

Curating Frame: 日常風景を作品化する自律移動型額縁

油井俊哉^{†1} 橋田朋子^{†1}

概要: 美術用品は何気ない日用品であっても作品だと勘違いさせ鑑賞させる効果を持ち、再解釈の機会を与えることに利用できる。また、機械に壁のシミを顔だと判断するような人とは少しずれた機械の解釈を伝えることも人に新しい発見を与える手がかかりになるだろう。本システム, Curating Frame は日常を再認識させるため、額縁やキャプションといった美術用品を自律型機械に変え、風景を展示するように切り取り名前をつけて日常を作品化していくシステムである。具体的には、壁や窓沿いを自律移動することで動的に風景を切り取っていく額縁と、額縁内の風景を識別しタイトルとして表示していくキャプションによって実現される。

Curating Frame: Autonomous Moving Picture Frame for Museumizing Everyday Landscapes

TOSHIYA YUI^{†1} TOMOKO HASHIDA^{†1}

Abstract: Art supplies can be used to give opportunities for reinterpretation by generating misunderstanding and by fostering appreciation that even casual daily objects are works of art. Also, communicating the interpretation of a machine that is slightly different from what is expected may be a clue that gives a new interpretation to the person. A Curating Frame is a system that turns art environments such as picture frames and captions into autonomous machines by extracting landscapes, displaying them, and giving them a title to create works of art. Specifically, it is realized by a frame that dynamically extracts landscapes by autonomously moving along walls and windows and by a caption that identifies a landscape within a picture frame and displays it as a title.

1. はじめに

現在, “歩きスマホ” が社会問題に挙がるほど私たちはデジタルメディアにばかり夢中になり, 実世界の日常風景に目を向け発見を得る機会が少なくなっている。このような現状に対し筆者らは, 何気ない日常風景を新たな視点で捉えることを目的にした, 実世界の対象に自然と目を向けて考えさせる仕組みが必要であると考え。そのため, その場にあるものを何か深い意味を持った作品と思わせてしまう美術用品や美術環境の効果に着目した。例えば, メガネを展示会場の床に置いたところ作品と誤解され, 観客が撮影をするまでに至った例が最近報告されるなど[1], 美術用品や美術環境はあらゆるものを作品化する効果を持っていることが示唆される。

また鑑賞対象となった日常風景に対し, 人が発見を得る手がかかりとなりうる仕掛けも重要と考えられる。その一つの方向性として, 画像認識における機械の誤認識がまるで見立てのように感じられる現象に注目する。機械が “あごひげ” を “携帯電話” に解釈した例[2]や, 顔交換アプリにて人の顔と顔ではない物体の箇所とが交換された例に見られるように, 誤った認識結果であっても正解に近い関係性を感じられると, あたかも機械が見立てを行ったかのように見える。一般的な作品でも対象自体を直接示すタイトル

ではなく見立てをタイトルとして用いることで解釈に幅を持たせることはよく行われている。そこで人の想像力を掻き立て再解釈を促す仕掛けとして, 機械の誤認識に基づく見立てを作品タイトルとして示すことが考えられる。

以上の背景を受け本研究では, 日常風景に自然と目を向け鑑賞の対象化をさせる美術用品として特に象徴的な額縁とキャプションを用い, さらに誤認識に基づく機械の見立てを作品タイトルに入れることで日常を作品化する自律移動型額縁 Curating Frame (図 1) を提案する。Curating Frame は展示企画を行うキュレータのように額縁自ら作品となるものを探しにいき, キャプション自らタイトルをつける機械である。具体的にはあらゆる対象に対して移動し風景を切り取っていく空洞の額縁と, 誤認識が起きやすく設計された機械の認識を風景のタイトルとして表示するディスプレイ型のキャプションのシステムである。本稿では先行事例, システム, 精度実験と動作確認, 議論, まとめと今後の展望について述べる。

^{†1} 早稲田大学
Waseda University



図 1 Curating Frame

Figure 1 Curating Frame.

