

深度情報を用いた空間認識による行動分析システムの開発

横山寛和^{*1}, 松原正也^{*2}

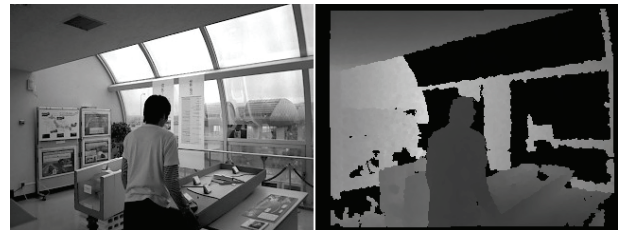
博物館展示スペースに距離カメラを設置し、カメラからの映像を画像処理することで観覧者の行動動線を自動分析するシステムを開発した。長時間連続で観覧者の行動動線を定量的に記録することで、行動分析を自動的に容易にしたことに加え、科学的根拠に基づいた展示評価を期待できる。本システムを用いて展示スペースにおける観覧者の行動分析を行った。その結果、展示物沿いに集まる行動動線や密度分布、立ち止まり時間比率を得ることができた。

〈キーワード〉画像処理, 人物追跡, システム開発, 距離画像, 行動分析, 博物館評価

1はじめに

博物館では、展示スペースの観覧者の動きに着目した行動分析が行われている。観覧者の観覧時間や行動動線などを記録・分析することで、展示スペースに対する観覧者のフィードバックを得ることは、展示手法や展示構成、配置などを検討する上での根拠となる。観覧者の行動分析により、観覧者の興味や見学方法、展示のメッセージが伝わったか否かを評価できる。展示の評価から現状を知ることは、課題を見つけ出し、新しい展示案などを導き出すことにつながるからである。たとえば野村(1993)は観覧者の行動動線を目視で記録することで、観覧者の行動パターンを導いた³⁾。また、朴ら(2005)は行動動線からハンズ・オン展示の有効性を検証した⁴⁾。原、松久ら(1982, 1984)は行動動線を記録し、観覧者の行動予測のための法則かを試みた^{1,2)}。江水(2011)は、展示の配置によって観覧者の行動に違いが見られたことを報告しており、行動分析による評価は、博物館を充実させる上で有用であった⁹⁾。

こうした観覧者の行動分析は、博物館のスタッフが、観覧者を個別に目視追跡することで記録している。これは、運営及びスタッフにとっては負担が大きいことや、一部の観覧者のみしか対象にできないことに課題があると言える。佐々木(2012)は、博物館活動についての調査で現場から「より簡便で、効率的な評価活動が



必要である」「作業量が多く、仕組みが複雑すぎる」

図1 カメラによる展示室の画像(左:光学画像, 右:距離画像)

と声があり、評価活動そのものが負担になっていると指摘している¹¹⁾。これからの博物館の在り方に関する検討協力者会議(2007)では、博物館の評価のためだけに過度に博物館の労力が割かれることは適当ではないと指摘されているが、現状の行動分析はスタッフへの負担が大きい⁷⁾。

本研究では深度情報を用いた空間認識による自動行動分析システムを開発した。これにより導入コストが低く、スタッフと観覧者の双方に負担を掛けずに長時間連続での行動動線の記録を実現した。システムは観覧者の座標とその時間で行動動線を記録する。そのため定量的なデータに基づいた行動分析につなげることが可能となる。長時間連続で記録された定量的な大量のデータは、大数の法則から信頼性の向上にもつながる。目視による行動分析と比較して、科学的根拠に基づいた展示スペースの検討につながることを期待できる。

*1 ヤフー株式会社

*2 岐阜大学総合情報メディアセンター

2 システムの概要

展示スペースに距離カメラを設置し、定点から撮影する。距離カメラは赤外線を用いて測距可能なカメラであり、画像の各ピクセルには色情報ではなく深度情報が記憶される(図1)。距離カメラとして、Microsoft社のKinect及びASUS社のXtion Pro Liveを用いる。これらのカメラは屋内を約80cmから9mまでの範囲を測距できる。

