

室内栽培システムにおける強化学習を用いた栽培手法の提案

高田 裕也^{1,a)} 中沢 実¹ 西川 幸延¹

概要: 現在の農業は天候やその他の環境変化の影響を受けやすく、就業人口と基幹的農業従事者が共に年々減少している等の問題がある。この問題の解決案として植物工場や室内栽培システムの運用がある。これは収支の点から大量生産するために規模を大きくするなど、効率の良い栽培条件で栽培する必要がある。そのため本論文では、室内で栽培でき、効率的な栽培条件の探索をするシステムの提案を行う。栽培条件の探索に関して機械学習手法が挙げられるが、大量のデータセットを必要とする、対象とする作物のデータがあるとは限らない等の問題がある。また栽培環境を変化させてから作物に影響が反映されるまでの時間は明確になっていない。そこで遅延報酬の概念がある強化学習手法を用いることによって効率的な栽培条件の探索を目指す。

1. はじめに

現在の農業は天候等の環境の変化に影響を受けやすい、基幹的労働従事者が年々減少している [1] 等の問題がある。この問題の解決案の一つとして植物工場の運用がある。しかし植物工場の問題点として、設備を設置する必要があるため、ある程度広い空間を確保しなければならない。さらに、収支の点から大量生産するために規模を大きくし、効率の良い栽培条件で栽培する必要がある。そこで本研究では省スペースで運用可能かつ効率の良い栽培条件を生成する自動水耕栽培システムの提案を行う。栽培条件の生成に関して大量のデータセットを用いる機械学習手法が挙げられる。しかし対象とする作物のデータがあるとは限らない等の問題がある。また栽培環境を変化させてから作物に影響が反映されるまでの時間は明確になっていない。そこで遅延報酬の概念がある強化学習手法を用いることによって効率的な栽培条件の生成を目指す。これによって逐次制御、少サンプル環境下での実行、栽培条件を変化させてから状態としての評価値が出るまでの不明確なタイムラグを考慮した栽培が可能であると考えられる。また「作物の状態を評価し強化学習に用いて栽培条件を設定する」ことによって利用者が農業に関する知識・経験がない人でも栽培することが可能である。

本論文では大規模なデータセットが無くとも栽培条件を探索する自動水耕栽培システムの提案を行う。2章では関連研究を挙げ、3章ではシステム全体構成、4章では提案手

法、5章で実験、6章で本論文のまとめを挙げる。

2. 関連研究

2.1 Personal Food Computer[2]

Personal Food Computer(PFC)はMITのOpenAgricultureプロジェクトで開発された自動水耕栽培システムである。PFCのプログラムや回路図、作成手順、栽培レシピはオープンソースとして公開されている。センサはpHセンサ、ElectricalConductivity(EC)センサ、温湿度センサ、CO₂センサ、水位センサを搭載し、光量、温度、pH、培養液の養分の濃度を制御できる。またセンサ、アクチュエータの追加や用意できないセンサを疑似センサとして認識させる機能がある。栽培条件の自動生成機能はなく、あらかじめ設定した栽培条件にしたがって栽培を行う。

2.2 農作物の自動灌水制御に向けたニューラルネットワークを用いたQ学習[3]

難波らによる研究では小松菜の栽培において灌水行動を最適化する。植物の状態は時間変化することから同じ状態は存在せず全ての状態に対してQテーブルを作成することが困難である。また植物の状態は連続量であるため状態をQテーブルで表すことが困難である。これらの問題を解決するために、Qテーブルを関数近似するQ学習を採用している。またメモリにエージェントが経験したサンプルの内容を保存し、内容をランダムに抽出して再利用するExperience Replayという手法を用いて学習の効率化とデータ数増加を図っている。1試行は2週間、ステップ数は7ステップ、報酬は時刻差分時の植物の背丈、行動は灌水行動4パターン、行動を起こした時から次の状態に遷移する

¹ 金沢工業大学
7-1 Ogaigaoka Nonoichi Ishikawa 921-8812, Japan

^{a)} b6800635@planet.kanazawa-it.ac.jp

ときは2日後としている。実験ではQ学習での栽培は植物の背丈や葉の広がりを入力値としているが、人の手による栽培はロックワールの乾き具合等Q学習の入力値以外の要素を考慮して灌水している。そのため人の手による栽培から得られたデータは現時点での最良解に過ぎず、強化学習によって学習した結果をもとに実際に栽培を繰り返すことでエージェントが選択する行動の良し悪しを評価することが必要であると論じている。

