

# 可視画像・熱画像対に基づく屋外シーンの影の識別

藤田 和也<sup>1,a)</sup> 川原 僚<sup>2,b)</sup> 岡部 孝弘<sup>2,c)</sup>

**概要：**シーン中に存在する影は、建物などに起因する継続的な影と撮影者などに起因する一時的な影の2種類に分類できる。このうち、撮影者などに起因する一時的な影は、撮影前のシーンには存在しない余計な影であるため、検出・除去することが望まれる。しかしながら、一時的な影と継続的な影は明るさも色も同一であるため、従来手法では識別することができない。そこで本稿では、これら二つの影領域では温度が異なることに着目して、可視画像に加えて熱画像を用いることで、一時的な影と継続的な影を識別する手法を提案する。具体的には、可視画像から検出した影領域を、温度情報、および、領域分割に基づいて、一時的な影と継続的な影に識別する。屋外シーンで撮影した可視画像・熱画像対を用いた実験を行い、提案手法の有効性を示す。

**キーワード：**影の識別, 熱画像, 領域分割, グラフカット

## 1. はじめに

一般に、人や建物などの物体により光源からの光が遮られることによって影が生じる。それらは、人や移動物体によってできる一時的な影と、建物や木などによってできる継続的な影の2種類に分類できる。写真を撮影する際に、撮影者などの一時的な影がシーンに写りこんでしまう場合がある。このような影は、撮影前のシーンには存在していなかった余計な影であるため、画像中から検出・除去することが望まれる。そこで、本研究では、太陽光下で撮影された画像から、一時的な影を検出する手法を提案する。

画像中の影は物体検出やセグメンテーションなどの画像処理において精度を著しく低下させる場合がある [1], [2]。また、検出した影のマスクを利用することでオブジェクトの再照明 [3], カメラ位置の推定 [4] などシーンを理解する重要な手がかりとなる。そのため、影の検出・除去は画像処理分野において重要な課題の1つである。影を検出・除去する手法は数多く存在し、これまでに色 [5], [6], [7] や明るさ [8], [9], [10] などに基づいた手法が提案されている。また、近年では機械学習を用いた手法 [11], [12], [13], [14] が高い精度を記録している。しかし、一時的な影と継続的な影は色や明るさで区別できないため、これらの手法では一時的な影の検出を行うことは困難である。また、補助情

報として深度を加えた RGB-D 画像を用いる手法 [15] も提案されている。深度情報により、複数物体にまたがっている影や半影領域に対して頑健な推定を行うことが出来るが、色や明るさと同様に深度情報では一時的な影と継続的な影の識別を行うことはできない。

そこで、本稿では可視画像・熱画像対を用いることで一時的な影と継続的な影の識別を行う手法を提案する。提案手法では、建造物などに起因する継続的な影領域は影のない領域と比較して温度が低くなるのに対して、撮影者などに起因する一時的な影領域は温度低下が起きていないため、影のない領域と同じ振舞いをすることに着目する。具体的には、既存の機械学習ベースの手法で検出した影領域を、温度情報に基づいて一時的な影、継続的な影、一時的な影と継続的な影が隣接する場合に分類する。一時的な影と継続的な影の両方を含む影領域については、領域分割を行うことで一時的な影のみの検出を実現する。実画像を用いた実験を行い、提案手法の有効性を示す。特に、提案手法により、可視画像のみでは不可能な継続的な影と一時的な影の識別が可能となることを示す。

## 2. 関連研究

### 2.1 影の検出・除去

影を検出・除去する手法は数多く存在し、従来は色や明るさに基づく手法が提案されている。Gong ら [5] は手入力での影領域と非影領域の初期値を与えることで、初期値の色情報に基づく影レイヤーの推定を行う手法を提案している。また、Finlayson ら [8], [9] は明るさを除去した色度画

