

ウェアラブルカメラを用いた目の動きによる夜間呼び出しシステムの開発 Development of Night-time Calling System by Eye Movement using wearable Camera

坂本 一樹[†] 齊藤 剛史[†] 伊藤和幸[‡]
Kazuki Sakamoto Takeshi Saitoh Kazuyuki Itoh

1. はじめに

筋萎縮性側索硬化症 (ALS) 患者や多系統萎縮症 (MSA) 患者、筋ジストロフィー患者は重度の運動機能障害のため意思を伝えることが困難になる。患者から介護士などの医療従事者への呼びかけの一つとしてナースコールが利用されている。ナースコールの子機には入力として押しボタンスイッチなどが用いられるが、ALS 患者のように筋肉を動かすことが困難な人に対しては残存機能に応じた入力装置を用意しなければならない。また、介護士が抱える問題として夜間の介護が挙げられる。特に在宅療養の場合は、夜間の様子を常に確認できる人がいないため、ALS 患者が助けを必要としている場合でも気づけないことがある。アイスイッチ[1]は、目の動きを用いて家電製品の ON/OFF やナースコールを使用することができる。夜間での使用も可能であるが、細かい眼球運動を検出することは困難である。また、アイスイッチはアームを含めた装置をベッドサイドに固定する必要があり、患者の位置が動くたびに位置を修正する必要がある。

本研究では、Convolutional Neural Network (CNN) を用いた瞳孔中心検出をもとに ALS の身体機能として残りやすい目の動きを用いた呼び出しシステムを開発する。夜間に運用することを想定し、近赤外線カメラを用いて目画像を撮影することで、暗所でも瞳孔中心の検出を可能にする。

2. 関連研究

これまでに様々なナースコールシステムが開発されている。Ongenaer [2]は、現在のコンテキストに基づいて呼び出しの優先度を評価し、最も適切な介護者を割り当てる、オントロジーベースのナースコールシステムを開発した。Unluturk [3]はナースコールシステムソフトウェア、無線通話システムソフトウェア、ロケーションシステムソフトウェアおよび通信プロトコルの統合システムを開発した。このシステムにより、患者の担当看護師が対応できない場合に、他の対応可能な看護師を呼び出す。Klemets と Toussaint[4]は、呼び出しの理由を看護師が識別できるナースコールシステムを提案した。スイッチなどの機器ではなく、システム全体の概念が提案されている。

画像ベースの眼球運動の検出に関しては、目の動きではなく注視点を推定する製品が提供されている。使用されるデバイスは、非ウェアラブルとウェアラブルの二つのタイプに大別できる。前者は、Tobii Pro Nano[5]や Tobii Pro Fusion[6]など、ディスプレイに接続するモニターベースのアイトラッカがある。後者は Tobii Pro Glasses 3[7]やガゾウ社の GPE3[8]など、メガネ型のフレームに取り付けられた小型カメラ付きのデバイスである。眼球運動の検出に関連す

[†]九州工業大学 Kyusyu Institute of Technology

[‡]国立障害者リハビリテーションセンター研究所

National Rehabilitation Center for Persons with Disabilities

る研究は既存のアイトラッカを使用して視線を分析する研究[9][10][11]と目や瞳孔中心の検出方法を提案する研究[12][13][14][15]に分けられる。

3. 開発システム

3.1 構成機器

本システムは、図1に示すように、ウェアラブルカメラ、小型コンピュータ、リレー制御器、ナースコールで構成されている。本システムは夜間での運用を想定しているため、標準的なカラーカメラや照明を利用できない。そこで、近赤外線 LED と近赤外線カメラを搭載したウェアラブルカメラを使用する。ウェアラブルカメラはユーザの頭の動きの影響を受けず、安定して目の画像を撮影することができる。

本システムでは、目画像を処理するためのコンピュータが必要となるが、ベッドサイドに大型のコンピュータを設置することができない。また、有線・無線ネットワークの使用も制限される。そこで、GPU を搭載した小型コンピュータを使用する。

小型コンピュータから直接、連続した信号を送るとナースコールは信号を受信する度に介護士を呼び出す。そこで、リレー制御器を使用し、誤送信による呼び出しを防ぐ。



図1：システムの構成

3.2 瞳孔中心検出

本システムでは、Chinsatit と Saitoh[13]によって提案された CNN ベースの瞳孔中心検出方法を適用する。この手法では、図2に示すように二つの CNN モデルを使用する。最初のモデルで目状態を分類し、二番目のモデルで瞳孔中心点を推定する。

