

Fusion of Thermal and Visible Videos for Heat Trace Detection on Touched Areas

大阪大学大学院情報科学研究科
情報ネットワーク学専攻 村田研究室
野口憲人

2023/2/9

研究背景 | マルチバンド画像処理

1

●多様なセンサを使用した画像認識技術の発展

- 可視カメラだけでなくサーマルカメラや近赤外線カメラなど比較的安価に入手可能になった
- 目的に応じてセンサを選択することで課題解決に貢献している
 - RGB カメラによる人検知, 部品の傷検知
 - サーマルカメラによる体温計測, 機械の異常検知

●マルチバンド画像処理への期待

- 異なるバンドのセンサを組み合わせることで認識精度が向上^[1,2]
- 単一バンドでは検出が難しかった物体の検出を可能に^[3]

研究背景 | 接触感染

2

●感染症における感染経路の一つとして接触感染

- 感染者が触れた物体表面にウイルスが一定期間生存^[4,5]
 - 新型コロナウイルス (COVID-19): 最大 3 日間
 - 重症急性呼吸器症候群 (SARS): 最大 9 日間
 - 中東呼吸器症候群 (MERS): 最大 2 日間
- 歴史的に 3 ~ 10 年スパンでパンデミックが繰り返し発生
 - SARS: 2002 年
 - 新型インフルエンザ (A/H1N1): 2009 年
 - MERS: 2012 年
 - COVID-19: 2020 年

接触された箇所を早急に特定し可視化することで警告したり,
早急な清掃を促したりするシステムの開発が求められる

研究背景 | 熱痕跡

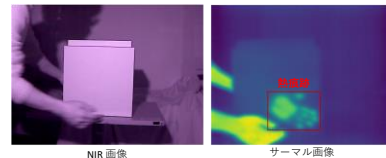
3

●接触箇所に出現する熱情報

- 人が物体に接触することで体温が物体へ伝達され, 物体表面に熱が残る (熱痕跡)
- マルチバンド画像処理の例として適している
 - サーマル画像には出現するが, RGB 画像や近赤外線 (NIR) 画像には出現しない

●熱痕跡を用いた研究例

- 所属グループ*はサーマル・NIR/RGB 画像を用いた熱痕跡可視化システムを開発^[6,7]
- 熱痕跡を用いたセキュリティやヒューマンインターフェースに関する応用例^[8,9]



NIR 画像

サーマル画像

* NTT コミュニケーション科学基礎研究所 (連携講座)

研究背景 | 熱痕跡の性質

4

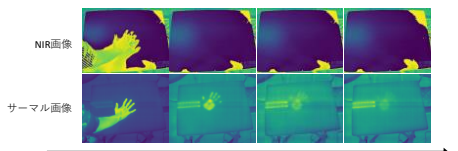
●時間的な性質

- 時間経過とともに熱が拡散することで減衰する
- 出現時に触れる行為を伴うため, 人が撮影領域に入ったりする

●空間的な性質

- 熱痕跡の出現後に熱痕跡の近くに人が存在する
- 特に, 出現直後は熱痕跡と手が至近距離に存在する

熱痕跡の検出には, 時間情報および空間情報がともに重要である可能性



NIR画像

サーマル画像

研究目的・アプローチ

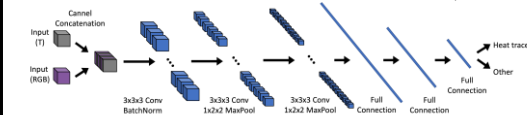
5

研究目的

- 深層学習を用いたマルチバンド画像処理による熱痕跡の識別
- 時間情報・空間情報の有効性の調査

●サーマル動画と RGB 動画を用いた熱痕跡の識別手法を提案

- サーマル画像と RGB 画像をチャネル結合し, RGB 画像として入力
- 時間・空間特徴量についても比較を行うことでより適した検出モデルを提案
 - 時間特徴量: 熱痕跡の消失特性, 発生過程における人の出入り
 - 空間特徴量: 熱痕跡および人・手の位置関係
- CNN を時間方向に次元を拡張した 3DCNN を構築
 - 連続する複数フレーム (動画) を入力可能にする
 - 熱痕跡の出現~消失を見つけられるかについて調査する
- テスト用データセットの入力における熱痕跡有無の分類における Accuracy により評価