2. 先行事例

2.1 美術用品を加え日常を作品化した事例

美術用品があるとその場にあるものを鑑賞すべき対象とってしまう、人の先入観を活用した先行事例が複数ある。The Vitrine Project[3]では日用品にガラスケースと、キャプションをつけることで作品化をし、日常はデザインで溢れていることを伝えている。また、watch your street[4]ではストリートにある人工物に対して額縁やキャプション、什器を備え付けることで道ゆく人に注意をむけてもらうための展覧会を行なっている。さらに、REAL LIFE INSTAGRAM[5]ではストリートに貼られたインスタグラムの投稿を模したボードの作品である。この作品は、本来ならば写真が入る部分がくり抜かれ、透過し色がつくフィルターをそこに施し、ボードの奥の日常風景を切り取っている。これらの事例では作品化させる美術用品は動かないのに対し、筆者らが提案するシステム Curating Frame は自律的に動き、より多くの対象を作品化していく。

2.2 機械の解釈を伝え日常を再解釈させる事例

機械による認識で時折発生する機械と人間との認識のズレをあえてそのまま利用し楽しませたり、機械の見立てのように捉えて活用する仕組み・作品も増えてきている。“背面ストライプの浦島太郎 日本昔話 Remix2” [6]では、日本の昔話を日本語から英語、英語から日本語へと2度の機械翻訳を通してできた文章からなる機械翻訳の解釈のズレを用いた絵本である。また、Shinseungback Kimyonghun の作品である Cloud Face[7]では顔検出アルゴリズムによって顔と認識された雲の写真の複数載せている。さらに、recognition[8]は、Tate のコレクションデータベースを用いて、アート作品とニュース画像の類似性を計算し、二者を関連づけるように表示する作品である。これらの作品では、機械に解釈させる対象は人が指示や用意をしており、実世界を探索することはなかった。それに対して、本システムでは実世界を動き回ることによって機械が解釈する対象を探索していく点が異なる。

2.3 自律移動型ロボットによる作品化の事例

自律移動型のロボットを用いて日常の風景や人を作品として捉えるものとして、Move Instinctively[9]がある。これは、複数のルンバの上にカメラを乗せ鑑賞者の不意な表情を撮り映像として表示していく作品である。また、鑑賞者の目の位置に合わせてロボットアームを用いて額縁を動かし、額縁の中にある磁性流体でその鑑賞者の目を反映した Eye Catcher[10]がある。これらの作品では映像データや磁性流体を用いており、作品化するものを一度別のメディアに落とし込んでいる。それに対して Curating Frame では直接実世界の風景を素材としてそのまま切り取ることで日常を直接的に作品化させる。

3. システム

提案するシステムは、自律移動し風景を切り取る額縁と、風景に対して少しずれた機械の解釈によるタイトルをつけて表示させるキャプションにより、日常を作品化する仕組みである。将来的には外で動かすシステムを目指す、今回はその前段階として室内の壁や窓沿いにて認識対象まで移動するシステムを提案する。額縁とキャプションのより具体的な要件を下記に挙げる。

- (1) 額縁：壁に認識対象が存在する場合、その壁の風景まで自律的に移動し留まる
- (2) キャプション：風景から誤りを含みやすいよう識別しタイトルを動的表示する

提案システムは(1)を達成する自律移動額縁機構と(2)を実現する動的表示キャプション機構の2機構によって構成する。本システムの全体像を図2に、システムの構成図を図3に示す。

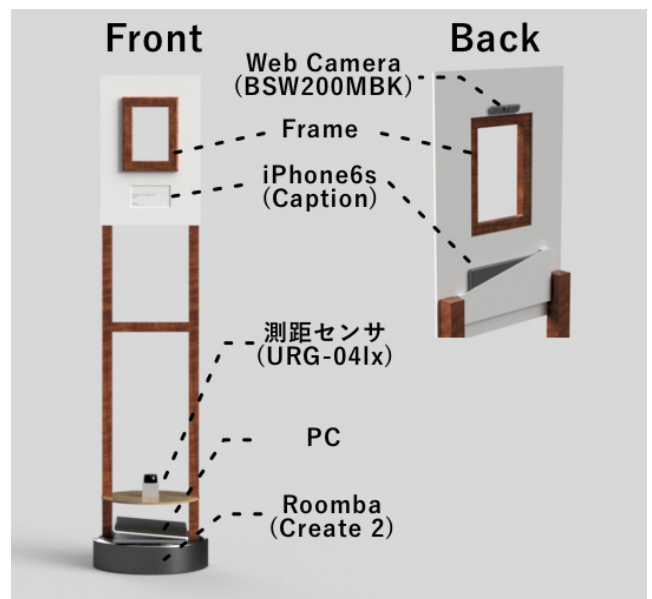


図 2 システム全体図

Figure 2 Overall of the System.

