

セマンティック・セグメンテーションにより抽出した

景観構成要素に基づく視点のクラスタリング

—大阪環状線の車窓シーケンス景観を対象として

CLUSTERING VIEWPOINTS BASED ON LANDSCAPE COMPONENTS

EXTRACTED BY SEMANTIC SEGMENTATION:

THE SEQUENCE LANDSCAPE FROM A TRAIN WINDOW OF THE OSAKA LOOP LINE

環境図形科学研究室 土井 亜香里

Spatial & Graphic Science Lab. Akari DOI

車窓景観はその沿線地域や街の印象を与えてくれる可能性があるが、従来のシーケンス景観の研究手法では多量の画像の分析を行うことが難しい。そこで画像系 Deep Learning の一種であるセマンティック・セグメンテーションを利用した手法を提案し、自動抽出した景観構成要素が分析に利用できることを示す。またシーケンス景観を空間で分割し、視点とマッチングした上で、構成比によりその視点のクラスタリングを行い、景観の連続性や地域の変化について検討する。

The landscape from a train window may give an impression of the area along the railway, but it is difficult to analyze a large number of images using the conventional research method of sequence landscape. Therefore, the purpose of this paper is proposing a method that uses semantic segmentation, a kind of image-based deep learning, and showing that automatically extracted landscape components can be used for analysis. Moreover, we divide the sequence landscape spatially and match it with the viewpoints. Further we cluster the viewpoints based on the ratio of landscape components and examine the continuity of the landscapes and the changes of area.

1. はじめに

我々が車窓から眺める風景である「車窓景観」は、具体的な街の区切り線こそないものの、電車の通る線路沿いの様々な街の雰囲気や印象、イメージを与えてくれるものであり、その地域を想像する一種のきっかけである。車窓景観は多くの構成要素を含み^[1]、たとえばその緑視率や天空率の変化は、沿線の街並みの変化と関係があると思われる。そのため沿線の地域のイメージを考えると、その車窓景観を指標のひとつとできるのではないかと。

しかし、連続するシーケンス景観の特徴や景観の構成要素との関係を定量的に分析した研究は少ない。また、従来の景観研究では構成要素を人力で分割するものが多く、多量の画像に対しては実施が難しい。

そこで本研究では景観構成要素の抽出をセマンティック・セグメンテーションにより自動化する。それを鉄道沿線の都市景観を電車から眺めた際の、視点移動を伴う「車窓シーケンス景観」に対し実施することで、手法の有意性を確認する。また、その構成比の

結果をもとに視点をクラスタリングすることで、景観の連続性や地域変化の把握を目指す。

対象路線として、JR 西日本の大阪環状線を選定した。大阪環状線は全 19 駅、総延長 21.7km の路線で、大阪内の都市景観を一望できる唯一の地上環状路線である。路線図(駅名)を図 1 に示す。商業地から住宅地まで、様々な地域を通過するため、ビル、住宅街、緑地等の様々な景観を含んでいる。

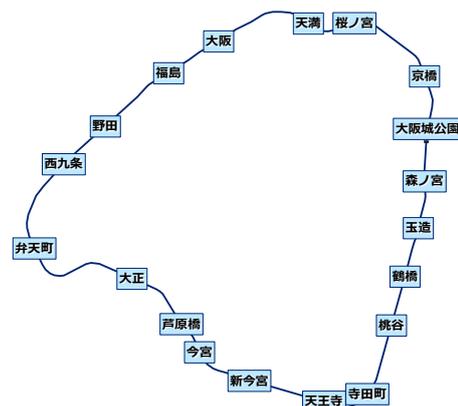


図 1 大阪環状線の路線図

2. 撮影手法

動画の撮影は1日目：2019年3月11日，2日目：6月3日に行い，そのうち2日目中に撮影したものを利用した．撮影時の環状線乗車に際しては，以下の乗車券をそれぞれ購入し使用した．どちらも330円である．

乗車券（外回り）「天王寺→天王寺」

経由：関西・大阪環状・西九条・大阪環状・京橋

乗車券（内回り）「天王寺→天王寺」

経由：大阪環状・京橋・大阪環状・西九条・関西

本撮影である2回目の撮影について表1に示す．

天候は曇天～晴れであり，撮影機材はSONY FDR-AX700（4Kビデオカメラ），raynox HDP-7880ES（広角0.79倍レンズ），窓の映り込み防止のための自作フードである．自作フードは段ボールで型枠を作り，内部に黒画用紙を貼ることで反射光を吸収できるようにした．

