Feature coding method based on shared weights support vector data description for face recognition

Mengmeng Liao, Yunjie Li* and Meiguo Gao

School of Information and Electronics, Beijing Institute of Technology, Beijing, Beijing, 100081, China
mmliao16@fudan.edu.cn, liyunjie@bit.edu.cn, meiguo_g@bit.edu.cn

Abstract. In this paper, we propose a feature coding method based on shared weights support vector data description (FCM-SWSVDD). The proposed process of FCM-SWSVDD is as follows. By considering the density information of the clusters and introducing the weighting learning, we propose an improved support vector data description (SVDD), named shared weights support vector data description (SWSVDD). SWSVDD can obtain the cluster center and cluster radius more accurately. Incorporating SWSVDD and the triangle coding into the same feature coding learning process, FCM-SWSVDD is proposed. After the features of those images are extracted by using FCM-SWSVDD, a sparse representation classifier is used to classify those features. Experimental results show that the performance of the proposed method exceeds many methods.

1. Introduction
Unsupervised feature learning is widely studied by researchers because it does not require a large number of sample labels. The classical unsupervised feature learning methods include Bag of Words (BoW) [1], Vector of Linearly Aggregated Descriptors (VLAD) [2], Coates’s method [3], etc. For example, Coates’s method can not only encode more rich information but also maintain spatial information. However, these methods have some common shortcomings. These methods only consider the distance between the patch and the cluster centers when encoding each patch. This is not enough for encoding. When the distances between the cluster centers and the patch are equal, the different coding effects of each cluster on the patch cannot be reflected. In other words, the density information of clusters should be considered in the feature encoding process. In this paper, we propose a feature coding method based on shared weights support vector data description (FCM-SWSVDD) by considering the density information of the clusters.

2. Support vector data description (SVDD)
Support vector data description (SVDD) [4] is an important data description method, which can describe the target data set in a hyper-spherical shape. SVDD can find the most compact ball for each cluster. The ball includes all the data points in the cluster. That is to say, SVDD can obtain the cluster center information and cluster radius of each cluster. Inspired by this idea, we intend to encode those patches using SVDD.

However, the radius and center of the ball obtained by SVDD are not very accurate due to the lack of consideration of the density distribution of each cluster. The contribution of each data point in each cluster should not be the same. Data in high-density regions is more important than data in low-density regions because the data which comes from the high-density region is more likely to be included in the
sphere than the data which comes from the low-density region. Therefore, different data points should be treated differently in model learning. In other words, each data point should be assigned an appropriate weight for model learning. In view of the above problem of SVDD, we propose a shared weights support vector data description (SWSVDD). Based on SWSVDD, combined with the triangle coding, a feature coding method based on shared weights support vector data description (FCM-SWSVDD) is proposed in this paper.

3. Feature coding method based on shared weights support vector data description (FCM-SWSVDD)

In this section, we will introduce SWSVDD and FCM-SWSVDD in turn.

3.1. Shared weights support vector data description (SWSVDD)

To obtain the more accurate center and radius of the ball, an improved SVDD called shared weights support vector data description (SWSVDD) is proposed in this paper.

Suppose we have a data set \( \Psi = \{ \mathbf{x}_m | m = 1, 2, ..., N \} \), \( \mathbf{x}_m \) is the \( m \) th sample point. Therefore, the model of SWSVDD can be written as

\[
\begin{align*}
\min_{r, \varepsilon} & \quad r^2 + \zeta \sum_{m=1}^{N} \omega_m \mathbf{e}_m^2 \\
\text{s. t.} & \quad \| \mathbf{x}_m - \mathbf{f} \|^2 \leq r^2 + \varepsilon, \quad \varepsilon_m \geq 0, \quad a = \frac{1}{N} \sum_{m=1}^{N} \mathbf{e}_m
\end{align*}
\]

(1)

where \( \omega_m \) is the weight coefficient of \( \mathbf{x}_m \). The value of \( \zeta \) is set to 0.3.

Next, we first derive the weight coefficient of each data point and then solve the equation (1).

We use the DBSCAN [5] to divide the entire data set into \( C \) clusters. Suppose the data set of the \( k \) th cluster is \( \{ x^k_1, x^k_2, ..., x^k_{\eta_k} \} \), where \( x^k_i = [v^k_{i1}, v^k_{i2}, ..., v^k_{iD}] \in \mathbb{R}^D \) is the \( i \) th data point in the cluster, \( i = 1, 2, ..., \eta_k \).

The average distance between two points in the cluster \( k \) is denoted as \( A_k \).

If \( \eta_k > 1 \), then

\[
A_k = \frac{2}{\eta_k (\eta_k - 1)} \sum_{i=1}^{\eta_k} \sum_{j=i+1}^{\eta_k} d(x^k_i, x^k_j)
\]

(2)

where \( d(x^k_i, x^k_j) = |v^k_{ij} - v^k_{ij}| + |v^k_{ij} - v^k_{ij}| + ... + |v^k_{ij} - v^k_{ij}| \).

