

レーザーレンジスキャナを用いたビデオ画像内 動体位置検出システムの一検討

石 芳正¹ 川上 朋也^{1,2} 松本 哲¹ 義久 智樹¹ 寺西 裕一^{1,3} 下條 真司¹

概要：IoT 環境ではデータの相互円滑利用のため、カメラから得られたビデオ画像に応じて、メタデータを付加するといった事前処理を加えることがある。しかしながら、一般に利用されるカメラデバイスは一般的な PC と比較して処理能力が劣ることが多く、メタデータを得るためにビデオ画像解析を行うといった高負荷な処理は困難な場合が多い。カメラデバイスと接続されたエッジコンピューティング対応ルータなどの、所望の計算処理を行えるエッジデバイスにより通信経路上でメタデータの付加処理を行う事も考えられるが、高負荷な処理を担うには不十分な場合もある。そこで、高負荷な処理に代替する外部の異なるセンサー情報を用いることで、エッジデバイス上で事前処理を処理可能とする手法を提案する。その事例として、本研究ではレーザーレンジスキャナを動体検出に適用し、カメラと連携するエッジデバイスにおいてビデオ画像中の動体位置をメタデータとして付与する撮影システムについて述べる。

YOSHIMASA ISHI¹ TOMOYA KAWAKAMI^{1,2} SATORU MATSUMOTO¹ TOMOKI YOSHIHISA¹
YUUCHI TERANISHI^{1,3} SHINJI SHIMOJO¹

1 はじめに

IoT (Internet of Things) や M2M (Machine to Machine) という言葉が一般誌や報道においても頻繁に用いられるようになり、様々なデバイスをネットワークに接続し相互に情報をやりとりすることで、これまで個別に動作していた機器では実現困難なサービスの実現を目指す技術開発に注目が集まっている。これらの IoT 環境下では、各種センサより得られたデータを円滑に利用するため、メタデータが付与される場合がある。これらメタデータは、センサの ID や仕様情報、設置位置情報といった事前指定可能な固定的なデータをメタデータする機会が多いが、センサから得られた観測データをリアルタイムに処理することで、動的にメタデータを生成して付与する場合も考えられる。

例えば、大規模ビルに設置されている加速度センサにおいては、各軸の加速度から算出した振動の方向ベクトルを

付与することにより、異常振動の早期検出が期待できる。また、交差点の交通監視カメラにおいては、信号機の各ランプの点灯状態をメタデータとしてカメラ画像に付与することで、交通事故発生時の事故要因調査の省力化が期待できる。特にカメラから得られるビデオ映像においては、解析処理により様々な情報を得ることができる。主要な道路に設置されている自動車ナンバー自動読取装置 (N システム) では、通過した車両のナンバープレートを解析し、その番号と通過時刻を記録することにより、捜査対象となる車両の移動経路の特定に活用されている。屋内に設置されるカメラにおいては、複数のカメラ映像を解析することで歩行者の動線や流量を求め、避難誘導に活かすシステムが提案されている [1]。また、歩行者の動線追跡は、不審者の検知や購買者の興味・関心の分析への応用が期待できることから、カメラにレーザーレンジスキャナを組み合わせることで安定した追跡を行うシステムが提案されている [2,3]。

しかしながら、これらのシステムでは厳密な解析処理を前提としているため、解析処理に際して高性能なコンピュータが必要となる。IoT 環境下において設置される、一般的なカメラ機器それ自体は一般的な PC と比較して処理能力が劣ることが多く、画像解析処理のように高負荷な処理を行う事は困難となる。カメラデバイスと接続される

¹ 大阪大学サイバーメディアセンター
Cybermedia Center, Osaka University,
Ibaraki, Osaka 567-0047, Japan

² 奈良先端科学技術大学院大学情報科学研究科
Graduate School of Information Science, Nara Institute of Science
and Technology,
Ikoma, Nara 630-0192, Japan

³ 国立研究開発法人情報通信研究機構
National Institute of Information and Communications Technology,
Koganei, Tokyo 184-8795, Japan