2.3 完全制御型植物工場における最適制御則探索のためのベイズ最適化を用いる手法の一検討 [4]

簡易型のPFCにおいて栽培手法をベイズ最適化を用いて探索する研究である。葉面積の即日の成長率が過去の状態や栽培手法に依存せずに、現時点での状態と栽培手法によって決定されるというモデルを仮定し、ベイズ最適化を用いる手法の提案である。変化させる条件は概日日射時間である。それぞれの条件を確率的に当たりが出るスロットマシン(バンディット)問題として捉える。報酬は1日あたりの葉面積の増大値である。1ステップは1日であり、6日に渡り条件を変化させる。実験の結果、即日の成長率は以前の成長率に影響を受けるような結果になった。これにより任意の時点の作物の状態は過去の環境に影響を受けることが示された。

3. システム構成

本システムは実際に作物を栽培する栽培システムと栽培条件を設定するサーバの2つから成り立つ。システム全体構成図を図1に示す。栽培システムは撮影した画像ファイルをサーバに送信する。サーバは受信した画像から作物の活性度を算出した後、栽培アルゴリズムにより次に設定する栽培手法を決定し栽培システムに送信する。栽培システムの説明を本節で行い、サーバで実行される処理についての説明を4節で行う。

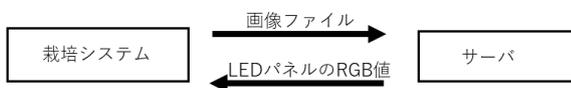


図1 システム全体構成図
 Fig. 1 System configurations

栽培システムはMITのOpenAgricultureプロジェクトが開発しているPFCを参考にして製作した。製作した栽培システムを図2栽培システム構成図を図3に示す。

システムに搭載されるセンサ群は温湿度センサ [5], 水温センサ [6], CO2センサ [7], 水位センサ [8], pHセンサ [9], ECセンサ [10] である。またカメラは可視光タイプの pi-camera と NoIR タイプの picamera を搭載している。アクチュエータ群はシステム内の温度制御用として冷暖房ペ

