



|                  |  |
|------------------|--|
| Title            | 画像の感情推定のためのラベル逆量子化を導入した正準相関分析  |
| Author(s)        | 齊藤, 直輝; 前田, 圭介; 小川, 貴弘; 浅水, 仁; 長谷山, 美紀   |
| Citation         | 電子情報通信学会論文誌 D, J106-D(5), 337-348<br><a href="https://doi.org/10.14923/transinfj.2022jdp7029">https://doi.org/10.14923/transinfj.2022jdp7029</a> |
| Issue Date       | 2023-05-01   |
| Doc URL          | <a href="http://hdl.handle.net/2115/89341">http://hdl.handle.net/2115/89341</a>  |
| Type             | article  |
| File Information | j106-d_5_337.pdf   |



[Instructions for use](#)

# 画像の感情推定のためのラベル逆量子化を導入した正準相関分析

斉藤 直輝<sup>†a)</sup>      前田 圭介<sup>††</sup>      小川 貴弘<sup>††</sup>      浅水 仁<sup>†††</sup>  
 長谷山美紀<sup>††</sup>

## Canonical Correlation Analysis Introducing Label Dequantization for Visual Emotion Recognition

Naoki SAITO<sup>†a)</sup>, Keisuke MAEDA<sup>††</sup>, Takahiro OGAWA<sup>††</sup>, Satoshi ASAMIZU<sup>†††</sup>,  
 and Miki HASEYAMA<sup>††</sup>

あらまし 本論文では、画像の感情推定のためのラベル逆量子化を導入した正準相関分析である Supervised Multi-view Canonical Correlation Analysis via Cyclic Label Dequantization (sMVCCA-CLD) を提案する。ラベルから算出される特徴量 (ラベル特徴量) の次元数は他の特徴量と比較して小さいため、従来の CCA では、構築する空間の次元数の低下により特徴量間の相関関係の表現が困難となる問題点が存在する。そこで、sMVCCA-CLD では、ラベル特徴量の次元数をラベル逆量子化により増加させながら、特徴量間の相関を最大化することで、次元数の制約を受けない共通潜在空間の構築を可能とする。更に、感情が円環状に配置されることを考慮してラベル逆量子化を行うことで、感情推定に適した共通潜在空間の構築を可能とする。以上で構築された空間に射影された新たな特徴量を利用することで、高精度な感情推定が可能となる。

キーワード 感情推定, 正準相関分析, ラベル逆量子化

### 1. ま え が き

情報通信技術の発達及びモバイル端末の普及により、SNS 等には大量かつ多様な画像がアップロードされている。これらの画像は、閲覧した利用者に感情を誘起させることが報告されている [1], [2]。画像等の視覚的な情報により誘起される感情は、マーケティング等の様々な分野における意思決定に影響を与えるため、その推定について研究が進められている [3]~[5]。

一般に Web 上の画像からは複数種類の特徴量を取得することが可能である。例えば、画像共有サービスにアップロードされた画像に対しては、画像から直接導

出可能な特徴量や、画像をアップロードする際に付与された説明文からその意味を表す特徴量を取得することができる。ここで、感情推定に複数種類の特徴量を統合的に利用することで、高精度化が実現可能であることが報告されており、様々な手法が提案されている [6]~[8]。特に、深層学習に基づく手法 [9] は、その推定精度の高さから注目されている。深層学習に基づく感情推定手法には、高精度な推定を実現するモデルを構築するために大量の訓練データが必要となる。しかしながら、画像を閲覧した利用者個人の感情を推定する際、一般的には訓練データとなる画像と感情の組合せを大量に収集することは困難であるため、深層学習に基づく手法では、推定に適した複数種類の特徴量の統合が困難となる可能性が存在する。そこで、少量の学習データで高精度な感情推定を実現するために、異なる種類の特徴量 (異種特徴量) を統合した特徴量表現を取得する方法として、正準相関分析 (Canonical Correlation Analysis: CCA) [10] に基づく手法が種々提案されている [11]~[14]。CCA は二種類の特徴量間の相関を最大化することで、それらの共通潜在空間を構築することができる。また、三種類以上の特徴量に対して共通

<sup>†</sup>北海道大学総合 IR 室, 札幌市

Office of Institutional Research, Hokkaido University, Sapporo-shi, 060-0808 Japan

<sup>††</sup>北海道大学大学院情報科学研究院, 札幌市

Faculty of Information Science and Technology, Hokkaido University, Sapporo-shi, 060-0813 Japan

<sup>†††</sup>釧路工業高等専門学校創造工学科, 釧路市

Department of Creative Engineering, National Institute of Technology, Kushiro College, Kushiro-shi, 084-0916 Japan

a) E-mail: saito@lmd.ist.hokudai.ac.jp

DOI:10.14923/transinfj.2022JDP7029

潜在空間を構築可能な Multi-view CCA (MVCCA) [15] やクラス情報を空間の構築に利用可能となるように拡張した supervised MVCCA (sMVCCA) [16], 深層学習の枠組みを導入した Deep CCA [17] が提案されている。感情推定では、これらの CCA により構築された共通潜在空間へ元の特微量を射影することで得られる新たな特微量を利用している。これにより、異種特微量を統合的に利用し、高精度な感情推定の実現が期待される。

一方, CCA による特微量統合には, 共通潜在空間の最大次元数が利用する全ての特微量の中での最小次元数と等しくなるという制約が存在する。特に, sMVCCA 等で利用する感情のラベルより得られる特微量 (ラベル特微量) は, 画像に付与されているラベルに対応する要素を “1”, それ以外の要素を “0” とすることで, 推定対象の感情の種類 (クラス数) と同じ次元数の one-hot ベクトルとして算出されるため, 他の特微量と比較して次元数が著しく小さくなる。したがって, ラベル特微量を利用した CCA では, 共通潜在空間の次元数が低下し, 特微量間の相関関係を十分に表現することが困難となる可能性が存在する。更に, 大量の画像に対して, 多種類の感情ラベルを付与することには多大な労力を要するため, 推定対象の感情の種類を増やすことによるラベル特微量の次元数増加は困難である。

そこで, 以前に我々は次元数の制約を解消するために Supervised Fractional-order Embedding MVCCA (SFEMCCA) [18] にラベル特微量に対する逆量子化 (ラベル逆量子化) を導入した SFEMCCA via Ordinal Label Dequantization (SFEMCCA-OLD) [19] を提案した。SFEMCCA-OLD では, ラベル逆量子化処理により, 相関を最大化する特微量のうち, ラベル特微量の次元数を増加させることで, 共通潜在空間の次元数も増加させ, CCA における次元数の制約の問題を解決可能とした。ここで, 高次元空間へ元の特微量を射影して, その空間上で CCA を行う Kernel CCA (KCCA) [20], [21] が提案されている。KCCA は通常の CCA では表現が困難な非線形相関を扱うことを目的として, 表現能力の高い高次元空間で線形 CCA の処理を行う。一方で, ラベル逆量子化を導入した CCA は, ラベル情報を特微量として扱う CCA において, ラベル情報の種類が共通潜在空間の次元数に依存するという制約を解消することを目的としている。したがって, KCCA とラベル逆量子化を導入した CCA は, 対象としている問題が大きく異なるため, ラベル特微量

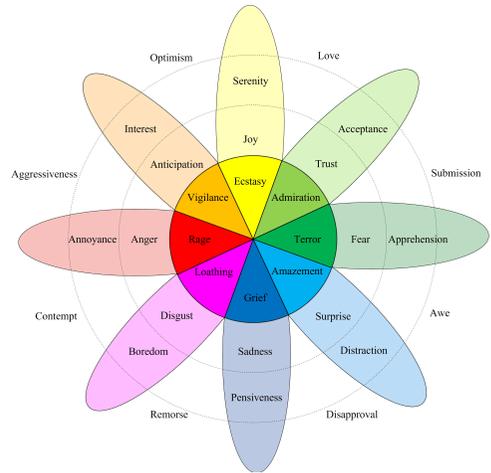


図1 Plutchikの感情円環モデル [22]  
Fig. 1 Plutchik's wheel model of emotions [22].