カメラは車内進行方向左側の車窓に対し垂直，床面からの高さ135cmに設置し撮影した．これは起立時と着座時の視線の中間を意図している．

撮影した動画はTMPGEnc Video Mastering Works 6を利用し，0.1秒ごとの静止画として抽出した（図2右部）．

撮影時にGeo TrackerをつかってGPSの位置座標を記録した．これを秒単位でメモした発着時刻と照らし合わせることで，後述の時空間マッチング手法から画像撮影時の位置を復元できるようにした．

表1 2日目：2019年6月3日の撮影概要

実際の撮影時刻	電車の発着時刻	
9:44~10:30	9:44~10:28	外回り
10:41~11:22	10:41~11:22	内回り →解析に利用
11:42~12:03 12:03~12:26	11:42~12:26	外回り
12:26~13:11	12:27~13:11	外回り →解析に利用
13:12~13:14 13:20~13:22	不明	内回り 撮影失敗
14:41~15:26	14:42~15:26	外回り
15:54~16:38	15:55~16:37	内回り



図2 撮影のイメージと動画の分割

3. 静止画の時空間マッチング

従来のシーケンス景観研究では，撮影開始からの秒数を計測し，等時間間隔でシーケンス景観を分割することで分析を行っている．例えば竹内^[2]は，時速30kmで走行する普通乗用車の助手席にカメラを設置し撮影した60秒間の沿道シーケンス景観を2秒間隔31枚の静止画に分割することで分析を行っている．

しかし，本研究では電車車内から撮影するため，速度変化が生じ，車窓景観を等間隔に分割するのが難しい．またクラスタリングに伴い，沿線との地域マッチングをすることで傾向を見たいと考えた．そのため本研究では長さをベースとして，電車がそれぞれの駅で発車した地点から20m間隔でそれぞれの区間を区切り，そこに撮影した画像を紐づけることとする．またそのために，電車の速度モデルを仮定し，0.1秒区切りの連番静止画から20m間隔にあたる画像を抽出する．概要を図3に示す．

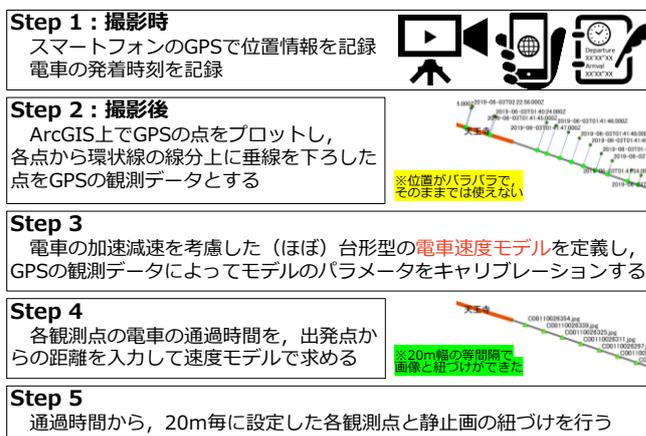


図3 静止画の時空間マッチングの概要

3.1. 時空間マッチングの概要

動画は撮影開始時間が記録されているので，その時間とフレームレートから切り出した任意の静止画の撮影時刻がわかる．一方GPSで取得した情報は，3次元の位置情報と1秒単位の世界標準時である．したがって，時刻で静止画と紐づけることができれば，静止画の撮影位置を特定することができる．

しかしGPSの位置情報には無視できない測定誤差と欠損（電車停車時にはデータを取得しないため）があるため，直接紐づけることは難しい．

そこで時刻による電車の加速→一定速度→二段階減速という速度変化を台形状の区分線形関数としてモデル化し（電車速度モデル），GPSの観測データを用いてモデルのパラメータのキャリブレーションを行う．そして得られたモデルを使って，任意の地点から等間隔に観測点を設定した際に，電車がそこを通過する時刻を推計し，各観測点に対応した静止画を抽出する．

3.2. 電車速度モデル

図 4 に電車速度モデルの概念図を示す. この速度モデルは任意の駅を出発して次の駅に停車するまでの速度変化を表現する. グラフの横軸は経過時刻 (秒), 縦軸は各時刻における速度を表す. また, 当該の駅を電車が出発する時刻を原点にとる.