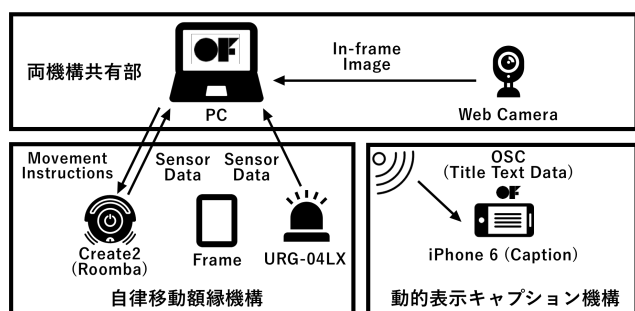


図 3 システム構成図

Figure 3 System Configuration Diagram.

3.1 自律移動額縁機構

自律移動額縁機構は壁沿いにおいて機械の認識対象まで自律移動し停止をする機構である。それを実現するため、自己位置推定と目的位置・角度への移動に基づく自律走行と機械の認識対象を探る関心領域取得の二つを主に行う。

3.1.1 構成要素

本システムで構成されるハードウェアは額縁、自律移動をするロボットであるルンバ (Create 2, iRobot)、角度分解能 0.36° で測距範囲 240° 、4m までの距離を測定できる測距センサ (URG-04LX, 北陽電機)、額縁の風景を捉えるため Web カメラ (BSW200MBK, Buffalo)、測距センサ用モバイルバッテリーと PC である。このうち Web カメラと PC は自律移動額縁機構のみでなく、動的表示キャプション機構でも用いる両機構共有部である。また、ソフトウェアは opneFrameworks, 測距計を動かすためのアドオンである ofxURG[11], openCV を使用した。

3.1.2 自律走行

主に自律移動額縁機構内のルンバと測距センサを用いて自己位置推定と目的位置・角度への移動を行い、自律走行機能を実現する。

ルンバはシステム全体を移動させるため、ルンバ上部に全システムを乗せ、測距センサはルンバの中心軸上に収まるように配置した。額縁は風景を覗き込めるようにするため、額縁内部を空洞にして額縁サイズに合わせた穴の空いている白いボードに取り付けた。

ルンバの移動は Create2 Open Interface[12]を用いて PC とシリアル通信で前進・後進・左回転・右回転・停止といった命令を出し制御している。また、車輪のモータを PID 制御することで目的角度、目的距離までの移動 (ナビゲーション) が可能である。

さらに自己位置推定や緊急停止を行うためルンバのセンサデータをストリームで取得した。センサデータとして取得したのはモータの Encoder Count, 衝突を判定する Bump, タイヤが浮いていないかを見る Wheel Drop などのデータである。Encoder Count を用いて使用開始時からのルンバの移動距離や回転角度を算出し自己位置推定をする。加えて、測距センサを用いて壁までの距離・ルンバと壁の角度を算

出することで自己位置推定の補完と壁に沿った走行が可能となった。

3.1.3 関心領域取得

自律移動額縁機構では風景に対して決めた関心領域に額縁が一旦止まるように設計する。そのため自律移動額縁機構内の Web カメラを用い、額縁内のみならず額縁外を含めた全体の風景を広く撮影しつつ、関心領域を探してその町域の中心まで留まる機能を実現する。