を利用した CCA における問題の解決にラベル逆量子化の導入は妥当であると考えられる。

先行研究 [19] における推定対象は, 関心の度合いなど順序性のあるラベルであるため, ラベル特微量のベクトルにおいて先頭と最後尾の要素に対応するラベルは全く異なるものとして扱い, ラベル逆量子化を行っている。一方で, 感情は円環状に配置して各感情の関連性を表す「円環モデル」として表現される。例えば, 図1に示す Plutchikの感情円環モデル [22] では, 喜び-悲しみ (joy-sadness), 怒り-恐れ (anger-fear), 嫌悪-信頼 (disgust-trust) 及び驚き-期待 (surprise-anticipation) の4組8感情について類似したものが隣接し, 相反するものが対面するように円環状に配置されている。感情をラベル特微量として表現したとき, 先頭と最後尾の要素は円環上で隣り合う感情であるため, 「類似している感情」として取り扱う必要がある。したがって CCA を用いた感情推定においては, ラベル特微量の先頭と最後尾の要素に対応する感情は高い関連性を有するものであることを考慮したラベル逆量子化を CCA に導入する必要がある。

そこで本論文では, 感情等の円環状で表現されるラベル (円環ラベル) の逆量子化を導入した CCA である sMVCCA via Cyclic Label Dequantization (sMVCCA-CLD) を提案する。sMVCCA-CLD では, ラベル特微量の先頭と最後尾の要素が関連のあるラベルであることを考慮して, 共通潜在空間の次元数の制約を解消するためのラベル逆量子化を行う。これにより, 円環ラ

ベルである感情の推定に適した共通潜在空間を構築可能となるため、推定精度の向上が期待できる。本論文で提案する手法は、先行研究で提案した手法 [19] について、円環ラベルを扱えるように定式化したものである。円環状に配置され順序性をもつラベルは感情のみではなく色や DNA など多様に存在するため、本論文の提案手法は様々な推定問題への活用が期待できる。したがって提案手法は、以前我々が提案した順序性をもつラベルの逆量子化を導入した CCA と比較して非常に高い拡張性を有する。

本論文の構成は以下のとおりである。まず 2. では、感情推定のために拡張する sMVCCA について説明する。次に、3. では、画像の感情の推定精度向上を目的に構築した sMVCCA-CLD について説明する。4. では、感情推定における sMVCCA-CLD の有効性を確認するため、画像の感情を推定する実験を行い、その結果について考察を行う。最後に、5. にまとめを示す。

## 2. sMVCCA

本章では、sMVCCA について説明する。今、 $N$  個のサンプルに対する  $M$  種類の特徴量を  $\mathbf{X}_m = [\mathbf{x}_{m,1}, \mathbf{x}_{m,2}, \dots, \mathbf{x}_{m,N}] \in \mathbb{R}^{D_m \times N}$  ( $m \in \{m_1, m_2, \dots, m_M\}$ ,  $D_m$  は各特徴量の次元数)、及び one-hot ベクトルからなるラベル特徴量を  $\mathbf{l}_i \in \mathbb{R}^{D_l}$  ( $i = 1, 2, \dots, N$ ) とする。ただし、 $m_1, m_2, \dots, m_M$  は各特徴量の種類を表す記号、及び  $l$  はラベル特徴量を表す記号である。また、 $D_l$  は推定対象の感情のクラス数を表し、 $\mathbf{X}_m$  は各行について中心化されているものとする。sMVCCA では、共分散行列  $\mathbf{C}_{mn} \in \mathbb{R}^{D_m \times D_n}$  ( $m, n \in \{m_1, m_2, \dots, m_M, l\}$ ) を次式により算出する。

$$\mathbf{C}_{mn} = \mathbf{X}_m \mathbf{X}_n^\top \quad (1)$$

ただし、 $\mathbf{X}_m$  及び  $\mathbf{X}_n$  において、 $m, n = l$  のときに表される特徴量  $\mathbf{X}_l$  はラベル特徴量であり、 $\mathbf{X}_l = [\mathbf{l}_1, \mathbf{l}_2, \dots, \mathbf{l}_N] \in \mathbb{R}^{D_l \times N}$  と定義される。また、 $m$  または  $n$  に該当する記号が  $M$  種類の特徴量とラベル特徴量の  $M + 1$  個存在するため、共分散行列  $\mathbf{C}_{mn}$  は  $(M + 1)^2$  個存在する。更に、直積空間  $W = \mathbb{R}^{D_{m_1}} \times \mathbb{R}^{D_{m_2}} \times \dots \times \mathbb{R}^{D_{m_M}} \times \mathbb{R}^{D_l}$ 、及びタプル  $\mathcal{W} = (\mathbf{w}_{m_1}, \dots, \mathbf{w}_{m_M}, \mathbf{w}_l)$  を定義するとき、共分散行列  $\mathbf{C}_{mn}$  を用いて、次式の最適化問題を解くことで、共通潜在空間への射影ベクトル  $(\hat{\mathbf{w}}_{m_1}, \dots, \hat{\mathbf{w}}_{m_M}, \hat{\mathbf{w}}_l)$  を算出する。

$$\begin{aligned} & (\hat{\mathbf{w}}_{m_1}, \dots, \hat{\mathbf{w}}_{m_M}, \hat{\mathbf{w}}_l) \\ & = \arg \max_{\mathcal{W} \in W} \left( \sum_m \sum_{n \neq m} \mathbf{w}_m^\top \mathbf{C}_{mn} \mathbf{w}_n \right) \quad (2) \\ & \text{s.t. } \mathbf{w}_m^\top \mathbf{C}_{mm} \mathbf{w}_m = 1 \end{aligned}$$

以上により、sMVCCA は異種特徴量を利用して、ラベル情報を直接反映させた共通潜在空間の構築を可能とする。

## 3. sMVCCA-CLD

本章では、提案手法である sMVCCA-CLD について説明する。今、逆量子化されたラベル特徴量 (逆量子化ラベル特徴量) を  $\mathbf{X}_L = [\mathbf{l}_1^{(\text{deq})}, \mathbf{l}_2^{(\text{deq})}, \dots, \mathbf{l}_N^{(\text{deq})}] \in \mathbb{R}^{D_L \times N}$  ( $D_L = 2D_l$ ) として定義する。sMVCCA-CLD では、この逆量子化ラベル特徴量を最適化することで、ラベル特徴量の次元数が 2 倍になるように逆量子化する。この逆量子化を繰り返すことで、ラベル特徴量の次元数を  $D_l \times 2^r$  ( $r$  は逆量子化を行う回数) 次元まで拡張可能とする。

上記の逆量子化ラベル特徴量に対して、共分散行列  $\mathbf{C}_{mn} \in \mathbb{R}^{D_m \times D_n}$  ( $m, n \in \{m_1, m_2, \dots, m_M, L\}$ ) を前章で説明した sMVCCA と同様に算出する。直積空間  $X \in W \times \mathbb{R}^{D_L \times N}$ 、及びタプル  $\mathcal{X} = (\mathbf{w}_{m_1}, \dots, \mathbf{w}_{m_M}, \mathbf{w}_l, \mathbf{X}_L)$  を定義するとき、sMVCCA-CLD では、ラベル逆量子化により、特徴量間の相関を最大化しつつ、従来の CCA における次元数の制約の影響を受けず、円環ラベルの推定に対応した射影の算出を実現するため、次式の目的関数を構築する。

$$\begin{aligned} & (\hat{\mathbf{w}}_{m_1}, \dots, \hat{\mathbf{w}}_{m_M}, \hat{\mathbf{w}}_{m_L}, \hat{\mathbf{X}}_L) \\ & = \arg \max_{\mathcal{X} \in X} \left( \sum_{m \in S} \sum_{n \in \{S | n \neq m\}} \mathbf{w}_m^\top \mathbf{C}_{mn} \mathbf{w}_n \right) \quad (3) \\ & \text{s.t. } \mathbf{w}_m^\top \mathbf{C}_{mm} \mathbf{w}_m = 1, \end{aligned}$$

$$\begin{aligned} & \|\mathbf{l}_i^{(\text{deq})}\|_0 = \|\mathbf{l}_i^{(\text{deq})}\|_1 = 1 \quad \forall i, \\ & Q\{\mathbf{l}_i^{(\text{deq})}\} = \mathbf{l}_i \quad \forall i \end{aligned}$$

ただし、 $S := \{m_1, m_2, \dots, m_M, L\}$  である。また、共分散行列  $\mathbf{C}_{mn} \in \mathbb{R}^{D_m \times D_n}$  は次式により定義される。

$$\mathbf{C}_{mn} = \begin{cases} \mathbf{X}_m \mathbf{X}_n^\top & \text{if } m, n \neq L \\ \mathbf{X}_m \hat{\mathbf{X}}_L^\top & \text{if } m \neq L \text{ and } n = L \\ \hat{\mathbf{X}}_L \mathbf{X}_n^\top & \text{if } m = L \text{ and } n \neq L \\ \hat{\mathbf{X}}_L \hat{\mathbf{X}}_L^\top & \text{if } m = n = L \end{cases} \quad (4)$$

