Compressive-sensing-based Human Action Recognition

Jun Jiang
School of Computer Science
Southwest Petroleum University
Chengdu, China
toddjiang@163.com

Gao Li
State Key Laboratory of Oil and Gas Reservoir Geology and Development
Southwest Petroleum University
Chengdu, China
lgmichael@263.net

***Abstract*—A new compressive sensing based dimensionality reduction method is proposed for human action recognition, in which a novel hybrid random matrix (HRM) is constructed and is proved to satisfy the restricted isometry property. It projects the high-dimensional features into a low-dimensional space via the HRM. Then the low-dimensional features are used for classification. Experimental results demonstrate that the proposed method is effective and efficient in human action recognition, and is on par with or better than the state-of-the-art methods.**

Keywords—compressive sensing; hybrid fusion matrix; action recognition

I. INTRODUCTION

In recent years, human action recognition has been extensively studied in computer vision community in that it finds applications in various areas including intelligent video surveillance, human-computer interaction, and smart robotics. Many action recognition methods use complex high-dimensional action features which might lead to unacceptable computational time or storage. To alleviate this problem, dimensionality reduction is adopted to produce low-dimensional features. Various types of dimensionality reduction techniques, such as principal component analysis (PCA), linear discriminant analysis (LDA), multidimensional scaling (MDS), isometric feature map (ISOMAP), and local linear embedding (LLE), have been proposed in the last several decades [1]. However, most of these methods need a data-dependent training process and hence result in limited generalization. In addition, they require high computational complexity as they involve time-consuming operations. Recently, compressive sensing (CS) has provided a new way of compressing signals [2, 3], and has been applied to image processing [4].

In this paper, we propose a compressive sensing with hybrid random matrix (CSHRM) method for recognizing human actions, where a novel measurement matrix is constructed and is shown to satisfy the restricted isometry property (RIP). As shown in Fig. 1, our method consists of three major components, i.e., feature extraction, dimensionality reduction, and classification. Experimental results demonstrate that the proposed method is effective and efficient to reduce the dimensions of the features, and performs on par with or better than the representative state-of-the-art action recognition methods.

II. PROPOSED METHOD

A. *Background Theory*

Given a vector x in R^N and a finite-dimensional operator Φ, namely a $M \times N$ measurement matrix in CS, we can obtain a vector y in $\mathrm{R}^M \left(M << N\right)$, such that

$$\mathbf{y} = \Phi\mathbf{x}\,. \tag{1}$$

Furthermore, if Φ satisfies the RIP, and **x** is K-sparse, then x can be exactly reconstructed from **y** with overwhelming probability. Φ is said to satisfy the RIP if there exists a constant $\delta \in (0,1)$ such that

$$(1-\delta)\|\mathbf{x}\|_2^2 \le \|\Phi\mathbf{x}\|_2^2 \le (1+\delta)\|\mathbf{x}\|_2^2 \tag{2}$$

holds for all K-sparse vectors in R^N [2]. The RIP implies that the metric structure is almost invariant in much lower

Fig 1. Main components of the proposed method.

dimensional space, and thus the input data can have a compact representation without losing information.

B. *Hybrid random matrix*

Two widely used measurement matrices, i.e., random Gaussian matrix (RGM) and random Bernoulli matrix (RBM), are presented in [5]. Based on these two matrices, we construct a novel measurement matrix called hybrid random matrix (HRM). The HRM, denoted as Φ_H , is defined as

$$\Phi_H = (1-\lambda)\Phi_G + \lambda\Phi_B , \qquad (3)$$

where $\lambda \in [0,1]$, Φ_G denotes the RGM, Φ_B denotes the RBM, and Φ_H, Φ_G, $\Phi_B \in R^{M\times N}$. Note that the RGM and RBM are two special cases of the HRM when $\lambda = 0,1$, respectively.

Proposition 1. Given $\delta \in (0,1)$, there exist positive constants c_1, c_2 depending on δ such that Φ_H satisfies the RIP with probability $\geq 1-2e^{-c_2 M}$ if $M \geq c_1 K \log(N/K)$.