実験手順

6

1. サーマル動画 (T) およびRGB動画 (V) の撮影

- ・ シンプルな条件で比較できるように白色壁面を接触対象とした

2. アノテーション

- ・ 矩形で熱痕跡を囲むことでアノテーションを行った
- ・ データセット作成後に「熱痕跡」または「その他」の2クラスにラベル付け

3. キャリブレーション

- ・ カメラの撮影位置から生じる位置ずれを射影変換により整合

4. 動画の時空間分割におけるデータセット作成

- ・ 空間領域の制限: 60 x 60 (narrow) または 200 x 200 (wide) で指定 (後述)
- ・ 時間方向の分割: フレームサイズを 1 (2D) または 10 (3D) に固定して分割

5. モデルの訓練・推論

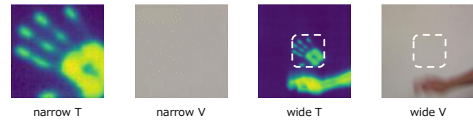
- ・ 訓練データセットにおけるラベル別データ数の割合が概ね等しくなるよう調整
- ・ テスト用に撮影した動画をテストデータセットとして使用
- ・ マルチバンド化された画像を入力する場合は入力前にチャンネル結合 (TV)

実験概要 | データセットに用いる画像

7

● 入力に用いる空間領域を指定

- ・ narrow: 60 x 60 画像全体に出現する熱痕跡が識別できるか
- ・ wide: 200 x 200 熱痕跡の発生過程 (人の腕や体の出入) が識別に効果があるか
 - GPUメモリの制約のため, wide 画像は 60 x 60 に解像度を下げて使用する
 - ・ GPU: Geforce RTX 3090 (Memory size: 24GB)

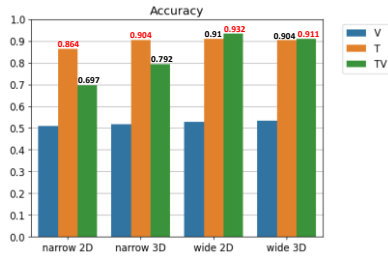


narrow および wide をネットワークの入力とした場合に、使用するバンド・時間情報が熱痕跡の識別に与える影響をそれぞれ調査する

結果・考察 | 使用バンドにおける比較 (T vs V vs TV)

8

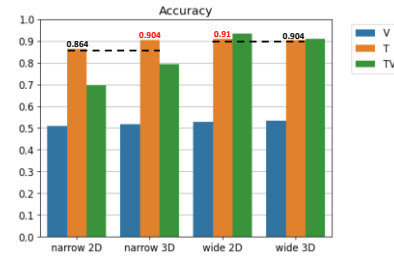
- ・ narrow では Tの方が高くなった一方, wide では TVの方が僅かに高くなった
 - narrow では撮影範囲が狭いため, 人・手の出現時間が非常に短く空間情報が効果的に活用されていない
 - wide では十分に人や手が出現しており, Vを補助的に用いることで周辺状況も含わせて学習できる可能性
- ・ Vのみでは熱痕跡の情報を含まないため, 非常に精度が低くなった



結果・考察 | 時間特徴量における比較 (2D vs 3D)

9

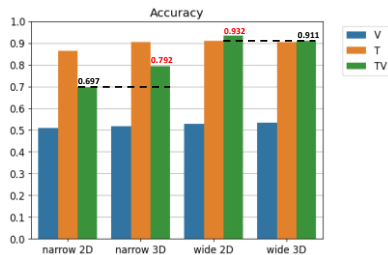
- ・ narrow では 3Dが高くなった一方, wide では 2Dの方が僅かに高くなった
 - narrow では熱痕跡の領域を事前に特定できているため, 時間情報により識別が容易になった
 - wide では熱痕跡が画像に比べて小さいため, 減衰する特徴を捉えきれない可能性



結果・考察 | 時間特徴量における比較 (2D vs 3D)

10

- ・ narrow では 3Dが高くなった一方, wide では 2Dの方が僅かに高くなった
 - narrow では熱痕跡の領域を事前に特定できているため, 時間情報により識別が容易になった
 - wide では熱痕跡が画像に比べて小さいため, 減衰する特徴を捉えきれない可能性



まとめ・今後の課題

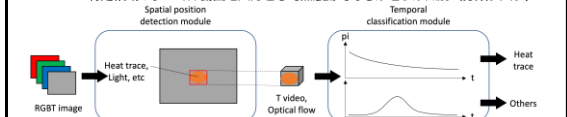
11

● サーマル動画とRGB動画を組み合わせた熱痕跡識別手法を提案

- ・ マルチバンド化・3DCNNを用いることで使用バンド・時間情報・空間情報の有効性を確認
 - 熱痕跡が画像全体に出現する場合には, 動画を入力とした時間情報が有効
 - 人が出現するような広い領域を用いた場合には, マルチバンド化された入力がある