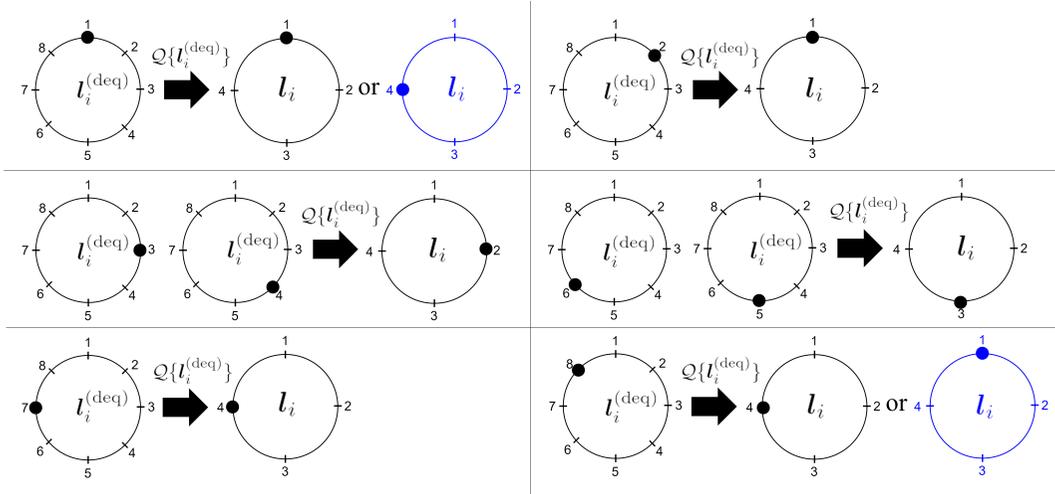


図2 ラベル特徴量が 8 次元の場合の量子化結果一覧。円環状のラベル推定に対応するため、青色で示している部分で、ラベル特徴量の先頭と最後尾の要素を類似しているラベルに対応する要素として量子化している。

Fig. 2 List of the label quantization results of the case in which  $l_i^{(deq)}$  is an eight-dimensional label feature vector. These blue parts show that the first and last elements of  $l_i^{(deq)}$  are regarded as the similar labels for realization of the cyclic label recognition.

$Q\{\cdot\}$  は入力された one-hot ベクトルの量子化を行う関数であり、その次元数を  $\frac{1}{2}$  倍に縮小する。例えば、ラベル特徴量  $l_i^{(deq)}$  が 8 次元の場合の量子化結果の一覧を図 2 に示す。  $Q\{\cdot\}$  では、円環上で隣接するラベル同士を一つのラベルとみなすことで量子化する。このとき、ラベル特徴量の先頭と最後尾の要素を類似しているラベルに対応する要素として扱う。量子化  $Q\{\cdot\}$  を用いた制約により、逆量子化ラベル特徴量  $l_i^{(deq)}$  の値が “1” となる範囲を元のラベル特徴量  $l_i$  の値により制限することができる。 sMVCCA-CLD では、ラベル逆量子化により、推定に利用する特徴量の次元数を適切に増加させ、表現能力を高くすることで、感情推定性能の向上が期待される。ここで、式 (3) の制約付き最適化問題は、「sMVCCA に基づく射影の更新」及び「ラベル特徴量の更新」を交互にあらかじめ決定した反復回数 ( $T$  回) 繰り返し、得られた  $T$  個の射影の中から感情推定に適した「射影の選択」を行うことで解くことができる。 sMVCCA-CLD による感情推定に適した共通潜在空間構築の処理の流れを Algorithm 1 に示す。以降では、sMVCCA-CLD における「sMVCCA に基づく射影の更新」、「逆量子化ラベル特徴量の更新」、及び「射影の選択」について、処理の詳細をそれぞれ説明する。

### Algorithm 1 sMVCCA-CLD

**Input:**  $x_{\tilde{m},i}$  ( $\tilde{m} \in \{m_1, m_2, \dots, m_M\}$ ), feature vectors.  
 $l_i$ ,  $D_l$ -dimensional cyclical label feature vectors.

**Output:** Projection matrices  $\hat{W}_{\tilde{m}}^{(opt)}$ .

Initialization of dequantized label feature vectors  $l_i^{(deq)} \in \mathbb{R}^{2D_l}$ .

**for**  $t = 1 \dots T$  **do**

Calculation of  $t$ -th projection matrix  $\hat{W}^{(t)}$  via sMVCCA in subsection 3.1.

Calculation of dequantized label feature vectors  $l_i^{(deq)} \in \mathbb{R}^{2D_l}$  in subsection 3.2.

**end for**

Selection of optimal projection matrices  $\hat{W}_{\tilde{m}}^{(opt)}$  in subsection 3.3.

### 3.1 sMVCCA に基づく射影の更新

sMVCCA-CLD では、逆量子化ラベル特徴量  $\hat{X}_L$  を固定し、次式の最適化問題を解くことで射影を更新する。

$$\begin{aligned}
 & (\hat{w}_{m_1}, \dots, \hat{w}_{m_M}, \hat{w}_L) \\
 & = \arg \max_{W \in \mathcal{W}} \left( \sum_{m \in S} \sum_{n \in \{S | n \neq m\}} w_m^\top C_{mn} w_n \right) \quad (5) \\
 & \text{s.t. } w_m^\top C_{mm} w_m = 1
 \end{aligned}$$

なお、元のラベル特徴量  $X_L$  を構成するベクトル  $l_i$  ( $i = 1, 2, \dots, N$ ) において要素が 1 となっている次元を  $D_1$  とするとき、式 (5) の最適化問題における逆量子化

ラベル特徴量  $\mathbf{l}_i^{(\text{deq})}$  の初期値は、 $2 \times D_1$  次元目の要素のみが 1 となっている one-hot ベクトルとする。例えば、元のラベル特徴量が  $\mathbf{l}_i = [1, 0]^T$  であるとき、逆量子化ラベル特徴量の初期値は  $\mathbf{l}_i^{(\text{deq})} = [0, 1, 0, 0]^T$  となる。式 (5) は sMVCCA の目的関数と等価であり、最終的に次式の一般化固有値問題に帰着される。

$$\begin{aligned}
 & \begin{bmatrix} \mathbf{0} & C_{m_1 m_2} & \cdots & C_{m_1 m_M} & C_{m_1 L} \\ C_{m_2 m_1} & \mathbf{0} & & C_{m_2 m_M} & C_{m_2 L} \\ \vdots & & \ddots & & \vdots \\ C_{m_M m_1} & C_{m_M m_2} & & \mathbf{0} & C_{m_M L} \\ C_{L m_1} & C_{L m_2} & \cdots & C_{L m_M} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{w}_{m_1} \\ \mathbf{w}_{m_2} \\ \vdots \\ \mathbf{w}_{m_M} \\ \mathbf{w}_L \end{bmatrix} \\
 = \eta & \begin{bmatrix} C_{m_1 m_1} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & C_{m_2 m_2} & & \mathbf{0} & \mathbf{0} \\ \vdots & & \ddots & & \vdots \\ \mathbf{0} & \mathbf{0} & & C_{m_M m_M} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & C_{LL} \end{bmatrix} \begin{bmatrix} \mathbf{w}_{m_1} \\ \mathbf{w}_{m_2} \\ \vdots \\ \mathbf{w}_{m_M} \\ \mathbf{w}_L \end{bmatrix} \quad (6)
 \end{aligned}$$

sMVCCA-CLD では、式 (6) を解くことで得られる固有値  $\eta$  ( $= \eta_1, \eta_2, \dots, \eta_D$ ) に対応する固有ベクトルを射影  $\{\hat{\mathbf{w}}_{m,1}, \hat{\mathbf{w}}_{m,2}, \dots, \hat{\mathbf{w}}_{m,D}\}$  として算出する。ただし、 $D \leq \min\{D_m\}$  であり、多くの場合、 $\min\{D_m\} = D_L$  となる傾向があるため  $D$  は  $D_L$  に依存する。sMVCCA-CLD では、次節で説明する処理によってラベル特徴量を逆量子化することで、 $D_L$  を増加させることが可能となり、次元数の制約の問題を解決可能とする。

### 3.2 ラベル特徴量の更新

sMVCCA-CLD では、次式の最適化問題を解くことで、逆量子化ラベル特徴量  $\mathbf{X}_L$  を更新する。

$$\begin{aligned}
 \hat{\mathbf{X}}_L &= \arg \max_{\mathbf{X}_L} \left( \sum_{d=1}^D \sum_{m \in \bar{S}} \sum_{n \in \{S | n \neq m\}} \hat{\mathbf{w}}_{m,d}^\top \mathbf{X}_m \mathbf{X}_n^\top \hat{\mathbf{w}}_{n,d} \right) \\
 \text{s.t. } & \|\mathbf{l}_i^{(\text{deq})}\|_0 = \|\mathbf{l}_i^{(\text{deq})}\|_1 = 1 \quad \forall i, \\
 & Q\{\mathbf{l}_i^{(\text{deq})}\} = \mathbf{l}_i \quad \forall i \quad (7)
 \end{aligned}$$