時刻 t における電車の速度 $v_i(t)$ は式(1)により推計される. また GPS のデータ (平面直角座標系の 6 系) に従った際の h における電車の速度 v_h は式(2)で表される.

$$v_i(t) = a_i t + b_i \quad \text{式(1)}$$

$i = 1, \dots, 4$

直線のパラメータ (a_i, b_i) は, t が

時間 $[0, t_2], [t_2, t_3], [t_3, t_4], [t_4, t_5]$ の

いずれの区間に入るかにより,

l_1 から l_4 までの線分のいずれかを示す

p : 折れ線の頂点の集合

l : 折れ線の集合

$$v_h = \frac{\sqrt{(x_g - x_h)^2 + (y_g - y_h)^2}}{t_h - t_g} \quad \text{式(2)}$$

連続した 2 時点の GPS データを時刻が早いほうから順番に $g, h \in G$, その時の時刻をそれぞれ $t_g, t_h (t_g < t_h)$ と表す

G : GPS の観測データ集合

(x_g, y_g) : $g \in G$ におけるその平面座標

(t_g, v_g) : $g \in G$ の GPS データにおける時刻と電車の速度のペア (図 4 の赤 ●)

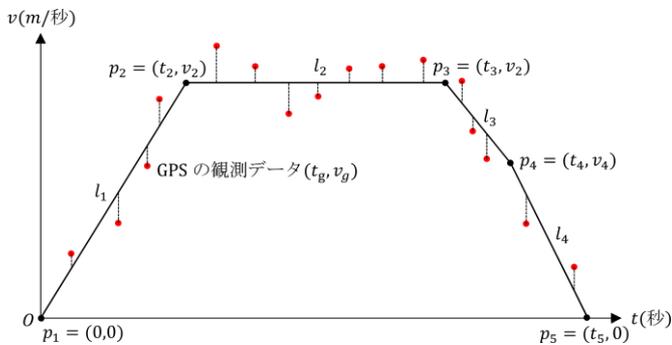


図 4 電車速度モデルと GPS プロット

3.3. 速度モデルのパラメータのキャリブレーション

基本的にはこれらの誤差の和を最小化するようにして決定される. すなわち, 式(3)の最適化問題を解いて決定される. ただし, 概形が台形になることを優先するため, 必ずしも誤差最小時のパラメータを用いるわけではない.

$$\text{Min.} \sum_{g \in G} (v(t_g) - v_g)^2 + (dsum - dis(g_s, g_e))^2 \quad \text{式(3)}$$

$$\text{Sub.} \quad t_2 < t_3 \leq t_4$$

$$v_2 = v_3$$

p_2, p_3, p_4 は速度モデルの概形がおよそ台形モデルになるように制約される

$g_s, g_e \in G$: 電車の発車位置, 停車位置

3.4. 距離から地点の通過時間を求める

$0 \leq dis(u) < dis(t_2)$ の時

$$dis(u) = \int_0^u v_1(t) dt$$

$dis(t_2) \leq dis(u) < dis(t_3)$ の時

$$dis(u) = \int_0^{t_2} v_1(t) dt + \int_{t_2}^u v_2(t) dt$$

$dis(t_3) \leq dis(u) < dis(t_4)$ の時

$$dis(u) = \int_0^{t_2} v_1(t) dt + \int_{t_2}^{t_3} v_2(t) dt + \int_{t_3}^u v_3(t) dt$$

$dis(t_4) \leq dis(u) < dis(t_5)$ の時

$$dis(u) = \int_0^{t_2} v_1(t) dt + \int_{t_2}^{t_3} v_2(t) dt + \int_{t_3}^{t_4} v_3(t) dt + \int_{t_4}^u v_4(t) dt \quad \text{式(4)}$$

$dis(u)$: 電車が出発してから u 秒後の出発点 g_s からの経路上の距離

ここで, $dis(u)$ は 20m 等間隔で設置されるため, 定数としてあらかじめ定まっている. したがって式(4)の積分を実施し, u に関し解くことで, 電車が距離 $dis(u)$ を通過する際の時刻 u を求めることができる.

