

全周パノラマ画像を用いた 離散的視覚情報の定量分析手法

セマンティックセグメンテーションによる空間要素の領域抽出の精度検証

副田和哉¹, 山田猛矢²

¹第一工科大学 工学部 建築デザイン学科 (〒899-4395 鹿児島県霧島市国分中央 1-10-2) k-soeda@daiichi-koudai.ac.jp

²第一工科大学 工学部 情報電子システム工学科 (〒899-4395 鹿児島県霧島市国分中央 1-10-2) t-yamada@daiichi-koudai.ac.jp

Quantification Analysis Method of Changes of Discretely Visual Information Using 360° Panorama Image

Kazuya SOEDA¹, Takeshi YAMADA²

¹Department of Architecture Design, Daiichi Institute of Technology

(1-10-2 Kokubuchuo, Kirishima City, Kagoshima, 〒899-4395) k-soeda@daiichi-koudai.ac.jp

²Department of Department of Information and Electronic Systems Engineering, Daiichi Institute of Technology

(1-10-2 Kokubuchuo, Kirishima City, Kagoshima, 〒899-4395) t-yamada@daiichi-koudai.ac.jp

Abstract: We have proposed an analysis method using 360° Panorama images to treat them quantitatively as visual space information from design theory. However, painting the captured images is an artificial one, and the larger the number of target images, the greater the workload. Therefore, to create a territory extraction model and automatically paint the images, we constructed a U-net model using semantic segmentation, one of the deep learning methods. As a result, we confirmed that the model is adequate for quantitative analysis if the conditions are the same as those of the teacher data.

Key words: VR Photography, 360° Panorama Image, Deep Learning, Semantic Segmentation, Visual Information

1. はじめに

本研究はこれまで定性的に扱われてきた空間の見え方を、意匠論の観点から空間の視覚情報として定量的に扱う研究である。空間体験はすべて独自のだから記述・分析は不可能だとするような考え方をとらず、空間体験には視覚情報のように定量化が可能な側面があるとすれば、そこで記述・分析される動的視覚情報は変化する秩序を持った時間の一部として構造〈ゲシュタルト〉を持ち得ていることがわかる。もし実空間である建築作品での空間体験を定量化することができるならば、第三者と空間の議論をより具体的に展開することも可能になる。そこで一般人にとってもより視覚的に分かりやすく取り組みやすい手法として、安価に購入でき今後も普及させやすい360度の撮影が可能な全天球カメラ^{注1)}で撮影した全天球画像に、円筒図法を援用した画像変換を行い、体験する人の視点から空間の視覚情報を把握し、その離散値としての視覚情報を定量化する定量分析手法を提案することを全体構想とする。筆者らはこれまでの研究^{注2)}で、全周パノラマ画像を用いた離散的視覚情報の変化の定量分析手法(以下 QDP 手法)と題した定量分析手法^{注2)}を提案

し、対象モデルにおける既往の定性的な指摘事項を参照としながら、より具体化された空間的特徴を指摘することで、QDP 手法の有効性を検証してきた。QDP 手法は、正積円筒図法によって描かれた面積情報が正確な全周パノラマ画像を用い、空間の全周を占める各空間要素を塗り分け、その面積比を立体角・立体角量として定量化するものである。

本研究の試みは、客観的に定量化ができ、結果の信頼性も高いと言える。しかしながら、例えば撮影した画像を構成要素ごとに分類する作業は、1枚1枚の画像をCADや画像加工ソフトを用いて塗り分けている。これは人の手に依存した作業であり主観的な側面が拭えない。また対象となる画像の枚数が多いほど作業的負担が大きいのも事実である。移動を伴う視覚情報の変化、即ち空間体験における時間性を内包した空間の見え方の変化である視覚的变化について、それぞれの視点から単独に検討するのではなく、空間体験の連続として捉えることが必要であるというアプローチだが、その定量分析手法やツールが十分に整備されているとは言い難い。