Proof. We first show

$$\mathrm{E}\left(\|\Phi_H \mathbf{x}\|_2^2\right) = \|\mathbf{x}\|_2^2 . \qquad (4)$$

Without loss of generality, we can assume $\|\Phi_G \mathbf{x}\|_2 \leq \|\Phi_G \mathbf{x}\|_2$, then

$$\mathrm{E}\left(\|\Phi_H \mathbf{x}\|_2^2\right) = \mathrm{E}\left(\|(1-\lambda)\Phi_G \mathbf{x} + \lambda\Phi_B \mathbf{x}\|_2^2\right) \leq \mathrm{E}\left(\left(\|(1-\lambda)\Phi_G \mathbf{x}\|_2 + \|\lambda\Phi_B \mathbf{x}\|_2\right)^2\right) \leq \|\mathbf{x}\|_2^2 , \qquad (5)$$

and we can also get

$$\mathrm{E}\left(\|\Phi_H \mathbf{x}\|_2^2\right) = \mathrm{E}\left(\|(1-\lambda)\Phi_G \mathbf{x} + \lambda\Phi_B \mathbf{x}\|_2^2\right) \geq \mathrm{E}\left(\left(\|(1-\lambda)\Phi_G \mathbf{x}\|_2 - \|\lambda\Phi_B \mathbf{x}\|_2\right)^2\right) \geq \|\mathbf{x}\|_2^2 , \qquad (6)$$

because $\mathrm{E}\left(\|\Phi_G \mathbf{x}\|_2^2\right) = \|\mathbf{x}\|_2^2$, $\mathrm{E}\left(\|\Phi_B \mathbf{x}\|_2^2\right) = \|\mathbf{x}\|_2^2$ according to [5].

Combining (5) and (6), we obtain $\mathrm{E}\left(\|\Phi_H \mathbf{x}\|_2^2\right) = \|\mathbf{x}\|_2^2$.

Since Φ_B obeys the concentration inequality [6], we have

$$\Pr\left(\left|\|\Phi_H \mathbf{x}\|_2^2 - \|\mathbf{x}\|_2^2\right| \geq \varepsilon\|\mathbf{x}\|_2^2\right) = \Pr\left(\left|\|(1-\lambda)\Phi_G \mathbf{x} + \lambda\Phi_B \mathbf{x}\|_2^2 - \|\mathbf{x}\|_2^2\right| \geq \varepsilon\|\mathbf{x}\|_2^2\right) \leq \Pr\left(\left|\left(\|\Phi_B \mathbf{x}\|_2^2 - \|\mathbf{x}\|_2^2\right)\right| \geq \varepsilon\|\mathbf{x}\|_2^2\right) \leq e^{-dM} , \qquad (7)$$

where d is a positive constant and depends only on δ .

Hence, by theorem 5.2 in [6], we can derive the proposition.

C. *Action recognition*

First, we use the algorithm presented in [7] to form raw features. Then each original high-dimensional feature $\mathbf{x}$ is transformed to the low-dimensional feature $\mathbf{y}$ by

$$\mathbf{y} = \Phi_H \mathbf{x} . \qquad (8)$$

In the classification step, $\mathbf{y}$ is fed into the Support Vector Machine (SVM) for training.

III. EXPERIMENTAL RESULTS

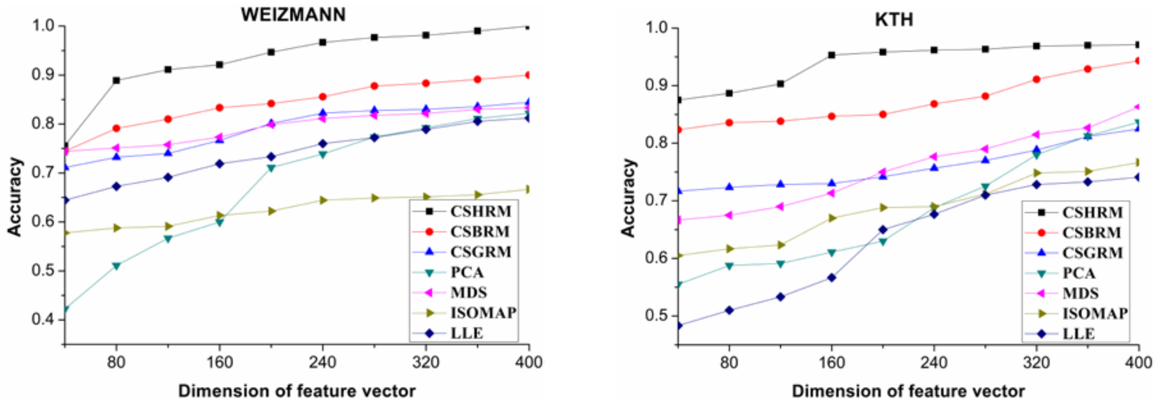
We use two benchmark datasets, i.e. WEIZMANN and KTH [7], to evaluate our method, and conduct the experiments using the leave-one-out cross-validation (LOOCV) strategy. All methods are implemented in Matlab 2011 on a PC with a single Intel Core2 2.3 GHz processor and 2 GB memory.