<sup>1</sup> 九州工業大学 大学院情報工学府 先端情報工学専攻

<sup>2</sup> 九州工業大学 大学院情報工学研究院 知能情報工学研究系

a) k.fujita@pluto.ai.kyutech.ac.jp

b) kawahara@vision.ist.i.kyoto-u.ac.jp

c) okabe@ai.kyutech.ac.jp

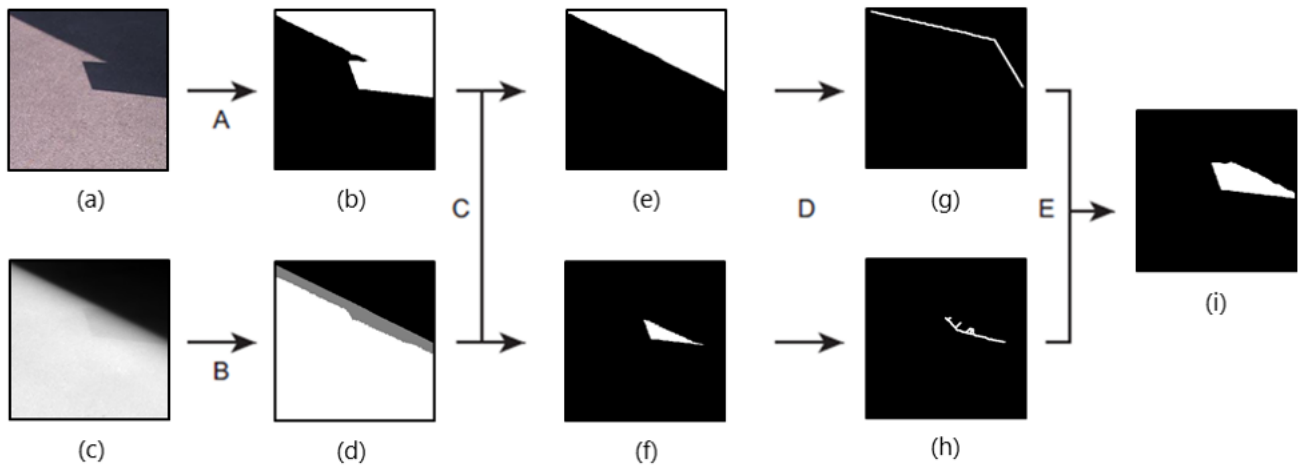


図1 提案手法の流れ: (a) 可視画像, (b) 機械学習手法で検出した影領域, (c) 温度情報の CSV データ, (d) 温度の3値化画像, (e) 背景の候補領域, (f) 前景の候補領域, (g) 細線化した背景領域, (h) 細線化した前景領域, (i) 一時的な影の識別結果

像を作成し、勾配領域でエッジ比較をすることで影の除去を行う手法を提案している。

これらの単一の RGB 画像を用いる手法では、複数光源に物体が照らされている場合や影が画像中に多く存在する場合などの複雑なシーンに対して頑健な影の除去は困難である。Xiao ら [15] は、深度情報を追加した RGB-D 画像を用いて、類似の法線、空間位置、色度を持つ画素は同じ色を持つという仮定の下で頑健に影の除去を行う手法を提案している。

このような手法は複数光源下など複雑な画像では機能しない、手入力が必要であるといったように、手法に制限があることが多い。そこで近年は自動で影の検出・除去を行う学習ベースの手法が多く提案されている。Khan ら [11] は最初に畳み込みネットワークを用いた影の検出手法を提案し、従来の色や明るさに基づく手法と比べて高い影領域の検出精度を示している。また、Nguyen ら [12] は感度パラメータを導入した条件付き敵対的生成ネットワーク scGAN を用いることで、多くの画像に適応可能な影の検出手法を提案している。Nguyen らの手法は影の検出のみであったが、影の検出と除去を同時に学習し実行する GAN も提案されており、Wang ら [13] は影の検出を行う GAN と影の除去を行う GAN を組み合わせた構造をもつ ST-CGAN を提案している。また、Ding ら [14] は ST-CGAN に LSTM を用いて再帰的に影の検出と除去を行うことで、より高精度な影の検出・除去精度を示している。