これに対し、本報告では深層学習を用いることで、各空間要素が空間の視覚情報を占める領域を抽出し、自動的に塗り分けることを目指す。近年、コンピュータの演算性能向上や新たな計算式の適応による精度向上によって、深層学習と呼ばれるアルゴリズムが、画像認識の分野においても多く使用されるようになった。特にセマンティック・セグメンテーションは、検出対象とする領域を、教師画像として正しい学習の模範的な画像を作成し、それらを読み込み学習することによって、検出したい領域を自動的に出力することが可能となる²⁾。この手法を用いれば、各空間要素が空間の視覚情報を占める領域を抽出し、効率的に塗り分けを行い、定量化することが可能となる。

一方で、建築における深層学習の活用は、BIMを背景とする施工での精度検証や工程管理、維持管理における保守点検で用いられることが多い。本研究のように既に存在する実空間としての建築空間を、深層学習を用いて、意匠論の観点から記述・分析しようとするアプローチは、未だ少ないのが現状である^{注3)}。

このように一連のプロセスをコンピューテーショナル化することで、空間要素の抽出基準がより明確化され、さらなる簡便化も期待できる。ひとたび空間要素の塗り分けを可能とする領域抽出モデルを作成すれば、同じ建築空間でも条件の異なる場合（天候や季節のような時間軸を踏まえた）の空間体験を効率よく大量に記述・分析できる可能性がある。したがって、この抽出プロセスをコンピューテーショナル化できるなら本QDP手法を、より取り組みやすい手法として、一般化できると考える。

そこで対象モデル（楨文彦設計の岩崎美術館）において複数回にわたる鑑賞ルートを撮影し、その1つを教師データとする深層学習の方法論を述べ、作成した学習モデルを他の画像に適用した結果を示し、現時点での達成点や課題について述べる。

2. 目的

本報告では、QDP手法を簡便化する1つのアプローチとして、特に構成要素に着目し、深層学習を用いたセマンティックセグメンテーションにより、画像の各領域をコンピューテーショナルに塗り分ける領域抽出モデルを作成し、その精度検証を行うことを目的とする。

3. 対象モデル

岩崎美術館は、クライアントである岩崎與八郎のプライベート・コレクション（主に絵画作品）を常設展示する為、楨文彦が設計し、1978年に竣工した私設美術館である。敷地の高低差に合わせて配置された4つの展示室が大きく繋がった一室空間である。鉄筋コンクリート造の地下1階地上1階建であり、延床面積1,114m²建築面積1,347m²の規模である（Photo1）。

岩崎美術館は、壁面に絵画作品を常設展示しており鑑賞ルート順を定義しやすいことや、絵画作品の性質上その領域を抽出し易いと判断し、セマンティック・セグメンテーションによる画像領域抽出の第一歩として先行的なモデルになり得ると考えた。



Photo1 Iwasaki Museum image

4. 研究の方法

4.1. 研究の流れ

対象モデルにおいて全天球カメラ「RICOH THETA V」を用いた動画撮影を行う。セマンティック・セグメンテーション^{注4)}を用いたU-net^{注5)}による領域抽出モデルを作成するため、データセットの作成を行う。撮影した全天球動画をRICOH THETAアプリを用いて正距円筒図法で表示する動画として書き出す。書き出した各動画を動画編集ソフトで動画の時間を調整した後、1秒ごとの静止画として書き出す。書き出した画像を構成要素ごとに手動で塗り分け、教師データを作成する。塗り分けた教師画像と基にしたオリジナル画像を用いてU-netで構成要素を塗り分ける領域抽出モデルを構築する。その後、構築した学習済みモデルを用いて、他の動画から書き出した画像データを塗り分け、この画像が正しく塗り分けられているかを検証する。塗り分けた画像と教師データの各割合の平均二乗誤差を検証し、さらにフーリエ変換した周波数から建築空間の動的視覚情報が内包する特徴量を得られるか検証する。

4.2. 調査概要

まず対象モデルとする岩崎美術館の本館で全天球カメラを用いた全天球動画の撮影を行う^{注6)}。その際、人が立った時の視線の高さ(1600mm)で全天球動画の撮影を行い、カメラ本体の向きは観察者の進行方向に統一した。また今回の調査では、カメラ設定をオートにして撮影を行った^{注7)}。岩崎美術館の本館において、玄関ホールから左回りに展示物の前を通過し、戻ってくる鑑賞ルート（Fig.1）で撮影を行い（各動画の撮影時間は2分30秒～3分30秒程度）、計4本の動画を得た。この動画を動画加工ソフトを用いて3分間の動画に調整した後、1秒間隔で静止画への書き出しを行う^{注8)}。すなわち、1本の動画から180枚の静止画を得ることができ、計720枚の画像データを得た。そのうちの180枚については、教師データとして使用する。