距離画像を画像処理し、観覧者の抽出と追跡をリアルタイムに行う。距離画像とは、深度情報を画像の濃淡に置き換えて表示した画像である。距離画像のみを使用することで、正確で演算負荷の少ない自動行動システムを実現する。光学画像は、照明のオン・オフ、日光の時間変化などの外乱影響を受けやすい。また、イルミネーションなど動的な光を使った展示である場合や展示スペース全体の照明が暗いと言った場合は、画像処理が難しい。しかし、赤外線によって撮影される距離画像はこうした影響を受けにくいからである。また、画像自体が画像の平面XYに加え、濃淡が奥行き、すなわちZ方向も合わせた3次元座標を持っている特徴がある。

まず、距離カメラから距離画像を10fpsで入力する。深度情報は16bitの値で入力されるが、画像処理の都合上8bitに丸めて距離画像を作成する。これにより最大36mmの誤差を予想できるが、微小であり無視できる。次に、前景抽出・背景除去を行う。距離画像から背景領域を削除し、前景領域(人物領域)のみを抽出する。本研究ではKim(2005)らのコードブック法を適用することで高速で正確な処理を実現した。

そして、前景距離画像にラベリングを行い、個々の人物をそれぞれ抽出する。ここで、距離画像は平面座標XYに加え、Z方向すなわち奥行き方向も併せた3次元座標を持っていることに着目した距離階層化ラベリングを提案・実現した。Z方向に約50cm毎にレイヤー化した座標空間を用いてラベリングする。これにより人物がXY座標で重なってしまい、連続になった画像においても、それぞれ人物を個別に抽出することができる。また、このとき検出した物体の画像上の大きさが300px以下の物体をノイズとして扱う補正も行う。

最後に、フレーム間での同一人物をマッチングし、動線をつなぐ。認識したが人物しなかった人物は新規に入

室した人物と見なし、見失った人物は2秒待ってから退室として扱う。新たに入室した人物には固有の識別番号を割り当てる。退室後、滞在時間2秒以上の人物の行動動線を記録する。滞在時間が極端に短い人物はノイズであると判断する。

本システムは、これらの一連の処理をリアルタイムに行う。結果を外部からモニタリングし、即時的なフィードバックなどにつなげられる。

3 距離画像の前景抽出

前景抽出の手法

距離画像から人物認識を行うための前処理として前景抽出を行う。入力画像から人物領域のみの前景画像を生成する。

前景抽出は一般的に、開展前など人物がいない状態、すなわち背景を学習しておき、観測には学習した背景をもとに入力画像から前景と背景の領域を分離することで行う。

森田ら(2005)は、背景の動的更新が可能な光学画像を用いた前景抽出を遠隔管理システムに実装した⁵⁾。Kimら(2005)は画像の各ピクセルに背景モデルを記録したコードブックを作成することで背景学習を行い、観測時には各画素の値について該当するコードブックを参照することで前景抽出を行った⁶⁾。本論文では、この手法をコードブックと呼ぶことにする。Enriqueら(2013)はコードブック法を距離画像に拡張し、前景抽出を行うDECBC法を提案した¹²⁾。Garyら(2009)はRGBカラーではなく、YUVカラーを用いたコードブックで処理の高速化を図った⁸⁾。

本研究での前景抽出法

本研究ではGaryらのコードブック法を参考に、深度情報のみによる前景抽出を行う。色情報は使用せず、深度情報のみに限ることで外乱の影響を抑え精度を高めると同時に、高速な処理を実現した。DECBC法のように色情報と深度情報を合わせた処理は、演算量が増大する問題がある。また、色情報に含まれる外乱の影響を受けてしまうことが予想できる。

まず、人物のいない展示スペースを(数十～数百フレーム撮影し、背景学習を行う。このときピクセル毎に、

表1 研究室での各前景抽出法の精度 F と平均処理時間 T [msec]

抽出方法	F										T msec
	1	2	3	4	5	6	7	8	9	10	
CStat	0.49	0.44	0.47	0.73	0.62	0.56	0.54	0.39	0.50	0.49	2.3
DStat	0.84	0.73	0.76	0.91	0.90	0.90	0.84	0.75	0.60	0.83	1.5
Kim	0.43	0.38	0.50	0.70	0.66	0.67	0.60	0.28	0.50	0.38	9.2
YUV	0.36	0.65	0.61	0.82	0.76	0.52	0.67	0.50	0.71	0.36	2.9
YUVD	0.47	0.72	0.67	0.87	0.84	0.61	0.76	0.63	0.76	0.47	567.7
DECB	0.67	0.44	0.62	0.87	0.82	0.82	0.66	0.36	0.58	0.64	10.1
本研究	0.91	0.91	0.85	0.93	0.95	0.94	0.93	0.89	0.88	0.91	1.1

