

# LASSO に基づく運転手表情データからの特徴抽出

## Feature Selection from Driver's Facial Expression Data Based on LASSO

大浦聡一郎\* 藤田隆吾\*\* 松川徹\*\*\* 西郷浩人\*\*\* 鈴木英之進\*\*\*  
(九州大学\*システム生命科学府 \*\*システム情報科学府 \*\*\*システム情報科学研究院)

### 1 序論

顔表情に現れる「苛立ち」を感じていると決定づける特徴は何であろうか。Paul Ekman らによる Facial Action Coding System[1] によると、人の幸福・悲しみ・恐怖・怒り・嫌悪・驚きの 6 つの感情を表す表情は文化・言語・性別などによらず普遍であるとされている。これら以外の表情が人類に普遍であるかどうかは明らかでない。森山らは運転者の「苛立ち」のような攻撃的な感情は危険運転を引き起こす一因であるとし、表情データから感情を検出する研究を行った [2]。また、野宮らは主成分分析を用いた表情における重要な特徴の選択に関して研究を行った [3]。これらの先行研究から「苛立ち」を決定づける数少ない要素で構成される特徴群の選択を行うことの重要性が伺える。

本実験では、自動車運転者の表情を「苛立ち」と「それ以外」に分類する 2 値分類学習問題に、LASSO に基づくロジスティック回帰学習を適用し、重要な特徴群を選択した。

実験の結果、眉間や唇間の距離などの特徴が選択され、これらが人間の直観に合致していることを確認した。

### 2 LASSO の紹介

入力は運転者顔画像から得られた  $p \times N$  特徴行列  $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)$ ,  $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{ip})^T$  と  $N$  次元目的変数ベクトル  $\mathbf{y} = (y_1, y_2, \dots, y_N)^T$  であり、出力は切片  $\beta_0$  と  $p$  次元回帰係数ベクトル  $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^T$  である。ここで、 $N$  はサンプル数、 $p$  は特徴数である。

線形回帰モデルを次のように仮定する。

$$y_i = \beta_0 + \sum_{j=1}^p x_{ij} \beta_j + e_i. \quad (1)$$

ここで、 $e_i$  は誤差項である。最小二乗推定では、以下の式によりパラメータ推定を行う。

$$\hat{\boldsymbol{\beta}} = \arg \min_{\beta_0, \boldsymbol{\beta}} \sum_{i=1}^N (y_i - \beta_0 - \sum_{j=1}^p x_{ij} \beta_j)^2. \quad (2)$$

本研究で用いた LASSO (Least Absolute Shrinkage and Selection Operator) [4, 5] は以下の式で表される。

$$\boldsymbol{\beta}^* = \arg \min_{\beta_0, \boldsymbol{\beta}} \sum_{i=1}^N (y_i - \beta_0 - \sum_{j=1}^p x_{ij} \beta_j)^2 \text{ s.t. } \|\boldsymbol{\beta}\|_1 \leq t. \quad (3)$$

別の表現では、

$$\boldsymbol{\beta}^* = \arg \min_{\beta_0, \boldsymbol{\beta}} (\|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|_2^2 + \lambda \|\boldsymbol{\beta}\|_1), \quad (4)$$

となる。 $\|\boldsymbol{\beta}\|_1$  は  $\boldsymbol{\beta}$  の  $\ell_1$  ノルムであり、以下の式で定義される。

$$\|\boldsymbol{\beta}\|_1 = \sum_{j=1}^p |\beta_j|. \quad (5)$$

式 (3) より LASSO は  $\ell_1$  ノルムに関する制約条件を付けた最小二乗推定であることが分かる。この制約条件は  $p=2$  の場合図 1 のようになり、疎性が生じて  $\beta_1$  の値は 0 と推定される。従って、LASSO を適用することにより、回帰とともに重要なパラメータ選択を行うことが可能となる。

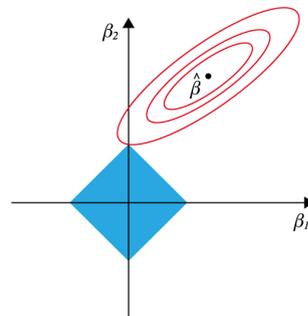


図 1:  $p=2$  のときの制約条件と推定値 ( $\hat{\boldsymbol{\beta}}$  は最小二乗推定の解、青いひし形が制約条件、赤い線と制約条件の交点が LASSO による推定値である。この場合  $\beta_1$  の値は 0 と推定され、最小二乗推定と比べるとスパースな解が得られていることが分かる。)