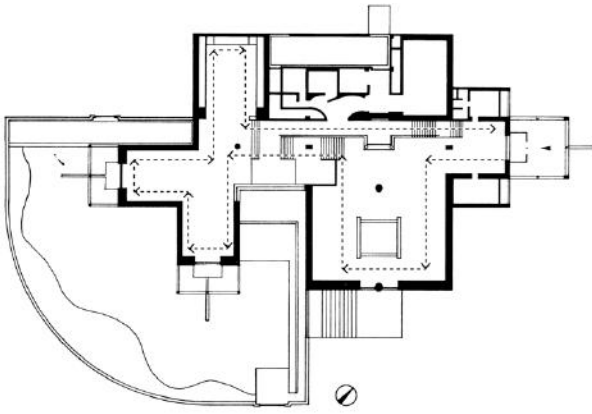


Fig.1 Iwasak Museum plan

5. 領域抽出モデルの作成

5.1. 教師データの作成

得られた画像を手動で塗り分けを行う。今回は、空間体験における視覚情報の意味ごとの分類である構成要素に注目をする。〈床・開口部〉や〈壁・天井〉、〈展示物〉に、ノイズ（撮影者）を加えた4つの構成要素ごとに塗り分け^{注9)}、教師データを作成する。本研究で扱うU-netモデルは、教師データとして採用する際、インデックスカラー画像のPNGカラーパレットのインデックス値をクラスとして扱うため、要素を特定の色のみに塗り分ける必要がある。教師データは、ナンバリングされたインデックスカラー画像と呼ばれる画像データである。こうして塗り分けた180枚の画像データを教師データとして採用する (Fig.2)。この教師データを〈teach : auto root am1〉とする。

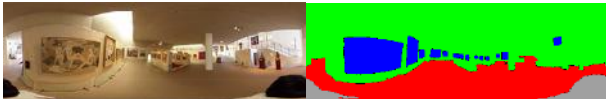


Fig.2 Example of Teacher Data

5.2. 学習モデルの作成と精度検証

セマンティック・セグメンテーションの手法の1つとして、本研究ではU-netによる領域抽出モデルを作成する。

学習する際の画像のサイズは、その性質上すべて同じサイズである必要がある。本研究での動作環境でプログラムを動かす場合、著しく動作が重くなってしまうため、384×192のサイズにリサイズして学習させることにした。尚、本研究はQDP手法の更なる簡便化、一般化を目的としている為、市販で手に入る一般的なスペックのコンピューターを採用している^{注10)}。

教師データの基にした画像データ180枚を訓練用とテスト用データで9:1に分け、構築したU-netで学習を行い、モデル作成を行った。評価指標のAccuracy (正解率) にはPixel Wise Accuracy (式 (1))、Loss (損失関数) には交差エントロピー誤差であるCross Entropy Error (式 (2)) を用いた (Fig.3)。学習済み

モデルにテスト用データを用いて精度検証を行った結果、96.8の正答率が得られた。Lossについても訓練用データとテスト用データに開きはあるが、どちらも0.4未満であり高い精度が確認できた (Fig.4)。

$$(\text{Pixel-wise accuracy}) = \frac{1}{N} \sum_{i=1}^k n_{ii} \quad (1)$$

N : 全画素数, k : クラス数, n_{ii} : 各クラスの正解画素数

$$E = - \sum_{i=1}^N t_i \log y_i \quad (2)$$

N : 全画素数,

t_i : 正解ラベル (正解クラスだけが1で、他は0 (one-hot 表現)),

y_i : 学習モデルからの出力 (確率)