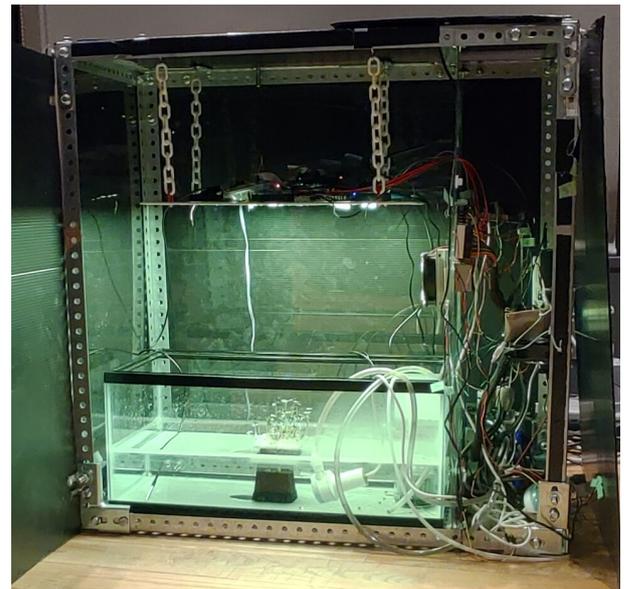


図2 製作した栽培システム
 Fig. 2 Incubator system

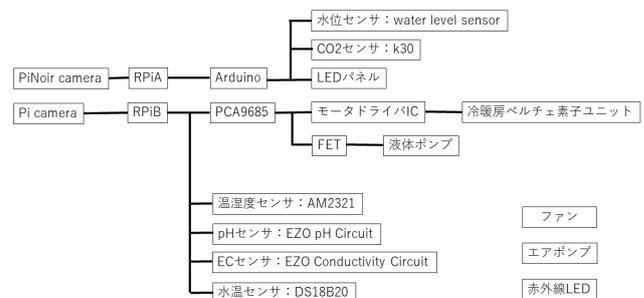


図3 栽培システム構成図
 Fig. 3 Incubator system configurations

ルチェ素子ユニット、水槽内へ液体を投入するための液体ポンプを搭載している。光量制御として8×8LEDパネル [11] を8個搭載している。システムは光量を制御するため黒色プラスチック段ボールにて覆っている。システム全体への電力はATX電源を用いて電源供給を行う。水位センサ, CO2センサ, LEDパネル制御をArduinoにて、システム内のそのほかのセンサ群, アクチュエータ群, カメラの制御は2つのRaspberryPiにて行う。センサ群とアクチュエータ群の制御をArduinoで行っている点について、RaspberryPiのGPIOポートは入力にはデジタル値のみでありアナログ値はサポートしていない。CO2センサはI2CまたはUARTで通信するが、RaspberryPiではうまく読み取れなかったためArduinoで読み取る。

システムは30分間隔でセンサ群からデータを取得し、温度と水量を調節する。またLEDパネルを白色にした後、画像を撮影しサーバに送信する。2台のRaspberryPiは同じタイミングで撮影する必要があるため、crontabによる定期実行によって同期をとる。

RaspberryPiで制御せずに常時稼働しているアクチュ

エータとして、ファン、エアポンプ、赤外線 LED[12] がある。ファンはモータドライバ IC と冷暖房ペルチェ素子ユニットの冷却のため、エアポンプは水槽内の液体の溶存酸素量増加のため、赤外線 LED は赤外線画像を撮影するために搭載する。

4. 提案手法

本節ではサーバで実行される処理についての説明を行う。栽培システムから受信した画像から作物の評価をする画像処理手法の説明を 4.1 節、評価値を元に栽培条件を生成するアルゴリズムの説明を 4.2 節で行う。

4.1 画像処理手法

作物の評価は Normalized Difference Vegetation Index (NDVI) を用いる [13]。NDVI とは植物による光の反射の特徴を生かし植物の量や活性を表す植生指標の一つである。高解像度地球観測衛星データを用いて森林の活動状況や干ばつ等のリモートセンシング [14] に使われるが、デジタルカメラを用いて近距離で撮影しても植物指標を算出することができる [15]。また植物がストレスを受けている状態では近赤外線の反射量が下がり、健康な状態では反射量が多くなる [16]。

NDVI 値は (1) 式によって画素ごとに求められる。NIR は赤外線カメラ画像の RGB 値の内の R 値であり、R は通常のカメラ画像の RGB 値の内の R 値である。NDVI 値のとりうる範囲は $-1 \sim +1$ であり $+1$ に近づくほど活性が高く健康であると言える。

$$NDVI = \frac{NIR - R}{NIR + R} \quad (1)$$

またスコア化しやすいように (2) 式の処理を行う [13]。これにより値の範囲が $-1 \sim +1$ から $0 \sim 200$ になる。本手法ではさらにスコア化しやすいように整数型にキャストする処理を行う。

$$NDVI' = (NDVI + 1.0) * 100 \quad (2)$$

ここで、NDVI 値が増大する要因は成長による葉面積の増加と、活性が悪い葉が良くなった場合である。NDVI 値が大きいほどスコア値が高く、また別時点で撮影した画像の NDVI 値と比較しやすいように (3) 式で計算する。 i が NDVI' 値、 $count(NDVI'_i)$ が NDVI' 値 i である総画素数である。

$$NDVI_{sum} = \sum_{0 \leq i \leq 200} i * count(NDVI'_i) \quad (3)$$

栽培システムでは、NDVI 画像を生成するために piCamera と piNoirCamera を用いる。ホワイトバランスの設定を

「自動」にして撮影すると、色合いが撮影するたびに変動する。これを解決するために、「自動」かつ白色下で複数回撮影した時のホワイトバランスゲインの最頻度値を元に手動で設定する。2 台の ISO 感度・シャッタースピード等の条件は同じである。画像処理を行う前の piCamera での撮影画像を図 4 に示す。2 枚の画像から NDVI 値を算出するために以下の処理を行う。

- (i). 2 枚の画像の特徴点の算出とマッチング
- (ii). 画像から栽培している領域を切り出す
- (iii). (1),(2) 式により画素毎に NDVI 値算出
- (iv). 整数型へキャスト
- (v). カラーマスクにより葉の領域を抽出
- (vi). (3) 式により NDVI スコアを算出