ここで、上記の制約付き最適化問題は射影  $\hat{\mathbf{w}}_{m,d}$  及び  $\hat{\mathbf{w}}_{n,d}$  を前節の「sMVCCA に基づく射影の更新」で算出された値に固定して解くため、式 (7) は次式のように書き換えられる。

$$\begin{aligned}
 \hat{\mathbf{X}}_L &= \arg \max_{\mathbf{X}_L} \left( \sum_{d=1}^D \sum_{\bar{m} \in \bar{S}} \hat{\mathbf{w}}_{\bar{m},d}^\top \mathbf{X}_{\bar{m}} \mathbf{X}_L^\top \hat{\mathbf{w}}_{L,d} \right) \\
 \text{s.t. } & \|\mathbf{l}_i^{(\text{deq})}\|_0 = \|\mathbf{l}_i^{(\text{deq})}\|_1 = 1 \quad \forall i, \\
 & Q\{\mathbf{l}_i^{(\text{deq})}\} = \mathbf{l}_i \quad \forall i \quad (8)
 \end{aligned}$$

ただし、 $\bar{S} = S \setminus L$  である。更に、式 (8) の  $\sum_{\bar{m} \in \bar{S}} \hat{\mathbf{w}}_{\bar{m},d}^\top \mathbf{X}_{\bar{m}} \mathbf{X}_L^\top \hat{\mathbf{w}}_{L,d}$  は次式のように展開できる。

$$\begin{aligned}
 & \sum_{\bar{m} \in \bar{S}} \hat{\mathbf{w}}_{\bar{m},d}^\top \mathbf{X}_{\bar{m}} \mathbf{X}_L^\top \hat{\mathbf{w}}_{L,d} \\
 &= \sum_{\bar{m} \in \bar{S}} \hat{\mathbf{w}}_{\bar{m},d}^\top \mathbf{x}_{\bar{m},1} \mathbf{l}_1^{(\text{deq})\top} \hat{\mathbf{w}}_{L,d} + \cdots \\
 &+ \sum_{\bar{m} \in \bar{S}} \hat{\mathbf{w}}_{\bar{m},d}^\top \mathbf{x}_{\bar{m},i} \mathbf{l}_i^{(\text{deq})\top} \hat{\mathbf{w}}_{L,d} + \\
 &\cdots + \sum_{\bar{m} \in \bar{S}} \hat{\mathbf{w}}_{\bar{m},d}^\top \mathbf{x}_{\bar{m},N} \mathbf{l}_N^{(\text{deq})\top} \hat{\mathbf{w}}_{L,d} \quad (9)
 \end{aligned}$$

ここで、式 (9) の第  $i$  項の最大化を考える。逆量子化の対象となる元のラベル特徴量  $\mathbf{l}_i$  において、先頭若しくは最後尾の要素が“1”ではない場合は、 $\hat{\mathbf{w}}_{\bar{m},d}$  を固定しているため、 $\sum_{\bar{m} \in \bar{S}} \hat{\mathbf{w}}_{\bar{m},d}^\top \mathbf{x}_{\bar{m},i}$  は定数となる。したがって、 $\sum_{\bar{m} \in \bar{S}} \hat{\mathbf{w}}_{\bar{m},d}^\top \mathbf{x}_{\bar{m},i} \mathbf{l}_i^{(\text{deq})\top} \hat{\mathbf{w}}_{L,d}$  の最大化は  $\sum_{\bar{m} \in \bar{S}} \hat{\mathbf{w}}_{\bar{m},d}^\top \mathbf{x}_{\bar{m},i}$  の値の正負に応じて  $\mathbf{l}_i^{(\text{deq})\top} \hat{\mathbf{w}}_{L,d}$  を最大化あるいは最小化する  $\mathbf{l}_i^{(\text{deq})}$  を決定することで実現可能である。ここで、 $\|\mathbf{l}_i^{(\text{deq})}\|_0 = \|\mathbf{l}_i^{(\text{deq})}\|_1 = 1$  の制約が存在するため、 $\mathbf{l}_i^{(\text{deq})}$  は one-hot ベクトルとなる。更に、 $Q\{\mathbf{l}_i^{(\text{deq})}\} = \mathbf{l}_i$  の制約により、ラベル特徴量  $\mathbf{l}_i$  の  $p$  ( $= 2, 3, \dots, D_1 - 1$ ) 番目の要素が“1”であるとき、逆量子化ラベル特徴量の  $(2p)$  あるいは  $(2p-1)$  番目の要素が“1”、他の要素は“0”となる。したがって、 $\mathbf{l}_i$  の  $p$  番目の要素が“1”であるとき、 $\sum_{\bar{m} \in \bar{S}} \hat{\mathbf{w}}_{\bar{m},d}^\top \mathbf{x}_{\bar{m},i} \mathbf{l}_i^{(\text{deq})\top} \hat{\mathbf{w}}_{L,d}$  の最大値は、定数  $\sum_{\bar{m} \in \bar{S}} \hat{\mathbf{w}}_{\bar{m},d}^\top \mathbf{x}_{\bar{m},i}$  の値の正負に応じて射影  $\hat{\mathbf{w}}_{L,d}$  の  $(2p)$  あるいは  $(2p-1)$  番目の要素から最大あるいは最小の要素を選択することに等価である。具体的に、ラベル特徴量  $\mathbf{l}_i$  の  $p$  番目の要素が“1”かつ  $\sum_{\bar{m} \in \bar{S}} \hat{\mathbf{w}}_{\bar{m},d}^\top \mathbf{x}_{\bar{m},i} \geq 0$  であるとき、 $(2p)$  及び  $(2p-1)$  番目の要素から、最大の要素を選択する。選択した要素と対応する  $\mathbf{l}_i^{(\text{deq})}$  の要素は“1”となり、他の要素は“0”となる。

次に、元のラベル特徴量  $\mathbf{l}_i$  の先頭若しくは最後尾の要素が“1”である場合について説明する。先頭若しくは最後尾の要素が“1”ではない場合と同様に、 $\hat{\mathbf{w}}_{\bar{m},d}$

**Algorithm 2** 逆量子化ラベル特徴量の更新

**Input:**  $\hat{\boldsymbol{w}}_{m,d}$ , projection vectors.  
 $\boldsymbol{x}_{\bar{m},i}$ , feature vectors.  
 $\boldsymbol{l}_i$ , cyclical label feature vectors.  
**Output:**  $\boldsymbol{l}_i^{(\text{deq})}$ , a dequantized feature vector.

**if** The first element of  $\boldsymbol{l}_i$  is “1”, **then**  
  **if**  $\sum_{\bar{m} \in \bar{S}} \hat{\boldsymbol{w}}_{m,d}^\top \boldsymbol{x}_{\bar{m},i} \geq 0$ , **then**  
    The index is obtained by selecting the largest element from the first, second or last element in  $\hat{\boldsymbol{w}}_{L,d}$ .  
  **else**  
    The index is obtained by selecting the smallest element from the first, second or last element in  $\hat{\boldsymbol{w}}_{L,d}$ .  
  **end if**  
**else if** The last element of  $\boldsymbol{l}_i$  is “1”, **then**  
  **if**  $\sum_{\bar{m} \in \bar{S}} \hat{\boldsymbol{w}}_{m,d}^\top \boldsymbol{x}_{\bar{m},i} \geq 0$ , **then**  
    The index is obtained by selecting the largest element from the  $(D_l - 1)$ -th, last or first element in  $\hat{\boldsymbol{w}}_{L,d}$ .  
  **else**  
    The index is obtained by selecting the smallest element from the  $(D_l - 1)$ -th, last or first element in  $\hat{\boldsymbol{w}}_{L,d}$ .  
  **end if**  
**else if** The  $p$ -th element of  $\boldsymbol{l}_i$  is “1”, **then**  
  **if**  $\sum_{\bar{m} \in \bar{S}} \hat{\boldsymbol{w}}_{m,d}^\top \boldsymbol{x}_{\bar{m},i} \geq 0$ , **then**  
    The index is obtained by selecting the larger element from the  $(2p)$ -th or  $(2p - 1)$ -th element in  $\hat{\boldsymbol{w}}_{L,d}$ .  
  **else**  
    The index is obtained by selecting the smaller element from the  $(2p)$ -th or  $(2p - 1)$ -th element in  $\hat{\boldsymbol{w}}_{L,d}$ .  
  **end if**  
  Set the element of  $\boldsymbol{l}_i^{(\text{deq})}$  corresponding to the obtained index to “1” and set the other elements to “0”.  
**end if**  
**return**  $\boldsymbol{l}_i^{(\text{deq})}$