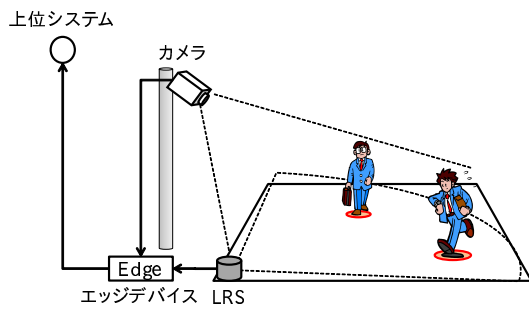


図 1: 想定システム構成

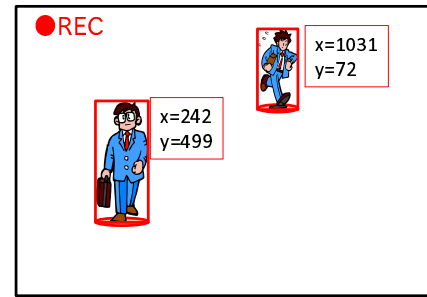


図 2: 想定出力イメージ

エッジコンピューティング対応ルータなどの、所望の計算処理を行えるエッジデバイスにより通信経路上で処理を行う事も考えられるが、画像解析処理のように高負荷な処理を担うには不十分な場合もある。

そこで、カメラ映像に加えて外部の異なるセンサー情報を用いることで、高負荷な処理に代替する情報を取得し、エッジデバイス上にて処理を処理可能とする手法を提案する。その一事例として、本研究ではレーザレンジスキャナを動体検出に適用し、カメラと連携するエッジデバイスにおいてビデオ画像中の動体位置をメタデータとして付与する撮影システムについて述べる。

2 動体位置検出システム

2.1 システム構成

図 1 は、想定しているシステム構成を示している。ある範囲を撮影するネットワークカメラ 1 台と、そのカメラの撮影範囲を計測範囲に収めるレーザレンジスキャナ 1 台が 1 つのエッジデバイスにより制御される。ネットワークカメラの撮影範囲は固定されたものとし、パン、チルト、ズームなどの撮影範囲が変わる動作には対応しないものとする。ネットワークカメラとレーザレンジスキャナが取得したデータは、このエッジデバイスを介して上位システムに伝送される。

検出対象とする動体は、主としてカメラ撮影範囲内を歩き来する歩行者を想定しており、検出された動体の位置をカメラ映像上で強調表示するものとする(図 2)。また、同時に検出された動体の数やカメラ画像上の位置などの情報をメタデータとして得る。

2.2 レーザレンジスキャナによる動体検出

レーザレンジスキャナは、レーザレンジスキャナより照射されたレーザ光が対象物により反射され、レーザレンジスキャナに戻るまでの伝搬時間よりスキャナと対象物との間の距離を非接触で測定する距離計測器である。同様にレーザ光を用いる距離計測器としてレーザレンジファインダがあるが、レーザレンジファインダが利用者が指示する

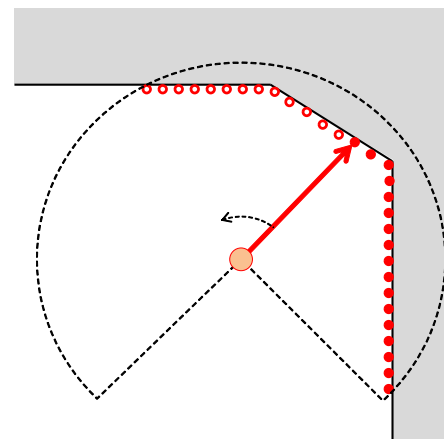


図 3: レーザレンジスキャナによるスキャン例

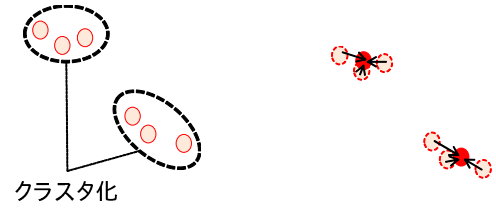
特定の対象物との距離計測を目的としていることに対し、レーザレンジスキャナは内部の測距部を周期的に回転させながら計測することにより、レーザレンジスキャナを中心とする円周方向の距離情報を点群として得ることを目的としている。