(1),(2) 式によって NDVI 値を算出するために、同一地点の画素を参照する必要がある。しかし、本システムはカメラを撮影面に対して平行にそれぞれ設置しているため、同一の撮影場所で撮影されたものではない。この状態で処理を行うと対象とする作物の NDVI 値が正しく算出できない。これを解決するために 2 つの画像の特徴点を算出し特徴点マッチングを行うことによって、対象画素の合致を行う。また水面の反射等によって対象とする作物以外の場所も葉の活性があると判断されるため、作物部分の抽出を行う必要がある。画像から対象とする作物の部分だけを抽出するために、切り出しとカラーマスクによって作物の部分だけを抽出する。

領域抽出した後 NDVI' 値を算出すると、マスクによって 0 になった画素が (2) 式によって NDVI' 値が 100 になってしまうためカラーマスクの処理を最後に行う。図 4 から栽培している領域を切り出した画像を図 5 に示す。また画素毎に NDVI' 値を求めた後にカラーマスクを行った画像を図 6 に示す。

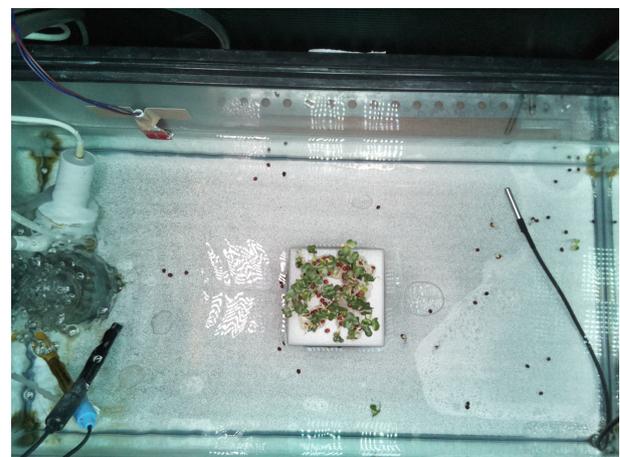


図 4 piCamera 撮影画像
Fig. 4 piCamera image



図 5 栽培領域抽出画像

Fig. 5 Cultivation area extraction image

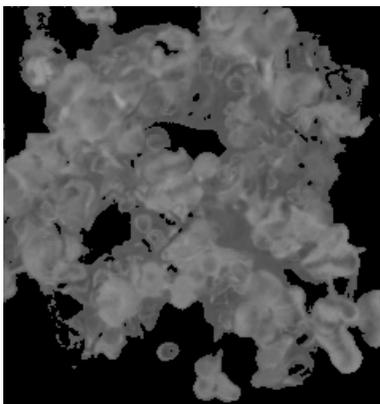


図 6 NDVI グレースケール画像

Fig. 6 NDVI grayscale image

4.2 栽培アルゴリズム

栽培アルゴリズムは強化学習手法の Q 学習を採用する。Q 学習とはある環境状態で行動を選択する価値を学習する手法である [17]。行動主体（エージェント）が置かれている状況（状態）とエージェントが環境に対して行うはたらきかけの種類（行動）の組に対して推定価値を (4) 式にて更新する。S は状態、A は行動、R は報酬値、 α が学習率、 γ が時間割引率である。はたらきかける対象（環境）が未知である場合でも計算することができる特徴がある。また価値更新式と行動を決定するための方策は異なる。

$$Q(S_t, A_t) \leftarrow (1 - \alpha) Q(S_t, A_t) + \alpha(R_{t+1} + \gamma \max_{a' \in A(s')} Q(S_{t+1}, a')) \quad (4)$$

栽培アルゴリズムで生成される栽培条件について、温度や水温とすると設定した値になるまで時間がかかる問題がある。また Q 学習で学習する際、連続値をどこで区切るかという問題が出てくる。そのため本アルゴリズムでの行動 a を色とする。対象とする色は光の三原色とその組み合わせの「黒（消灯）」、「赤」、「青」、「緑」、「赤緑」、「赤青」、「青緑」、「白」である。それぞれ RGB 値で設定し、最大値「255」か最小値「0」とする。