しかし、一時的な影と継続的な影は色や明るさ、深度では区別できないため、上記に示したような従来手法では一時的な影と継続的な影を識別することは困難である。そこで、提案手法では、同じ材質上において、継続的な影領域と一時的な影領域の温度が異なることに着目し、シーンの温度情報を画素単位で保持する熱画像を補助情報として用

いることで、一時的な影と継続的な影の識別を可能とする。

## 2.2 熱画像を用いた画像処理

熱画像は、物体の反射光を捉える可視画像とは異なり、物体が放射する遠赤外線を二次元的に測定したものであり、シーンの温度情報を画素単位で保持している。そのため、温度情報が重要な手がかりとなる非破壊検査や物体検出などに活用されている。Davis ら [16] は熱画像から初期の関心領域を求め、勾配領域で可視画像と熱画像の関心領域内のエッジを合成することによって従来の背景差分法よりも頑健に物体検出を行う手法を提案している。これにより、昼夜問わず監視カメラ映像から人の検出を行うことが可能である。また、Bulanon ら [17] は、可視画像と熱画像の情報を統合することで、時間帯による果物の温度変化に着目した、自動収穫システムのための頑健な果物の検出手法を提案している。さらに、機械学習を用いた手法にも熱画像を用いる研究は存在するが、公開されているデータの少なさが問題としてあげられる。そこで、Zhang ら [18] は、通常の可視画像のみを用いて学習を行った顔認識手法に熱画像を適用できるように、顔を撮影した熱画像を対応する可視画像に変換する敵対生成ネットワーク (TV-GAN) を提案している。このように、可視画像だけでは物体検出が困難な状況での熱画像の有効性が示されている。

本研究では、可視画像・熱画像対を用いた画像処理技術の新たな応用として、可視画像のみでは識別困難な、一時的な影と継続的な影の識別手法を提案する。

## 3. 提案手法

提案手法の概略を図1に示す。本手法では、従来の可視画像のみを用いた影の検出手法により得た影領域を領域ごとに温度情報に基づき、一時的な影、継続的な影、一時的

表 1 提案手法の基本的な考え方

		熱画像	
		高い	低い
可視画像	明るい	一時的な影 (前景)	継続的な影 (背景)
	暗い	非影	非影

な影と継続的な影の両方を含む領域の 3 パターンに識別する。

提案手法では、可視画像で影でありかつ熱画像で温度が高い領域は前景、可視画像で影でありかつ熱画像で温度が低い領域を背景として、手入力を必要とせず自動で初期ラベルを求める。また、この初期領域と従来の可視画像のみを用いた影検出の結果領域毎に共通部分を取得することで影の識別を行う。影領域が一時的な影の初期ラベルのみを含む場合は、従来の影の検出手法による検出結果をそのまま用いるが、単一の領域に両方の影が含まれる場合はグラフカットの枠組みで領域分割を行うことで一時的な影のみを検出する。以下では、初期ラベルの取得方法と、一時的な影と継続的な影の識別方法について、具体的な処理を説明する。

### 3.1 前景・背景の候補領域の取得

可視画像と温度情報を組み合わせることで、前景・背景の候補領域の取得を行う。まず、可視画像は影の検出を、熱画像は高温領域、低温領域の取得を行う。可視画像は既存手法により影の検出を行う (図 1A)。今回、影の検出には ST-CGAN を [13] 用いた。熱画像の温度分布は非影領域、真影領域、半影領域の 3 つのクラスに分類できると考えることができるため、3 値化を行うことで、温度の低い領域、高い領域を取得する (図 1B)。今回、3 値化には大津の閾値を用いた [19]。