If \( \eta_k = 1 \), then

\[
A_k = \frac{1}{\eta_k} \sum_{i=1}^{\eta_k} \sum_{j=i+1}^{\eta_k} d(x^k_i, x^k_j)
\]

(3)

Generally, the distances between data points within the same cluster are far less than the distances between data points within different clusters. Hence, we assume that those data points within the same cluster have the same weight coefficient. The weight coefficient of each data point in the \( k \) th cluster is denoted as \( \omega(k) \), and its expression is as follows.

\[
\omega(k) = 1 - \frac{\sqrt{A_k}}{\sum_{i=1}^{C} \sqrt{A_i}}
\]

(4)

The Lagrange function of equation (1) can be written as
\[
\Gamma(r, \epsilon, \alpha, \beta) = r^2 + 2 \sum_{i=1}^{N} \alpha_i \xi_i + \sum_{i=1}^{N} \beta_i \epsilon - \frac{1}{N} \sum_{m=1}^{N} \left( -r^2 - \xi_n \right) - \sum_{m=1}^{N} \beta_m \epsilon
\]

where \(\alpha_m\) and \(\beta_m\) are the corresponding Lagrange multipliers.

Let \(\frac{\partial \Gamma}{\partial r} = 0\) and \(\frac{\partial \Gamma}{\partial \epsilon_m} = 0\), we can obtain the dual function of equation (5).

\[
\begin{align*}
\left\{ \begin{array}{l}
\max \sum_m \alpha_m \langle \bar{x}_m, \bar{x}_n \rangle - \frac{2}{N} \sum_m \alpha_m \langle \bar{x}_m, \bar{x}_n \rangle \\
\text{s.t.} \sum_m \alpha_m = 1, \alpha_m \in [0, \xi_{\alpha_m}], m = 1, 2, \ldots, N
\end{array} \right.
\end{align*}
\]

Equation (6) can be reformulated as

\[
\begin{align*}
\left\{ \begin{array}{l}
\min -\frac{2}{N} \alpha^T \epsilon - \alpha^T u \\
\text{s.t.} \alpha^T e = 1, \alpha_m \in [0, \xi_{\alpha_m}], m = 1, 2, \ldots, N
\end{array} \right.
\end{align*}
\]

where \(g = (\langle \bar{x}_m, \bar{x}_n \rangle)_{N \times N}, u = (\langle \bar{x}_m, \bar{x}_n \rangle)_{N \times 1}, e = (1, 1, 1, \ldots, 1)^T\).

For equation (7), the value of \(\alpha\) can be obtained by using the linear programming method. Then the value of \(r\) can be obtained by

\[
r^2 = \bar{x}_m \cdot \bar{x}_n - 2 \sum_{m,n=1}^{V} \alpha_m \langle \bar{x}_m, \bar{x}_n \rangle + \sum_{m,n=1}^{V} \alpha_m \alpha_n \langle \bar{x}_m \cdot \bar{x}_n \rangle
\]

where \(\bar{x}_m\) is the support vector, \(V\) is the support vector set.

Figure 1 shows the sketch map of SWSVDD.

3.2. FCM-SWSVDD

For any patch \(q_j\) of an image \(q\), \(j = 1, 2, \ldots, N\), it can be encoded as

\[
F(q_j) = [F_1(q_j) \quad F_2(q_j) \quad \cdots \quad F_C(q_j)]^T
\]

where \(F_i(q_j) = [F_{i,1}(q_j) \quad F_{i,2}(q_j) \quad \cdots \quad F_{i,C}(q_j)]\), \(i = 1, 2, \ldots, C\), \(F_{i,1}(q_j) = \max \{0, g(s) - s_i(q_j)\}\), \(s_i(q_j) = \|q_j - c_i\|_2\) is the distance from patch \(q_j\) to the centroid of the \(i\)th SVDD ball, \(g(s)\) is the mean value of \(s_1, s_2, \ldots, s_C\). \(F_{i,2}(q_j) = \max \{0, H(m) - m_i(q_j)\}\), where \(m_i(q_j) = \|q_j - R_i\|_2\) is the distance from the patch \(q_j\) to the surface of the \(i\)th SVDD ball, \(H(m)\) is the mean value of \(m_1, m_2, \ldots, m_C\).

Figure 2 shows the encoding process of each patch by using SWSVDD.
3.3. FCM-SWSVDD for face recognition

The process of FCM-SWSVDD for face recognition is as follows. We firstly use SWSVDD to obtain the cluster center and cluster radius of each cluster. Then, each image is divided into several patches by sliding a window with size $5 \times 5$. After that, we use FCM-SWSVDD to encode each patch and obtain a feature. We concatenate the encoded features of all patches in each image to obtain the features of the image. Finally, we use the sparse representation classifier (SRC) to classify those features.

4. Experimental results and analysis

In this section, the FEI database and UMIST database are used in the experiment to test the performance of FCM-SWSVDD.