● 今後の課題

- ・ 実環境を想定したデータセットを用いて実験
 - さまざまな背景・物体を接触対象に設定
 - 接触せず手を伸ばしたり, カメラ前を横切ったりするフェイクを含んだ動画を撮影
- ・ この知見に基づいた2つのモジュールからなるモデルの検証
 1. マルチバンド画像を入力とした熱痕跡の可能性のある領域を特定 (検出タスク)
 2. 特定領域のサーマル動画をを入力として熱痕跡であるかどうか識別 (分類タスク)



参考文献

12

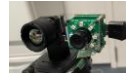
- [1] Q. G. Alexander, V. Hoskere, Y. Narazaki, A. Maxwell, and B. F. Spencer, "Fusion of thermal and RGB images for automated deep learning based crack detection in civil infrastructure," *AI in Civil Engineering*, vol. 1, no. 1, p. 3, 2022.
- [2] S. Spath, A. Gon Calves, B. Rigault, S. Suzuki, M. Bouazzi, Y. Matsuo, and H. Prendinger, "Deep learning with RGB and thermal images onboard a drone for monitoring operations," *Journal of Field Robotics*, vol. 39, pp. 840–868, 2022.
- [3] D. Huo, J. Wang, Y. Qian, and Y.-H. Yang, "Glass segmentation with RGB-thermal image pairs," *arXiv*, 2204.05453, 2022.
- [4] N. van Doremalen, T. Bushmaker, D. H. Morris, M. G. Holbrook, A. Gamble, B. N. Williamson, A. Tamin, J. L. Harcourt, N. J. Thornburg, S. I. Gerber, J. O. Lloyd-Smith, E. de Wit, and V. J. Munster, "Aerosol and surface stability of SARS-CoV-2 as compared with SARS-CoV-1," *New England Journal of Medicine*, vol. 382, no. 16, pp. 1564–1567, 2020.
- [5] G. Kampf, D. Todt, S. Pfander, and E. Steinmann, "Persistence of coronaviruses on inanimate surfaces and their inactivation with biocidal agents," *Journal of Hospital Infection*, vol. 104, no. 3, pp. 246–251, 2020.
- [6] Y. Shirai, Y. Kishino, Y. Yanagisawa, K. Ohara, S. Mizutani, and T. Suyama, "Alertable Surfaces: an actual environment that alert virus attachment (in Japanese)," in *Proceedings of Workshop on Interactive System and Software (WISS 2020)*, 2020.
- [7] Y. Kishino, Y. Shirai, Y. Yanagisawa, H. Sugawara, K. Ohara and S. Mizutani, "Alertable Surfaces: an actual environment that alert virus attachment by recognizing human touch and sanitization (in Japanese)," in *Proceedings of Workshop on Interactive System and Software (WISS 2022)*, 2022.
- [8] Y. Abdelrahman, M. Khamis, S. Schneegass, and F. Alt, "Stay cool! understanding thermal attacks on mobile-based user authentication," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*, 2017, p. 3751–3763.
- [9] A. Sahami Shirazi, Y. Abdelrahman, N. Henze, S. Schneegass, M. Khalilbeigi, and A. Schmidt, "Exploiting thermal reflection for interactive systems," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*, 2014, p. 3483–3492.

付録① 実験概要 | 撮影環境

13

●撮影環境の概要

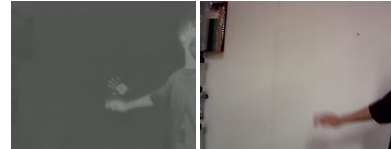
- 熱痕跡可視化システム^[6,7]で用いたカメラを使用
- 撮影されたサーマル動画を間引くことにより RGB カメラとのフレームレートを整合



●接触対象

- 背景情報による識別精度への影響を低減するため白色壁面を対象に設定

Camera	Band	Image size [pixels]	Frame rate [fps]
FLIR Boson 640	NWIR	512 × 640	60
GoPro MCM-320	Visible	480 × 640	30