クラスタリングの都合上, 外回りと内回りの撮影位置を合わせる必要があるため, 各内回り始点を各外回り終点にずらす位置合わせを行っている.

4. セマンティック・セグメンテーションによる分析

セマンティック・セグメンテーション (SS) は画像系 Deep Learning の一種で, 画素 (ピクセル) 単位でそれが何かを認識するタスクを行う手法のことである. 自動運転や医療用画像処理等の用途で使用される. 前者では車両, 歩行者, 交通標識, 歩道, 道路などの特徴を識別する必要があるため, セグメントとしてそれらが用意されているモデルが多い.

4.1. セマンティック・セグメンテーション: 条件設定

DeepLab v3+ (xception71_dpc_cityscapes_trainval)^[3] の Tensorflow によるモデルの実装を利用した. また, データセットは主に Cityscapes^[4] を利用し (図 5), 水域が含まれる画像に対してのみ ADE20K^[5] を実行, 該当部分を重ね合わせるという処理を行った.

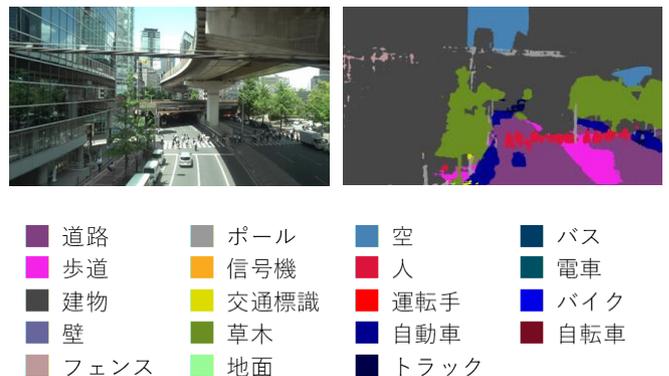


図 5 Cityscapes の SS の例 (上段左: 元画像, 上段右: SS, 下段: Cityscapes の分類)

4.2. セマンティック・セグメンテーション: 水域追加

データセット: Cityscapes の SS は、8 割以上の精度があるものの、水域の分類が存在しないため、川や海などの水域が道路や壁に分類されてしまう。そのためデータセット: ADE20K の SS を利用して、水域の部分だけ補完することを考える。ADE20K は精度がよくないので、水に属すると思われる属性（水、海、川、噴水、滝、湖）を全て水域として、二値化するように実施し（図 6）、Cityscapes に補完した（図 7）。

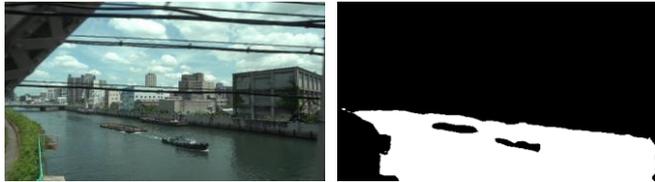


図 6 ADE20K の SS の例（左：元画像 右：SS）

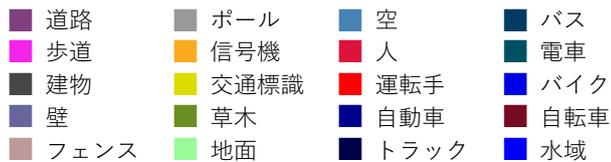
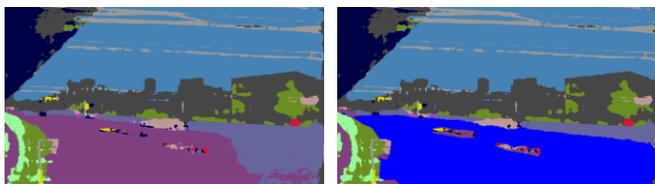


図 7 SS の比較（上段左：Cityscapes, 上段右：水域補完後, 下段：本研究で使用する SS の分類）

4.3. セマンティック・セグメンテーション: 結果

SS から得られたピクセル数から景観構成要素の面積比を算出し、その推移の例を図 8 に示す。また構成要素のうち、6 種類（空、建物、草木、道路、歩道、水域）の構成割合を色の濃度とみなし Processing を用いて可視化した（図 9）。その際、面積割合は各構成要素内で [0, 1] に正規化(Normalization)している。