を固定しているため、 $\sum_{\bar{m} \in \bar{S}} \hat{\boldsymbol{w}}_{m,d}^\top \boldsymbol{x}_{\bar{m},i}$  は定数となり、 $\sum_{\bar{m} \in \bar{S}} \hat{\boldsymbol{w}}_{m,d}^\top \boldsymbol{x}_{\bar{m},i}$  の値の正負に応じて  $\boldsymbol{l}_i^{(\text{deq})}$  の  $\hat{\boldsymbol{w}}_{L,d}$  を最大化あるいは最小化する  $\boldsymbol{l}_i^{(\text{deq})}$  を決定する。ここで、ラベル特徴量  $\boldsymbol{l}_i$  の先頭の要素が “1” であるとき、逆量子化ラベル特徴量の 1, 2,  $D_l$  番目のいずれかの要素のみ “1” となる。したがって、 $\boldsymbol{l}_i$  の先頭の要素が “1” であるとき、 $\sum_{\bar{m} \in \bar{S}} \hat{\boldsymbol{w}}_{m,d}^\top \boldsymbol{x}_{\bar{m},i} \boldsymbol{l}_i^{(\text{deq})}$  の最大値は、定数  $\sum_{\bar{m} \in \bar{S}} \hat{\boldsymbol{w}}_{m,d}^\top \boldsymbol{x}_{\bar{m},i}$  の値の正負に応じて射影  $\hat{\boldsymbol{w}}_{L,d}$  の 1, 2, 及び  $D_l$  番目の要素から最大あるいは最小の要素を選択することに等価である。一方で、ラベル特徴量  $\boldsymbol{l}_i$  の最後尾の要素が “1” であるとき、逆量子化ラベル特徴量の  $D_l - 1$ ,  $D_l$ , 及び 1 番目のいずれかの要素のみ “1” となる。そこで、最後尾の要素が “1” であるときは、 $\sum_{\bar{m} \in \bar{S}} \hat{\boldsymbol{w}}_{m,d}^\top \boldsymbol{x}_{\bar{m},i}$  の正負に応じて  $\hat{\boldsymbol{w}}_{L,d}$  の  $D_l - 1$ ,  $D_l$ , 及び 1 番目の要素から最大あるいは最小の要素を選択する。以上の処理の詳細を Algorithm 2 に示す。sMVCCA-CLD では、射影とラベル特徴量の更新を交互に  $T$  回反復し、次節で説明する射影の選択を行うことで、ラベル特徴量の次元数を 2 倍に拡張可能とする。これを目標の次元数に達するまで繰り返すことでラベルの逆量子化を実現する。

**3.3 射影の選択**

sMVCCA-CLD で扱うスパース最適化問題は凸問題とならないため、その収束は保証されていない。したがって、射影とラベル特徴量の更新を反復することで得られた射影から、感情の推定に適した射影の選択を行う。まず、 $t (= 1, 2, \dots, T)$  回目の反復で得られた射影行列を次式により定義する。

$$\hat{\boldsymbol{W}}_{\bar{m}}^{(t)} = [\hat{\boldsymbol{w}}_{\bar{m},1}^{(t)}, \hat{\boldsymbol{w}}_{\bar{m},2}^{(t)}, \dots, \hat{\boldsymbol{w}}_{\bar{m},D}^{(t)}] \in \mathbb{R}^{D_{\bar{m}} \times D} \quad (10)$$

ただし、 $\hat{\boldsymbol{w}}_{\bar{m},d}^{(t)}$  は  $t$  回目の反復で得られる射影ベクトルである。得られた行列の中から、推定に適した射影行列  $\hat{\boldsymbol{W}}_{\bar{m}}^{(t_{\text{opt}})}$  を次式により選択する。

$$t_{\text{opt}} = \arg \max_t \left( \sum_{\bar{m} \in \bar{S}} \sum_{\bar{n} \in \{\bar{S} | \bar{n} \neq \bar{m}\}} \text{Tr}(\hat{\boldsymbol{W}}_{\bar{m}}^{(t)\top} \boldsymbol{X}_{\bar{m}} \boldsymbol{X}_{\bar{n}}^\top \hat{\boldsymbol{W}}_{\bar{n}}^{(t)}) \right) \quad (11)$$

sMVCCA-CLD では、得られた射影行列  $\hat{\boldsymbol{W}}_{\bar{m}}^{(t_{\text{opt}})}$  を用いることで次式により、各特徴量を共通潜在空間へ射影することが可能となる。

$$\hat{\boldsymbol{X}}_{\bar{m}}^{(t_{\text{opt}})} = \hat{\boldsymbol{W}}_{\bar{m}}^{(t_{\text{opt})\top} \boldsymbol{X}_{\bar{m}} \in \mathbb{R}^{D \times N} \quad (12)$$

以上より、次元数の制約を受けない共通潜在空間の構築を実現する。

**4. 実 験**

本章では、sMVCCA-CLD による画像の感情推定の有効性を検証するために行った実験の結果について説明する。本実験では、Emotion-6 (E6) データセット [23] 及び Flickr-Instagram (FI) データセット [24] の二種類のデータセットについてそれぞれ感情推定を行った。E6 データセットでは、Plutchik の感情円環モデル [22] における joy, love, fear, surprise, sadness, 及び anger の六つの感情のいずれかを誘起する計 8,017 枚<sup>(注1)</sup> の画像を用いた。一方で、FI データセットでは、図 3 に示す Mikel の感情円環モデル [25] における Fear, Sadness, Disgust, Anger, Amusement, Contentment, Awe 及び Excitement の八つの感情のいずれかを誘起する計 21,829 枚<sup>(注2)</sup> の画像を用いた。本実験では、こ

(注1)：データセットに含まれる 8,350 枚の画像の中で、Web 上でアクセス可能な画像のみを用いた。

(注2)：データセットに含まれる 23,308 枚の画像の中で、Web 上でアクセス可能な画像のみを用いた。



図3 Mikelの感情円環モデル [25]  
Fig. 3 Mikel's wheel from the psychological model [25].

これらのデータセットに対して、5分割交差検証により訓練サンプルとテストサンプルに分割し、評価を行った。各データセットに含まれる画像から誘起される感情を推定するため、以下のとおりに算出した画像特徴量及びキャプション特徴量の二種類の特徴量を利用した。

**画像特徴量 (1,280 次元)**

特徴量を算出する画像を 224×224 画素にリサイズし、ImageNet で学習済みの畳み込みニューラルネットワーク [26] の一つである EfficientNet-B0 [27] へ画像を入力した際に最終層の直前の層から出力される 1,280 個の数値を並べたベクトルを画像特徴量とした。

**キャプション特徴量 (164 次元)**

MSCOCO'14 データセット [28] を用いて学習済みの image captioning モデルである show, attend and tell モデル [29] において、ImageNet により学習された ResNet-101 [30] に基づくエンコーダより出力されるサイズが 14×14×2,048 のテンソルを行列として表現した 196×2,048 の行列に対して、フィルタサイズを 50×50 とする平均値プーリングを行うことで 164 次元のキャプション特徴量を算出した。

これらの特徴量に対して、sMVCCA-CLD を用いることで射影を求め、画像特徴量を共通潜在空間へ射影することで新たな画像特徴量を算出した。本実験では、最適化時の反復回数を実験的に  $T = 50$  と設定した。

本実験では、sMVCCA-CLD の有効性を検証するため、CCA [10], DCCA [31], LDCCA [32], DLPCA [33], sMVCCA [16], GrMCCA [34], LapMCCA [35], DeepCCA [17], 及び sMVCCA-OLD [19] の 9 種類の比較手法を用いた。CCA は最も基本的な特徴量統合手法であり、比較手法として用いた。また、DCCA 及び LDCCA は、判別分析に基づいてクラス情報を用いる手法であ

るため、クラス情報を特徴量の一つとして直接用いることの有効性を検証するために用いた。DLPCA は映像や生体情報等の異種特徴量を用いたユーザに特化したアプリケーションに関する研究において有効性が認められており [36], [37], 本実験の設定と一致しているため比較手法として用いた。sMVCCA はラベルの逆量子化を導入する有効性を検証するために比較手法として用いた。更に、相関構造を利用する手法と幾何構造を利用する手法を組み合わせた GrMCCA 及びデータの近傍構造を保存した共通潜在空間を構築することでデータの幾何構造を考慮可能とする LapMCCA は、異種特徴量を統合する CCA のベンチマークとして用いられるため [38], 本実験の比較手法として採用した。また、DeepCCA は、近年注目を集めている深層学習を用いた手法であるため、比較手法として用いた。ただし、本実験で用いた DeepCCA のモデルは 3 層の隠れ層をもつ。更に、このモデルは rmsprop により最適化され、バッチサイズは 32, エポック数は 200 及び学習率は 0.01 とした。sMVCCA-OLD は感情が円環状に配置されることを考慮してラベル逆量子化を行う有効性を検証するために比較手法として用いた。これらの比較手法に加え、目的関数の最大化によるラベル逆量子化処理により、推定性能が向上していることを検証するため、sMVCCA-CLD において、無作為の次元に "1" を入れることでラベル逆量子化を行う手法をベースラインとした。