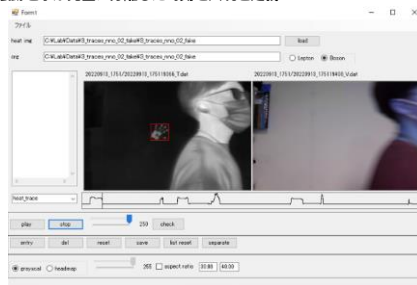


付録② 実験概要 | アノテーション・出現時刻の定義

14

●矩形による熱痕跡のラベル付け

- アノテーションツールを用いて目視で矩形により熱痕跡を囲む
- 熱痕跡と手が完全に分離した時刻を出現と定義

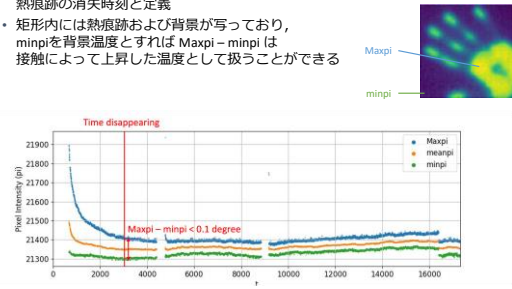


付録③ 実験概要 | 熱痕跡の消失時刻の定義

15

●矩形内の温度差における閾値を用いて消失時刻を定義

- 矩形内の最大温度 (Maxpi) と最小温度 (minpi) の差が0.1度を下回った時刻を熱痕跡の消失時刻と定義
- 矩形内には熱痕跡および背景が写っており, minpiを背景温度とすれば $Maxpi - minpi$ は接触によって上昇した温度として扱うことができる

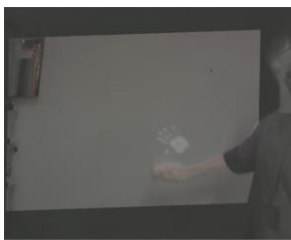


付録④ 実験概要 | キャリブレーション

16

●射影変換による位置合わせ

- センサ間の物理的な距離から発生する位置ずれの整合
- 対応する点を手動で3点以上設定し, 拡大縮小・剪断・回転・上下左右を表現する射影行列を算出
- RGB画像に対して射影行列との積を取ることで画像を変換



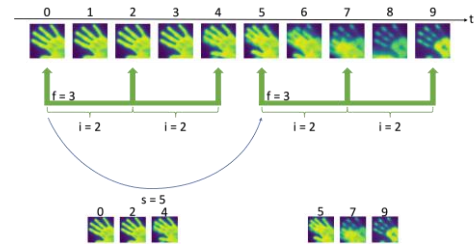
付録⑤ 実験概要 | 時間方向の分割

17

●時間長を固定して時間方向に分割

- 本実験では2Dと3Dの比較を行う
- フレームサイズ f , フレーム間隔 i のウィンドウをスライドサイズ s ずつずらして収集

$f=5, s=5, i=2$ の例



付録⑥ 実験概要 | データセット作成 (2)

18

● 入力データの時間長を指定した動画の分割

- 本実験では 2D と 3D の比較を行うため、入力には画像単体と動画を使用
- フレームサイズ f 、フレーム間隔 i のウィンドウをスライドサイズ s ずつずらして収集
- 本実験では、以下の表のように各パラメータを固定して時間方向に分割

	3D	2D
フレームサイズ f	10	1
スライドサイズ s	5	5
フレーム間隔 i	2	-

付録⑦ 実験概要 | 学習・推論

19

● ネットワークに用いたパラメータ

	2D	3D
チャンネルサイズ	32, 64, 128	64, 128, 256
カーネルサイズ	3x3	3x3x3
プーリングサイズ	2x2	1x2x2

● 最終的に得られたデータセット内容

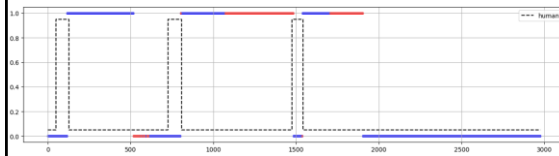
CNN	f	bs	size	Train		Test	
				T	F	T	F
2D	1	256	narrow	22733	10371	1542	1438
			wide	19331	13173	1542	1438
3D	10	256	narrow	22630	10424	1529	1448
			wide	19822	13232	1529	1448

T: heat trace, F: others

付録⑧ | テストデータセットに対する推論結果の例 20

● narrow 3D TV における推論結果

- 縦軸：
 - 0: false
 - 1: True
- 横軸：
 - 時系列順に並べた入力データのindex



熱痕跡は連続的な温度変化のため消失前後で誤りが多く発生する一方、出現は十分に識別ができています