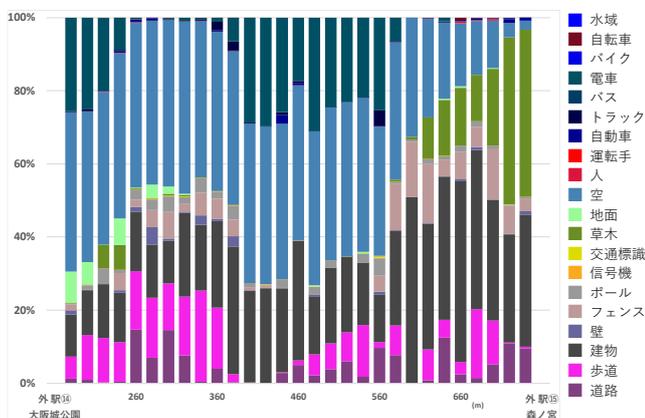


図 8 景観構成要素の面積比の推移の例
「外回り, 大阪城公園—森ノ宮」

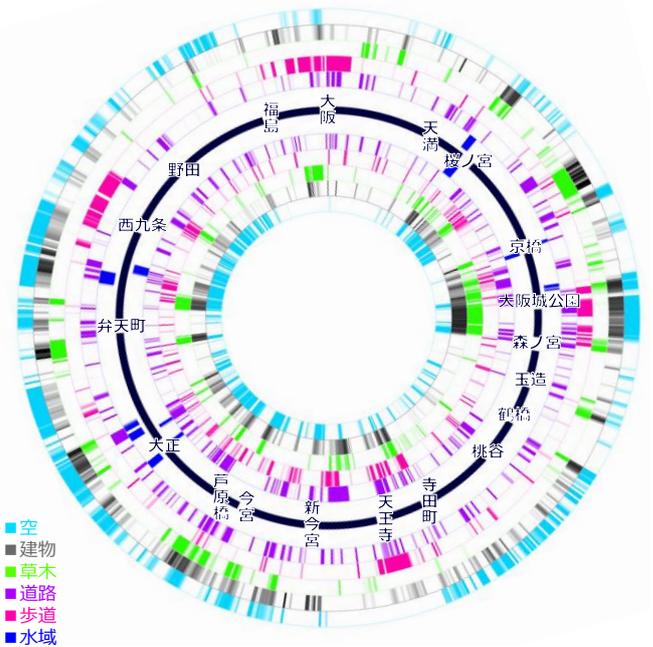


図 9 6 種類の景観構成要素の濃度的可視化

全体として建物の構成比の多い区間、空の構成比の多い区間、草木の構成比の多い区間がそれぞれ顕著で、その上で道路・歩道、水域が該当する部分にあるように見える。

5. 画像のクラスタリング

本研究では、SS の画像を元を得た景観構成要素の面積比のデータをクラスタリングすることで、景観の連続性や構成要素の変化のタイミング等を把握、分析する。これにより、電車から車窓景観を眺めるときの景色の変遷、つながり方を包括的に見ることができると考える。クラスタリングとは、様々な性質のものがまじりあった集団から、定義した類似度を元に、似た性質を持つクラスタを作る方法のことである。

本研究では階層的手法（階層クラスタ分析）で分類し、分類に用いる類似度（対象間の距離）の定義はユークリッド距離、クラスタ間の距離の測定方法は Ward 法を用いる。

なお、階層的手法ではクラスタリング後にクラスタ数を決定する必要があるが、本研究ではデータに Box-Cox 変換を施した後、Upper Tail 法により最適と思われるクラスタ数を求め、それを利用する。また、データをばらばらのものとして扱う隣接制約なしのクラスタリングと、景観として隣り合う、連続するものとして、ある一定区間を同一クラスタとみなす隣接制約ありのクラスタリングを考える。

なお、全体としては面積が少なくても、景観に大きな影響を与える構成要素がありうることを考慮し、景観構成要素は各項目内で平均が 0、分散が 1 の正規分布になるように標準化 (Standardize) している。