次に、影の検出結果 (図 1(b)) と 3 値化した熱画像 (図 1(d)) から、一時的な影を前景、継続的な影を背景として、候補領域の取得を行う (図 1C)。表 1 の条件に従い、前景は図 1(b) で影である領域のうち、図 1(d) で温度が高い領域、背景は図 1(b) で影である領域のうち、図 1(d) で温度が低い領域の共通部分から候補領域を取得する。境界部分の信頼性が低い画素を排除するために初期ラベルには前景・背景の候補領域を細線化したものを用いる (図 1D)。

### 3.2 影の識別

影の検出結果 (図 1(b)) と前景・背景の候補領域 (図 1(g),(h)) の共通領域を取得することで影領域ごとに影を、一時的な影、継続的な影、一時的な影と継続的な影の両方を含む領域の 3 パターンに識別する (図 1E)。具体的には、前景の候補領域との共通部分のみ存在する領域は一時的な影、背景の候補領域との共通部分のみ存在する領域は継続的な影、両方との共通領域が存在する領域は一時的な影と

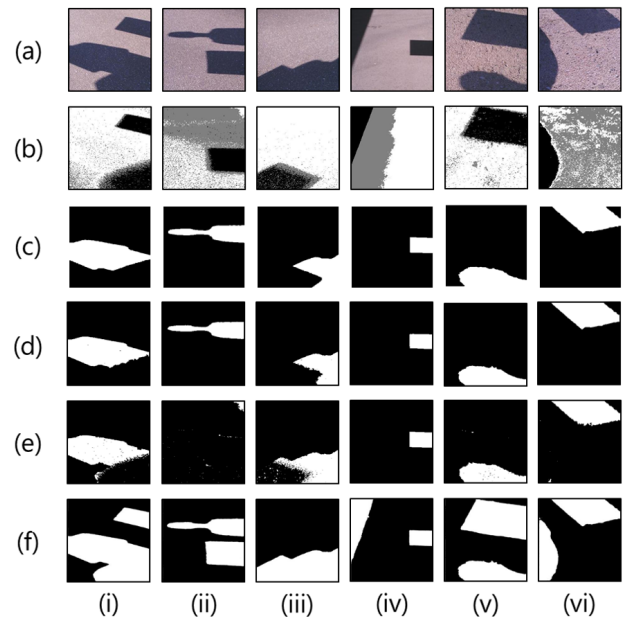


図 2 実験結果: (a) 可視画像, (b) 熱画像, (c) 真値, (d) 提案手法の実験結果, (e) 我々の従来手法 [21] の実験結果, (f) ST-CGAN [13] の検出結果, (i), (ii), (iii), (iv) はアスファルト, (v), (vi) は砂のシーン

表 2 適合率と再現率

		(i)	(ii)	(iii)	(iv)	(v)	(vi)	平均
(d)	適合率	0.997	0.999	0.990	1.000	0.992	0.999	0.996
	再現率	0.971	0.958	0.937	0.938	0.981	0.977	0.960
(e)	適合率	0.967	0.933	0.717	1.000	0.995	0.992	0.934
	再現率	0.962	0.502	0.986	0.983	0.950	0.963	0.891

継続的な影が隣接する場合である。最後に 3 パターンに識別した結果を用いて、画像中の一時的な影の取得を行う。一時的な影のみの場合、影の検出結果の領域をそのまま一時的な影の検出結果として用いる。単一の領域に両方の影が含まれる場合は、影の検出結果の領域をそのまま利用することができないため、可視画像と熱画像を合わせた 4 バンドの画像と図 1(g),(h) を用いて、領域内に対してグラフカットの枠組みで領域分割を行うことで、図 1(i) に示すような一時的な影領域のみを取得する。今回、領域分割手法には LazySnapping [20] を用いた。

## 4. 実験

### 4.1 一時的な影の検出

提案手法の有効性を示すために撮影者等の影が写り込んだ実画像を用いて実験を行った。今回は、晴れた屋外で一般的なシーンに落とされる影を撮影した。実験に用いた FLIR 社の T510 は、解像度  $240 \times 320$  の可視画像と、その画像の画素毎に温度情報を記録した CSV データを出力することが可能なカメラである。影の検出手法に用いた ST-CGAN [13] は  $256 \times 256$  の RGB 画像を入力とするため、バイリニア補間を行い拡大した後にトリミングを行っ