背景モデルを記録するコードブックを作成する。背景学習時、毎回の深度情報からそのピクセルが背景に取り得る範囲を示すコードワードをコードブックに追加する。もしくは既に該当するコードワードが存在していた場合は、そのコードワードを更新する。ピクセル毎に背景として取り得る領域をコードブックに学習する。観測時は、入力される深度情報からそのピクセルのコードブックのうち、いずれかのコードワードに属していたら背景と判断し、そうでなければ前景すなわち人物領域として判断する。ただし、深度情報の入力値 0 は距離ではなく測距エラーであることを示す値であるため学習及び判定に反映させない。

評価

無作為に抽出した 10 枚の画像の処理について、評価

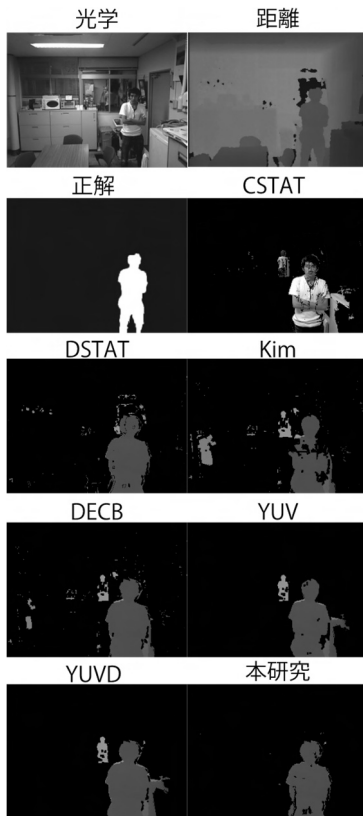


図2 研究室における前景抽出結果

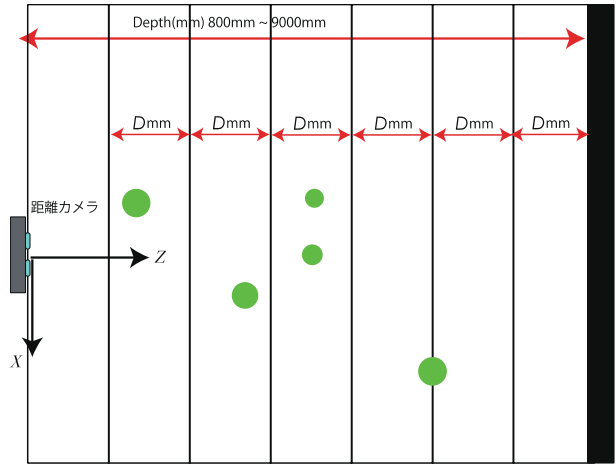


図3 Z軸(深度方向)での階層分割

指標 F 値と平均処理時間 T を算出した (表 1)。 F が大きいほど正確である。図 2 は生成された前景画像を示す。ただし、CStat, DStat は森田ら (2005) の手法に光学画像および距離画像を入力した処理、YUV は Gary (2009) の手法で、YUVD はそれに深度情報も合わせた処理を示す。

本研究の手法は、他の手法と比較して高い F 値を示した。また、平均処理時間も他のコードブック法 (Kim, YUV, DECB) と比べて最短時間であった。深度情報のみを使用したコードブック法により、より正確で高速な前景抽出を実現した。

4 距離階層化ラベリング

前景距離画像にラベリングを行い、個々の人物をそれぞれ抽出する。本研究では距離階層化ラベリングを提案・開発した。

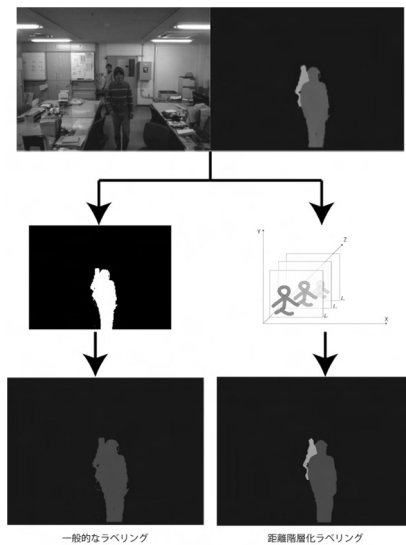


図4 通常のラベリングと距離階層化ラベリング

レイヤー化

距離階層化ラベリングは、距離画像が平面 XY に加え、濃淡が奥行き、すなわち Z 方向もあわせた 3 次元座標を持つことに着目し、濃淡により画像を分割することで Z 方向にレイヤー化する(図 3)。レイヤー毎に二値化し、ラベリングを行う。複数の人物が平面 XY で連続、すなわち人物が重なった画像においても Z 方向にレイヤー化した座標空間により、それぞれの人物を個別に抽出できる(図 4)。パラメータ調整の結果、Z 方向を $D=50\text{cm}$ でレイヤー化する。

