

## 映像コンテンツ視聴時における興味行動推定

橋本遼（金沢大学）・南保英孝（金沢大学）

## 1. まえがき

将来、自動運転の自動車が急速に普及する中で、自分で運転する必要がないため運転中に映像コンテンツを楽しむことが考えられる。そこで、映像コンテンツを視聴する人の視線情報からそのコンテンツに対する興味度を推定することで、より興味のあるコンテンツへと更新するなど、その人ならではの付加情報を提供できないかと考えた。本研究では、映像コンテンツを視聴している人がどこに興味をもって見ているかを推定する機械学習モデルを構築し、推定精度 80%以上を研究目標とする。

## 2. 興味行動の定義

関連研究から、人がある対象を見つけて、その対象をよく知ろうとするときに 1.5 秒ほどの視線の滞留があることが分かり、その 1.5 秒を閾値として視線の動きを 4 つの興味行動に分類し興味の状態を推定することにした。視線の動きから定めた 4 つの興味行動を以下に示す。

表 1: 視線の動きによる興味行動の定義

興味状態 Interest lv.	lv.1: 無反応	lv.2: 注目対象模索状態	lv.3: 詳細情報獲得状態	
興味ラベル Labels	①非注視 Not Look (Not Attention)	②探索 Search (Looking Around)	③追従 Follow	④注視 Look (Attention)
振る舞い Behaviors	見ていない	キョロキョロして模索する状態	動く対象を見続ける状態	動かない対象を見続ける状態
視野範囲 Viewfield ±5[deg]				

## 3. 手法

推定モデルの学習アルゴリズムとして時系列データの学習や分類に強い Long Short-Term Memory (以下 LSTM) を用いる。LSTM は深層学習の一種であり、時間的な相関関係から情報を抽出することができる。視線の動きを時系列データと捉えて処理できるのではないかと考えた。

## 4. 実験

学習に用いるデータは映像コンテンツを視聴している人のデータ約 19000 秒で、一部を検証データとして利用する。特徴量として鼻と両目の座標情報 (x、y) を用いる。

推定の対象とする興味行動は「その他」除いた「探索」「追従」「注視」の 3 クラスとする。また、異なる興味行動間の動きの細かな違いを判別するために 2 段階推定を適用する。1 段階目で、動きのある動作 (探索・追従) と動きの少ない動作 (注視) を推定する。2 段階目で、動きのある動作のうち、「探索」か「追従」の 2 クラスを推定する。

## 5. 結果

LSTM による検証データの推定結果の混合行列を以下に示す。

縦軸が正解クラスで、横軸が LSTM による推定結果である。1 段階目による推定精度が 82%、2 段階目による推定精度が 75% となった。

表 2: 1 段階目混合行列

	推定結果	
	追従・探索	注視
追従・探索	73	12
注視	28	109

表 3: 2 段階目混合行列

	推定結果		
	追従	探索	注視
追従	0	2	12
探索	12	59	1
注視	9	19	108

## 6. 考察

1 段階目の分類では研究目標である分類精度 80% を達成することができたが、2 段階目では、「追従」の分類精度が悪く 80% を下回ってしまい研究目標を達成することができなかった。

「追従」を識別することができていないため、新たなパラメータの追加や学習手法の検討が必要であると考えられる。また、今回の実験で用いた座標情報に加えて、映像コンテンツを視聴している人と画面との距離も考慮することで精度の改善が期待できるのではないかと考えられる。

## 7. まとめ

本研究では、LSTM を用いて映像コンテンツを視聴している人の興味行動推定を行った。学習モデルの精度は 1 段階目の分類では 80% 以上を達成することができたが 2 段階目では達成することができなかった。

精度向上のための新たな手法として、RNN や CNN を用いず Attention のみを用いた機械学習モデル Transformer を使った興味度推定を行いたいと考えている。