

COUNTフォレストを用いた 群衆密度の推定技術

雑踏でも人数を高精度かつ リアルタイムに推定

次世代のスマートコミュニティを目指す、安全・安心で快適な社会を実現するためには、人物の行動センシング技術が重要になります。特に、おおよしの人が集まる場所での群衆の行動分析は、映像監視システムなどへの応用に不可欠な技術です。従来の人数推定手法では、各人物の全身又は上半身を個別に検出していたため、複数の人が互いに重なり合っている場合には未検出となる人が発生し、大幅に少ない人数を出力してしまうという問題がありました。

そこで東芝は、群衆の行動分析を行う際の前提となる群衆密度の推定技術において、複数の人が互いに重なり合っている場合でも、高精度かつ少ないメモリー量で人数を推定できる手法を開発しました。

技術的背景

安全・安心で快適な社会を実現するためには、人物の行動をセンシングする技術が重要になります。群衆密度の推定技術は、画像中の人物の密度を推定して人数をカウントする技術(図1)であり、群衆の行動分析につながります。近年この技術の利用が、映像監視システムや、防災、マーケティングなどに広がってきています。例えば、防犯カメラへの応用では、異常な混雑が生じた場合に転倒事故などを検知できるようになり、監視員の労力を大幅に低減できます。また、マーケティングへの応用では、店の買い物客の混雑度を検出することで購入履歴以外の顧客行動も

分析できます。

群衆密度を推定する 従来手法と開発手法

従来の人数推定手法では、各人物の全身又は上半身を個別に検出していたため、人どうしが互いに重なっている場合には未検出となる人が発生し、人数の推定を誤ってしまうという問題がありました。東芝は、個々の人物ではなく、群衆を分析対象とすることで、この問題を解決しました(図2)。群衆を分析する手法は既にありましたが、人数推定が難しいために精度が悪く、密度を直接推定するので多くのメモリー量と計算量が必要になるという問題がありました。

今回開発した手法では、あらかじめ

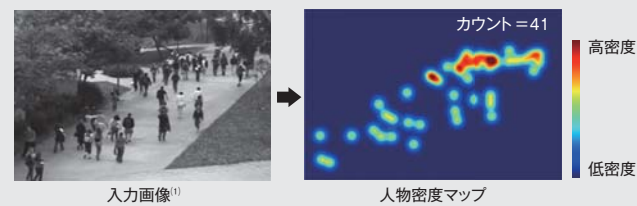


図1. 群衆密度の推定 — 群衆を撮影した画像から、群衆密度と人数を推定します。

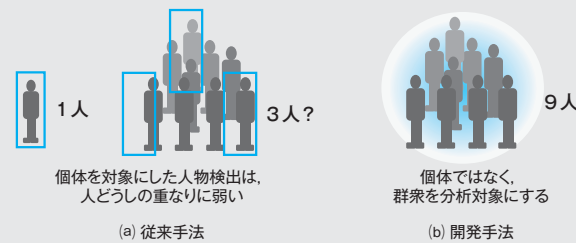


図2. 従来手法と開発手法における分析対象の違い — 開発手法では、個々の人物ではなく、群衆を分析対象とします。

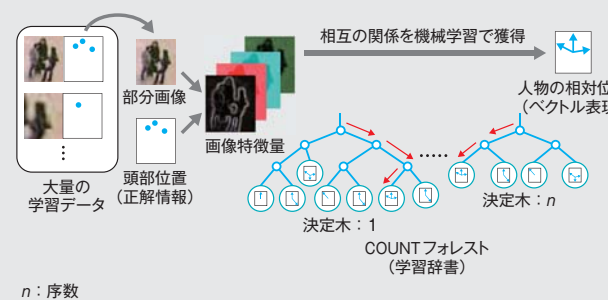


図3. 今回開発した学習手法 — COUNTフォレストを用いて、部分領域の画像特徴量からそこに含まれる人物の相対位置を出力する変換を学習します。

教示された人物の位置に基づいて、局所的な部分領域の画像特徴量からそこに含まれる人物の相対位置を獲得し、その結果を用いて群衆密度の推定を行います。人物の相対位置の情報量は小さいため、少ないメモリー量で高速処理が可能になります。

群衆の画像を小さく切り分けて、そこに含まれる人物の相対位置を出力する変換を、機械学習で獲得します。学習段階では、まず人物位置が既知の多数の画像を用いて部分画像から輝度や色などの特徴量を抽出し、その特徴量の値に応じて人物位置を求める決定木を学習します。決定木とは、データの特徴量を用いた簡単なルールで分岐を作り、特徴空間を分割することで判別や