図 3 は、レーザレンジスキャナにより取得できる計測情報の模式図を示している。中央の円はレーザレンジスキャナ、破線で示した範囲は、レーザレンジスキャナにより計測可能な範囲を示しており、レーザレンジスキャナから伸びた赤矢印はレーザレンジスキャナより照射されたレーザ光を示している。レーザ光は 1 ステップずつ照射角度を変えて照射され、図の例では黒実線で示した障害物に照射されたレーザ光により赤丸で示した計測点に対する距離が得られている。続く白抜き赤丸は、これから計測される計測点を示している。

レーザレンジスキャナの一例として、北陽電機株式会社の UTM-30LX-EW [4] の仕様を表 1 に示す。UTM-30LX-EW では、1 回のスキャンによりスキャナを中心とした 270 度、半径 30 m の範囲に存在する 1081 点 (270° / ステップ角 0.25° + 端点 1) の距離情報が得られる。スキャンは 25 ミリ秒間隔、即ち毎秒 40 回行われるため、計測範囲内の

表 1: UTM-30LX-EW の仕様 (北陽電機株式会社資料より抜粋)

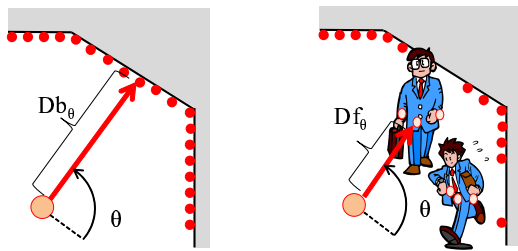
項目	仕様
レーザー光源	半導体レーザー 905nm
計測距離	0.1 ~ 30m
測距精度	± 30mm (0.1~10m), ± 50mm (10~30m)
走査角度	270°
角度分解能	ステップ角 0.25°
走査時間	25ms / scan
外寸・重量	W62 × D62 × H87.5 (mm), 300g
インターフェース	Ethernet 100BASE-TX
電源・消費電流	DC12V, 700mA (定常動作時)



(a) 検出点のクラスタリング

(b) クラスタ中心の算出

図 5: 動体位置の算出



(a) 背景距離登録

(b) 動体の検出

図 4: 背景差分法による動体の検出

状態変化をリアルタイムに取得できる。

カメラの撮影範囲に対応するため、レーザーレンジキャナはスキャン面が地面と平行となるよう設置し、歩行者の胴体部を捕らえる方法 [5, 6, 7] を用いる。スキャン範囲内における動体検出は、これらの先行研究と同様に背景差分法を用いる。

背景差分法では、まずレーザーレンジキャナと背景となる障害物との間に動体が存在しない状態で計測範囲内を距離計測し、レーザーレンジキャナと背景となる障害物までの距離を記録する。この時、レーザー光の照射角度 θ に対応する距離を Db_θ とする (図 4a)。背景距離を記録した後、レーザーレンジキャナは計測範囲内を距離計測を繰り返す行い、照射角度 θ に応じた距離 Df_θ を得る。計測範囲に動体が現れた場合、レーザーレンジキャナとその動体までの距離が計測されるため、 $Df_\theta < Db_\theta$ となる計測点が現れる (図 4b)。このように $Df_\theta < Db_\theta$ を満たす動体の侵入によるものと考えられる計測点を検出点と呼ぶ。得られた検出点は、位置情報がレーザーレンジキャナの位置を基準とする極座標で表現されており、点間の距離などが扱いたげないため、ここでレーザーレンジキャナの位置を原点とする直交座標系に変換する。

検出点は、レーザーレンジキャナと動体の距離に依存するが、一般に一つの動体に対し、複数個の検出点が得られることとなる。このため、得られた検出点の集合に対しクラスタリング処理を行い、各動体に相当する点群を得る

(図 5a)。それら点群の中心点を算出し、動体の推定位置とする (図 5b)。

以上により得られた動体位置は、レーザーレンジキャナを基準としたスキャン水平面上での位置であるため、カメラ映像の座標とは適合しない。このため、カメラの設置位置を基準とする座標系に変換する。この変換処理は、一般的な 3D CG におけるワールド座標変換、ビュー座標変換に相当し、動体位置座標と変換行列の積により変換処理を行える。得られたビュー座標に対し、射影変換を行いカメラ映像上の座標を得る。これらの座標変換処理は、カメラとレーザーレンジキャナが固定的に設置されていることから、変換行列には事前に算出した行列が利用できるため、都度算出は不要となる。カメラ映像中の動体の数は、一般にカメラ映像の画素数より十分に小さいことから、カメラ映像に対して画像解析を行うより低負荷での動体位置検出が期待できる。