分類モデルのアーキテクチャは AlexNet[16]を使用する。このモデルの出力は目の状態を表す「開眼」と「閉眼」の二つである。目を閉じている際に瞳孔中心点を検出することはできないので、次の処理はスキップされる。

回帰モデルには、ConvNet[17]を使用する。このモデルは瞳孔中心点(P_x, P_y)を出力する。

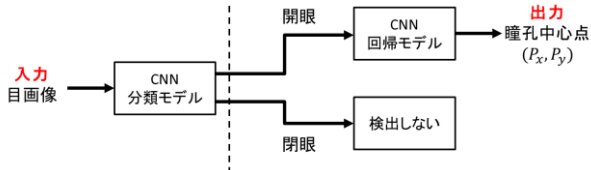


図2：二つのCNNモデルの構成

3.3 システムの動作

本システムの動作中の画面を図3に示す。左上のウィンドウでは、目画像の上に目状態や瞳孔中心点などの情報を表示している。右上のウィンドウは検出された瞳孔中心点の垂直方向のグラフ、左下のウィンドウは水平方向のグラフであり、リアルタイムで描画される。

本システムでは、瞳孔の周辺にしきい値を設けておき、ユーザの瞳孔中心がしきい値を超えるように目を動かすことでリレー制御器へ信号を送信する。

図4に目画像と表示される情報の例を示す。ここで、緑の円は検出された瞳孔中心であり、瞳孔周辺の長方形はしきい値である。一番左端の図は正面を向いており、二番目、三番目の図はそれぞれ左上のしきい値を超えた例である。図の右端は目を閉じた時の例である。

患者の中には、目を素早く動かすことができるユーザとゆっくりと目を動かすユーザがいる。前者の場合、リレー制御器への信号は、瞳孔中心がしきい値を超えた直後にコンピュータから送信する単発モードを導入する。瞳孔中心の検出誤差によりしきい値を超えない場合でも、再度眼球を動かすことで誤検出への対策ができる。一方、後者の場合、しきい値を超えている間に信号を送信し続ける連続モードを導入する。本システムではこの二つのモードやしきい値を手動で設定する。これにより、様々な症状の患者へ対応できる。

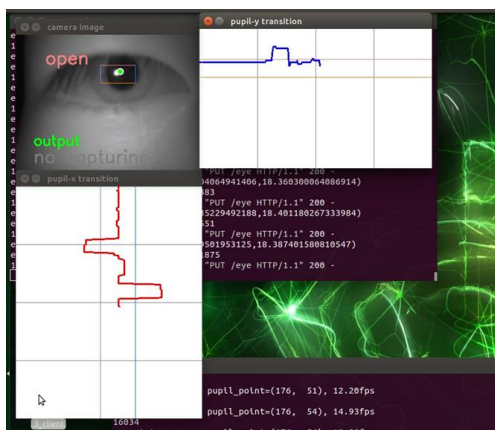


図3：システム動作中のメイン画面

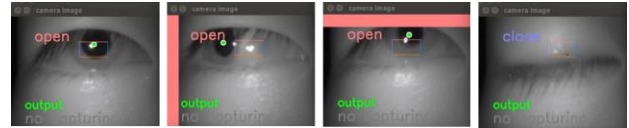


図4：目画像と表示される情報の例

4. 評価実験

4.1 瞳孔中心検出

4.1.1 データセット

本研究では主に介護士、ALS患者の目画像を使用した。表1に被験者毎の目画像の収集枚数を示す。被験者は正面と上下左右に目を動かし目画像を撮影した。被験者毎に収集した枚数が異なるので、各被験者から400枚の目画像を選定し実験に使用する。

目画像は1280×720 [pixels]で撮影され、処理時間を短縮するために120×80[pixels]に変更される。目状態と瞳孔中心の真値は目視で与え、目状態は瞳孔が50%以上見える場合は「開眼」、それ以外は「閉眼」として二つのラベルを付けた。

表1：被験者毎の目画像の収集枚数

ID	被験者情報	性別	収集枚数[枚]
s1	介護士	男	464
s2		男	419
s3		男	480
s4		男	1,087
s5	患者	男	561
s6		女	1,767
s7		男	1,859
合計			6,637