**FEI database.** This database mainly involves pose variations. Figure 3 shows some samples in the database.

![Figure 3. Some samples of the FEI database.](image)

For the images of each category in the database, we randomly select half of them for training, and the rest for testing. Table 1 shows the recognition rates of various algorithms on the FEI database.

| Method      | Recognition rate |
|-------------|------------------|
| LMP [6]     | 87.46            |
| SRC [7]     | 79.12            |
| RDCDL [8]   | 89.49            |
| RRNN [9]    | 85.33            |
| EqM [10]    | 88.15            |
| KONPDA [11] | 86.99            |
| MRARC [12]  | 90.08            |
| FCM-SWSVDD  | **92.12**        |

From Table 1, we can see that the recognition rate of FCM-SWSVDD is 92.12%, which is higher than the recognition rate of other algorithms in the table. RRNN is a deep learning method. Although it has been pre-trained, its recognition rate is still 7% lower than that of FCM-SWSVDD. The reason
for this result is as follows. The large number of training samples used for pre-training is quite different from the test samples.

**UMIST database.** This database mainly involves pose variations and expression variations. Figure 4 shows some samples in the database.

![Figure 4. Some samples of the UMIST database.](image)

For the images of each category in the database, we randomly select 10 images for training, and the rest for testing. Table 2 shows the recognition rates of various algorithms on the UMIST database.

Table 2. Recognition rates (%) of various algorithms on the UMIST database.

| Method       | Recognition rate |
|--------------|------------------|
| LMP [6]      | 89.95            |
| SRC [7]      | 85.49            |
| RDCDL [8]    | 96.12            |
| RRNN [9]     | 94.58            |
| EqM [10]     | 96.11            |
| KONPDA [11]  | 96.08            |
| MRARC [12]   | 96.76            |
| FCM-SWSVDD   | 98.36            |

It can be seen from Table 2 that the recognition rate of FCM-SWSVDD is 98.36%, which is the highest recognition rate in the table. It can be found that the advantage of FCM-SWSVDD on the UMIST database is not as great as that of the FEI database. This is due to the small variations involved in the UMIST database, so other algorithms can better deal with the recognition problem in this situation.

5. **Conclusion**

By considering the density of clusters and introducing the weighting learning, we propose an improved SVDD, named SWSVDD. SWSVDD can obtain the cluster center information and cluster radius information of each cluster more accurately. Based on SWSVDD, an unsupervised coding method is proposed, which is called FCM-SWSVDD. FCM-SWSVDD not only considers the distance from the patch to the center of each cluster but also considers the distance from the patch to the surface of the SVDD sphere of each cluster. FCM-SWSVDD was used for face recognition and achieved good experimental results.

**Acknowledgments**

This work was supported in part by the Post-doctoral Innovative Talent Support Program (Grant no. BX20200048) and the National Natural Science Foundation of China (Grant no. 61976019).

**References**

[1] Chen, T., Yap, K.H. (2014) Discriminative BoW framework for mobile landmark recognition. IEEE Transactions on Cybernetics, 44: 695-706.

[2] Jégou, H., Perronnin, F., Douze, M., Sánchez, J., Pérez, P., Schmid, C. (2012) Aggregating local image descriptors into compact codes. IEEE Transactions on Pattern Analysis and Machine Intelligence, 34:1704-1716.

[3] Coates, A., Lee, H., Ng, A.Y. (2011) An analysis of single-layer networks in unsupervised feature learning. Journal of Machine Learning Research, 15: 215-223.
[4] David, M.J., Robert, P.W.D. (2004) Support vector data description. Machine Learning, 54: 45-66.

[5] Ester, M., Kriegel, H.P., Sander, J., Xu, X.W. (1996) A density-based algorithm for discovering clusters in large spatial databases with noise. [C]//Proceedings of 2nd International Conference on Knowledge Discovery and Data Mining, 96:226-231.

[6] Yang, W.K., Zhang, X., Li, J. (2020) A local multiple patterns feature descriptor for face recognition. Neurocomputing, 373: 109-122.

[7] Wright, J., Yang, A.Y., Ganesh, A., Sastry, S.S., Ma, Y. (2009) Robust face recognition via sparse representation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 31:210-227.

[8] Lin, G.J., Yang, M., Yang, J., Shen, L.L., Xie, W.C. (2018) Robust, discriminative and comprehensive dictionary learning for face recognition. Pattern Recognition, 81:341-356.

[9] Li, Y., Zheng, W.M., Cui, Z., Zhang, T. (2018) Face recognition based on recurrent regression neural network. Neurocomputing, 297: 50-58.

[10] Sun, J.N., Yang, W.M., Xue, J.H., Liao, Q.M. (2020) An equalized margin loss for face recognition. IEEE Transactions on Multimedia, 22: 2833-2843.

[11] Hu, L.C., Zhang, W.S. (2020) Orthogonal neighborhood preserving discriminant analysis with patch embedding for face recognition. Pattern Recognition, 106: 107450:1-14.

[12] Wang, Y.L., Tang, Y.Y., Li, L.Q., Chen, H. (2020) Modal regression-based atomic representation for robust face recognition and reconstruction. IEEE Transactions on Cybernetics, 50: 4393-4405.