2.3 カメラ映像解析とレーザーレンジキャナの連携による動体検出

レーザーレンジキャナは、障害物に反射されたレーザー光により測定するため、障害物によりレーザー光が遮られる範囲の情報は取得できない。これはオクルージョン問題と呼ばれており、図 6 では、図中 O で示した範囲はレーザーレンジキャナからは不可視の領域となるため、そこに動体が存在し、カメラにより撮影されていたとしても検出することはできない。オクルージョン問題に対しては、複数のレーザーレンジキャナを別個に設置することにより、あるレーザーレンジキャナの不可視領域にあっても、他のレーザーレンジキャナの検出結果と合わせることでオクルージョンの影響を抑制する研究は行われている [2, 7]。しかしながら、本研究ではレーザーレンジキャナを 1 台のみとしているため、これらの手法は採用できない。

そこで、この不可視領域はカメラ映像の画像解析により動体検知を行う。レーザーレンジキャナの不可視領域のみを画像解析の対象とすることにより、動体が存在する可能性がある範囲のみを画像解析することとなるため、カメラ映像全体を処理する場合と比較し、処理負荷は抑制される

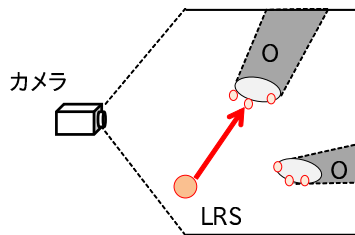


図 6: オクルージョン問題

と期待できる。また、レーザレンジスキャナによって動体が検出されていない期間は画像解析処理を省略することも可能となる。

画像解析の対象を不可視領域のみとするには、図 6 O で示した不可視領域がカメラ映像上に占める領域を求める必要がある。これには、カメラの撮影範囲に相当する仮想の 3 次元空間に不可視領域に相当する仮想オクルージョンエリアを設置し、その仮想オクルージョンエリアがカメラ映像上に占める範囲を求めることを行う。

まず、各クラスタの検出点をレーザレンジスキャナによるスキャン方向順に走査し、レーザレンジスキャナから見た場合のクラスタの端点とその座標を求める。レーザ照射方向に従い、レーザレンジスキャナ設置位置とクラスタの端点座標を結ぶ直線をカメラの撮影範囲の端まで延長することで、スキャン平面上において図 6 O で示した不可視領域を得ることができる。次に、この不可視領域を、スキャン平面に対し垂直方向（鉛直方向）に掃引することで 3 次元空間での仮想オクルージョンエリアを得る。仮想オクルージョンエリアの下限は地面とし、上限は事前設定した 2 m 程度の高さを与える。最後に、得られた仮想オクルージョンエリアに対して、2.2 節で述べた座標変換と射影変換を行う事でカメラ映像上の仮想オクルージョンエリアを得ることができる（図 7）。

画像解析処理は一般に矩形領域を対象とするため、得られたカメラ映像上の仮想オクルージョンエリアを包含する矩形を求め、画像解析による動体検出はこれらの矩形領域を処理対象領域とする（図 8）。

3 プロトタイプシステムの事前性能評価

エッジデバイスに Raspberry Pi 3 の利用を想定し、基礎的な性能評価として Raspberry Pi 3 のカメラ映像受信性能と、エッジデバイス上でレーザレンジスキャナによる動体検出処理を行った場合に受信性能に与える影響を調査した。

3.1 評価環境

先述の通り、エッジデバイスとして Raspberry Pi 3 を使用する。Raspberry Pi 3 の基本仕様は表 2 の通り、最大 1.2GHz で動作する ARMv8 コアを 4 コア持ち、1 GiB のメモリを