1 試行にかかる時間は作物の植えから収穫までの期間ではなく、栽培システムでのデータ取得間隔である 30 分とすることで、試行時間の短縮と試行回数の増加を行う。

状態の定義について、NDVI 面積や葉面積等の作物由来の状態にすると、同じ状態が出現することが稀であることや行動の問題と同じく連続値をどこで区切るかの問題が出てくるため行動と同じく色について状態 s とする。また試行 t_{t+1} 回時の a_{t+1} を試行 t_{t+2} 回時の s_{t+2} とする。これにより 1 試行分ではあるが直前の試行の行動と現時点での行動のある時刻での植物の状態は過去の環境に影響を受けることを考慮できる。

行動を決定するための方策は ϵ -greedy 方策とする。 ϵ -greedy 方策は確率 ϵ でランダムな行動を、確率 $1 - \epsilon$ で状態 s において一番価値が高い行動を行う。 ϵ の値が高いほどランダムな行動をすることになり、今までの経験をあまり活用せずに探索に重点を置くような行動になる。 ϵ の値が低いほど、あまり探索を行わず今までの経験を活用するような行動になる。ハイパーパラメータである確率 ϵ 、学習率 α 、時間割引率 γ の値は試行回数が多く取れない、学習が早く収束することを考慮してそれぞれ 0.5, 0.9, 0.9 とする。報酬は現試行時の NDVI スコアと 1 試行前の NDVI スコアの差分値とする。

サーバで実行する栽培アルゴリズムの流れを Algorithm1 に示す。Q テーブルの初期状態はすべての要素が 0 である。この状態で Q 学習を実行すると最初に実行した行動を多く選択しがちになり、探索がうまくいかない。これを避けるために Q テーブルを乱数で初期化してから Q 学習を行う。

Algorithm 1: 栽培アルゴリズム実行の流れ

```

行動  $a$  をランダムで決定
Q テーブルを乱数で初期化
栽培システムからの画像を受信し NDVI スコア計算
while True do
  if 栽培システムからの画像を受信 then
1   画像から NDVI スコアと、報酬  $r$  を計算
2   Q テーブル更新
3   状態  $s \leftarrow$  行動  $a$ 
4   行動  $a \leftarrow \epsilon$ -greedy()
5   行動  $a$  を栽培システムに送信
  end
end

```

5. 実験

5.1 実験方法

栽培方法はスポンジを培地とする固形培地耕を用いる。また栽培する作物は収穫期間が 1 週間から 10 日のカイワレ大根を用いる。スポンジ上にカイワレ大根の種を 100 粒

重ならないように蒔き、発芽するまで温度を一定に保ち暗所状態で栽培する。発芽した後、本アルゴリズムを適用して5日栽培する。

5.2 実験結果

試行回数毎の行動と NDVIsum 値を図7に示す。

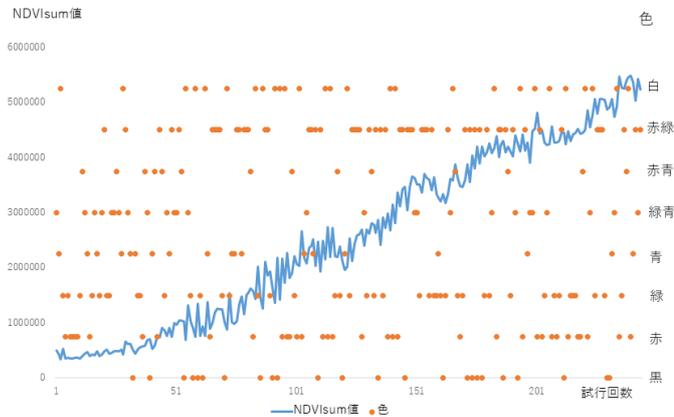


図7 試行回数と行動, NDVIsum 値の関係

Fig. 7 Relationship between number of trials, actions and NDVIsum value

試行回数が0から50回あたりまでは行動が全行動がまんべんなく選択されている。試行回数が60を超えたあたりから赤緑が選択される回数が増え、逆に青や赤青はほとんど選択されていない。

各行動の選択回数を図8に示す。赤緑が最も選択されている回数が多く、次いで緑が選択されている回数が多かった。栽培終了時のQテーブルの内容を表1に示す。行名が状態 s , 列名が行動 a である。状態 s : 青, 行動 a : 赤緑のときに最も高い値となっている。