隣接物体の結合と面積フィルタ

ラベリング後、補正処理として隣接物体の結合と面積フィルタを行う。隣接物体の結合は、画像内のひとつの物体が複数のレイヤーにまたがって分割される場合を考慮して、ラベル行列について隣接している異なる物体同士を検出し、物体間の Z 方向の距離が約 10cm 以下であれば結合する。レイヤー化は各ピクセルの濃淡値のみで分割したため、ひとつの物体が複数のレイヤーにまたがってしまうことが予想され、これを補整する。隣接物体の検出はラベル行列に対してラブラシアンフィルタを施して行う。面積フィルタは、検出した物体で、そのピクセル数が 400px 以下と小さい面積の物体はノイズとして除去する。画像上で、人物は面積が大きく、ノイズは小さいと予想できるからである。これにより誤検出を減らす。

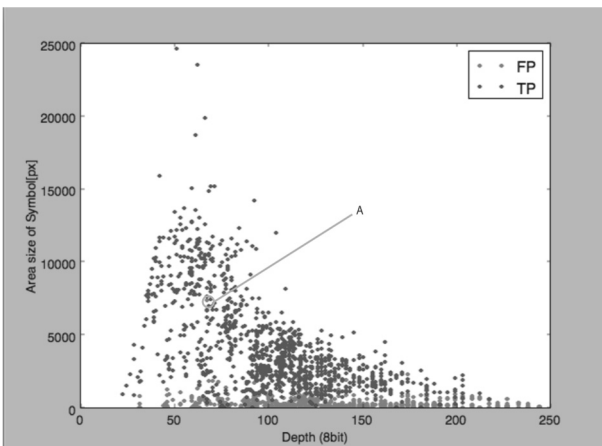


図 5 検出した物体(TP, FP)の距離と面積

また、ノイズの面積は検出した距離に依存せず、ほぼ一定であった。図 5 は、横軸に距離、縦軸に面積をとり、検出した人物(TP, 青点)とノイズ(FP, 赤点)を表示してい

る。ノイズはいずれの距離においても一定以下になっていると言え、しきい値は一意に決定することが望ましいと言える。ただし、A で示した面積の大きい FP は 1 人の人物を 2 分割したものである。これはサンプル 2006 個中の 1 個と微少であり、無視できると判断する。結合距離しきい値 10cm および面積フィルタしきい値 400px は FROC 分析により決定した値である。

5 フレーム間での同一人物追跡

抽出した人物をフレーム間において時間軸方向での結合を行う。

方法

フレーム間でのマッチングには人物 P, Q 間の距離 d を式 1 で導き、最短距離の人物でかつ $d < 0.2$ であれば同一人物と見なす。ただし、 P_x, P_y, Q_x, Q_y は P, Q の画像上での位置を示し、 P_z, Q_z は実座標系での距離を示す。それぞれ取り得る最大値で正規化している。

$$d = \sqrt{\left(\frac{P_x - Q_x}{320}\right)^2 + \left(\frac{P_y - Q_y}{240}\right)^2 + \left(\frac{P_z - Q_z}{9000}\right)^2} \quad (1)$$

なお、観測時間が 2 秒以下であった人物はノイズとして扱い記録しない。また、一時的に見失う場面を考慮して、見失った人物は不在度をカウントアップし、2 秒以上待ってから退室として扱う。また、見失った人物の中でも最後に確認した座標が他の人物と近接している場合は、人物どうしのすれ違いや重なり合いであるとして、位置を保持する。退出後、その人物の各々の実座標系での座標点とその時間 (X, Y, Z, t) を記録する。

