.6%



元動画：1h27m11s



6.8%



重要：4倍  
不要：8倍

カラーヒストグラムのみ：12m47s

$$\text{圧縮率} = \frac{\text{キュレーション動画}}{\text{元動画}}$$



重要：2倍  
不要：32倍  
停止時：128倍

提案手法の一部：5m57s

# 結論と今後の課題



- ✓ ドライブレコーダーを用いた観光動画キュレーションシステムの提案を行なった。
- ✓ 提案手法のアルゴリズムの一部を用いて実際に動画要約を行なった。

## 今後の課題

- ・オブジェクト、綺麗な景色の検出アルゴリズムの検討
- ・シーンの再生速度の検討
- ・評価尺度の検討

## 今後の展望

- ・金谷の研究(歩行時の観光動画要約)との統合



# Appendix

# NIMA : Neural Image Assessment



美しさの評価はAVA: A large-scale database for aesthetic visual analysis - IEEE Conference Publicationを用いる。

画像の品質に関する評価はTID2013を用いる。



(a)



(b)



(c)



(d)



(e)



(f)



(g)



(h)



(i)



(j)



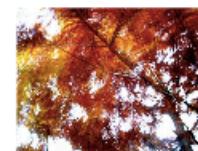
(k)



(l)



(m)



(n)



(o)

# 今後の課題(Next Step)



解析(Analyze)

Color Histogram

Velocity

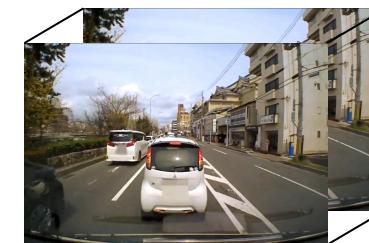
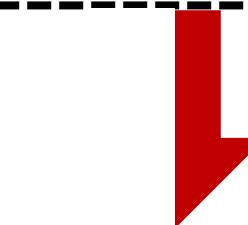
Point of Interest

Person Detection

Object Detection



Original Movie



Curation Movie

Notification



Evaluation