本実験では、CCA により得られた画像及びキャプション特徴量に対する射影の中で、「画像特徴量に対する射影」のみを利用した。これにより、画像特徴量を共通潜在空間へ射影することで得られた新たな特徴量を  $K$ -Nearest Neighbor ( $K$ -NN) アルゴリズム [39] による推定器へ入力することで感情を推定した。なお、 $K$ -NN アルゴリズムにおける近傍数は  $K = 50$  とした。本実験では、以下に示す感情円環上での誤差の平均 (Mean Absolute Cyclic Error: MACE), 平均絶対誤差 (Mean Absolute Error: MAE), 及び Classification Accuracy (以降、Accuracy) を評価指標とした。

$$MACE = \frac{1}{N} \sum_{i=1}^N \min(d_i, |N_{\text{label}} - d_i|) \quad (13)$$

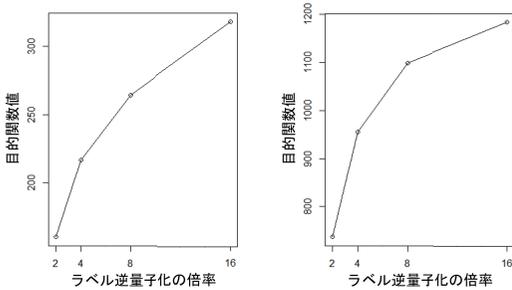
$$MAE = \frac{1}{N} \sum_{i=1}^N |d_i| \quad (14)$$

$$Accuracy = \frac{N_{\text{correct}}}{N} \quad (15)$$

表 1 提案手法及び比較手法による感情推定結果の評価指標

Table 1 Evaluation indices of the emotion recognition results obtained by the proposed method and the comparative methods.

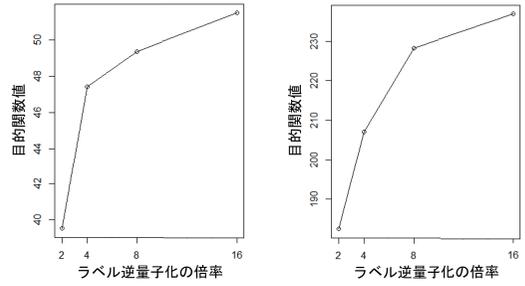
|                 | E6  |              |             |              | FI  |              |              |              |
|-----------------|-----|--------------|-------------|--------------|-----|--------------|--------------|--------------|
|                 | 次元数 | MACE(↓)      | MAE(↓)      | Accuracy(↑)  | 次元数 | MACE(↓)      | MAE(↓)       | Accuracy(↑)  |
| CCA [10]        | 41  | 1.01         | 1.41        | 0.444        | 82  | 0.933        | 1.04         | 0.592        |
| DCCA [31]       | 6   | 1.08         | 1.48        | 0.369        | 8   | 1.09         | 1.19         | 0.480        |
| LDCCA [32]      | 41  | 1.03         | 1.44        | 0.440        | 82  | 0.920        | 1.04         | 0.595        |
| DLPCCA [33]     | 41  | 1.05         | 1.43        | 0.426        | 82  | 0.952        | 1.06         | 0.577        |
| sMVCCA [16]     | 6   | 0.999        | 1.37        | 0.449        | 8   | 0.965        | 1.07         | 0.581        |
| GrMVCCA [34]    | 82  | 1.01         | 1.41        | 0.449        | 82  | 0.931        | 1.05         | 0.589        |
| LapMVCCA [35]   | 41  | 1.01         | 1.42        | 0.443        | 82  | 0.930        | 1.05         | 0.591        |
| DeepCCA [17]    | 41  | 1.05         | 1.46        | 0.434        | 41  | 1.15         | 1.53         | 0.320        |
| sMVCCA-OLD [19] | 24  | 0.996        | 1.38        | 0.446        | 64  | 0.895        | 0.996        | <b>0.605</b> |
| ベースライン          | 24  | 1.05         | 1.46        | 0.427        | 64  | 0.954        | 1.07         | 0.580        |
| sMVCCA-CLD      | 96  | <b>0.980</b> | <b>1.36</b> | <b>0.459</b> | 32  | <b>0.891</b> | <b>0.990</b> | <b>0.605</b> |



(a) E6 データセット (b) FI データセット

図 4 sMVCCA-CLD による逆量子化ステップごとの訓練データに対する目的関数値

Fig. 4 Object function value of sMVCCA-CLD for training data at each label dequantization step.



(a) E6 データセット (b) FI データセット

図 5 sMVCCA-CLD による逆量子化ステップごとのテストデータに対する目的関数値

Fig. 5 Object function value of sMVCCA-CLD for test data at each label dequantization step.

ただし、 $N_{\text{label}}$ 、及び  $N_{\text{correct}}$  はそれぞれ推定対象の感情の種類及び正しく推定された画像数を表し、 $d_i$  ( $i = 1, 2, \dots, N$ ;  $N$  はデータセット内の全画像数) は  $i$  番目の画像における推定されたラベルと真のラベルの感情円環上で時計回りに離れているラベル数を表す。MACE の  $|N_{\text{label}} - d_i|$  は、感情円環上で反時計回りに離れているラベル数を表すため、式 (13) で求められる MACE は、推定結果が感情円環上において平均的にどの程度離れているかを表す。MACE 及び MAE は値が小さいほど、推定誤差が小さい結果であることを示し、Accuracy は大きいほど、推定結果と実際の感情の種類が一致している画像数が多いことを表す。

本実験では、以下に示す「次元数がラベルの種類に依存しない手法 (CCA, LDCCA, DLPCCA, GrMVCCA, LapMVCCA, DeepCCA)」、 「次元数がラベルの種類に依存する手法 (DCCA, sMVCCA)」、及び「ラベル特徴量の逆量子化を行う手法 (sMVCCA-OLD, ベースライ

表 2 ラベル特徴量  $I_i$  の両端の要素に対応する感情ラベルに属するサンプルの推定結果の MACE

Table 2 MACE of the recognition results of the samples belonging to emotion labels that correspond to elements at both ends of label features  $I_i$ .

|                 | E6           | FI          |
|-----------------|--------------|-------------|
| sMVCCA-OLD [19] | 0.842        | 1.35        |
| sMVCCA-CLD      | <b>0.706</b> | <b>1.28</b> |

ン, sMVCCA-CLD)」の三つに分けて共通潜在空間の次元数を設定した。

(1) 次元数がラベルの種類に依存しない手法 (CCA, LDCCA, DLPCCA, GrMVCCA, LapMVCCA, DeepCCA)

設定可能な共通潜在空間の最大次元数は、キャプション特徴量の次元数と同じ 164 次元であるため、{164, 123, 82, 41} の 4 種類の次元を設定し、 $K$ -NN アルゴリズムによる推定結果で MACE が最も低い値になった次元数を採用した。

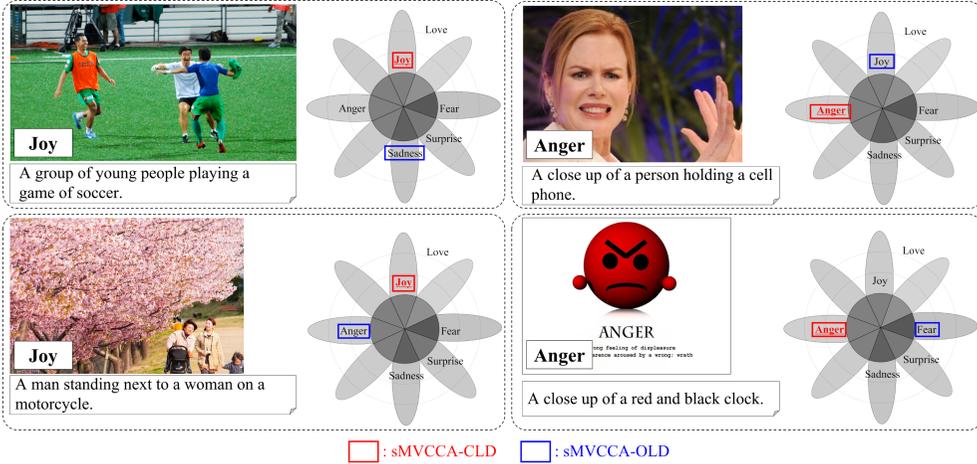


図6 E6 データセットにおける推定結果の例  
 Fig. 6 Example of the emotion recognition on E6 dataset.