We compare our CSHRM with PCA, MDS, ISOMAP, LLE, CSRGM (compressive sensing with RGM), and CSRBM (compressive sensing with RBM). In the feature extraction phase, we use the same settings as in [7] and thus obtain the 2500-dimensional features. The λ of Φ_H is emprically set to be 0.6. The numbers of neighbors for ISOMAP and LLE are set to be 5 and 12, respectively. Finally, in the classification phase, the radial basis function kernel is chosen for the SVM. For fairness, the optimal parameters of all algorithms are selected to achieve the best performance, and the results are obtained by averaging over 20 independent realizations.

Fig.2 shows the accuracy comparisons of all seven dimensionality reduction methods for different dimensions. We can see that the CSHRM outperforms the other algorithms in dimensions ranging from 40 to 400 on both datasets. The best accuracies achieved by the CSHRM are 99.0% on the WEIZMANN and 97.1% on the KTH with dimension 400. Table 1 presents the computational time costs of various algorithms in dimension 400. Since the feature extraction is the same for all algorithms, we compare merely the dimensionality reduction and classification (training/testing) costs of them. It can be seen that the CSHRM costs the second least, which benefits from the simple and training-free dimension-reduced operation. Moreover, we compare our action recognition approach with the state-of-the-art methods [7] in Table 2. Clearly, our approach is superior to all the other methods on the KTH, and is the second best on the WEIZMANN.

TABLE I. . Computational time costs of seven dimensionality reduction methods on two datasets with dimension 400.

Method	Time cost (seconds)	
	WEIZMANN	*KTH*
PCA	1.5	15.6
MDS	4.4	12.9
ISOMAP	11.4	24.7
LLE	10.9	33.2
CSBRM	0.4	9.5
CSGRM	0.9	17.1
CSHRM	**0.5**	**9.8**



(a) (b)

Fig 2. Recognition accuracy of seven dimensionality reduction methods for different dimensions on two datasets.(a)WEIZMANN (b) KTH.

TABLE II. . Comparison of recognition accuracy with the state-of-the-art methods [7]

Method	Accuracy (%)	
	WEIZMANN	*KTH*
Bregonzio et al.	96.66	94.33
Sun et al.	97.80	94.00
Ikizler et al.	-	94.00
Lin et al.	-	93.43
Wang et al.	100	92.51
Liu et al.	-	92.30
Kläser et al.	84.30	91.40
Niebles et al.	90.00	83.30
Dollar et al.	85.20	81.17
Liu et al.	-	94.16
Zhao et al.	-	91.17
Savarese et al.	-	86.83
CSHRM	**99.20**	**95.10**

IV. CONCLUSION

In this paper, we have presented the CSHRM for human action recognition. Specifically, a novel HRM is constructed and is shown to obey the RIP. Experimental results demonstrate that our CSHRM outperforms the conventional dimensionality reduction algorithms in terms of efficiency and effectiveness. Furthermore, our action recognition method can give competitively good or even better performance when compared with the state-of-the-art methods.

REFERENCES

[1] C. Burges, Dimension Reduction: A Guided Tour, Foundations and Trends in Machine Learning, vol.2, no.4, pp. 275-365, 2010.

[2] E. Cande`s and M. Wakin,, An Introduction To Compressive Sampling, IEEE Signal Process. Mag., vol.25, no.2, pp. 21–30, 2008.

[3] D. Donoho, Compressed sensing, IEEE Trans. Inf. Theory, vol.52, no.2, pp. 1289–1306, 2006.

[4] A. Wang, L. Liu, B. Zeng and H. Bai, Progressive image coding based on an adaptive block compressed sensing, IEICE Electron. Express, vol.8, no.8, pp. 575-581, 2011.

[5] D. Achlioptas, Database-friendly random projections: Johnson-Lindenstrauss with binary coins, J. Comput. Syst. Sci., vol.66, no.4, pp. 671–687, 2003.

[6] R. Baraniuk, M. Davenport, R. DeVore and M. Wakin, A simple proof of the restricted isometry property for random matrices, Constr. Approx., vol.28, no.3, pp. 253–263, 2008.

[7] M. Bregonzio, T. Xiang and S. Gong, Fusing appearance and distribution information of interest points for action recognition, Pattern Recogn., vol.45, no.3, pp. 1220–1234, 2012..

.