5.1. クラスタリング：ユークリッド距離

最も一般的で日常的に用いられる距離の概念であり、 n 次元（パラメータ数 n ）の距離 $d(x, y)$ は座標 $X(x_1, \dots, x_n), Y(y_1, \dots, y_n)$ を用いて式(5)で表される^[6]。

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + \dots + (x_n - y_n)^2} \quad \text{式(5)}$$

5.2. クラスタリング：階層クラスタ分析

最も似ている組み合わせから順番にクラスタにしていき、最終的に全サンプルが併合された、1つのクラスタにまとまった状態まで行う手法である。途中のプロセスを階層として表すことができるため、横軸をサンプル、縦軸を類似度（距離）にとったデンドログラムを作成することができる（図 10）^[7]。

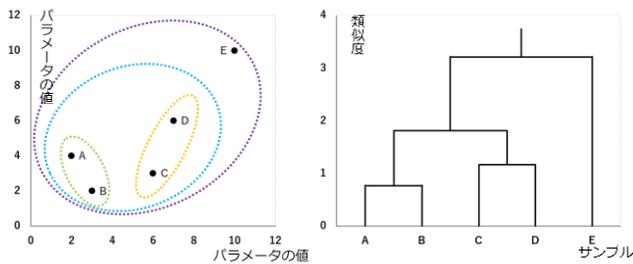


図 10 パラメータ数 2 のときの階層クラスタ分析のイメージとデンドログラム

5.3. クラスタリング：Ward 法

2つのクラスタを結合したと仮定したとき、それにより移動したクラスタの重心とクラスタ内の各サンプルとの距離の二乗和と、元々の2つのクラスタ内での重心とそれぞれのサンプルとの距離の二乗和の差が最小となるクラスタ同士を結合する手法である（図 11）。この差を情報ロス量といい、式(4)のように表す^[8]。

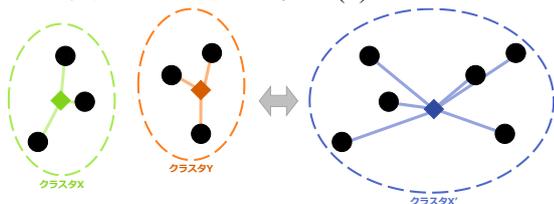


図 11 Ward 法のイメージ

$$\Delta = L(X') - L(X) - L(Y) \quad \text{式(4)}$$

Δ : 情報ロス量

$L(A)$: クラスタ A の重心と各サンプルの距離の二乗和
 X' : $X \cup Y$

5.4. クラスタリング：Upper Tail 法と Box-Cox 変換

Upper Tail 法は、階層クラスタ分析におけるクラスタ数決定法の 1 つで、統計的な停止規則を用いてクラスタ数を求めるものである。大きさ n の標本に対してクラスタを生成するための距離 α は $n-1$ 個である。クラスタ数 1 の分割時の距離 α_1 から降順に、クラスタ数 $n-1$ の分割時の α_{n-1} まで存在し、 α の平均と標準偏差から有意な α を導くことでクラスタ数が決定される。

停止規則は、 $j=1$ から始め、式(5)を満たすまで j を増加させることであり、停止した j が最適なクラスタ数となる。 k は 2~4 の間で適宜選定する^[9]。

本研究では、サンプルの存在割合が 98% になるときの k の値を逆算し、そこから最適なクラスタ数を計算する。したがって今回、 $k = 2.326\dots$ である。

$$\alpha_j \leq \alpha + ks_\alpha \quad \text{式(5)}$$

$\bar{\alpha}$: α の分布の平均 $\bar{\alpha}$

s_α : α 不偏分散の平方根

Upper Tail 法では標本集団が正規分布に従うデータを扱うが、今回の素データは正規分布に従わないため、データを正規分布に変換することで Upper Tail 法を適用できるようにする必要がある。

正規分布に変換する方法として、Box-Cox 変換を用いる。元の値 x は式(6)によって y に変換され、パラメータ λ は変換後の y の分布が最も正規分布に近くなるように、プロファイル尤度法で最適化される^[10]。

$$y^\lambda = \begin{cases} \frac{x^\lambda - 1}{\lambda} & (\lambda \neq 0) \\ \log(x_i) & (\lambda = 0) \end{cases} \quad \text{式(6)}$$