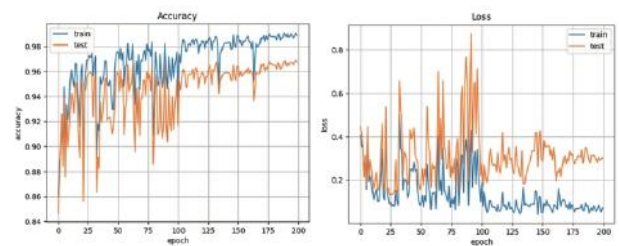


Fig.3 Verification of The Accuracy and Loss of The Learning Model

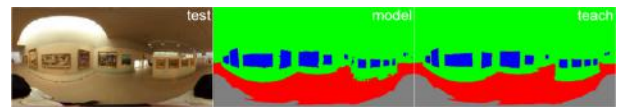


Fig.4 Example of a Learning Model

6. 領域抽出モデルの適用および精度検証

6.1. 検証方法について

ここからは前節で構築した学習済みモデルを用いて、他の動画から書き出した画像の塗り分けを行う。正積円筒図法へ変換し、各要素の面積比 (立体角量) を抽出し、精度検証を行う。抽出した数値について、教師データ及び学習済みモデルが塗り分けた各面積比の平均二乗誤差を求め、正しく塗り分けられているか検証する。さらに、これらの時間変動をフーリエ変換し周波数の構造と、その相関係数を比較することで建築空間としての動的視覚情報を内包する特徴量が得られるかを検証する。今回は教師データと同じ条件である午前撮影した画像データ2つをそれぞれ〈video 1 : auto root am2〉〈video 2 : auto root am3〉とし、条件の異なる午後撮影した画像データの1つを〈video 3 : auto root pml〉として比較する。教師データは、先述した〈teach : auto root am1〉である。

6.2. 検証方法について

平均二乗誤差 (式 (3)) を求めることで学習モデルを用いて塗り分けた画像データと教師データとの要素の面積比の誤差を検証する。

$$E = \frac{1}{N} \sum_{i=1}^N (y_i - t_i)^2 \quad (3)$$

E : 平均二乗誤差 (MSE : Mean Square Error)

t_i : 学習モデルを用いて塗り分けた地点 (No. i) での構成要素の割合
y_i : 教師データにおける地点 (No. i) での構成要素の割合

〈床・開口部〉や〈壁・天井〉、〈展示物〉の3つの要素の面積比 (立体角量) を比較することで検証を行った。〈video 1〉と〈video 2〉と〈teach〉について比較すると、教師データと同じ条件である場合は、近似する値を多く確認することができる (Fig.5)。特に〈video 2〉の面積比は、教師データと特に近似していることが確認できる。〈video 1〉についても1枚1枚の面積比を比較すると教師データのものと差があるように見えるが、同じような波形が少しずれていることが確認できる。これは動画を撮影した際の歩くスピードの違いによる差が生じているためだと考えられる。

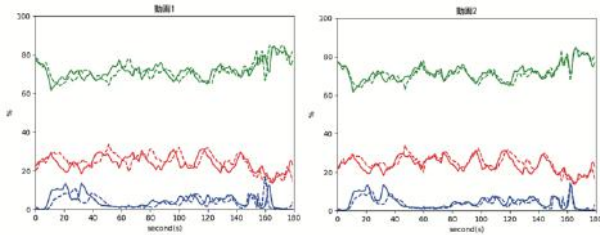


Fig.5 Time variability of space components (Video1 and Video2)

この2つの画像データの各要素の割合の平均二乗誤差は、それぞれ3前後であり、教師データとの塗り分けに大きな隔たりがないことが確認できた (Table. 1)。

Table.1 Verification of Mean Square Error

平均二乗誤差	video1	video2	video3
床・開口部	3.697	2.575	9.726
壁・天井	3.494	2.155	11.308
展示物	3.246	2.170	4.139

一方で、教師データと異なる条件である〈video 3 : auto root pml〉は、教師データから逸脱している値が多いことが確認できた (Fig.6)。各要素の割合の平均二乗誤差は〈video 3〉では平均して約8の誤差が出ていることが確認できた (Table.1)。

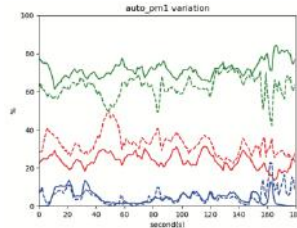


Fig.6 Time variability of space components (Video3)