4.1.2 実験条件

最適な条件を提案するためROI抽出、Data Augmentation (DA)、Fine Tuning (FT)を組み合わせた実験条件を設けた。

本研究で収集した目画像には背景が映り込んでいるため目周辺のROI抽出を行う。本研究においては、ROIの抽出方法としてアスペクト比を固定せずに抽出する方法 (ROI1) と固定して抽出する方法 (ROI2) の2通りで実験する。アスペクト比を固定する場合はCNNへの入力画像と同じ3:2で固定する。

ユーザの瞬きに対してフレーム間差分 $I_{\Delta t}$ を取得し、しきい値処理により2値画像 $I_{\Delta t_{th}}$ を作成する。1フレーム前の画像と現フレーム画像をそれぞれ図5(a)、図5(b)に示す。またその時の $I_{\Delta t_{th}}$ を図5(c)に示す。 $I_{\Delta t_{th}}$ を蓄積し、2値化処理を施した後、ラベリング処理により最も大きい領域のみを抽出する。その後、図5(d)の赤枠もしくは緑枠に示すようにROIが一定のサイズ以下になるまでROIを狭める。ここで、赤枠はROI1、緑枠はROI2の領域を表す。元の目画像からROIを抽出した図5(e)のROI1または図5(f)のROI2をCNNモデルへの入力画像とする。

本実験で使用できる目画像の枚数が少ないためDAにより学習用目画像を増やす。拡大・縮小、平行移動、回転、

輝度値の補正を組み合わせ、1枚の目画像につき4枚ずつ画像を生成した。図6に生成した目画像を示す。

FTについては、6名の男性、3名の女性、合計9名の学生から異なるウェアラブルカメラで1,980枚ずつ、合計17,820枚の画像で収集した。これらを用いてCNNモデルを学習させた重みを初期値として、介護士、患者の目画像を学習させた。

2種類のROI、DA、FTを組み合わせた8条件で実験を行った。学習は患者1名をテストデータとし、残り6名を学習データとした一人抜き法を適用した。

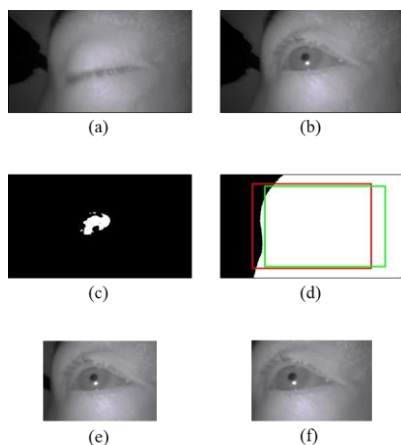


図 5 : 二つの ROI 抽出法

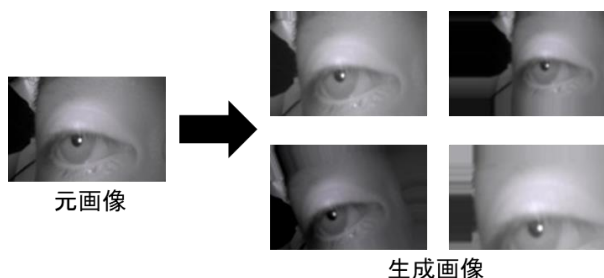


図 6 : Data Augmentation による目画像の生成

4.1.3 実験結果

実験結果を表2に示す。目状態の分類については、全体的に閉眼より開眼に対する精度が高い。これは閉眼の学習データが少ないためと考えられる。ROI1を適用した場合の精度が最も高く、82.1%だった。瞳孔中心検出については、ROI1+DA+FTを適用した場合に平均誤差が最小となり、1.17pixelだった。図7に瞳孔中心の真値(緑)と検出結果(赤)を描画した画像を示す。図7中の(a), (b), (c), (d)の誤差はそれぞれ2.47, 2.63, 7.22, 8.82pixelである。

表 2 : 目状態の分類と瞳孔中心検出の結果

実験条件	目状態の分類[%]			平均誤差 [pixel]
	開眼	閉眼	平均	
ROI1	96.8	67.5	82.1	1.54
ROI2	96.8	59.6	78.2	1.52
ROI1+DA	96.9	38.0	67.5	1.36
ROI2+DA	93.2	57.2	75.2	2.14
ROI1+FT	96.6	52.7	74.7	1.36
ROI2+FT	97.4	45.2	71.3	1.48
ROI1+DA+FT	95.9	39.7	67.8	1.17
ROI2+DA+FT	96.1	37.0	66.6	2.27