求めた最適クラスタ数を表 2 に示す。

表 2 条件設定と $k = 2.326$ のときのクラスタ数

	外回り	内回り	外・内
隣接制約なし	11	10	13
隣接制約あり	16	28	10

図 12 はクラスタリング結果の例である。およそ A, C, E, G, I の 5 クラスタに分けられる。それらの主な構成要素比（標準化済）の平均値を表 3, 表 4 に示す。

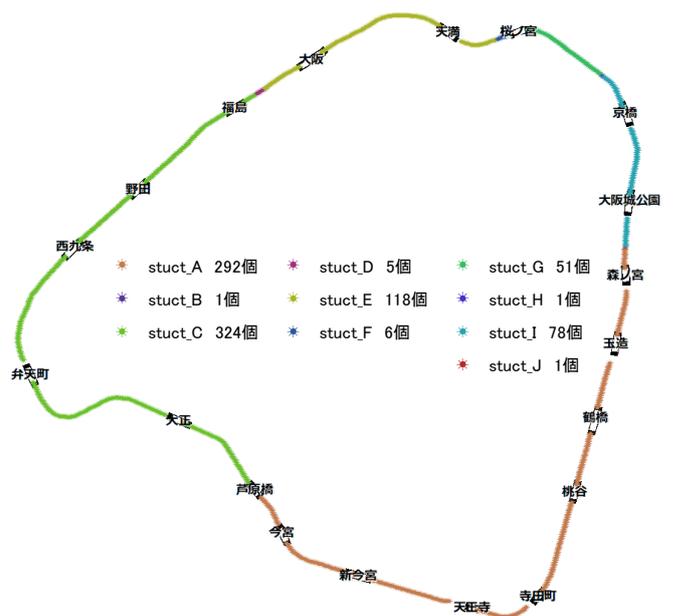


図 12 クラスタリング結果の例 (外・内回り 隣接制約あり)

表 3 構成要素比の平均値 (外回り, 抜粋)

クラス 個数	stuct_A 292 個	stuct_C 324 個	stuct_E 118 個	stuct_G 51 個	stuct_I 78 個
外 道路	0.06	-0.08	0.02	0.02	0.1
外 歩道	-0.08	-0.15	0.6	-0.31	0.27
外 建物	0.15	-0.11	0.41	-0.01	-0.5
外 壁	0.09	-0.02	-0.19	-0.04	0.12
外 草木	-0.09	-0.05	-0.14	0.73	0.23
外 地面	0.14	-0.18	-0.15	-0.01	0.37
外 空	-0.15	0.39	-0.64	-0.36	0.17
外 水域	-0.17	0.15	-0.04	-0.17	0

表 4 構成要素比の平均値 (内回り, 抜粋)

クラス 個数	stuct_A 292 個	stuct_C 324 個	stuct_E 118 個	stuct_G 51 個	stuct_I 78 個
内 道路	0.07	-0.01	-0.32	-0.03	0.28
内 歩道	0.08	0.02	-0.29	0.19	-0.14
内 建物	0.13	-0.07	0.67	-0.36	-0.88
内 壁	0.11	-0.07	-0.16	-0.11	0.2
内 草木	-0.17	-0.13	-0.03	0.14	1.18
内 地面	-0.03	-0.18	0.03	-0.24	1
内 空	-0.16	0.47	-0.9	0.76	-0.49
内 水域	-0.14	0.02	-0.13	-0.14	0.13

表 3, 表 4 の構成要素比は構成要素ごとに比較し、赤に近いほど(クラス間で、相対的に)値が大きく、黄色は中央的、緑に近いほど値が小さいことを示す。

それぞれのクラスの特徴を以下に示す。

- A (森ノ宮—天王寺—芦原橋) は草木や水域が少なく、建物、壁がやや多い。中央的な区間である。
- C (芦原橋—西九条—福島) は空と水域が多く(特に外、環状線全体から見て西側)、草木と地面が少ない。
- E (福島—大阪—桜ノ宮) は外・内両側に建物が多く、外側(環状線北側)に歩道が多い。空が少ない。
- G (桜ノ宮—京橋) の外側は草木が多く、歩道が少ない。内側は歩道と空が多く、草木もやや多い。
- I (京橋—森ノ宮) は草木、地面、道路、水域が多い。特に内側(環状線内・西側)に草木と地面が目立つ。その分、建物の占める割合が少なめである。