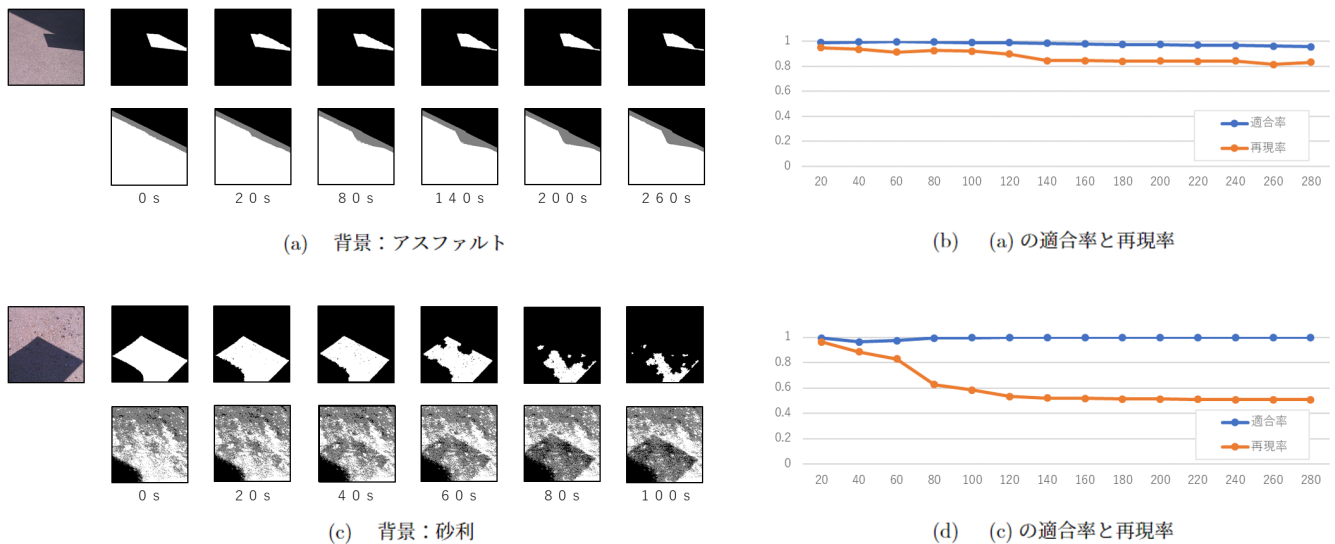


図 3 時系列データの実験結果

た画像を実験に用いた。

可視画像の影の検出結果を図 1(b) に、大津の閾値 [19] により 3 値化を行った結果を図 1(d) に示す。これらの結果を用いて前景・背景の候補領域を取得した結果を図 1(e),(f) に、細線化した結果を図 1(g), (h) に示す。最後に、図 1(a), (c), (b), (g), (h) を用いて影の識別を行った結果を図 1(i) に示す。図 1(a), (i) を比較すると、一時的な影と継続的な影を正しく識別できていることが確認できる。

また、その他の画像に対して実験した結果を図 2 に示す。図 2(e) に用いた我々の従来手法は、今回の提案手法と同様に図 1(g), (h) を取得して、可視画像と熱画像を合わせた 4 バンドの画像に対し、画像全体で領域分割を行うものである。また、適合率と再現率を今回の提案手法 (図 2(d)) と我々の従来手法 [21] (図 2(e)) に対し求めた結果を表 2 に示す。図 2(c) と図 2(d) を比較すると、どの画像に対しても正しく一時的な影のみが検出できていることが確認できる。

また、図 2(d) と図 2(e) を比較すると、画像 (i), (ii), (iii) に対しては我々の従来手法よりも今回の提案手法の方が正しく一時的な影の検出ができていることが確認できる。また、画像 (iv), (v), (vi) は我々の従来手法と同程度の検出結果を得られていることが確認できる。表 2 から、同様の結果が得られていることが確認できる。これらの結果から、今回の提案手法が定性的、定量的に良好な結果が得られていることが確認できる。