6-3. フーリエ変換による精度検証

次に時間変動をフーリエ変換し (式 (4))、その周波数の構造を比較することで近似性を見る。その相関係数を用いることで (式 (5))、建築空間の動的視覚情報を内包する特徴量が得られるかどうか検証する。

$$F_k = \sum_{t=0}^{n-1} f_t e^{i \frac{2\pi k t}{n}} \quad (4)$$

F_k : 離散フーリエ変換

f_t : 空間構成要素の時間変動 (サンプリング間隔1秒)

それぞれのフーリエ変換後のグラフを教師データのものと比較すると、〈video 1〉と〈video 2〉については

$$\rho_{FF} = \frac{1}{\sigma_F \sigma_{\mathcal{F}}} \frac{1}{n} \sum_{k=1}^n (\mathcal{F}_k - m_{\mathcal{F}})(F_k - m_F) \quad (5)$$

ρ_{FF} : 相関係数

σ_F : 教師データの空間構成要素割合のフーリエ変換における標準偏差

σ_ℱ : 学習済みモデルから出力された空間構成要素のフーリエ変換における標準偏差

ℱ_k : 教師データの空間構成要素のフーリエ変換

F_k : 学習済みモデルから出力された空間構成要素のフーリエ変換

m_ℱ : 教師データの空間構成要素割合のフーリエ変換における平均値

m_F : 学習済みモデルから出力された空間構成要素のフーリエ変換における平均値

n : サンプル数

どの要素も教師データのものと一致しており、〈video 3〉では視対象以外の要素は教師データのグラフと異なる部分があることが確認できた (Fig.7)。

さらに相関係数について注目すると、〈video 1〉と〈video 2〉はどちらも0.85以上あり、強い相関を得ることができている。〈video 2〉は全要素が0.92以上と非常に強い相関を得ることができた。〈video 3〉は0.7を下回っている要素も確認できた (Table.2)。

Table.2 Verification of Correlation Coefficient

相関係数	video1	video2	video3
床・開口部	0.859	0.945	0.699
壁・天井	0.898	0.928	0.523
展示物	0.916	0.956	0.882

