

視線ジェスチャ検出とそのユーザインターフェースへの応用

庄林 祐太†

藤江 真也†

†千葉工業大学工学部未来ロボティクス学科

1 はじめに

ユーザの意図的な視線の動きを視線ジェスチャとして認識する手法、およびそのユーザインターフェースへの応用について検討する。

本来、視線はユーザが操作対象の情報を視覚的に得るために向けているものである。そのため、特に視線ジェスチャを情報機器操作に利用する際には、ジェスチャを意図した視線動作と意図していない視線動作（情報を得るための動き）を区別する必要がある [1]。このような問題を避けるために、ジェスチャ検出をする範囲を特定の範囲（画面の端に配置された点など）に限定するといった方法が提案されている [2]。

本研究ではこの問題に対し、視線ジェスチャ認識の開始と終了をそれぞれ、意図的に目の閉じる動作と一定時間の視線の滞留によって検出することにする。また、視線ジェスチャは瞬きを行った点を起点とし、対象となる動作を行ったかどうかを判定することで認識する。こうすることで、誤検出を最小限にとどめながら、視線の動きも少ないことでユーザの負担が軽い効率的なインターフェースを実現する。

2 視線ジェスチャ検出

2.1 対象ジェスチャ

認識対象ジェスチャとして、L字を描くような直線の組み合わせが少なく、簡単なものを採用した。図1に、対象ジェスチャを示す。

図中には、起点を基準として左側に動作するもののみを示した。本稿で述べる実験についても左側に限定したものである。実際にシステムを運用する際には、これらのデータを反転させたものを利用することによって右側のジェスチャにも対応することが可能である。

2.2 視線情報の取得

本研究では、視線情報の検出に Tobii Technology 社が開発した Tobii EyeX[3] という視線検出装置を使用する。本装置は、PC 画面のフレームの下側につけて使用する。視線情報として、画面上の視線位置、装置に対

する眼球位置を取得することができる。また、眼球位置から目の開閉を取得することも可能である。

2.3 視線ジェスチャ開始・終了検出

ジェスチャに無関係な視線動作を誤検出することは、誤動作の原因となるため避けなければならない。本研究では、ジェスチャの開始と終了に関してはユーザが特定の動作を行うことで、開始と終了をシステムに明示的に指示する方式をとる。

ジェスチャの開始については、ユーザが片目を閉じたときをその合図と解釈する。片目を閉じるという動作を意図しないで行うことは極稀であることから、有用であると考えた。ジェスチャの終了については、視線が一定時間、一定の範囲内に滞留することをその合図と解釈する。直前までジェスチャに対応する動きをしていたものが停止している状態を検出することは合理的であると考えた。

2.4 視線ジェスチャ認識

視線検出装置で時系列的に検出される視線情報を利用した視線ジェスチャ認識手法について述べる。

ジェスチャの開始が検出された点を原点とする視線位置を x_t, y_t (t はジェスチャ開始後のフレーム数) とする。実際には、固視微動（眼球が常に細かに揺れている）の影響で、センサから直接得られる視線位置にはブレが含まれる。そのため、 x_t と y_t は、このブレを取り除く前処理を施している。さらに、これらのフレーム間差分 $\Delta x_t = x_t - x_{t-1}$, $\Delta y_t = y_t - y_{t-1}$ を加えた4次元のベクトルを特徴量とする。

識別器としては、ニューラル・ネットワーク、特に、時系列情報を扱うことができる RNN (Recurrent Neural Network) を用いる。4次元の特徴量を入力層、100素

