2ZE-06

高齢者の住居内行動把握のための動画記録システム

酒勾 暁史 †
青山学院大学 †
社会情報学部

宮治 裕 ‡
青山学院大学 ‡
社会情報学部

1. はじめに

近年，計算機能力の向上とデータセットの構築 [1] により，機械学習手法を用いて動画像からキャプション生成が可能になりつつある．キャプション生成技術とは，画像もしくは動画像を入力すると，その説明文を生成することである．この技術を活用し，人物の行動を写した動画像を処理をすることにより，行動を自然言語で記録することが可能となる．

動画像キャプション生成技術を人物の行動記録に利用することは，行動がテキストで記録され，プライバシの問題解決につながる．また，通知の自動化など，家族による確認コストが大きく減ることも可能となる．

上記の動画像キャプション生成技術による日常記録システムを実際に使用するためには，人物の行動を記録した動画が取得できることが前提である．

全方位カメラは，一度に全方向の情報を取得でき，一部屋一台の設置のみで全体の様子を記録が可能である．このことから，特殊なセンサ機器ではなく，設置が容易で，人体には不侵食・不可侵である 360 度カメラからの画像を用いて高齢者の行動を取得することが有効である．

以上より，本研究では人間の行動や動作を把握するために必要な動画像を記録し続けるシステムの構築を目的とする．本研究では，360 度全天球カメラで室内全体の様子を把握し，人物検出をおこなったうえで，画像の切り出し処理を行い，人物を中心とした動画像の記録をおこなう．

2. システム構成

本研究で構築したシステムの概要図を図 1 に示す．本研究で構築したシステムは，全天球カメラ RICOH THETA [2] から得られるフレームに対して， 人物検出手法である SSD（Single Shot MultiBox Detector）[3] を用いて人物の位置と大きさを推定する人物検出部，人物検出の結果を元に切出し範囲を求めて切出しを行なう人物切出し部の 2 つに分かれている．この 2 つは，画像処理 PC で動作する． 2 つの処理を経たフレームは，画像処理 PC 内で動画ファイルに記録される．

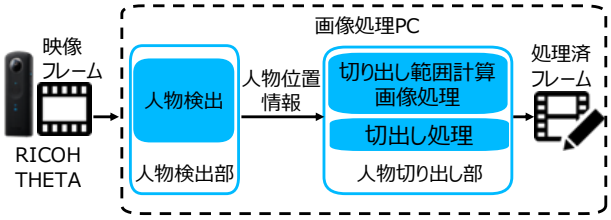


図 1: システム概要図



図 2: 拡大後の範囲（解像度 2,400 × 960）

2.1. 人物検出部

入力フレームを，重みファイルのモデルに合わせて，600 × 300 ピクセルの画像にリサイズし人物検出をおこなう．検出された物体のうち，信頼度（識別物体らしさ）が 10 %以上で person のラベルがつけられているものの座標情報（x 座標，y 座標，物体の幅，高さ）を取得した．なお，person ラベルの物体が複数検出された場合は，その中で最も信頼度の高いものの座標情報を使用する．その後，600 × 300 ピクセルのフレームで得られた物体検出の座標情報を，元のフレームサイズに対して対応するように拡大し，人物検出範囲の情報を得る．この座標情報を，次の人物切出し部へと引き継ぐ．

2.2. 人物切出し部

人物検出部からの座標情報を基に，フレーム内に収まる範囲で切り出す範囲を求め，拡大・切り出しをおこなう（図 2）．切出し後は，1,920 × 1,080 ピクセルのフレームにリサイズをおこなう（図 3）．

フレームの左端もしくは右端に人物が検出された場合は，横方向の拡張範囲がフレームサイズを超えてしまう．この場合， equirectangular 形式のフレームは，左端と右端を合わせると一周 360 度となるため，超過する側とは反対側から 480 ピクセル分を切出し，結合することで横方向にフレームサイズを拡大したうえで，検出範囲の左右を拡大し切り出しをおこなう（図 2）．

最後に，1,920 × 1,080 となるようにリサイズと切り出し処理が行われたフレームを，動画ファイルに追加して，該当フレームにおける処理を完了する．

Video capturing system for machine learning to detect human behavior
† Akifumi Sakawa, Aoyama Gakuin University
‡ Yutaka Miyaji, Aoyama Gakuin University

Copyright ©2019 Information Processing Society of Japan.
All Rights Reserved.