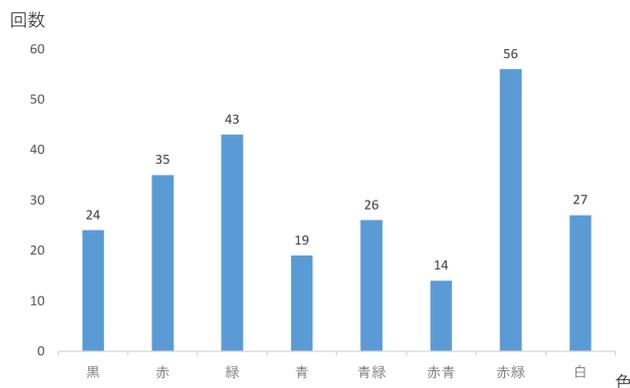


図8 各行動の選択回数

Fig. 8 Number of selections for each action

本手法を適用して栽培した時と白, 赤緑で栽培した時の NDVIsum 値の比較のグラフを図9に示す。本手法適用時と赤緑色栽培時を比較すると、本手法適用時は NDVIsum

値の細かい乱高下が起きているのに対して赤緑色栽培時は変化がなだらかである。また値の推移がほぼ同じである。白色栽培時は全体的に NDVIsum 値が他よりも高くなっている。また試行回数が200回付近で値が減少しているのは、成長したカイワレスプラウトの葉の一部が NDVIsum 値を計算する画像の領域から外れてしまったためである。

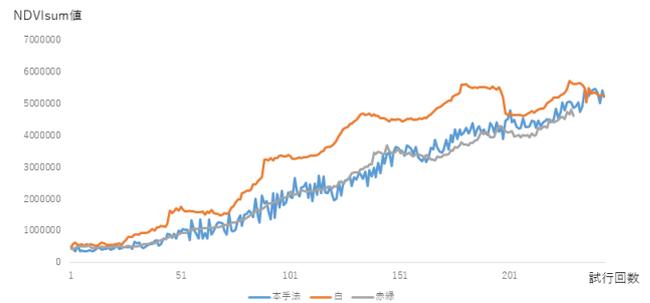


図9 本手法と「白」「赤緑」での栽培結果の比較

Fig. 9 Comparison of cultivation results of "white" and "red green" with this method

6. まとめ

本論文では強化学習の1つであるQ学習を実行しながら作物を自動栽培するシステムの提案を行った。本システムはPFC[2]を参考にして製作しているため、植物工場と比較して栽培条件を制御しやすい。また省スペースでも栽培可能なため、都会の中のデッドスペースでの活用が期待できる。1試行を種植えから収穫までの1回の栽培ではなく、30分区分切りとした時間を1試行とすることで試行回数の増加を図った。さらに、ある試行での行動を次の試行の状態とすることで、1試行前だけが影響を考慮できる。作物が環境の変化を受けてからセンサやカメラに反映されるまでの時間は不明であったが、本実験より1試行時間を30分とおいても、栽培時の色の変化における作物の反応がとれることがわかった。本実験では色を変化させてNDVI画像による作物の活性を観測したが、温度や湿度、二酸化炭素濃度等の変化では30分より長い時間をかけて作物の活性状況が反映される可能性がある。図7では赤緑の選択が多かったものの、 ϵ -greedy 方策で $\epsilon = 0.5$ としているため、結果が収束しているとは言えない。また図9では栽培光を赤緑で栽培するより白で栽培する方が良いという結果になった。このような結果になった原因として、作物の評価を任意の地点のNDVI値ではなくNDVIsum値としているため、葉面積が増大した時点の行動を良い行動として学習したためだと考える。

今後の課題として数百回程度の試行回数でも収束が見込める方策とハイパーパラメータを検討していく必要がある。また作物の評価に関してNDVI値に着目する場合、

表 1 栽培終了時の Q テーブル

Table 1 Qtable at the end of the experiment

行動 a \ 状態 s	黒	赤	緑	青	青緑	赤青	赤緑	白
黒	695872	17972	684170	578176	245	22226	599236	537864
赤	38623	676711	642615	377417	688289	493086	549611	668942
緑	682717	588780	670538	596593	547803	391698	574055	35010
青	17166	307294	727360	294257	770977	257490	1015460	525077
青緑	484786	32242	508310	71831	496581	79576	674346	853219
赤青	-55288	112218	496208	15144	131206	16580	19159	479418
赤緑	438986	379781	461431	315093	553376	731912	429541	29832
白	-16073	494729	322197	-177804	-140212	120105	12818	2374

NDVIsum 値を算出するのではなく葉の 1 画素を追従して NDVI 値を算出する必要がある。

参考文献

[1] 農業労働力に関する統計：農林水産省. <https://www.maff.go.jp/j/tokei/sihyo/data/08.html>. (Accessed on 01/07/2020).