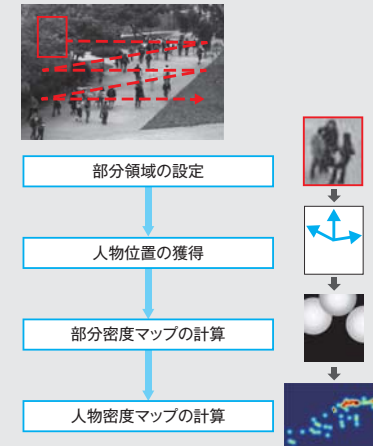


図4. 密度推定の処理フロー — 処理フローに基づいて、群衆密度を推定します。

表1. 開発手法と従来手法における人数推定精度の比較結果

手法	ショッピングモール映像データセット ⁽¹⁾			
	mae	mse	mde	
従来手法	RR ⁽¹⁾	3.59	19.0	0.110
	CA-RR ⁽²⁾	3.43	17.7	0.105
開発手法	2.50	10.0	0.080	

手法	キャンパス映像データセット ⁽¹⁾			
	mae	mse	mde	
従来手法	RR ⁽¹⁾	2.25	7.8	0.110
	CA-RR ⁽²⁾	2.07	6.9	0.102
開発手法	1.61	4.4	0.075	

RR : Ridge Regression
CA-RR : Cumulative Attribute Ridge Regression

回帰を行うモデルのことです。学習した決定木に部分画像を入力すると、人物の相対位置が出力されます(図3)。この方法は、多数の決定木を用いるためにランダムフォレストと呼ばれています。その特長は推定精度が高く処理速度も速いことで、当社は、人物の相対位置に特化した学習手法をCOUNT (CO-voting Uncertain Number of Targets) フォレストと名づけました⁽²⁾。

密度推定の処理フロー

密度推定の処理フローを図4に示します。まず、画像をスキャンして部分領域を設定し、それぞれの画像特徴量を抽出します。得られた画像特徴量を学習したCOUNTフォレストに入力し、各

領域での人物の存在位置を獲得します。次に、人物の存在確率を輝度で表した部分密度マップを計算するために、人物の各推定位置を中心とした2次元正規分布で部分密度マップを近似します。そして、各スキャン位置の部分密度マップを全画像に対して足し合わせて平均化した密度マップを算出します。この算出した密度マップが求めたい人物密度マップとなり、マップにおける密度の総和(群衆密度)が人数を表します。画像の特徴量から人物の密度推定を精度良く行うことで、人物どうしの重なりによらず、正確に人数をカウントできます。

評価実験

実験では、公開評価データのショッピングモール映像データセット⁽³⁾ (2,000枚、960×540画素)とキャンパス映像データセット⁽¹⁾ (2,000枚、238×158画素)を用いて、人数推定精度を評価しました。最初の800枚を学習用に、残りの1,200枚をテスト用に、それぞれ使用しました。また評価尺度として、平均絶対誤差mae、平均2乗誤差mse、及び平均標準誤差mdeの三つを用いました。

開発手法と従来手法における人数推定精度の比較結果を表1に示します。三つの評価尺度に対して開発手法がもっとも高い精度を実現しており、この手法の優れた性能が確認できました。

今後の展望

1台のカメラで広範囲を撮影した映像から、群衆の人数を高精度に計測できる群衆密度の推定技術を開発しました。この技術は、人が重なって映っていたり、非常に小さく映っていたりしても人物を見つけ出すことができ、公開評価データを用いた実験で優れた性能が得られることを確認しました。この

技術により、多くの人が滞留している場所を発見して通知することで監視の平準化と省力化が図れる他、滞留する場所を可視化することで混雑緩和策を検討できると考えられます。今後、この技術を、メディアインテリジェンス技術を融合して体系化したクラウドサービス“東芝コミュニケーションAI RECAI-US”へ搭載できるように、更なる開発を進めていきます。

文献

- (1) Statistical visual computing laboratory (SVCL). "Privacy Preserving Crowd Monitoring : Counting People without People Models or Tracking". SVCL. <http://www.svcl.ucsd.edu/projects/peoplecnt/>, (accessed 2017-03-17).
- (2) Pham, Q. V. et al. "COUNT Forest: CO-voting Uncertain Number of Targets using Random Forest for Crowd Density Estimation". 2015 IEEE International Conference on Computer Vision. Santiago, Chile, 2015-12, IEEE, 2015, p.3253 - 3261.
- (3) Loy, C. C. "Mall Dataset-crowd counting dataset". Chen Change Loy. <http://personal.ie.cuhk.edu.hk/~ccloy/downloads_mall_dataset.html>, (accessed 2017-03-17).
- (4) Chen, K. et al. "Feature Mining for Localized Crowd Counting". Proceedings of British Machine Vision Conference (BMVC) 2012, Guildford, UK, 2012-09, BMVC, 2012, p.21.1 - 21.11.
- (5) Chen, K. et al. "Cumulative Attribute Space for Age and Crowd Density Estimation". Proceedings of 2013 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2013), Portland, OR, 2013-06, CVPR, 2013, p.2467 - 2474.

ファン クオク ヴェト

技術統括部
研究開発センター
インタラクティブメディアラボラトリー研究主務