(2) 次元数がラベルの種類に依存する手法 (DCCA, sMVCCA)

設定可能な共通潜在空間の最大次元数は、E6 データセットでは6次元、FI データセットでは8次元となる。本実験では、1次元から最大次元まで1次元ずつ増加させながら共通潜在空間を構築し、K-NN アルゴリズムによる推定結果の MACE が最も低くなった次元数を採用した。

(3) ラベル特徴量の逆量子化を行う手法 (sMVCCA-OLD, ベースライン, sMVCCA-CLD)

ラベル特徴量の次元数を増加させることで、共通潜在空間の最大次元数を2, 4, 8, 及び16倍に拡張し、K-NN アルゴリズムによる推定結果で MACE が最も低い値となった次元数を採用した。

提案手法及び比較手法により構築された共通潜在空間の次元数及び推定結果の評価指標をデータセットごとに表1に示す。E6 及び FI データセット共に、提案手法である sMVCCA-CLD による感情推定結果が全ての評価指標で最良であることが確認できる。ここで、sMVCCA-OLD 及び sMVCCA-CLD により最良の推定結果を得るために構築された共通潜在空間の次元数が sMVCCA と比較して高くなっていることが確認できる。更に、推定結果の MACE も sMVCCA より減少している。これらの結果より、ラベル逆量子化により共通潜在空間の次元数の制約を解消する有効性が確認できる。これは、共通潜在空間の表現能力の低下を抑制することで、推定に必要な情報が欠落しなかったため

であると考えられる。また sMVCCA-CLD では、ラベルの円環構造も考慮しているため、MACE が最も低くなっている。sMVCCA-CLD 及び sMVCCA-OLD 以外の手法は、共通潜在空間を構築する際に、感情の順序性を考慮せずに「分類問題」として扱うため、順序性に関わる MACE 及び MAE の値が大きくなっていると考えられる。以降では、sMVCCA-CLD と各手法の評価指標を比較することで、提案手法の有効性について説明する。まず、CCA と比較して提案手法による推定誤差が小さいことが確認された。また、DCCA, LDCCA, 及び DLPCA よりも提案手法の誤差が小さいことから、共通潜在空間の構築にラベル特徴量を利用する有効性が確認された。更に、提案手法の誤差が sMVCCA の誤差を下回っていることから、ラベル逆量子化の枠組みを導入した有効性が確認できる。ラベルの種類数によって、共通潜在空間の次元数の制約を受ける CCA である DCCA 及び sMVCCA による推定結果の Accuracy が sMVCCA-CLD を下回っていることが確認できる。この結果より、円環構造を考慮したラベル逆量子化を sMVCCA に導入することで、感情推定における「分類問題」としての性能も向上していることが確認できる。

また、GrMCCA 及び LapMCCA と比較して、提案手法は高精度な感情推定が可能であることが確認された。更に、深層学習を用いた手法である DeepCCA における感情の推定誤差が提案手法よりも大きくなることが確認された。この原因として、本実験で使用した

二種類のデータセットにおけるサンプル数が深層学習による推定モデルの学習に十分ではなかったためであると考えられる。

sMVCCA-CLD の推定誤差が sMVCCA-OLD より小さいことが確認できる。ここで、E6 データセット及び FI データセットについて、図4及び5に sMVCCA-CLD においてラベル特徴量の次元数を 2, 4, 8, 及び 16 倍に増加させたとき、訓練データ及びテストデータにおける式 (3) に示した目的関数の値をプロットしたグラフをそれぞれ示す。これらのグラフより、E6 及び FI データセットの両方において、ラベル特徴量の次元数の増加に伴い、訓練及びテストデータにおける目的関数の値も増加していることが確認できる。

更に、各データセットにおいて、ラベル特徴量の先頭及び最後尾の要素に対応する感情ラベルのサンプルのみで算出した sMVCCA-CLD 及び sMVCCA-OLD による推定結果の MACE を表2に示す。表2より、ラベル特徴量の両端に対応する感情ラベル<sup>(注3)</sup>については、推定誤差が小さくなっていることが確認できる。また、図6に E6 データセットにおいて、ラベル特徴量の両端の要素に対応する感情である“joy”及び“anger”の画像とその推定結果の一例を示す。図6の結果より、sMVCCA-OLD で誤推定した画像に対して、sMVCCA-CLD では円環ラベルの逆量子化を導入することで、正しい感情の推定が可能となることが確認できる。

最後に、ラベル逆量子化処理を無作為に行うベースラインの評価指標を sMVCCA-CLD が大きく上回っていることから、目的関数の最大化により、ラベル逆量子化を行う有効性が確認できる。以上より、感情推定において円環ラベルに対する逆量子化を CCA に導入する有効性が確認された。

## 5. む す び

本論文では、ラベル逆量子化を導入した CCA を感情等の円環ラベルの推定に利用できるように拡張した sMVCCA-CLD を提案した。sMVCCA-CLD では、円環ラベルに対する逆量子化によりラベル特徴量の次元数を増加させることで、次元数の制約による共通潜在空間の表現能力が低下する問題を解決可能とした。二種類のデータセットを利用した実験の結果より、sMVCCA-CLD による画像の感情推定の有効性を確認

した。

ここで感情円環モデルの表現方法として、極座標を用いた手法が提案されている [40]。極座標による表現により、ラベル間の隣接関係を保持することが可能となり、特徴量間の相関関係を見出しやすくなることが期待できる。したがって、極座標によるラベル特徴量の表現の導入が今後の検討課題としてあげられる。

謝辞 本研究の一部は、JSPS 科研費 JP21H03456 の助成により行われた。

## 文 献

- [1] D. Joshi, R. Datta, E. Fedorovskaya, Q.-T. Luong, J.Z. Wang, J. Li, and J. Luo, “Aesthetics and emotions in images,” *IEEE Signal Processing Magazine*, vol.28, no.5, pp.94–115, 2011.
- [2] P.J. Lang, “A bio-informational theory of emotional imagery,” *Psychophysiology*, vol.16, no.6, pp.495–512, 1979.
- [3] J. Zhang, X. Liu, M. Chen, Q. Ye, and Z. Wang, “Image sentiment classification via multi-level sentiment region correlation analysis,” *Neurocomputing*, vol.469, pp.221–233, 2022.
- [4] H. Zhang, Y. Liu, D. Xu, K. He, G. Peng, Y. Yue, and R. Liu, “Learning multi-level representations for image emotion recognition in the deep convolutional network,” *Proc. Int. Conf. Graphics and Image Processing*, vol.1208321, p.16, 2022.
- [5] G. Chandrasekaran, N. Antoanela, G. Andrei, C. Monica, and J. Hemanth, “Visual sentiment analysis using deep learning models with social media data,” *Applied Sciences*, vol.12, no.3, p.1030, 2022.
- [6] M. Soleymani, D. Garcia, B. Jou, B. Schuller, S.-F. Chang, and M. Pantic, “A survey of multimodal sentiment analysis,” *Image and Vision Computing*, vol.65, pp.3–14, 2017.
- [7] M. Chen, L.-L. Zhang, X. Yu, and Y. Liu, “Weighted co-training for cross-domain image sentiment classification,” *J. Computer Science and Technology*, vol.32, no.4, pp.714–725, 2017.
- [8] X. Yang, S. Feng, D. Wang, and Y. Zhang, “Image-text multimodal emotion classification via multi-view attentional network,” *IEEE Trans. Multimedia*, vol.23, pp.4014–4026, 2020.
- [9] W. Guo, Y. Zhang, X. Cai, L. Meng, J. Yang, and X. Yuan, “Ld-man: Layout-driven multimodal attention network for online news sentiment recognition,” *IEEE Trans. Multimedia*, vol.23, pp.1785–1798, 2020.
- [10] H. Hotelling, “Relations between two sets of variates,” *Biometrika*, vol.28, no.3, pp.321–377, 1936.
- [11] S. Nemati, “Canonical correlation analysis for data fusion in multimodal emotion recognition,” *Proc. Int. Symposium on Telecommunications*, pp.676–681, 2018.
- [12] Y.-T. Lan, W. Liu, and B.-L. Lu, “Multimodal emotion recognition using deep generalized canonical correlation analysis with an attention mechanism,” *Proc. Int. Joint Conf. Neural Networks*, pp.1–6, 2020.
- [13] C. Guanghui and Z. Xiaoping, “Multi-modal emotion recognition by fusing correlation features of speech-visual,” *IEEE Signal Processing Letters*, vol.28, pp.533–537, 2021.
- [14] L. Chen, K. Wang, M. Li, M. Wu, W. Pedrycz, and K. Hirota,