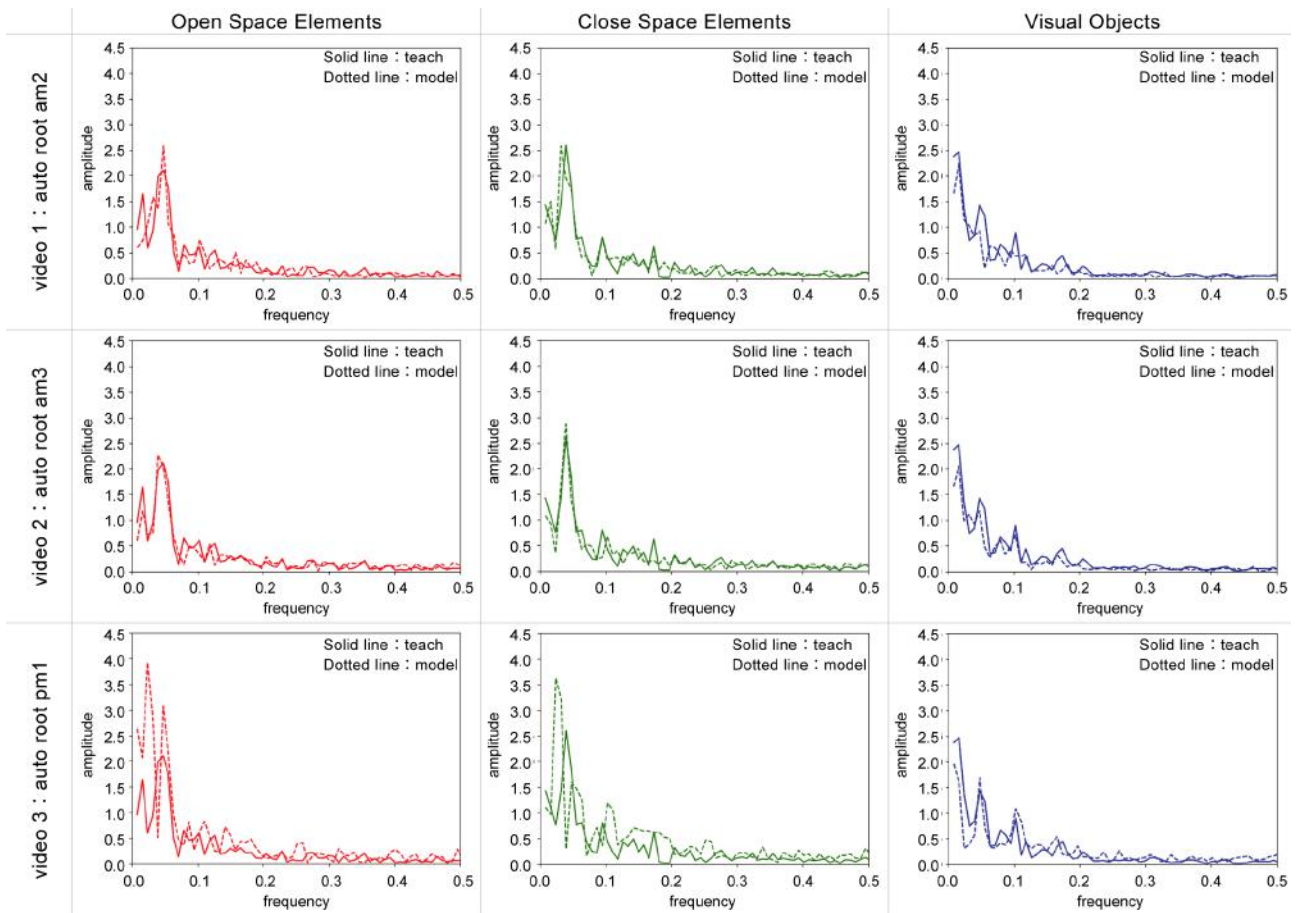


Fig.7 Graph of Fourier transform for each video and space element.

6.4. フーリエ変換による精度検証

検証実験を行った結果、学習モデルと同じ条件である午前中に撮影した動画であるならば、平均二乗誤差は3前後と低く正しく各要素を塗り分けることができた。さらにフーリエ変換後のグラフの相関係数も0.85以上あり強い相関を得ることができ、建築空間の動的視覚情報を内包する特徴量が得られることが確認できた。

以上のように、U-net を用いたセマンティック・セグメンテーションによる画像領域抽出は教師データと条件が同じであるならば、定量分析に有効であることを確認できた。

7. 本モデルの課題

今回、正しく塗り分けることが出来なかった画像データの1つ1つに注目すると塗り分けがうまくいっていないように見える。ただ大まかに比べてみると、全く塗り分けられていないわけではなく、数値的にも視対象は塗り分け可能であったことなど、ある程度は要素を識別できていることが確認できる (Fig.8)。

このことから、オリジナルデータとしての全周パノラマ画像や学習時にリサイズした画像の画素数や解像度といった情報量の影響によって詳細な塗り分けができ

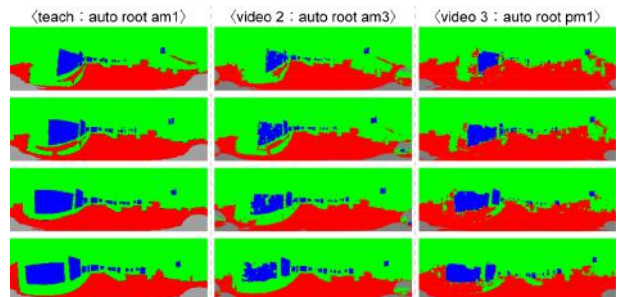


Fig.8 Example of paint separation

なかったことが考えられる。本研究では現在、一般的に普及しているコンピューターのスペックでも機能するような安価で簡易な取り組みやすい手法として一般化することを目指している。そのため今回の条件では、画像データそのものの情報量の問題と計算処理を行うコンピューターの問題が精度に影響している。またU-net モデルの特徴は、画像サイズの大きさの制限がなく、階層を深くすることで、今後さらにサイズの大きな画像でも学習させることが可能であることがあるが、今回は画素数や解像度といった画像サイズそのものの情報量を減らしたことにより、そのモデルとしての特徴が上手く活かされなかったことも考えられる。