Web カメラでは、額縁を正面から鑑賞者がのぞいた時の風景を捉えなければならない。そのため、額縁の裏上部に識別用の広角 Web カメラを設置した。このカメラでは額縁手前 50cm 地点で鑑賞者が見た風景であれば、額縁の奥へ 20cm 以上離れた位置にある対象が撮影可能である。カメラの位置は半径が約 17cm のルンバのほぼ中心軸にあり、ルンバの半径よりも内側に対象物が入ってくることは考えにくい。カメラの取得すべき範囲としては十分である。

また、今回額縁が移動し留まるための関心領域の条件として、壁の前に何かしらの物体がある、もしくは窓の景色があることとした。この関心領域の抽出をするためグレースケール化したカメラ画像における白壁ではない箇所を色の濃さで閾値を変え輪郭抽出をすることにより実現した。その関心領域の位置情報は自律走行する際のソフトウェアに送られ、額縁が対象に留まる際正確に額縁に収めるための位置調整に用いられる。また、数秒から数十秒間留まり展示したのちに、再度新たな風景を探しに額縁が移動して行くよう処理が施されている。このとどまる秒数に関しては動的表示キャプション機構で求める認識の評価によって変化するようプログラムされている。さらに、常に関心領域がカメラ画像全体に広がっている場合には、額縁はゆっくりと風景を展示して進んで行くように調節した。

3.2 動的表示キャプション機構

動的表示キャプション機構は、風景に対して行う認識に多少誤りを持たせることで、人の解釈と機械の解釈にズレを生じさせ、タイトルとして表示するキャプションの機構である。ハードウェアは両機構共有部である Web カメラ、PC に加え iPhone6s によって構成される。また認識用のソフトウェアは openFrameworks, openCV, Imagenet で学習された inception v3 を用いる。

本システムではより作品らしいタイトルにさせるため、人とズレた機械の解釈を押し出す工夫を画像処理と使用する識別のモデルで行なっている。一つ目に認識する風景画像のぼかしである。ここでは、機械にはっきりとした結果を見せるのではなく、特徴量を削減させるために風景画像にブラーをかけ解像度を下げた。これにより輪郭や形状があやふやになり、機械は少しズレた結果を出しやすくなる。また、二つ目としては認識モデルとして用いた inception v3 である。今回、認識させるためのモデルとして inception v3 の他にクラウドサービスである Computer Vision API

(Microsoft Azure)[13]や Google Cloud Vision API (Google Cloud Platform)[14]を検討した。ただ、クラウドサービスであると識別精度が高く、機械のズレが生じにくかった。そのため本仕組みでは認識精度が多少低い inception v3 を使用した。

これらの手法を用いて自律移動で目的対象まで留まり始めてから再度移動し始めるまでの間、風景に何が写っているのかの識別と表示を行う。カメラ画像は額縁を覗いた際の風景の大きさと合うようにカメラ画像上の領域を制限しブラーをかけた上で inception v3 にて識別をした。今回、識別するクラス数は 1001 であった。

また、本機構で用いる iPhone6s はキャプションを表示するディスプレイとして使用した。この iPhone は額縁が取り付けられているボードの下部裏面から取り付ける。このボードの iPhone 取り付け部には iPhone ディスプレイと同サイズの穴が空いており、正面から鑑賞者が見ると iPhone の縁はボードに隠されディスプレイ部のみが見えるように設計されている。

キャプションのタイトルとしては inception v3 から得られた評価が一番高かったものを使用する。そのタイトルは、PC 上で動く inception v3 のソフトウェアから iPhone6s に OSC 通信にてテキストで送られ、一般の美術展で用いられるようなキャプションのフォーマットに従い、タイトルを入れた画像を iPhone6s のディスプレイに表示する。

4. 精度実験・動作確認

本章では、自律移動額縁機構が目的対象へと留まる位置に対する精度実験を行なった。また留まった後、動的表示キャプション機構によりどのようなタイトルを出力するのかを確認した。

