

# Actor-Critic により感覚尺度を学習する照明制御システムの提案

中村 彰之

## 1 はじめに

従来、照明システムは複数の照明を一括にしか制御できなかった。現在は複数の照明をネットワーク化することで複数の照明を個別に制御できるように進歩している<sup>1)</sup>。これにより多様な光環境や知的な動作が実現可能となっている。しかし、機能が向上するに伴いその操作は複雑かつ多様なものとなり、従来の UI (User Interface) ではユーザの操作負担が大きなものとなっている。

このような問題に対して「少し明るく」や「とても暗く」といった感覚的に明るさを指示できる UI が有用であると考えられる。そこで、本稿では「少し」や「とても」という感覚的な尺度を Actor-Critic という強化学習手法を用いてユーザごとに学習することで、照明を感覚的に操作できるシステムについて検討する。

## 2 照明のネットワーク化

本研究では照明がネットワーク化したシステムを備えた実験室を構築している。その実験室を Fig. 1 に示す。



Fig.1 ネットワーク化した照明システムを備えた実験室 (出典：自作)

本実験室には Fig. 1 に示すような照明が 48 灯備わっており、それらの光度 (照明の放つ明るさの指標: cd) を個別に調節することが可能である。これにより多様な光環境や知的な動作が実現できる。

このような照明システムにおける UI として、照明の光度を個別に設定する形式の UI が考えられているが、複数の照明の光度を個別に調節する操作はユーザにとって大きな負担となる<sup>2) 3)</sup>。そこで、本稿では「少し明るく」や「とても暗く」といった感覚的に明るさを指示できる照明制御システムを提案する。

## 3 感覚的操作による照明制御システム

本システムにおいてユーザは現在の明るさに対する相対的な要求を入力する。現在では「とても明るく」「明るく」「少し明るく」「少し暗く」「暗く」「とても暗く」の 6 つの選択肢を用意することを想定している。本システムの要件を以下に示す。

- 照度制御  
照明の明るさをシステムが制御する際には、光度ではなくよりユーザ視点である照度 (物体を照らす明るさの指標: lx) に着目して制御すべきである。
- 感覚尺度の学習  
「とても」や「少し」といった感覚はユーザによってことなるため、ユーザごとに感覚尺度を学習する。

### 3.1 照度制御

#### 3.1.1 概要

本システムの照度制御では照度センサを用いることを前提とする。照度センサを用いた照度制御問題は、目標とする照度とセンサから取得できる照度の誤差を最小とする光度パターンを算出する最適化問題として考えられ、目的関数  $f$  は式 (1) のように定式化できる。

$$\begin{aligned} \min \quad & f(L) = (I - I_0)^2 \\ \text{subject to} \quad & L = \{ l \mid 0 \leq l \leq 1090 \} \end{aligned} \quad (1)$$

$L$ : A set of the luminance in all lightings (cd)

$l$ : The luminance of a certain lighting (cd)

$I$ : The illuminance of the sensor (lx)

$I_0$ : The target illuminance (lx)

$l$  は光度を示し、本実験室の照明器具の最大光度は 1090(cd) である。また、 $I$  はセンサから得られる照度、 $I_0$  は目標とする照度を示す。本システムではこの目的関数を準ニュートン法の代表的な手法である BFGS 法によって最適化する。

#### 3.1.2 実験

照度制御が実現できているかを確認するために Fig. 1 の実験室で照度制御実験を行った。48 灯の照明は 2 灯 1 組となっているため、1 組を 1 灯の照明と見なし照明数を 24 灯とした。目標照度を 2000lx として行った実験結果を Fig. 2 に示す。

Fig. 2 の縦軸は評価値、横軸はステップ数を示す。ステップ数を重ねる毎に評価値は減少し、3 ステップ目のセンサ照度は 1990lx となった。このことから BFGS 法による照度制御は実現できていると言える。

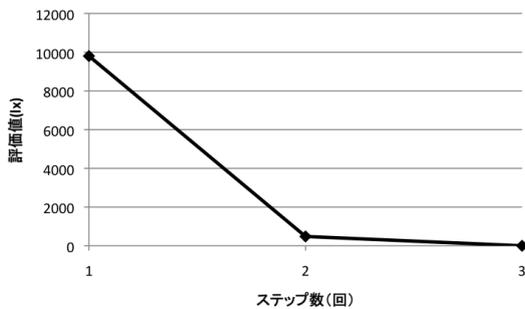


Fig.2 評価値の推移 (出典：自作)

