

擬似窓の映像と音の連動による快適性の検証

中山 知美[†] 佐藤 華和子^{††} 小野 景子[†][†]同志社大学理工学部 ^{††}同志社大学大学院理工学研究科

1 はじめに

窓の効用には、外との繋がりによる開放感の向上や外界情報の認知、窓からの良好な景観による疲労回復などの効果があると報告されている [1]。しかしながら、近年、都市化やビルの大規模化により、窓がないまたは窓からの景観が良好でないオフィス環境が増加している。そこで、窓の代替物としてディスプレイに窓を模した装飾を行った擬似窓に注目が集まっており、オフィス環境や無機質な空間となりやすい病院の集中治療室において導入が進んでいる [2]。加えて、ユーザの顔の動きに連動した擬似窓映像の移動、窓の開閉に連動した通風、太陽の光を模した照明などの追加要素により、擬似窓がもたらす開放感が高まることが報告されている [3]。また、擬似窓へ音を付与することで擬似窓の効用が向上することが報告されている [4]。

擬似窓の追加要素のうち、音は付与しやすく、適用効果が高いと考えられるが、実際のライブ音源は人にとって不快な音を含み、効用を低下させる場合があるため、効果的な擬似音源の制御が必要である。そこで本研究では、擬似窓に映る物体と音との制御の第一歩として、物体検出・認識手法である YOLO (You Only Look Once) [5] を用いて擬似窓に映写する映像の動きと付与する音の連動制御手法を提案する。

2 擬似窓の映像と音の連動手法の提案

本研究では擬似窓に映る物体のうち人物の検出を行い、人物の位置によって擬似窓に付与する音の音量を可変制御する手法を提案する。具体的には、人物が動画の中心へ近づくとつれて音量を上げ、遠ざかるにつれて音量を下げるように制御する。人物の検出には YOLO という物体検出・認識手法を用いる。YOLO は CNN を用いた手法であり、処理が単純で速く、リアルタイム処理に適している。

2.1 人物検出手法

YOLO を用いて人物を検出するためには、モデルのファインチューニングが必要である。ファインチュー

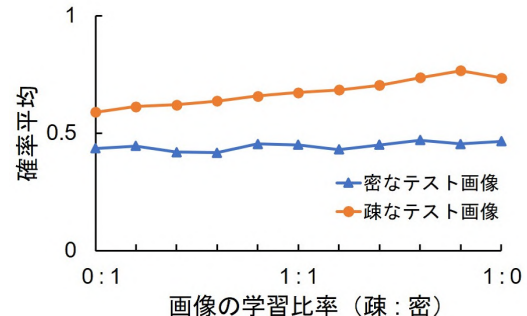


図 1: 学習データの比率による検出精度比較の結果

ニングにおいて、適切な学習データを選択するため、人が疎な画像と人が密な画像の学習比率を変更して精度比較を行った。人を囲む矩形同士が重なっていない画像を人が疎な画像、その他を人が密な画像と定義した。用いる学習データは 2 人の人物が写る画像 250 枚であり、人が疎な画像と密な画像の比率を 0.1 ずつ変更し学習した。テスト画像として人が疎な画像と密な画像を各 50 枚利用し、全検出物体のクラス確率の和を正解人物数で割った確率平均で精度を示す。クラス確率とは、識別した物体が人物である確率を示す。

学習データの比率による検出精度比較の結果を図 1 に示す。図 1 より、人が密な画像に対する検出精度は学習比率に相関がないことがわかる。一方で、人が疎な画像に対する検出精度は学習時に疎な画像の比率が多い方が高いことがわかる。よって、擬似窓に映る人物検出に用いるモデルは、全て人が疎な画像で学習を行った。

2.2 音の音量変更手法

擬似窓の映像に映る人物の位置に合わせて音を変更する手順を以下に示す。手順 1~3 は動画のフレーム毎に行い、4~6 は 0.5 秒毎に行う。 n は検出人数を示す。

1. YOLO で検出した人物の動画上の座標を取得
2. 動画の中心座標と人物座標の距離 d_i を算出
3. d_i の逆数の総和 $S_f = \sum_{i=0}^n \frac{1}{d_i}$ を算出
4. 0.5 秒間の S_f の最大値 S_s を取得
5. S_s をもとに音量変更倍率 R_s を算出
6. 音を R_s 倍の音量に変更

R_s の算出方法は、同時間の S_s のみを利用して R_s を算出する代入手法と、前後 2 秒の情報を用いて R_s を線形的に増加減少する線形変化手法を提案する。代入手法では、他の物体と重なる場合に人物検出が欠落

Collaborative Sound in Virtual Window

[†] Tomomi NAKAYAMA

(nakayama.tomomi@mikilab.doshisha.ac.jp)

^{††} Kanako SATO[†] Keiko ONODoshisha University ([†])Graduate School of Science and Technology, Doshisha University (^{††})