### 3 実験と考察

本実験では表情の特徴として顔のランドマーク間の距離を使用した。車載カメラにより撮影した運転者の顔画像から抽出した 68 個のランドマーク座標  $L(i) = (l_{xi}, l_{yi})$ ,  $i = 0, \dots, 67$  を使用して、感情を顔に表す際に変化すると予測される 11 個の特徴  $F(j)$ ,  $j = 1, \dots, 11$  を算出した。

$j$  番目の特徴  $F(j)$  は 2 個のランドマーク座標  $L(p) = (l_{xp}, l_{yp})$ ,  $L(q) = (l_{xq}, l_{yq})$  を用いて次のように計算される。

$$F(j) = \sqrt{(l_{xp} - l_{xq})^2 + (l_{yp} - l_{yq})^2}. \quad (6)$$

実際に使用した 11 個の特徴を表 3 に示す。

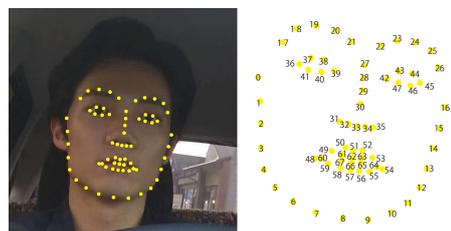


図 2: 運転者顔画像とランドマークの例

顔画像は「苛立ち」と「それ以外」の 2 種類をそれぞれ 100 枚、合計 200 枚用意した。これらの画像は約 59 分の自動車運転中の動画から抽出した 89,630 枚から選出されたものである。実験ではその 200 枚から無作為に 100 枚を選択して「苛立ち」表情の枚数の異なる画像群  $P_1, P_2, P_3$  を作成した。これは「苛立ち」が顔に現れやすい人とそうでない人の場合を模擬することにより、より普遍な特徴群を抽出する狙いがある。各画像群の「苛立ち」表情の含有率を  $I$  とすると、 $P_1, P_2, P_3$  はそれぞれ  $I=20\%$ ,  $I=50\%$ ,  $I=80\%$  である。画像群に含まれる顔画像の違いの結果への影響を

表 1: 使用した特徴とその番号

番号	特徴	備考
E1	左目の開き	L(43) と L(47) の距離
E2	右目の開き	L(38) と L(40) の距離
E3	眉間の距離	L(21) と L(22) の距離
E4	口の開き	L(62) と L(66) の距離
E5	左頬	L(31) と L(39) の距離
E6	右頬	L(35) と L(42) の距離
E7	左口角	L(6) と L(48) の距離
E8	右口角	L(10) と L(54) の距離
E9	鼻の下部と眉間の距離	L(27) と L(33) の距離
E10	左目と左眉の距離	L(19) と L(38) の距離
E11	右目と右眉の距離	L(24) と L(43) の距離

表 2: 各画像群の回帰係数の平均と順位

特徴	平均係数値			平均	順位
	$P_1$	$P_2$	$P_3$		
F1	-3.53	2.34	-3.51	3.13	2
F2	-0.26	0.37	0	0.21	8
F3	-1.64	0.65	-1.32	1.21	4
F4	5.95	-3.62	4.74	4.77	1
F5	0.04	0.09	-0.26	0.13	9
F6	0	0	0	0	11
F7	0.75	-1.22	1.13	1.03	6
F8	1.56	-0.14	1.13	0.94	7
F9	0	-0.13	-0.10	0.08	10
F10	-0.81	-0.94	-1.67	1.14	5
F11	-1.77	3.77	-1.73	2.42	3

考慮して,  $P_1, P_2, P_3$  において, それぞれ含まれる画像が異なるものを 5 個作成し, 平均値により評価を行う。

これらの画像群のランダムマークより算出された特徴行列と目的変数に対して, 統計分析フリーソフト“R”の“glmnet”を使用して LASSO を適用した。LASSO の出力例を図 3 に示す。縦軸は係数値, 横軸は罰則項の係数  $\lambda$  の対数値である。グラフ上部の数値は選択された特徴数を示す。