評価

展示スペースの観覧者 A~D の 4 人を追跡した(図 6)。人物 A は比較的長時間、展示スペースに滞在し、画面内

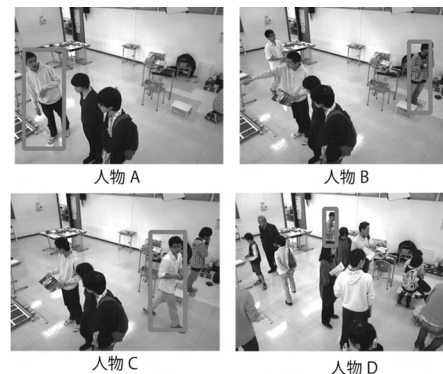


図 6 追跡対象とした人物

を広く動いた。人物Bは画面の隅を小さく移動した。人物Cは早歩きで展示スペースを横断した。人物Dはカメラから遠方の位置を移動した。

表2 人物追跡の評価結果

	正解数/全フレーム数	正解率 (%)
人物A	1339/1353	98.97
人物B	133/133	100.00
人物C	/6464	100.00
人物D	176/188	93.62

フレーム*i*の追跡対象人物*P*と次フレーム*i+1*の全ての人物との物体間距離を算出し、フレーム間での同一人物*P*の物体間距離が他の人物と比べて最短であれば、正しく物体追跡できたとと言える。*N*枚のテストフレームで試行し、その正解率を導いた(表2)。

いずれの人物においても高い正解率を記録した。人物Cの結果から早歩きで移動する観覧者でも追跡できたとと言える。一方人物Dは、カメラから離れているため人物同士の重なり合いの影響によりやや低い正解率になったと考えられる。

6 システムの検証

岐阜県・長良川河口堰管理所アクアプラザながらにて2014年4月1日から6月22日まで開催された巡回企画展「魚道展」(撮影日4月30日)、2013年11月2日から二日間開催された岐阜大学祭におけるロボコンサークルの展示「ロボ展」(撮影日11月3日)、2014年11月1日から二日間開催された岐阜大学祭における教育学部美術教育講座による「美術教育講座在学生作品展」(撮影日11月2日、以下「美術展」)においてXtion Pro Liveで定点から撮影し、本システムの検証を行った。なお開発・検証に用いたコンピュータはMacBook Air(Intel Core i7, 1.7GHz, 8GB RAM)と一般的なノートPCである。

行動動線の分析

本システムが記録した観覧者の時間毎の座標値から行動動線を再現した(図7)。各観覧者の個別もしくは複数の行動動線を視覚的に確認することができる。X座標とZ座標の値から展示スペースの床面座標系で表示しており、横軸がXで縦軸がZである。原点(X,Z)=(0,0)がカメラの位置する地点である。