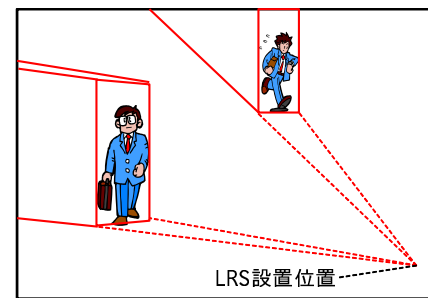


図 7: 仮想オクルージョンエリアの設定

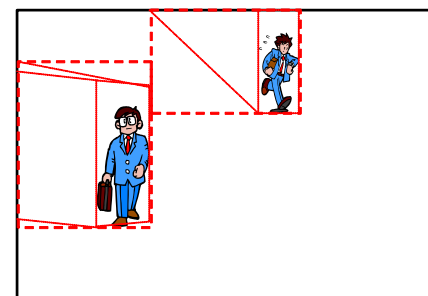


図 8: 包含矩形の算出

表 2: Raspberry Pi 3 の主な仕様

項目	仕様
CPU	Broadcom BCM2837 (ARMv8 1.2GHz 64-bit 4 core)
メモリ	1GiB LPDDR2 SDRAM
ストレージ	microSDHC カード
ネットワークインターフェース	Ethernet 100BASE-TX, WiFi 802.11 b/g/n

備える。レーザレンジスキャナは北陽電機 UTM-30LX-EW，ネットワークカメラに Panasonic DG-SP509 を利用する。動体検出処理は Raspberry Pi 3 上で実行し、ネットワークを介してレーザレンジスキャナから得られた距離情報より 2.2 節で述べた背景差分法により動体位置を順次算出する。ネットワークカメラが送出するカメラ映像のフォーマットは、エッジデバイスでの画像処理・画像加工を想定し、フレーム操作が容易な Motion JPEG over HTTP を用いる。

カメラ映像受信性能評価のため、ネットワークカメラの画像取得解像度と 1 秒あたりの撮影フレームレート（fps）を変更して映像ストリームを送信し、Raspberry Pi 3 上で動作させる受信クライアントが 10 秒間に受信できたフレーム数から 1 秒あたりの受信フレームレート（fps）を求める。画像取得解像度は 1920 × 1080，1280 × 720，640 × 360 の 3 種類を用いるものとし、各解像度において撮影フレームレート 10 fps，15 fps，30 fps として計測を

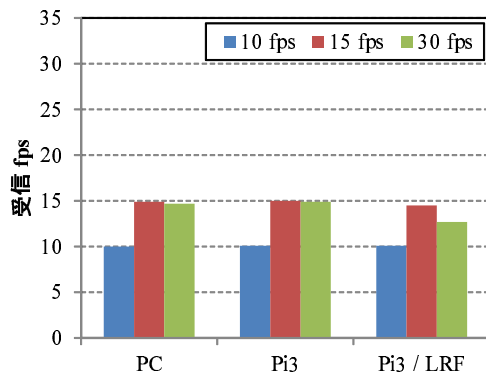


図 9: 解像度 1920 × 1080

行う。カメラの処理能力がボトルネックとなる場合も想定されるため、比較用に PC による映像受信も行った。

3.2 評価結果

図 9, 図 10), 図 11 は, それぞれ画像取得解像度 1920 × 1080, 1280 × 720, 640 × 360 で撮影した結果を表している。各グラフ横軸は, 映像受信環境を意味しており, PC は PC による受信時, Pi3 は Raspberry Pi 3 による受信時でレーザレンジスキャナによる動体検出処理なしの場合, Pi3 / LRF は Raspberry Pi 3 による受信時でレーザレンジスキャナによる動体検出処理をバックグラウンドで実行した場合となる。各グラフの縦軸は受信フレームレート, グラフの色の違いは撮影フレームレートを意味している。

解像度 1920 × 1080 の場合 (図 9) では, 撮影フレームレートを 30 fps とした場合, PC により受信した場合でも受信フレームレートが 15 fps となっていることから, 高解像度ではカメラの撮影能力の上限が 15 fps あたりとなっていることが分かる。また, Pi3 / LRF においては, 撮影フレームレート 15 fps, 30 fps において, Pi3 と比較して受信フレームレートがやや低くなっていることから, レーザレンジスキャナによる動体検出処理の処理負荷により受信処理が滞っていると考えられる。

解像度 1280 × 720 の場合 (図 10), 解像度 640 × 360 の場合 (図 11) では, PC にて受信した場合は送信フレームレート通りに受信できており, これらの解像度ではカメラの処理能力がボトルネックとなることはないと考えられる。また, Raspberry Pi 3 で受信した場合においても, ほぼ送信フレームレート通りに受信できていることが分かる。しかしながら, 解像度 1280 × 720, 送信フレームレート 30 fps では Pi3 / LRF において, Pi3 と比較して受信フレームレートわずかに (約 0.6 fps) 低くなっており, 動体検出処理の処理負荷の影響の可能性が疑われる。