すると、音が途切れる。一方で、線形変化手法では音量を緩やかに変更可能であるが、現在より先の情報を用いるため音の遅延が生じる。

2.3 速度算出手法

代入手法と線形変化手法の課題を解決し、人物検出の欠落を補間してリアルタイム性を高めた手法として速度算出手法を提案する。代入手法と線形変化手法は音量変更倍率 R_s の算出手法である一方で、速度算出手法は動画上の人物の座標を取得する際に、人物検出の欠落を補間する手法である。速度算出手法における人物位置の補間の手順を以下に示す。 $t = f$ を人物位置を補間する現在の時刻、 $t = f - 1$ を1フレーム前の時刻、 x_t を時刻 t における検出人物の座標とする。

1. $t = f - 1$ において人物検出した座標の近傍に $t = f$ においても検出座標があるか判断
2. 検出座標がなかった場合、 x_{f-1} の近傍に位置する x_{f-2} を取得
3. $s_1 = x_{f-1} - x_{f-2}$ を算出
4. x_{f-2} の近傍に位置する x_{f-3} を取得
5. $s_1 = x_{f-2} - x_{f-3}$ を算出
6. 4, 5 と同様に $s_1 \sim s_{30}$ を算出
7. $s_1 \sim s_{30}$ の平均値 s_{ave} を算出
8. $x_f = x_{f-1} + s_{ave}$ を推定の人物位置として算出

3 実際の映像を用いた人物の位置推定に関する実験

3.1 実験概要

実際に擬似窓に映写する映像と音を用いて音の連動実験を行う。映像は実験室周辺の屋外で撮影した5分間の映像を用いた。音は複数の会話が混在する教室で録音したものをを用いた。本実験では、代入手法と速度算出手法において、5分間の映像のうち音の途切れが生じる割合を比較する。音の途切れは、音の音量が0でない値から0に変化したフレームと定義する。

3.2 実験結果および考察

実際の映像における速度算出手法による人物位置の補間結果を図2に示す。ピンク色の矩形がYOLOによって検出した人物位置、青色の矩形が速度算出手法により補間した人物位置を示している。図2より、木の後ろに隠れた人物の位置補間が成功していることがわかる。

5分間の映像のうち音の途切れが生じたフレーム数をカウントした結果、代入手法は約4.5% (27フレーム)、速度算出手法は約2% (12フレーム) 存在した。また、目視により利用した映像に人物が映らなくなるタイミングをカウントすると約7回存在した。よって、実際の音の途切れは、代入手法は約3% (20フレーム)、速度算出手法は約0.8% (5フレーム) 存在したといえ

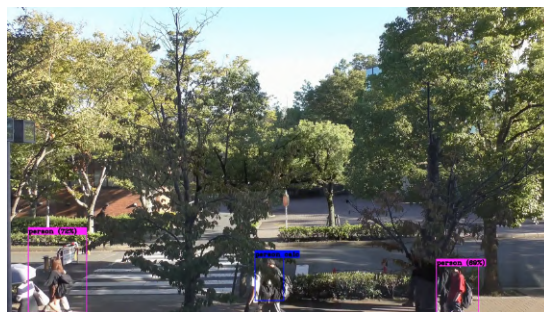


図2: 速度算出手法による人物位置の補間結果

る。したがって、速度算出手法は音の途切れを防ぐことが可能であるといえる。

以上より、提案した速度算出手法は、他物体に検出の対象物体が重なった場合のように検出が欠落した際に、位置を補間する手法として有効であるとわかる。

4 おわりに

本研究では、YOLOを用いて擬似窓に映写する映像の人物の動きと付与する音の連動手法を提案した。

また、YOLOの再学習において適切な学習データを選択するため、人が疎な画像と密な画像の学習比率を変更して精度比較を行った。精度比較の結果、疎な画像に対する検出精度は、学習時に疎な画像の比率が多い方が高いことが明らかとなった。加えて、動画における人物検出の欠落を補間する手法として速度算出手法を提案した。本手法は位置を補間し、音の途切れを防ぐ手法として有効であることを実証した。

参考文献

- [1] 武藤浩, 宇治川正人, 安岡正人, 平手小太郎, 山川昭次, 土田義郎: 窓の心理的効果とその代替可能性 地下オフィスの環境改善に関する実証的研究 その2, 日本建築学会計画系論文集, Vol. 60, No. 474, pp. 57-63 (1995).
- [2] Satoki Inoue, Eriko Takezawa et al.: Effective Medical Creation (EMC) — A New Approach to Improvement of Patient Management in the Standpoint of Hospital Room Environment, Open Journal of Anesthesiology, Vol. 10, pp. 409 - 421 (2020).
- [3] 會 珍: 開放感を高めるデジタル窓の開発, 北陸先端科学技術大学院大学知識科学研究科修士論文 (2017). <http://hdl.handle.net/10119/14109>
- [4] 久留 亜沙美, 三木 光範, 川合 由夏: 音を付与した擬似窓が執務者に与える心理的効用の検証, 同志社大学ハリス理化学研究報告, Vol. 60, No. 4, pp. 229 - 233 (2020).
- [5] Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi: You Only Look Once: Unified, Real-Time Object Detection, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 779 - 788 (2016).