6. 結果の考察

面積比推移、濃度的可視化、クラスタリングを総じて、以下に全体的な特徴をまとめる。

- 環状線西側では空と水域が目立つ。青や水色のイメージ。
- 環状線北側、大阪駅前後では建物が多く、空が少ない。外側には歩道の顕著な区間がある。
- 環状線北東部、桜ノ宮—京橋—大阪城公園—森ノ宮では草木、地面が多い。なお、SS の地面は緑地の意を含んでおり、緑のイメージが強い。
- 南側(森ノ宮から南、芦原橋から南)は建物、壁が多いため、北側に次いで、やや閉塞的である。灰色～建物色のイメージ。

7. まとめ

本研究では大阪環状線の車窓景観を対象に、動画でシーケンス景観を撮影し、それを静止画に分割・視点との時空間マッチングを行った。従来では時間的分割を行うことの多い景観研究だが、GPS と速度モデルを利用することで、空間的分割を行うことのできる可能性を提案できた。そして Cityscapes と ADE20K のデータセットを利用した DeepLab v3+ のセマンティック・セグメンテーションにより、景観構成要素の自動抽出・面積比の把握を行うことができることを示した。これは、従来の手法に対し、緑視率や天空率の自動計算への応用が期待できる。また、景観構成要素の面積比に対して階層クラスタ分析を行うことで、景観の連続性や視界・地域変化を把握した。大阪環状線沿線の街のイメージに関わっていると思われる。課題としては、今回の時空間マッチングの GPS は福島—大阪間で採りづらく、このような場合は速度モデルが望んだ形に近づかないため、改善の余地がある。また、クラス数についても精査が必要である。景観内の該当区域と実際の事物との比較を行えるとよい。

参考文献

- 山崎栄紀。「自然要素に着目した大阪環状線の車窓景観の魅力の解明」(URL: http://ulpd.envi.osakafu-u.ac.jp/osakafu-content/uploads/sites/261/2019Greanduation_yamazakieiki.pdf) 2019。(アクセス日: 2019年9月19日)
- 竹内稔, 藤本信義, 三橋伸夫。「シーケンス景観と連続シーン景観の評価構造分析—農山村地域における景観評価に関する研究 その1」, 日本建築学会計画系論文集 第475号, pp.119-128, 1995.
- L. C. Chen, et al. "Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation," In ECCV, 2018.
- M. Cordts, et al. "The Cityscapes Dataset for Semantic Urban Scene Understanding," In Computer Vision and Pattern Recognition (CVPR), 2016.
- Bolei Zhou, Hang Zhao, Xavier Puig, Sanja Fidler, Adela Barriuso and Antonio Torralba. "Scene Parsing through ADE20K Dataset." In Computer Vision and Pattern Recognition (CVPR), 2017.
- クラスタ分析の手法①(概要) | データ分析基礎知識 (URL: https://www.albert2005.co.jp/knowledge/data_mining/cluster/cluster_summary) 閲覧日: 2019年12月9日
- クラスタ分析の手法②(階層クラスタ分析) | データ分析基礎知識 (URL: https://www.albert2005.co.jp/knowledge/data_mining/cluster/hierarchical_clustering) 閲覧日: 2019年12月9日
- バイオインフォマティクスによる遺伝子発現解析 (URL: <https://www.slideshare.net/sesejun/ss-24923282>) 閲覧日: 2019年12月9日
- 志津綾香, 松田眞一「クラスタ分析におけるクラスタ数自動決定法の比較」アカデミア 情報理工学編 11, pp.17-34, 南山大学, 2011年, (URL: http://www.st.nanzan-u.ac.jp/info/nas/academia/vol_011pdf/11-017-034.pdf)
- Box-Cox 変換で変数を正規分布に近づける - About connecting the dots. (URL: <http://smrmkt.hatenablog.jp/entry/2013/06/22/110555>) 閲覧日: 2019年12月10日