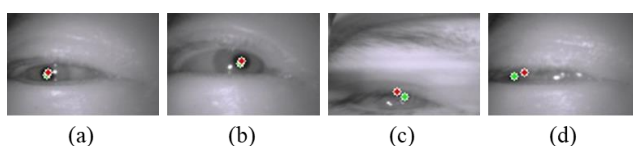


図 7 : 瞳孔中心検出結果

4.2 システム評価実験

4.2.1 実験方法

システムを評価するため、健常者5名を対象に被験者実験を行った。図8の左図に示すように、被験者はベッドに横たわり、約5分間にわたって実験した。指示以外で目を動かさないようにするため、図8の右図に示すように、壁に取り付けられたモニターを見つめ、正面を見続けるようにした。上下左右の音声指示を用意し、実験中、被験者は指示された方向に眼球を動かす。音声指示の時間と方向はランダムであり、被験者は12回の指示を受ける。



図 8 : 実験の様子

4.2.2 実験結果

実験は大学の研究室で夕方に室内灯を消灯して行った。実験時の明るさは照度計で測定した結果、最小値が3.92lx、最大値が16.7lx、平均値が9.9lxであった。

全ての実験で真陽性(TP)、偽陰性(FN)、偽陽性(FP)を数えた。また、適合率P、再現率R、F値Fを下式で計算した。

$$P = TP / (TP + FP)$$

$$R = TP / (TP + FN)$$

$$F = 2PR / (P + R)$$

表3に示すように、全被験者の適合率、再現率、F値はそれぞれ0.833、1.000、0.909だった。これは、目の動きの検出が成功したことを表す。一方で適合率は0.833となっ

ているが、これは S2 と S5 が瞬きをして目を閉じた瞬間に瞳孔中心検出に失敗したためであった。

図 9 に S4 の実験における瞳孔中心座標の時間推移を示す。横軸は時間に応じたフレーム数で、縦軸は x または y 座標である。赤い曲線は検出された瞳孔中心の座標で、緑と青の線は左右または上下のしきい値を表す。垂直のピンクの帯は目が閉じていることを表す。これらのグラフから、瞳孔中心が 12 回の呼び出しに応じてしきい値を超えていることが確認できる。

表 3 : システム評価実験の結果

指標	S1	S2	S3	S4	S5	合計
TP	12	12	12	12	12	60
FN	0	10	0	0	2	12
FP	0	0	0	0	0	0
P	1.000	0.545	1.000	1.000	0.857	0.833
R	1.000	1.000	1.000	1.000	1.000	1.000
F	1.000	0.706	1.000	1.000	0.923	0.909

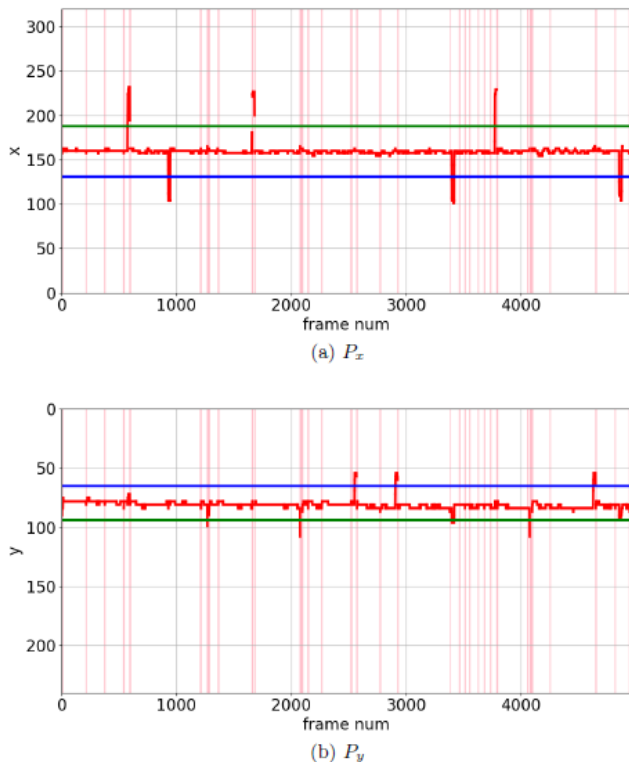


図 9 : S4 の瞳孔中心の時間推移

5. 結論

本研究では、目の動きによる夜間呼び出しシステムを開発した。開発システムを評価するため瞳孔中心検出と呼び出し実験の二つの実験を行った。実験の結果、瞳孔中心検出においては平均誤差 1.17pixel の高い検出精度を得た。呼び出し実験においては、健常者に対して実験し、高い成功率が得られた。今回は健常者を対象に実験したが、今後は患者を対象に実験する。また、瞳孔中心検出に関して、瞬きによる検出の失敗が見られるため、改善を図る。