以上の計測評価により, Raspberry Pi 3 においてレーザレンジスキャナによる動体検出処理を行いつつ映像を扱う

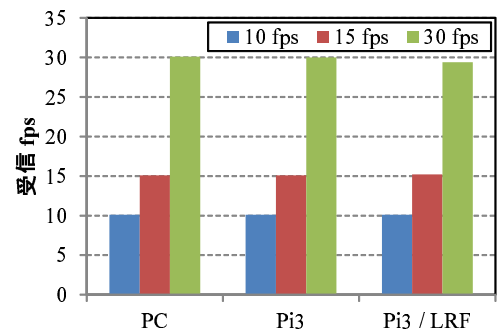


図 10: 解像度 1280 × 720

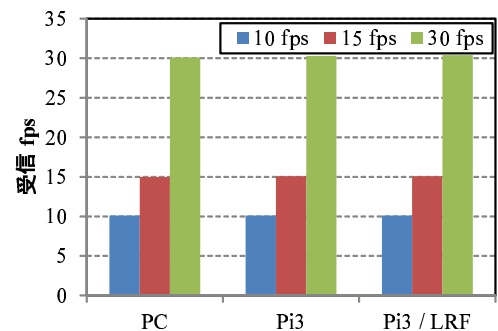


図 11: 解像度 640 × 360

場合, 1920 × 1080 の高解像度では 10 fps が実用上限となり, フレームレートが優先される場合は 1280 × 720 や 640 × 360 に解像度を落とす必要があることが判明した。

4 まとめ

本研究では, レーザレンジスキャナにより得られる動体検出情報を用いることで, カメラ映像内の動体位置検出に必要な映像解析範囲を制限し, 解析処理負荷を低減する手法について述べた。

今後は, 提案手法の実装を進め実環境下における有用性の評価を進める予定である。

謝辞 本研究の一部は, 文部科学省科学研究費補助金・基盤研究 (B)(15H02702), 挑戦的萌芽研究 (26540045) の研究助成によるものである。また, 事前評価に用いた動体検出処理用ソフトウェアは大阪大学大学院情報科学研究科廣森講師より提供を受けた。ここに記して謝意を表す。

参考文献

- [1] 寺西裕一, 地引昌弘, 西永 望: 映像解析による大規模避難誘導システムの設計, 情報処理学会研究報告, Vol. 2016-DPS-166, No. 12, pp. 1-18 (2016).
- [2] 中村克行, 趙 卉菁, 柴崎亮介: レーザスキャナと画像センサの融合による歩行者追跡, 第 11 回 画像センシングシ

- ンポジウム 講演論文集, pp. 177–180 (2005).
- [3] 小林貴訓, 杉村大輔, 関真規人, 平澤宏祐, 鈴木直彦, 鹿毛裕史, 杉本晃宏: 分散カメラとレーザ測域センサの統合によるエリア内人物追跡, 第 10 回 画像の認識・理解シンポジウム 講演論文集, pp. IS-2–23 (2007).
- [4] 北陽電機株式会社: 測域センサ UTM-30LX-EW, <http://www.hokuyo-aut.jp/02sensor/07scanner/download/products/utm-30lx-ew/>.
- [5] Fod, A., Howard, A. and Mataric, M. J.: Laser-Based People Tracking, *Proceedings of The IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3024–3029 (2002).
- [6] 和田悠佑, 山口弘純, 東野輝夫: レーザレンジスキャナによる歩行者軌跡と実人物のマッチング手法の提案, 情報処理学会マルチメディア, 分散, 協調とモバイルシンポジウム (DICOMO 2012) 論文集, pp. 1877–1888 (2012).
- [7] 上嶋祐紀, 藤田和久, 樋口雄大, 廣森聡仁, 山口弘純, 東野輝夫, 下條真司: ひとつがつながるなび一位置と気持ちと空間の共有, 情報処理学会マルチメディア, 分散, 協調とモバイルシンポジウム (DICOMO 2013) 論文集, pp. 2105–2115 (2013).