図 3: 切り出し処理後のフレーム（解像度 1,920 × 1,080）

3. 実験

本研究で構築した動画像を記録し続けるシステムの性能評価のために実験をおこなった．実験は，人物検出部と人物切出し部の 2 つに分けて性能を評価する．人物検出部の性能評価では，本システムの人物検出部によって人物が検知されたフレーム数を，総フレーム数で割ることによって正確に人物を検出できた割合を示す．人物切出し部の性能の評価では，この判断として，本システムの人物切出し部によって切り出されたフレームが， 物体検出アルゴリズムによって，人物が写っていると判断され，人間の行動を認識するのに適している動画となっているかを検証する．この際，1 フレームから検出される最大検出人数は 1 名とした．

3.1. 実験環境

2018 年 12 月に，学内の教室に一般家庭の室内に近い環境を用意して実験をおこなった．実験環境の広さは，約 30.72 平方メートルで，室内には，給湯スペース，中心部には机と椅子，テレビが設置してある．360 度カメラは，部屋の中心の机の上に設置した．

3.2. 実験方法

実験対象者は男子大学生 1 名とし実験をおこなった．実験は，着ている服の色が人物検出に影響を与える場合を考え，黒いパーカーとグレーのパーカーで，2 度同じ行動をおこなってもらった．実験対象者には，1. 入室，2. 水を汲む，3. テレビをみる，4. 食器洗い，5. 退出という順序で一連の行動をおこなってもらった．行動時間は，それぞれ約 2 分 30 秒間であった．

3.3. 実験 1: 人物検出部の性能評価実験

実験 1 の結果を表 1 に示す．人物検出部では，人物が正確に検出できたフレームが，52.39 %と 53.62 %に留まった．フレーム中で人物が小さくなると，人物として検出できていなかった．また，パーカーの色が床や壁・ドアの色と似ていたため，人物が背景と同化してしまったことも未検出の原因として挙げられる．

3.4. 実験 2 結果: 人物切出し部の性能評価実験

実験 2 として，実験 1 において得られたフレームに対して， SSD を用いて人物検出をおこなった．このとき，信頼度（検出物体らしさ）は，50 %以上に限定して検出対象とした．実験 2 の結果を，表 2 に示す．最大検出人数を 1 名として人物検出をおこなった場合，黒いパーカー着用時は 999 フレームに対して 989 フレーム（98.99 %）で人物が検出された．グレーのパーカー着用は，965 フレームに対して 964 フレーム（99.89 %）で人物が検出された．これは，実験 1 検出時の信頼度は 10 %から 65 %の範囲に分布することが多かったが，信頼度 90 %以上でも 9 割のフレームを検出していた．このことから，切り出し後の動画として十分な精度であったといえる．人物検出部検出時の信頼度が小さくても，人物のみが抽出・拡大されることで，実験 2 実施時は人物の信頼度が向上したといえる．

表 1: 実験 1 の結果

	総フレーム数	人物検出フレーム数（割合）
黒いパーカー着用時	1863	999（53.62%）
グレーのパーカー着用時	1842	965（52.39%）

表 2: 実験 2 の結果

	フレーム数	人物検出フレーム数
黒いパーカー着用時	999	989
グレーのパーカー着用時	965	964

4. おわりに

本研究では，人物動作の認識をおこなうことに向け，人物を検出し，切出し処理をした上で記録するシステムを構築した．検証をおこなった結果，人物検出部では，50 %程度の精度であった．人物切出し部の検証では，98 %以上の精度で人物が写っていると判断された．

本研究では， SSD を用いて人物検知をおこなったが，人物検出の精度は約 50 %ほどでしかなかった．人物検知の精度をより向上させるためには，SSD 以外の物体検出手法を用いることや各アルゴリズムのパラメータの検証と最適化が必要である．

参考文献

[1] Stair captions: Constructing a large-scale japanese image caption dataset. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, vol.2*, pp. 417–421, 2017.

[2] 庄原誠, 竹中博一. 全天球撮像デバイス “ricoh theta” の開発. 日本写真学会誌, Vol. 77, No. 3, pp. 234–237, 2014.

[3] Wei Liu et al. SSD: single shot multibox detector. *CoRR*, Vol. abs/1512.02325, , 2015.

Copyright ©2019 Information Processing Society of Japan.
All Rights Reserved.