#### 4.2 時間経過による一時的な影の検出精度

本手法では、一時的な影の検出に熱画像を用いている。そのため、撮影までの間にある程度時間が経過すると、一

時的な影領域の温度が低下し、継続的な影との識別が困難になると考えられる。そこで、5 分間に渡り 20 秒間隔で撮影した画像群を用いて、提案手法により一時的な影領域の検出を行い、時間経過による検出精度変化の検証実験を行った。真値画像との比較により適合率と再現率を求めることで時間経過による精度変化を求める。背景がアスファルトの場合と砂利の場合で、一時的な影と継続的な影が隣接している場合に対し実験を検出精度の比較を行う。前節の実験と同様に、撮影画像に対しバイリニア補間により拡大を行いトリミングを行った画像を実験に用いた。

背景がアスファルトの場合の実験結果を図 3(a) に、その時の適合率と再現率の変化を表 3(b) に示す。これらの結果から、120 秒経過後から再現率が 85 % 程度に減少したものの、それ以降も 300 秒経過時まで 80 % 以上を維持し、適合率はほとんど変化がないことがわかる。図 3(a) から、境界部分の一時的な影が多少削られてはいるものの概ね正しく検出ができていることが確認できる。これらの結果から、背景素材がアスファルトのとき、2 分以内であれば定性的にも定量的にも良好な結果を得ることが確認できた。

また、背景が砂利の場合の実験結果を図 3(c) に、その時の適合率と再現率の変化を表 3(d) に示す。表 3(d) から、60 秒経過時点で 80 %、80 秒経過時点で 60 % へと再現率が大きく低下していることが確認できる。図 3(c) から同様に、60 秒経過時点で一時的な影の一部が除去されており、80 秒経過時点では一時的な影の大半が除去されていることが確認できる。これは、一時的な影と継続的な影が隣接していない場合と同様に、一時的な影内の温度が低下していくことで、その部分も継続的な影として領域分割時に選択されてしまい除去されてしまったと考えられる。



これらの結果から、背景素材が砂利のとき、1分以内であれば定性的にも定量的にも良好な結果を得ることが確認できた。

アスファルト、砂利に対して実験を行った結果から、5分程度の時間経過に伴う温度変化では、適合率はあまり変化せず、再現率は時間経過に伴い下がる傾向にあり、その影響は背景素材に依存することがわかる。

## 5. おわりに

本稿では、可視画像と熱画像に基づく一時的な影と継続的な影の識別手法を提案した。提案手法では、継続的な影領域と一時的な影領域では温度が異なることに着目し、可視画像から既存の学習ベースの手法で検出した影領域を領域毎に、継続的な影、一時的な影、一時的な影と継続的な影が隣接している場合の3パターンに識別する。隣接している場合では影領域内に対して領域分割を行うことにより一時的な影のみの検出を行った。実画像を用いた実験を行い、提案手法により一時的な影のみの検出が可能となることを確認した。また、時間経過に伴う一時的な影の検出精度変化について検証実験を行った。

今回の提案手法は、晴れた屋外の背景画像が一様なシーンを対象としているため、複雑なシーンには対応することができない。これは、温度情報を全体で3値化し、温度の高い部分と低い部分を取得するという処理を行っているため、素材毎の温度の違いを考慮できないためである。そのため、より自然なシーンに対応するために、可視画像から背景素材を判別しそれを考慮した識別を行えるようにすることが今後の課題である。