### 3.2 感覚尺度の学習

本システムではユーザの明るさに対する感覚尺度を学習するために、強化学習の Actor-Critic を用いる。強化学習とはシステムが試行錯誤を繰り返すことで環境の状態に応じた適切な行動パターンを学習する手法である。強化学習では環境とエージェントという概念を用いている。環境は複数の状態を持つ対象問題であり、エージェントは対象問題で与えられた目的を果たすように学習を行う。本システムでは連続値の行動出力が必要であるため強化学習の中でも Actor-Critic を用いる。Actor-Critic の概念図を Fig. 3 に示す。

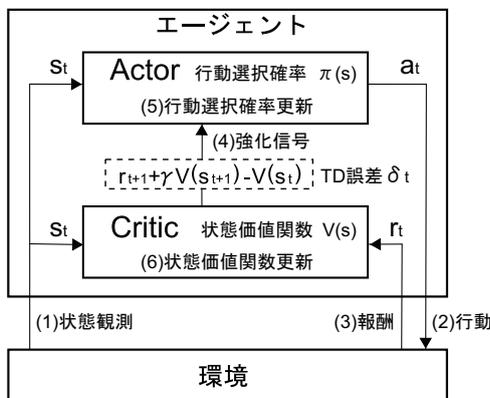


Fig.3 Actor-Critic の概念図 (参考文献<sup>4)</sup> より参照)

Fig. 3 に示すように、Actor-Critic におけるエージェントは行動を選択する actor 部と実行した行動を評価する critic 部により構成される。エージェントは (1)~(6) の工程を繰り返すことにより、どの状態の時にはどの行動を実行するべきかを学習する。以下に、Fig. 3 における各工程を説明する。

- (1) 状態観測  
エージェントが環境において状態  $s_t$  を観測する。
- (2) 行動  
actor が行動選択確率  $\pi(s_t)$  に従って行動  $a_t$  を実行する。ここで、行動選択確率として正規乱数などを用いることで連続値の行動出力が可能となる。
- (3) 報酬  
行動  $a_t$  によって状態が  $s_{t+1}$  に遷移し、その評価として critic に報酬  $r_t$  がスカラーの形で与えられる。

### (4) 強化信号

critic は次の状態  $s_{t+1}$  を観測し、actor への強化信号として TD(Temporal Difference) 誤差を計算する。TD 誤差  $\delta_t$  は状態価値関数  $V(s_t)$  を用いて以下のように表される。状態価値関数は各々の状態がどれだけ良い状態であるかを示す。

$$\delta_t = r_t + \gamma V(s_{t+1}) - V(s_t)$$

$\gamma$  は割引率と呼ばれ、 $0 \leq \gamma \leq 1$  の係数である。

### (5) 行動選択確率更新

TD 誤差  $\delta_t$  を用いて actor の行動選択確率  $\pi(s_t)$  を更新する。正規乱数を用いている場合は、 $a_t$  が選択され易くもしくは選択され難くなるようにその平均と標準偏差を変更する。

### (6) 状態価値関数更新

TD 誤差  $\delta_t$  を用いて critic の状態価値関数  $V(s_t)$  を更新する。

$$V(s_t) \leftarrow V(s_t) + \alpha(r_t + \gamma V(s_{t+1}) - V(s_t))$$

$\alpha$  は学習率と呼ばれ、 $0 \leq \alpha \leq 1$  の係数である<sup>4)</sup>。

本システムではユーザの「とても明るく」や「少し暗く」などの要求を受けた時に、目標照度をどれだけ変化させるかを上述した Actor-Critic により学習する。

## 4 今後の展望

照度制御に関しては、収束するまでの時間がより短縮されるようにパラメータを調整する。また室内に複数のユーザがいる状況を想定して、複数の照度センサを用いた照度制御を実現する。

感覚尺度の学習に関しては、まず Actor-Critic を実装し「とても明るく」や「少し暗く」などの要求がある状態にはどのような行動を選択するべきかを学習する。そして第二段階としてユーザの求める照度の変化量はその時の照度によって異なり得るという仮定の元、「とても明るく」や「少し暗く」の次元に加えて現在照度という次元も加味した学習を行う。

### 参考文献

- 1) Vipul Singhvi, Andreas Krause, Carlos Guestrin- James H. Garrett Jr, H. Scott Matthews : Intelligent light control using sensor networks, Proceedings of the 3rd international conference on Embedded networked sensor systems, pp.218-229, 2005.
- 2) Barry Brumitt, JJ Cadiz : "Let There Be Light" Examining Interfaces for Homes of the Future, Proceedings of Interact '01, 2001.
- 3) Krzysztof Gajos, Daniel S. Weld : SUPPLE - Automatically Generating User Interfaces-, Proceedings of the 9th international conference on Intelligent user interfaces, 2004.
- 4) 小林 (重) 研究室 - 強化学習  
<http://www.fe.dis.titech.ac.jp/research/rl/index.html>