[2] Overview personal food computer — mit media lab. <https://www.media.mit.edu/projects/personal-food-computer/overview/>. (Accessed on 12/23/2019).

[3] 難波脩人, 辻順平, 能登正人. 農作物の自動灌水制御に向けたニューラルネットワークを用いた Q 学習. 人工知能学会全国大会論文集, Vol. JSAI2019, pp. 1F3OS17a04–1F3OS17a04, 2019.

[4] 飯塚 達哉, 繁田 亮, 川原圭博. 完全制御型植物工場における最適栽培制御則探索のためのベイズ最適化を用いる手法の一検討. 信学技報, Vol. 118, No. 127, pp. 47–52, 2018.

[5] Am2321 datasheet. <https://datasheetspdf.com/pdf-file/770161/Aosong/AM2321/1>. (Accessed on 02/13/2020).

[6] Dfr0198 dfrobot — センサ、トランスデューサ — digikey. https://www.digikey.jp/product-detail/ja/dfrobot/DFR0198/1738-1311-ND/7597054?utm_adgroup=Temperature%20Sensors%20-%20Thermocouple%2C%20Temperature%20Probes&utm_source=google&utm_medium=cpc&utm_campaign=Google%20Shopping_Sensors%2C%20Transducers&utm_term=&productid=7597054&gclid=EAIaIQobChMivIvIyqj05wIVE7awCh2DNg9wEAQYAIBEGIVW_D_BwE. (Accessed on 02/13/2020).

[7] Senseair k30 | 製品一覧 | センサー. <http://senseair.jp/product/oem/67.html>. (Accessed on 02/06/2020).

[8] Water sensor module user's manual. https://www.fecegypt.com/uploads/dataSheet/1480850810_water.pdf. (Accessed on 02/13/2020).

[9] Ezo ph circuit — atlas scientific. https://www.atlas-scientific.com/product_pages/circuits/ezo_ph.html. (Accessed on 02/06/2020).

[10] Ezo e.c. circuit — atlas scientific. https://www.atlas-scientific.com/product_pages/circuits/ezo_ec.html. (Accessed on 02/06/2020).

[11] Adafruit neopixel neomatrix 8x8 - 64 rgb led pixel matrix id: 1487 - \$34.95 : Adafruit industries, unique & fun diy electronics and kits. <https://www.adafruit.com/product/1487>. (Accessed on 02/13/2020).

[12] 赤外線投光器キット(改): 組立キット 秋月電子通商-電子部品・ネット通販. <http://akizukidenshi.com/catalog/g/gK-00094/>. (Accessed on 02/13/2020).

[13] 植生指標データについて | 国土地理院. <https://www.gsi.go.jp/kankyochiri/ndvi.html>. (Accessed on 01/11/2020).

[14] 若森弘二, 市川ドルジュ, 小栗直宏. 衛星画像及び地上計測による穀物生育モニターリング. 日本リモートセンシング学会誌, Vol. 36, No. 5, pp. 507–509, 2016.

[15] 尾崎敬二. デジタルカメラ画像から導出の植物指標評価. 日本画像学会誌, Vol. 51, No. 2, pp. 118–124, 2012.

[16] mm3_presentation_20cybernetech_full.pdf. https://www.maff.go.jp/j/kanbo/needs/attach/pdf/mm3_presentation_20cybernetech_full.pdf. (Accessed on 01/14/2020).

[17] 牧野貴樹, 澁谷長史, 白川真一, 浅田稔, 麻生英樹, 荒井幸代, 飯間等, 伊藤真, 大倉和博, 黒江康明, 杉本徳和, 坪井祐太, 銅谷賢治, 前田新一, 松井藤五郎, 南泰浩, 宮崎和光, 目黒豊美, 森村哲郎, 森本淳, 保田俊行, 吉本潤一郎. これからの強化学習. 森北出版, 2016.