しかし上記の不十分さが本質的かつ空間的問題になるというのではない。本研究で採用した全天球カメラについても、現在開発改良の途上にあるものであり、今後オリジナルデータとしての全周パノラマ画像のダイナミックレンジや解像度のような精度の向上も期待できる。また、計算処理を行うコンピューターや深層学習のシステムについても日々その処理能力は向上しており、更なる一般化（簡便化や経済的負担の低減）も期待できる。こうしたことを考慮すると、今回の環境で生じた問題も自ずと解決されることが期待できる。また深層学習を用いた領域抽出も、今回採用したセマンティック・セグメンテーションの他にも異なるアプローチは存在しており、こうした別のアプローチを試みることで、その精度向上や適用範囲の拡大は十分に期待できる。このように学習方法によって異なる部分はあるものの、考え方やプロセスを他のアプローチに広範に適用することは可能である。それでもって抽出モデルとしての有効性は、各々の条件に即したものになる。

8. おわりに

横文彦設計の岩崎美術館を対象モデルに、画像データの収集・教師データの作成を行いU-netを用いた構成要素を塗り分ける領域抽出モデルを作成した。次にそのモデルを適用して正しく画像が塗り分けられているかを検証すべく、塗り分けた画像データと教師データの各要素の平均二乗誤差を確認し、さらにフーリエ変換した周波数から建築空間としての動的視覚情報を内包する特徴量が得られるかを確認し、定量分析手法としての有効性を検証した。

検証を行った結果、教師データと同じ条件であるならば、平均二乗誤差は3前後と正しく各要素を塗り分けることができた。さらにフーリエ変換後のグラフの相関係数も0.85以上あり強い相関を得ることができ、建築空間の動的視覚情報を内包する特徴量が得られることが確認できた。

以上のように、深層学習を用いることで、各空間要素が空間の視覚情報を占める領域を抽出し、自動的（コンピューテーショナル）に塗り分けることができるならば、これまでに取り組んできた既往研究¹⁾³⁾において塗り分けてきた画像データを教師データとして学習させることも可能となり、より対象を広げることできる。こうしたコンピューテーショナルな手法を確立することは、人為的な負担を減らすだけでなく、塗り分けの際の主観性を減らすことにもなり、客観的な定量分析手法としての精度の向上にも繋がる。本研究の成果は、これまで空間的にも時間的にもトリミングされた建築写真から定性的に語られるものであった建築空間を、より具体的に変化する環境として定量的に議論することを可能にするものである。今後、高度情報化していく建築において、空間を単なる情報の集積と捉

えるだけでなく、それら空間の情報が持つ意味について議論する際の足掛かりとなる建築の新たな意味を考える重要な視座になる。

謝辞

本研究を進めるにあたって、当時第一工業大学学生であった榎本雅貴・小川朝陽の両氏に、データの収集・モデルの作成について多くの協力をいただきました。また、現地での調査にご協力して頂いた岩崎美術館の方々に記して感謝申し上げます。

参考文献

- 1) 副田和哉, 平瀬有人, 三島伸雄: 全周パノラマ画像を用いた離散的視覚情報の変化の定量分析手法 パラッツォ型鑑賞空間を対象とするモデル的研究, 日本建築学会計画系論文集, 第86巻, 第780号, pp. 425-435, 2021. 2.
- 2) 安野貴人, 姜叡, 西名大作, 金子京平: 都市河川景観評価におけるDeep Learningの応用についての検討その1—景観構成要素の領域抽出手法の提案—, 日本建築学会大会学術講演梗概集, pp. 21-24, 2019. 7
- 3) 大橋征幹: セマンティック・セグメンテーションによる画像認識を用いた緑視率調査法に関する検討, 日本建築学会大会学術講演梗概集, pp. 739-740, 2019. 9
- 4) Olaf Ronneberger, Philipp Fischer, Thomas Brox: U-Net: Convolutional Networks for Biomedical Image Segmentation, pp1-8, 2015
- 5) 宮岸幸正, 材野博司: シークエンス景観における景観行動と空間の開放度・インパクト度との関係, 日本建築学会計画系論文報告集, 第440号, pp. 119-125, 1992. 10