謝辞

本研究は、JSPS 科研費 19KT0029 の助成によるものである。本研究のデータは、九州工業大学大学院情報工学研究院等における人を対象とする研究審査および国立障害者リハビリテーションセンター倫理審査委員会の承認を受けている。

参考文献

- [1] アイスイッチ. <http://www.emfasys.co.jp/index8f.html>
- [2] Ongenaes, F., Claeys, M., Kerckhove, W., Dupont, T., Verhoeve, P., Turck, F.: A self-learning nurse call system. *Computers in Biology and Medicine* 44, 110-123 (2014). <https://doi.org/10.1016/j.compbiomed.2013.10.014>
- [3] Unluturk, M.S., Ozcanhan, M.H., Dalkilic, G.: Improving communication among nurses and patients. *Computer Methods and Programs in Biomedicine* 120(2), 102-112 (2015). <https://doi.org/10.1016/j.cmpb.2015.04.004>
- [4] Klemets, J., Toussaint, P.: Does revealing contextual knowledge of the patient's intension help nurses' handling of nurse calls? *International Journal of Medical Informatics* 86, 1-9 (2016). <https://doi.org/10.1016/j.ijmedinf.2015.11.010>
- [5] Tobii Pro Nano. <https://www.tobiiipro.com/product-listing/nano/>
- [6] Tobii Pro Fusion. <https://www.tobiiipro.com/product-listing/fusion/>
- [7] Tobii Pro Glasses 3. <https://www.tobiiipro.com/product-listing/tobii-pro-glasses-3/>
- [8] GPE3. <https://www.gazo.co.jp>
- [9] Ohya, T., Morita, K., Yamashita, Y., Egami, C., Ishii, Y., Nagamitsu, S., Matsuishi, T.: Impaired exploratory eye movements in children with asperger's syndrome. *Brain and Development* 36(3), 241-247 (2014). <https://doi.org/10.1016/j.braindev.2013.04.005>
- [10] Wang, C.C., Hung, J.C.: Comparative analysis of advertising attention to facebook social network: Evidence from eye-movement data. *Computers in Human Behavior* 100, 192-208 (2019). <https://doi.org/10.1016/j.chb.2018.08.007>
- [11] 土井 俊央, 村田 厚生.: GUI での情報探索における画面の複雑さが反応時間・眼球運動特性に与える影響. *人間工学* Vol.54, No.6 pp. 236-247 (2018). <https://doi.org/10.5100/jie.54.236>
- [12] 吉岡 隆宏, 中島 哲, 小田切 純一, 富森 英樹, 福井 琢.: 眼鏡着用時における環境条件にロバストな瞳孔検出. *映像情報メディア学会技術報告* Vol.38, No.10 (2014). https://doi.org/10.11485/itetr.38.10.0_37
- [13] Chinsatit, W., Saitoh, T.: CNN-based pupil center detection for wearable gaze estimation system. *Applied Computational Intelligence and Soft Computing* 2017(2017). <https://doi.org/10.1155/2017/8718956>
- [14] Gou, C., Zhang, H., Wang, K., Wang, F.Y., Ji, Q.: Cascae learning from adversarial synthetic images for accurate pupil detection. *Pattern Recognition* 88, 584-594(2019). <https://doi.org/10.1016/j.patcog.2018.12.014>
- [15] Yiu, Y.H., Aboulatta, M., Raiser, T., Ophye, L., Flanagan, V.L., Eylonburg, P., ahimadi, S.A.: DeepVOG: Opensource pupil segmentation and gaze estimation in neuroscience using deep learning. *Journal of Neuroscience Methods* 324 (2019). <https://doi.org/10.1016/j.jneumeth.2019.05.016>
- [16] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinto. *Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems*, pp.1299-1312, 2016.
- [17] Pfister, T., Simonyan, K. Charles, J. Zisserman, A.: Deep convolutional neural networks for efficient pose estimation in gesture videos. In: *Asian Conference on Computer Vision*. pp. 538-552 (2015)