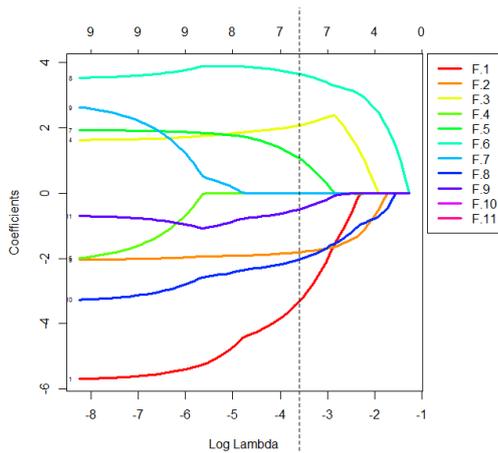


図 3: LASSO の出力例 ( $P_1$  のとき ( $I = 20\%$ ))

この出力結果に対し, one-standard-error rule[5] を適用して罰則項の係数  $\lambda$  を決定する。one-standard-error rule では, 交差検定により定まった二乗誤差が最小となる  $\lambda$  の標準誤差を超えない中で最小の  $\lambda$  を罰則項の係数として与える。ここでは 10 分割交差検定による評価の後, この手法を適用した。

画像群  $P_1, P_2, P_3$  に対し LASSO を適用して得られた係数値の平均値と順位を表 3 に示す。順位付けは  $P_1, P_2, P_3$  の平均係数値の平均値の絶対値の大きさにより行った。この結果から特に重要であると判断された上位 5 個の特徴は「口の開き」, 「目の開き (左)」, 「目と眉の距離 (右)」, 「眉間の距離」, 「目と眉の距離 (左)」であり, これは人間の直観に合致している。

LASSO により得られた回帰を 10 分割交差検定により評価した結果, 正答率は平均して  $P_1$  では約 94.5%,  $P_2$  では約 92.0%,  $P_3$  では約 92.4% であった。正答率から LASSO により選出された特徴群は「苛立ち」と「それ以外」の判別において有用であることが明らかになった。

$I = 80\%$  の場合が最も多様な「苛立ち」表情を含んでいるが,  $I = 20\%$  の場合と比較しても LASSO により抽出された特徴群に大きな変化がないことから, この特徴群は今回使用した「苛立ち」顔画像において共通しているといえる。

今回は 11 個の距離特徴という数が少なく, 単純な特徴を使用したため, 他の表情にも共通するであろう大まかな特徴群の抽出しかできていないが, 「苛立ち」表情に普遍な特徴が存在するならば, 手法の工夫次第ではその特徴を LASSO により抽出可能であると考えられる。

#### 4 結論

LASSO により「苛立ち」表情において人間の直観に合致する特徴群の抽出が可能であることが明らかになった。また交差検定により, この特徴群は「苛立ち」と「それ以外」の二値分類問題において有用であることが示された。謝辞

本研究を進めるにあたり, データ収集にご協力いただきました鈴木研究室の皆様, データ収集参加者の皆様から感謝いたします。本研究の一部は科研費 15K12100 の助成を受けています。

#### 参考文献

- [1] Paul Ekman & Wallace V. Friesen, “Constants across Cultures in the Face and Emotion”, *Journal of personality and social psychology*, 17(2), 124-129, 1971.
- [2] Tsuyoshi Moriyama, Khiaat Abdelaziz, & Noriko Shimomura, “Face Analysis of Aggressive Moods in Automobile Driving Using Mutual Subspace Method”, *21st International Conference on Pattern Recognition*, pp. 2898-2901, 2012.
- [3] 野宮浩揮, 宝珍輝尚, “顔特徴量の有用性推定に基づく特徴抽出による表情認識”, 知能と情報 (日本知能情報ファジィ学会誌), 23(2), 170-185, 2011.
- [4] Robert Tibshirani, “Regression Shrinkage and Selection via the Lasso”, *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1), 267-288, 1996.
- [5] Trevor Hastie, Robert Tibshirani & Martin Wainwright. *Statistical Learning with sparsity: The Lasso and Generalizations*, pp. 7-24, CRC Press, 2016.