**謝辞** 本研究の一部は、JSPS 科研費 JP20H00612 の助成を受けた。

## 参考文献

- [1] R. Cucchiara, C. Grana, and M. Piccardi, "Improving shadow suppression in moving object detection with hsv color information", In Proc. ITSC2001, pp.334–339, 2001.
- [2] I. Mikic, P. C. Cosman, and G. Kogut, "Moving shadow and object detection in traffic scenes", In Proc. ICPR2000, Vol.1, pp.321–324, 2000.
- [3] K. Karsch, K. Sunkavalli, and S. Hadap, "Automatic scene inference for 3d object compositing", ACM Transactions on Graphics, Vol.33, Issue.3, pp.1–15 2014.
- [4] I. Junejo, and H. Foroosh, "Estimating geotemporal location of stationary cameras using shadow trajectories", In Proc. ECCV2008, pp. 318–331, 2008.
- [5] H. Gong, and D. Cosker, "Interactive Shadow Removal and Ground Truth for Variable Scene Categories", In Proc. BMVC2014, BMVA Press, 2014.
- [6] X. Yu, G. Li, Z. Ying, and X. Guo, "A New Shadow Removal Method Using Color-Lines", In Proc. CAIP2017, pp.307–319, 2017.
- [7] X.Hu, L.Zhu, C.Fu, J.Qin, P.Heng, "Direction-aware Spatial Context Features for Shadow Detection" In Proc. CVPR2018, pp.7454–7462, 2018.
- [8] G. Finlayson, S. Hordley, M. Drew, and C. Lu, "On the Removal of Shadows from Images", IEEE Trans. PAMI, Vol.28, No.1, pp.59–68, 2006.
- [9] G. Finlayson, M. Drew, and C. Lu, "Entropy Minimization for Shadow Removal", IJCV, Vol.85, No.1, pp.35–57, 2009.
- [10] R. Guo, Q. Dai, and D. Hoiem, "Single-image Shadow Detection and Removal Using Paired Regions", In Proc. CVPR2011, pp.2030–2040, 2011.
- [11] S. Khan, M. Bennamoun, and F. Sohel, "Automatic shadow detection and removal from a single image", IEEE Trans. PAMI, Vol.38, Issue, 3, pp.431–446, 2016.
- [12] V. Nguyen, T. Vicente, and M. Zhao, "Shadow detection with conditional generative adversarial networks", In Proc. ICCV2017, pp. 4510–4518, 2017.
- [13] J. Wang, X. Li, and J. Yang, "Stacked conditional generative adversarial networks for jointly learning shadow detection and shadow removal" In Proc. CVPR2018, pp.1788–1797, 2018
- [14] B. Ding, C. Long, and L. Zhang, "ARGAN: attentive recurrent generative adversarial network for shadow detection and removal" In Proc. ICCV2019, pp. 10212–10221, 2019.
- [15] Y. Xiao, E. Tsougenis, and C. Tang, "Shadow Removal from Single RGB-D images", In Proc. CVPR2014, pp.3011–3018, 2014.
- [16] J. Davis, and V. Sharmam, "Background-subtraction Using Contour-based Fusion of Thermal and Visible Imagery", CVIU, Vol.106, No.2-3, pp.162–182, 2007.
- [17] D. Bulanon, T. Burks, and V. Alchanatis, "Image Fusion of Visible and Thermal Image for Fruit Detection", Biosystems Engineering, Vol. 103, No.1, pp.12–22, 2009.
- [18] T. Zhang, A. Wiliem, S. Yang, and B. Lovell. "TV-GAN: Generative adversarial network based thermal to visible face recognition.", In Proc. ICB, pp.174–181, 2018.
- [19] N.Otsu, "A Threshold Selection Method from Gray-Level Histograms", IEEE Trans Systems, Man, and Cybernetics. Vol.9, Issue 1, pp.62–66, 1979.
- [20] Y. Li, J. Sun, C.-K. Tang, and H.-Y. Shum, "Lazy Snapping", ACM Trans. Graphics, Vol.23, No.3, pp.303–308, 2004.
- [21] K. Fujita, R. Matsuoka, and T. Okabe, "Detecting Momentary Shadows from Visible and Thermal Image Pair", In Proc. European Signal Processing Conference, pp.606–610, 2021.