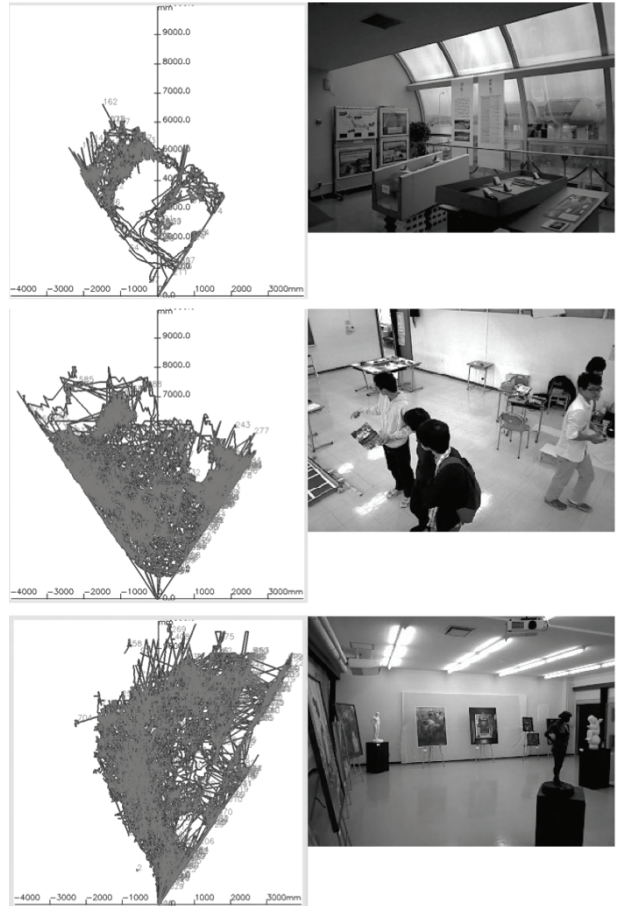


図7 約30分間の全観覧者の行動動線(上段:魚道展, 中段:ロボ展, 下段:美術展)とカメラの光学画像

観覧者の行動動線の表示は、行動分析において記録される基本的なデータであり、これまでの行動分析においても目視で記録されてきた。しかしながら、目視記録は、スタッフに負担が掛かることや、一部の観覧者のみしか対象にできないことに課題があった。本システムは、目視記録によるデータと等価な行動動線の記録を自動化し、長時間連続での記録を可能にした。スタッフの負担を軽減に加え、長時間連続して記録した行動動線から、時間を区切って行動動線を表示・分析することで多くの観覧者のデータを得ることが期待できる。

観覧者密度分布マップ

記録した各観覧者の座標値から展示スペースの観覧者密度マップを作成した(図8)。図から展示スペースで観覧者が密集しているところ、すなわち多くの人が立ち止まったり、通行したところ、もしくは疎なところを捉えることができる。展示スペースXZ平面を25cm×25cmのセルに区切り、セル毎の観覧者数の分布を色で表した。赤

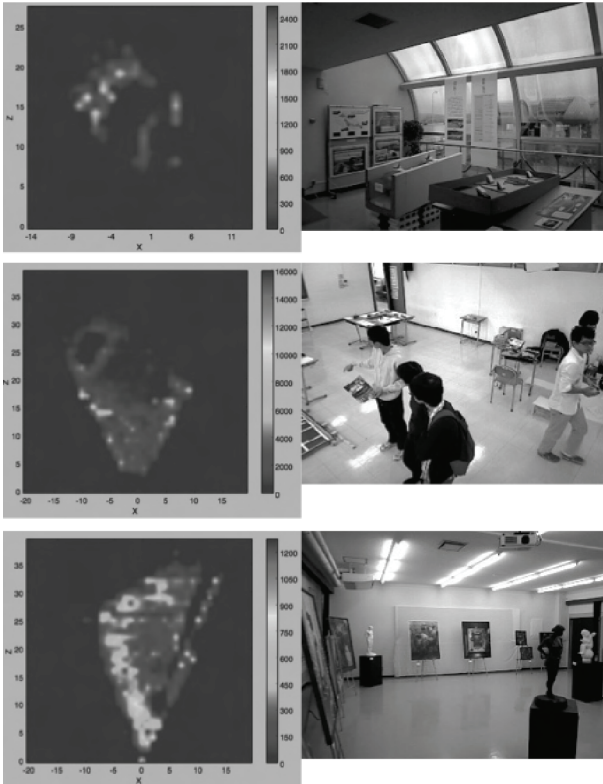


図8 観覧者密度マップとカメラの光学画像
(上段:魚道展, 中段:ロボ展, 下段:美術展)

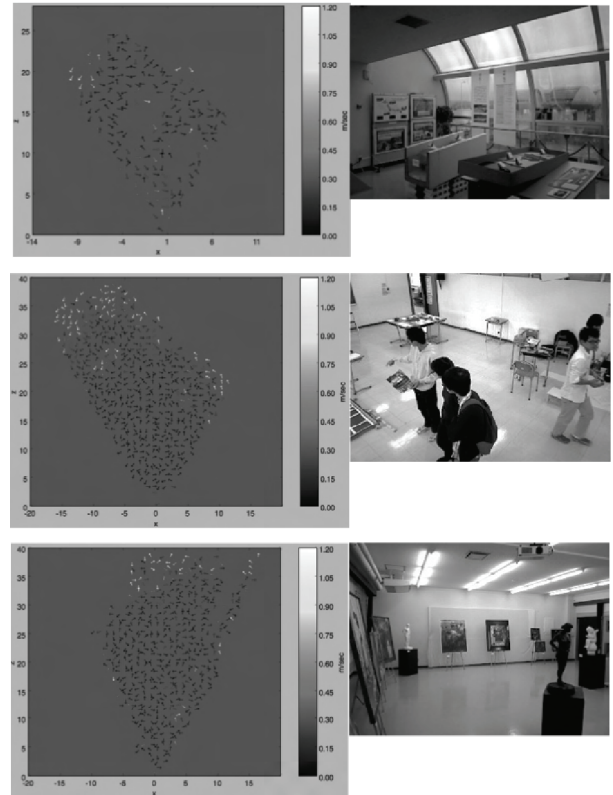


図9 移動速度ベクトル分布とカメラの光学画像
(上段:魚道展, 中段:ロボ展, 下段:美術展)