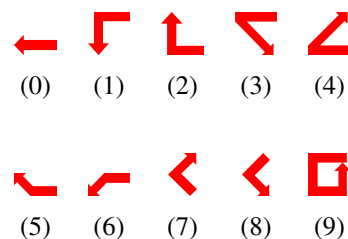


図1: 認識対象ジェスチャとその番号

Gaze gesture detection and its application to user interface

†Yuta SHOBAYASHI †Shinya FUJIE

†Chiba Institute of Technology

表 1: 認識結果の Confusion Matrix

		認 識 結 果									
		0	1	2	3	4	5	6	7	8	9
正	0	.803	0	0	.019	0	.058	.011	0	0	0
	1	.019	.673	0	.076	0	0	.211	.019	0	0
	2	.019	0	.692	0	.115	.173	0	0	0	0
	3	0	.204	0	.346	0	0	.040	0	.346	.061
	4	0	0	.220	.020	.280	.020	0	.420	.040	0
解	5	.055	0	.129	0	0	.796	0	0	.018	0
	6	.040	.183	0	0	.020	0	.714	0	.020	.020
	7	0	0	.036	0	.200	.018	0	.745	0	0
	8	0	.037	0	.111	.129	.018	0	0	.592	.111
	9	.038	.038	.019	.057	.057	.096	.038	.115	.115	.423

子の LSTM (Long short-term memory) を中間層とし、ジェスチャの異なり数である 10 素子の出力層を持つネットワークとなっている。

3 実験

3.1 実験条件

提案した視線ジェスチャ認識の評価を、収録したデータに対して行う。

被験者 10 人に、10 種類のジェスチャをそれぞれ約 5 回ずつ行わせた。ジェスチャの指示は、被験者が見ている画面の中心に図 1 に示した画像を数秒間表示することで行った。指示する順序はランダムである。こうして収集したデータを用いて、学習データに 9 人、テストデータに 1 人の配分で交差検証を 10 回行う。

3.2 結果と考察

結果を表 1 に示す。ジェスチャ(3)と(4)が、それぞれ(8)、(7)と混同していることがわかる。図 1 を見ると、これらのジェスチャは後半の部分が、共通した方向に動くことが分かる。表を見ると、その他についても似ている動作が混同していることがわかる。混同しやすいジェスチャの組み合わせが含まれないように、絞り込みを行うことが望まれるが、この点については今後の課題である。

4 ユーザインターフェース

提案した視線ジェスチャ認識手法を応用したユーザインターフェースについて検討する。まず、ハンズフリーで操作できる例として動画視聴を考える。動画視聴では、再生、停止、音量調整などが主な操作となる。視線ジェスチャをこれらの操作に割り当てることで、画面を見るだけで操作が可能となる。また、Web ブラウザの操作を考えた場合、ページ送り・戻し、検索欄へ



図 2: ポインタ。左からそれぞれ、視線ジェスチャ時のユーザの視線位置、通常時の視線位置、視線ジェスチャ開始時の原点を示す。

のカーソル移動などに視線ジェスチャを割り当てることで、キーボード、マウスとの併用して効率的な操作の実現が期待できる。

視線はジェスチャ入力以外にも利用されるため、システムがジェスチャ認識中かどうか分かることが重要である。本研究では図 2 で示すように、ポインタの見た目によりジェスチャ認識中かどうか、また、視線ジェスチャ開始時の点に色違いのポインタと同じ図形を置いておくことで、どのジェスチャを入力していることになっているかが理解しやすいようにした。

5 まとめ

情報機器の操作のための視線ジェスチャの認識手法と、それを利用したユーザインターフェースについて検討した。今後の課題としては、インタフェースの評価が挙げられる。想定したシステムを、実際に被験者に利用させ、その印象を評価する予定である。

参考文献

- [1] H. Drewes and A. Schmidt. Interacting with the computer using gaze gestures. *Proc. Conf. on Human-computer interaction'07*, pp. 475–488, 2007.
- [2] 高, 他. Eye Gesture 入力インタフェースシステムの開発. 信学技報, HIP, Vol. 112, No. 46, pp. 31–36, May 2012.
- [3] Tobii Technology Inc, Tobii EyeX. <http://www.tobii.com/xperience/developer/>.