4.1 自律移動額縁機構の精度実験

自律移動額縁機構では、壁の前に物体があるか認識したのち、その物体の位置まで移動して留まるように設計されている。このシステムの精度を調べるため壁に設置された対象の位置まで Curating Frame が正確に移動し停止できるのか、誤差を調べた。実験にて各対象に留まった時の額縁内の様子を図 4 に、目的対象物への移動・停止を行っている様子を図 5 に示す。ここでは壁に 50cm 感覚で壁に沿って置かれた 6 種類の対象（ボトル、花瓶、りんご、クッション、“真珠の耳飾りの少女”の絵画、“牛乳を注ぐ女”の絵画）に向かうように、壁から 30cm 離れた Curating Frame が自律的に移動し、目的対象の中心位置まで Curating Frame が留まることのできるのかを測定する。測定する誤差は、水平方向における認識対象の中心座標と額縁の中心座標の差とした。3 回壁に沿って走らせ、目的対象に対する平均誤差を出した。その結果、平均誤差は 2.5cm であった。これは鑑賞者が額縁を見る視点が中心軸から約 5 度ずれることとほぼ同義であるため、許容の範囲と言えるだろう。



図 4 目的対象物への停止後の様子とタイトルの結果
(a) ボトル (b) 花瓶 (c) りんご (d) クッション (e) 絵画
真珠の耳飾りの少女 (f) 絵画 牛乳を注ぐ女
Figure 4 Status of stop to objective object and result of title.
(a) Bottle (b) Vase (c) Apple (d) Cushion (e) Painting : Girl with
a Pearl Earring (f) Painting : The Milkmaid.



図 5 目的対象物へ移動・停止を行う様子

Figure 5 Experimental situation.

4.2 動的表示キャプション機構の出力結果

4.1 で述べた実験において、対象物に留まった後に動的表示キャプション機構がどのようなタイトルを出すのかを確認した。今回の実験において、認識の表示回数が一番多かったタイトルの結果が図 4 に示されている。また、認識対象と表示された全てのタイトルを表 1 に示す。

図 4 に示している実際の名前とタイトルの関係を見ていく。(a)ボトルには“冷蔵庫”とタイトルをつけられた。これは後ろの背景の壁が白色であることから、白物家電である冷蔵庫に機械が誤認識したと考えられる。また、(b)花瓶や(c)りんごではタイトルが実際の対象と一致した。(c)のりんごのタイトルでは筆者らが想定していた結果である“りんご”よりも詳しい青リンゴの品種“グラニースミス”を提示した。さらに、(d)クッションでは“バナナ”と認識されている。この結果はクッションの曲がっている形と黄

表 1 表示されたタイトル一覧

Table 1 List of displayed titles.

識別する対象	認識結果
花瓶	花瓶 / マラカス / 砂時計
ボトル	冷蔵庫 / ワインボトル / 香水
りんご	グラニースミス
クッション	バナナ / グラニースミス
真珠の耳飾りの少女	海水帽 / アカデミックガウン
牛乳を注ぐ女	ベーカリー / 祭壇 / レストラン

色よりこのタイトルが得られたと考えられる。また今回、(e)真珠の耳飾りの少女、(f)牛乳を注ぐ女 というようにすでにタイトルが付けられている二つの絵画に対しても Curating Frame にタイトルをつけた。その結果(e)真珠の耳飾りの少女では“海水帽”とタイトルがつけられ、青ターバンの部分が誤認識されたように感じられる。(f)牛乳を注ぐ女では“ベーカリー”という結果が得られた。これは、パンが絵画の中に写っていることから得られた結果だと考えられる。これら(a)~(f)までを検討してみたところ、動表示キャプション機構において目的とした少しズレた機械の解釈によるタイトルがつけられたと言えるだろう。

5. 議論

本システムでは額縁の奥の風景を作品として用いているが、それにより生じる問題や課題が3点ある。鑑賞者の視線と額縁の位置関係によって額縁に収まる風景の領域が変わる点、額縁と対象物の距離によって額縁の風景に合ったカメラ画像の領域指定範囲が変わってくる点、額縁の高さが変わらないため平行にしか風景を切り取れない点である。これらを解決する方法として、鑑賞者の位置を固定する工夫をすること、対象物までの距離を測りカメラ領域を制限すること、高さ変化する機構を導入することがある。

また現在のシステムでは道が荒い場所だと車輪がすべることによってエンコードによる自己位置推定がずれてしまう。このような場所で自走させるためにはエンコードに限らずより精度の高い LIDAR の使用や SLAM をもちいた正確なマップの作成、自己位置推定をする必要がある。