色に近いほど密な領域であり、青色に近いほど疎な領域である。

魚道展では、奥側に設置されたパネル展示の周辺が密になっており、観覧者が立ち止まったことが考えられる。ロボ展では、左側の緑色シートの展示物に沿って密になっており、認識の展示であったことが考えられる。美術展では、左側の壁の展示物に沿って密になっており、観覧者は壁に沿って立ち止まりながら見学したことが捉えられる。特に、美術展では壁沿いに密度の高い領域が集中していることが明確に現れている。一方、展示スペース中央の密度は壁沿いと比べて全体的に低く、観覧が少なかったことが考えられる。展示スペースの特定の領域に観覧者が集中することは、観覧順路の混雑につながりスムーズな観覧が難しくなる可能性がある。そこで、この結果から壁沿いの人気な作品の一部を中央に配置換えし、混雑を分散させるか、中央に新たな展示物を設置することでスペースを有効に活用するといった改善を提案することができる。

博物館の行動分析(観察調査)の指標のひとつに「展示

創始に立ち止まった観覧者の人数の割合の指数」を示すアトラクティブ・パワーがある¹⁰⁾。展示物周辺で密な領域は、立ち止まった観覧者の割合が高い、すなわちアトラクティブ・パワーが高いことが言える。密度分布図から展示スペース全領域のアトラクティブ・パワーを定量的に導出できる。どの展示物が人気かを判断する根拠となり、展示スペースを有効活用できるように配置換えの根拠やその評価につなげられる。

移動速度ベクトル分布図

記録した座標値から観覧車の移動速度ベクトル分布を作成した(図9)。矢印の向きが観覧者の移動方向を示し、色が速さを示している。展示スペースの人の流れと相対的に移動速度が速い領域と遅い領域を捉えることができる。図のカラーバーが示すように白色に近いほど速く歩いた領域で、黒色に近いほどゆっくり歩いた領域である。展示スペース XZ 平面を 25cm×25cm のセルに区切り、セル毎に移動ベクトルを表した。移動速度ベクトルは観覧者の座標値を時間微分することで導いた。

領域毎の移動速度は観覧者密度分布マップと同様に、

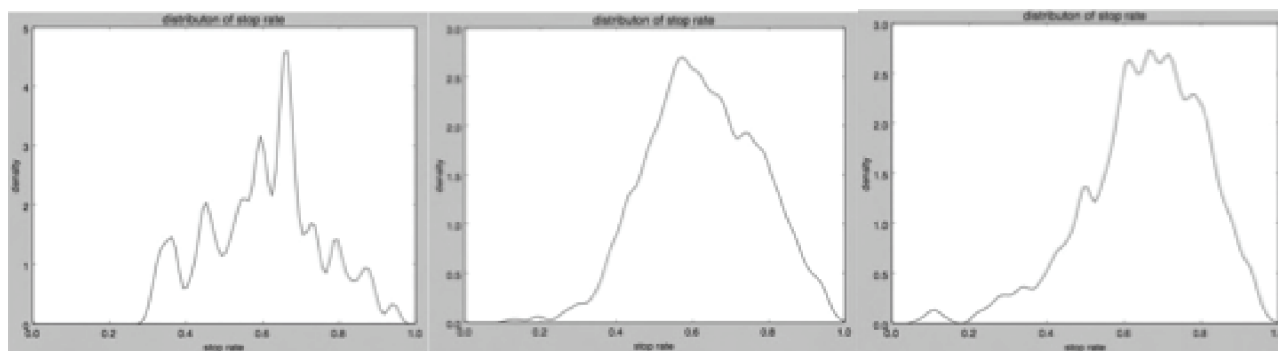


図 10 立ち止まり比率の度数分布(左から魚道展, ロボ展, 美術展)

アトラクティブ・パワーを測る根拠となる。矢印の向きは、観覧者の平均的な巡回路を知る根拠となる。矢印の向きと移動速度を合わせれば、入り口から出口まで、どのような動線をたどり、どこで立ち止まったのかを評価することにつながる。