注

注1) 撮影の際の操作性はもちろん（全天球カメラはピント調整なども必要なく写真撮影における特別なスキルは必要とせず）、これまで行われてきた複数のカメラ画像を重畳して天球画像を生成するといった専門性を必要とする画像加工（変換）も必要とせず、常に同じ変換プロセスで全天球画像から全周パノラマ画像（正距円筒図法）を生成することができる。この簡易な全天球カメラを用いることで、これまでより簡易に空間記述・分析を行うことが可能になると考えた。

注2) 離散的視覚情報とは各視点における個々の独立した視覚情報とする。動画を用いた分析であっても静止画として書き出すことによって、動画本来の時間的な連続性は失われてしまい、1枚1枚は独立した静止画的「視覚情報」になると考える。本研究ではこれを「離散的視覚情報」と呼び、空間体験を離散的視覚情報として抽出することで視点や移動の自由さが担保された空間の空間体験を分析するものである。

注3) 日本建築学会の計画系論文集等においても所謂、査読論文は管見する限り見受けられず、大会梗概集などにおいて、いくつか存在する限りである。こうした発表等も河川景観や緑景観に注目したものであり、本研究のような建築作品を対象としたものは管見する限り見受けられなかった。

注4) セマンティックセグメンテーションは、画像に写っている物体をピクセル単位でクラス分類して領域を抽出するものであり、矩形で大きさと位置を捉える物体検出と比較して、物体の輪郭を捉えることが出来る特長がある。

注5) U-net とは画像セグメンテーション(物体がどこにあるか)を推定するためのネットワークである。生物医科学の画像セグメンテーションを行うために2015年に発表された⁴⁾。左右対称でアルファベットの「U」に似た構造をしていることから「U-net」と呼ばれている。U-netの特徴は少ない教師データでも学習の効果が期待できるところであり、時には50以下のデータでも学習の効果が出る事が証明されている。さらに、他のセグメンテーションモデルなどと異なり全結合層と呼ばれる画像データを1つにまとめる処理がないため、画像サイズの大きさの制限がなく、今後さらにサイズの大きな画像でも学習させることが可能である。以上のことから本研究ではU-netモデルを用いたセマンティック・セグメンテーションによる学習を行った。

注6) 今回の撮影条件は〈日時：2020年11月21日(土)11:00-14:00, 場所：鹿児島県・指宿市, 天候：曇り〉であり、カメラ設定は、RICOH THETA Vを用いて以下の設定〈ピクセル数：幅5376・高さ2688, 絞り：f/2, 色空間：sRGB, ホワイトバランス：昼白色蛍光灯〉で撮影を行った。

注7) RICOH THETA V及びそのアプリケーションを用いて、天球状の画像を天頂角補正の上、正距円筒図法による全周パノラマ画像として、書き出すことが可能となる。また全天球カメラの選定について、いくつかある全天球カメラであるが各製品とも現在開発改良の途にあるものである。その為、突出して優れた機器はなくRICOH THETA Sについても、カメラそのもののスペックやカメラ本体の大きさなどを他の機器と比べても遜色はなく、本研究においてはなるべく容易に扱える操作性と身近に手に入る入手性、そして経済性を考慮し、今回の撮影に採用した。他にもハイスpekな装置などもあるが、その場合は高い費用を必要とし機器そのものも大きくなるため、現地での連続的な撮影には不向きであると考えた。

注8) 鑑賞空間における歩行速度が幾分遅い速度1,000mm/secであることや彫刻ギャラリー内の展示物が細かく配置されていることなどを考慮し1秒間隔を採用した。この間隔については今後の課題とした。

注9) 既往研究⁵⁾において“空間が開かれている度合いを開放度、目に強い刺激を与える度合いをインパクト度、それらを除く部分を閉鎖度”として抽出し、それらに関する構成要素の分類による開放性・閉鎖性、インパクト度及びその変化率という定量的指標を用いることで構成要素との関係が説明できると指摘されていることを踏まえたものである。

注10) 動作環境|OS：LINUX ubuntu 14.06, CPU：intel Xeon, メモリ：32GB, GPU：NVIDIA geforce GTX 1080 × 2 またコンピュータの処理能力を補う方法として計算環境のサーバーをレンタルすることも考えられたが、今回は比較的安価に、安定して使用できる手持ちのコンピュータ(PC)の範疇で、検証を行なった。