(注3) : E6 データセットでは、joy 及び anger, FI データセットでは、fear 及び excitement を表す。

- “K-means clustering-based kernel canonical correlation analysis for multimodal emotion recognition in human-robot interaction,” IEEE Trans. Industrial Electronics, pp.1–1, 2022.
- [15] N.M. Correa, T. Eichele, T. Adalı, Y.-O. Li, and V.D. Calhoun, “Multi-set canonical correlation analysis for the fusion of concurrent single trial ERP and functional MRI,” Neuroimage, vol.50, no.4, pp.1438–1445, 2010.
- [16] G. Lee, A. Singanamalli, H. Wang, M.D. Feldman, S.R. Master, N.N. Shih, E. Spangler, T. Rebbeck, J.E. Tomaszewski, and A. Madabhushi, “Supervised multi-view canonical correlation analysis (sMVCCA): Integrating histologic and proteomic features for predicting recurrent prostate cancer,” IEEE Trans. Medical Imaging, vol.34, no.1, pp.284–297, 2014.
- [17] G. Andrew, R. Arora, J. Bilmes, and K. Livescu, “Deep canonical correlation analysis,” Proc. Int. Conf. Machine Learning, pp.1247–1255, 2013.
- [18] Y. Ito, T. Ogawa, and M. Haseyama, “SFEMCCA: Supervised fractional-order embedding multiview canonical correlation analysis for video preference estimation,” Proc. Int. Conf. Acoustics, Speech and Signal Processing, pp.3086–3090, 2018.
- [19] M. Matsumoto, N. Saito, K. Maeda, T. Ogawa, and M. Haseyama, “Supervised fractional-order embedding multiview canonical correlation analysis via ordinal label dequantization for image interest estimation,” IEEE Access, vol.9, pp.21810–21822, 2021.
- [20] S. Akaho, “A kernel method for canonical correlation analysis,” arXiv preprint cs/0609071, pp.1–7, 2006.
- [21] K. Fukumizu, F.R. Bach, and A. Gretton, “Statistical consistency of kernel canonical correlation analysis,” J. Machine Learning Research, vol.8, no.2, pp.361–383, 2007.
- [22] R. Plutchik, “A general psychoevolutionary theory of emotion,” Theories of Emotion, pp.3–33, Elsevier, 1980.
- [23] R. Panda, J. Zhang, H. Li, J.-Y. Lee, X. Lu, and A.K. Roy-Chowdhury, “Contemplating visual emotions: Understanding and overcoming dataset bias,” Proc. European Conf. Computer Vision, pp.579–595, 2018.
- [24] Q. You, J. Luo, H. Jin, and J. Yang, “Building a large scale dataset for image emotion recognition: The fine print and the benchmark,” Proc. AAAI Conf. Artificial Intelligence, vol.30, pp.308–314, 2016.
- [25] J.A. Mikels, B.L. Fredrickson, G.R. Larkin, C.M. Lindberg, S.J. Maglio, and P.A. Reuter-Lorenz, “Emotional category data on images from the international affective picture system,” Behavior Research Methods, vol.37, no.4, pp.626–630, 2005.
- [26] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” Proc. IEEE, vol.86, no.11, pp.2278–2324, 1998.
- [27] M. Tan and Q. Le, “Efficientnet: Rethinking model scaling for convolutional neural networks,” Proc. Int. Conf. Machine Learning, pp.6105–6114, 2019.
- [28] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C.L. Zitnick, “Microsoft coco: Common objects in context,” Proc. European Conf. Computer Vision, pp.740–755, 2014.
- [29] K. Xu, J. Ba, R. Kiros, K. Cho, A. Courville, R. Salakhudinov, R. Zemel, and Y. Bengio, “Show, attend and tell: Neural image caption generation with visual attention,” Proc. Int. Conf. Machine Learning, pp.2048–2057, 2015.
- [30] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp.770–778, 2016.
- [31] T. Sun, S. Chen, J. Yang, and P. Shi, “A novel method of combined feature extraction for recognition,” Proc. IEEE Int. Conf. Data Mining, pp.1043–1048, 2008.
- [32] Y. Peng, D. Zhang, and J. Zhang, “A new canonical correlation analysis algorithm with local discrimination,” Neural Processing Letters, vol.31, no.1, pp.1–15, 2010.
- [33] X. Zhang, N. Guan, Z. Luo, and L. Lan, “Discriminative locality preserving canonical correlation analysis,” Proc. Chinese Conf. Pattern Recognition, pp.341–349, 2012.
- [34] Y.-H. Yuan and Q.-S. Sun, “Graph regularized multiset canonical correlations with applications to joint feature extraction,” Pattern Recognition, vol.47, no.12, pp.3907–3919, 2014.
- [35] Y.-H. Yuan, Y. Li, X.-B. Shen, Q.-S. Sun, and J.-L. Yang, “Laplacian multiset canonical correlations for multiview feature extraction and image recognition,” Multimedia Tools and Applications, vol.76, no.1, pp.731–755, 2017.
- [36] T. Matsui, N. Saito, T. Ogawa, S. Asamizu, and M. Haseyama, “Gaze-based visual feature extraction via dlcca for visual sentiment estimation,” Proc. Int. Workshop on Advanced Image Technology, vol.11049, p.110490Q, 2019.
- [37] M. Matsumoto, N. Saito, T. Ogawa, and M. Haseyama, “Interest estimation for images using eye gaze-based visual and text features via dlcca,” Proc. IEEE Global Conf. Life Sciences and Technologies, pp.3–4, 2020.
- [38] K. Maeda, S. Takahashi, T. Ogawa, and M. Haseyama, “Multi-feature fusion based on supervised multi-view multi-label canonical correlation projection,” Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing, pp.3936–3940, 2019.
- [39] T. Cover and P. Hart, “Nearest neighbor pattern classification,” IEEE Trans. Information Theory, vol.13, no.1, pp.21–27, 1967.
- [40] J. Yang, J. Li, L. Li, X. Wang, and X. Gao, “A circular-structured representation for visual emotion distribution learning,” Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp.4237–4246, 2021.

(2022年5月25日受付, 10月24日再受付,  
2023年1月13日早期公開)



斉藤 直輝 (正員)

2014 釧路高専専攻科了。2016 北海道大学大学院情報科学研究科修士課程了。2019 同大学院情報科学研究科博士後期課程了。2019 北海道大学大学院情報科学研究科博士研究員。2019 釧路高専創造工学科助教。2022 北海道大学総合 IR 室助教, 現在に至る。マルチモーダル信号処理・機械学習に関する研究に従事。IEEE 会員。博士 (情報科学)。



前田 圭介 (正員)

2015 北海道大学・工学部卒。2017 同大大学院情報科学研究科修士課程了。2018 日本学術振興会特別研究員 (DC2)。2019 同大大学院情報科学研究科博士後期課程了。2019 日本学術振興会特別研究員 (PD)。2020 同大総合 IR 室特任助教。2022 同大大学院情報科学研究科特任助教。現在に至る。画像・テキスト・生体情報などの異種データを用いたマルチモーダル信号処理・機械学習とその応用に関する研究に従事。IEEE 会員。博士 (情報科学)。



小川 貴弘 (正員)

2003 北海道大学・工学部卒。2005 同大大学院工学研究科修士課程了。2007 同大大学院情報科学研究科博士後期課程了。2008 同大大学院情報科学研究科助教。2016 同大大学院情報科学研究科准教授。2023 同大大学院情報科学研究科教授。現在に至る。2008 電子情報通信学会論文賞。ITE Transactions on Media Technology and Applications, Associate Editor 等を歴任。画像復元及び映像処理に関する研究に従事。映像情報メディア学会、ACM, IEEE 各会員。博士 (情報科学)。



浅水 仁 (正員)

1994 室蘭工大・工・材料物性卒。1996 同大大学院修士課程了。1999 北大大学院工学研究科博士課程了。同年 (株) 日立製作所入社。2003 釧路高専電子工学科助手。2004 同高専准教授。2014 同高専教授。現在に至る。デジタル信号処理及び画像処理の研究に従事。博士 (工学)。



長谷山美紀 (正員)

1988 北海道大学大学院工学研究科修士課程了。同大学助手、助教授 (1994-1995 アメリカワシントン大学客員准教授) を経て2006 同大学院情報科学研究科教授。画像・映像及び音響信号などマルチメディア信号処理の研究に従事。2008 電子情報通信学会論文賞。2015 映像情報メディア学会丹羽高柳賞業績賞。情報処理学会、映像情報メディア学会、IEEE 各会員。博士 (工学)。