立ち止まり比率の度数分布

記録した座標値、入室時間及び退室時間から展示スペースでの移動時間と達止まりの時間比率を算出し、全観覧者の立ち止まり時間比率の度数分布図を作成した(図10)。展示スペースの滞在時間の内、実際に展示物をじっくり見た時間と展示物間を移動したなどの時間の比率である。

観覧者の座標値を時間微分し、時間後との移動速度を求めた。立ち止まり中の移動速度は0、もしくは0に近い値であると言えるため、移動速度をk-means法により立ち止まり時間と移動中に2分類した。「立ち止まり時間率=立ち止まりに分類された速度の数/全移動速度の数」である。

博物館の行動分析(観察調査)の評価のひとつに「立ち止まり観覧者が費やした時間を、その展示を観覧するのに必要な時間に対する割合の指数」を示すホールディング・パワーがある¹⁰⁾。立ち止まり時間の比率が大きければ展示スペース全体のホールディング・パワーは高いと言える。立ち止まり時間率の分布もしくはその平均や標準偏差などはホールディング・パワーの根拠となる。立ち止まり時間率の導出は、展示スペース全体を定量的に評価することにつながる。

7 まとめ

本研究では、距離画像の画像処理による展示スペース

の自動行動分析システムを開発した。これにより、長時間にわたり定量的な行動動線の記録を実現した。これまでスタッフの負担となってきた目視による行動動線の記録を自動化した。そして、システムによる行動動線記録は定量的でかつ大量あるため、統計などを駆使した科学的根拠に基づいた博物館の分析・評価そして検討につながる事が期待できる。

本研究では、システムの定量的な記録からの評価方法として、これまでの目視で記録された行動動線と等価な再現に加え、観覧者密度マップ、移動速度ベクトルの分布、立ち止まり時間比率の度数分布の作成を提案した。

システムの画像処理の開発過程で、コードブック法に深度情報を適用することで正確で高速な前景抽出、人物が平面XYで連続であってもレイヤー化したZ方向を合わせた座標空間で個別に人物を識別できる距離階層化ラベリング、フレーム間での物体間距離および不在度や人物同士の重なりを考慮した物体追跡を提案および開発した。

参考文献

- 1) 原宏, 1982, 展示スペースに於ける鑑賞動線の分析に関する研究, 学術講演梗概集. 計画系 57(計画系), pp. 1109-1110
- 2) 松久 伸行, et al. 1984, 美術館に於ける鑑賞動線に関する研究, 建築雑誌. 建築年報 1984, p. 86
- 3) 野村東太, 1993, 博物館の展示・解説が来館者行為に与える影響: 博物館に関する建築計画的な研究V, 日本建築学会計画系論文報告集(445), pp. 73-81.
- 4) 朴鍾来 et al. 2005, 科学系博物館における展示手法と利用者の行動特徴からみた展示の分析, 日本建築

- 学会計画系論文集 (593), pp. 57-63
- 5) 森田真司 et al. 2005, 全方位画像センサを用いたネットワーク対応型遠隔監視システム, 電子情報通信学会論文誌. D-II, 情報・システム, II-パターン処理 J88-D-II(5), pp. 864-875
 - 6) Kyungnam Kim, et al. 2005, Real-time foreground-background segmentation using codebook model, Real-time Imaging, 11-3, pp. 167-256.
 - 7) これからの博物館の在り方に関する検討協力者会議, 2007, 新しい時代の博物館制度の在り方について (報告) 文部科学省 p. 13
 - 8) Gary Bradski, Adrian Kaehler, 松田晃一訳, 2009, 詳解 OpenCV, オライリージャパン, pp. 280-290
 - 9) 江水是仁, 2011, 展示の配置の違いと観覧集団の違いによる観覧行動の特性: 博物館のビジュアルスタディとしての観覧行動に関する研究その 6, 学術講演梗概集 E-1, 建築計画 I, 各種建物・地域施設, 設計方法, 構法計画, 人間工学, 計画基礎, pp. 81-82.
 - 10) 大堀哲, 水嶋英治, 2012, 博物館学 (2) 博物館展示論 *博物館教育論学文社, pp. 115-116
 - 11) 佐々木亨, 泰井良, 2012, 博物館評価を評価する: ODA の評価方法・枠組みと比較して, 日本ミュージアム・マネジメント学会研究紀要, 16, pp. 103-111
 - 12) Enrique J. Fernandez-Sanchez, et al. 2013, Background Subtraction Based on Color and Depth Using Active Sensors, Sensors, 13, pp. 8895-8915.