また、現在 inception v3 で識別を行なっているが、この識別では対象が画像領域内のどこにあるのかわからない。そのため Yolo[15]のような認識した対象の画像内における位置や大きさがリアルタイムに取得可能なものを使用し、関心領域の抽出も同時に行う方法もあるだろう。

6. 今後の展望とまとめ

本稿では、風景を切り取り名付けることで日常を作品化させる自律移動型額縁 Curating Frame について述べた。システムは自律移動額縁機構、動表示キャプション機構によって構成される。また対象への移動における精度実験を

行なったところ、壁に飾られた対象に対する額縁の位置の誤差が平均 2.5cm であった。それに合わせて、6 つの対象における Curating Frame が出したタイトルの結果を示した。

今後の展望としては、本システムを用いた実世界の検索システムがある。額縁に探してきてほしい対象を入力することで額縁が自律的に探し回ったのちに飾り、画像処理というバウンディングボックスのような役割を実世界でさせることができるだろう。他には、本システムの識別器として様々な画家が描いた絵風に似た風景を切り取り、その画家の名前をタイトルとして表紙する自律移動型額縁や、SNS に額縁が切り取った画像をあげ、フォロワーからの反応がよかったものを飾っていくように学習していく額縁のように幅を広げていきたい。

謝辞 本稿は、公益財団法人立石科学技術振興財団 研究助成 A (課題番号 2171018)による研究成果の一部である。

参考文献

- [1] creators : “A List of the Times People Mistook Totally Random Objects for Art”.
https://creators.vice.com/en_us/article/8x4pv5/a-list-of-the-times-people-mistook-totally-random-objects-for-art, (参照 2017-12-20).
- [2] Støj : <http://stoj.io/projects/an-algorithm-watching-a-movie-trailer>, (参照 2017-12-20)
- [3] Huskies : “The Vitrine Project”.
<http://www.irishdesign2015.ie/liminal/projects/the-vitrine-project/>, (参照 2017-12-20).
- [4] Till Maria Jürgens : “watch your street”.
<https://vimeo.com/185135788>, (参照 2017-12-20).
- [5] Bruno Ribeiro : “REAL LIFE INSTAGRAM”.
<http://nitcho.ws/REAL-LIFE-INSTAGRAM>, (参照 2017-12-20).
- [6] 文：原倫太郎，絵：原游，訳：翻訳ソフト，“背面ストライプの浦島太郎 日本昔話 Remix2”.
<http://www.hfj-ami.jp/remix/urashima.html>, (参照 2017-12-20).
- [7] Shinseungback Kimyonghun : “Cloud Face”.
http://ssbkyh.com/works/cloud_face/, (参照 2017-12-20)
- [8] “recognition “. <http://recognition.tate.org.uk/>, (参照 2017-12-20).
- [9] 山内沙也果 : “Move Instinctively”.
<https://www.youtube.com/watch?v=EuLzrVm7D4A&feature=youtu.be>, (参照 2017-12-20).
- [10] Lin (Charlie) Zhang : “Eye Catcher”,
<http://www.interactivearchitecture.org/lab-projects/eye-catcher>, (参照 2017-12-20).
- [11] ofxUrg <https://github.com/yusuketomoto/ofxUrg>, (参照 2017-12-20).
- [12] iRobot Create2 Open Interface
https://cdn-shop.adafruit.com/datasheets/create_2_Open_Interface_Spec.pdf, (参照 2017-12-20).
- [13] Microsoft : Computer Vision API
<https://azure.microsoft.com/ja-jp/services/cognitive-services/computer-vision/> (参照 2017-12-20).
- [14] Google : CLOUD VISION API
<https://cloud.google.com/vision/?hl=ja> (参照 2017-12-20).
- [15] “An algorithm watching a movie trailer” Joseph Redmon , Santosh Divvala, Ross Girshick, Ali Farhadi, You Only Look Once: Unified, Real-Time Object Detection. 2016 IEEE Conference on Computer Vision and Pattern Recognition