

## Evaluation of Physical Education Multimedia Teaching for Data Assimilation

<https://doi.org/10.3991/ijoe.v14i04.8366>

Lifu Xing

Ningbo Dahongying University, Ningbo, Zhejiang, China

188872807@qq.com

**Abstract**—The evaluation of physical education (PE) multimedia teaching refers to the prediction of physical education multimedia teaching quality in the absence of initial multimedia teaching information. Therefore, the evaluation method of PE multimedia teaching based on unsupervised feature learning has achieved good performance. But, its quality prediction accuracy decreases significantly with the reduction of the feature dimension. In order to overcome this defect, the author combines the active learning strategy with the unsupervised feature learning and proposes a kind of data assimilation framework to improve the discriminability of the representation of teaching features. The results show that the proposed method can enhance the accuracy of teaching quality prediction by 8%. Experiments show that, when feature dimension is relatively low, the proposed method can improve the teaching quality prediction accuracy by 8% compared with the method based on unsupervised feature learning. At the same time, the performance of the proposed method is superior to that of the other physical education multimedia teaching evaluation methods at present.

**Keywords**—active learning; dictionary learning; feature learning; teaching presentation; teaching quality evaluation

### 1 Introduction

With the popularization of teaching acquisition equipment such as mobile phones and cameras and the rapid development of social networks, teaching plays an increasingly important role in our daily life. However, due to the limitation of the performance of the teaching processing system, various kinds of distortion will be introduced in the process of teaching access, transmission and processing. The introduction of distortion will reduce the aesthetics of teaching, which, at the same time, will impeding people from obtaining information from the teaching. Therefore, it is necessary to study the teaching quality evaluation methods, so as to apply them to the monitoring of the teaching acquisition equipment performance, the optimal selection of teaching processing system and the optimization of teaching methods.

As people are the ultimate recipients of teaching, the subjective evaluation of the quality of teaching by people is the most reliable method [1]. However, the subjective quality evaluation method is time consuming, labor intensive and difficult to be inte-

grated into the teaching processing system. Therefore, the design of objective teaching quality evaluation method [2], so that the quality of teaching can be evaluated by computer automatically and accurately is of great significance in the field of teaching processing.

The existing objective teaching quality evaluation methods can be divided into two categories: the reference teaching quality evaluation method [2-3] and the physical education multimedia teaching evaluation method [4]. Between them, the former requires all or part of the initial teaching information to evaluate and test the teaching quality by comparing the difference between the test teaching and the initial teaching information. On the other hand, the evaluation of the multimedia teaching of the physical education is to predict the teaching quality of teaching without the initial teaching information. In the actual application, as the initial teaching information is usually not available, the study of physical education multimedia teaching evaluation will be conducive to the promotion of the teaching quality evaluation method in the practical application.

As the research on the subjective perception mechanism is not sufficiently in-depth, how to extract the teaching features that are highly related to the quality of teaching has become an important challenge for the evaluation of multimedia teaching in physical education. There are two kinds of teaching feature extraction methods in the existing physical education multimedia teaching evaluation: (1) Extraction of features based on the statistical characteristics of the natural scene [5-6]; (2) Unsupervised feature learning [7-8]. In the first method, it is assumed that the natural teaching has certain statistical characteristics. The distortion can affect these statistical characteristics, making the teaching seem "unnatural." Thus, based on the statistical characteristics of natural teaching, the features can be extracted, which in turn can be used to predict the quality of teaching. In literature [9], the statistical characteristics of the natural scene were first introduced into the evaluation of physical education multimedia teaching. Subsequently, the corresponding teaching statistical characteristics were extracted based on the statistical law of teaching in the pixel domain and teaching cosine transform coefficient, wavelet transform coefficient and Gabor transform coefficient to carry out the quality prediction.

In order to overcome the defects of the unsupervised feature learning method, Active Learning strategy [10] is introduced into the dictionary learning in this paper. And a data assimilation framework is put forward and applied to the evaluation of physical education multimedia teaching. In this framework, teaching block is first extracted from the training teaching, and the teaching block is pre-treated to obtain the local characteristic description of teaching. Subsequently, based on the active learning strategy, the characterization of the teaching block and the difference between the teaching blocks are taken into comprehensive consideration. The iterative selection of the optimal teaching block is performed to construct the dictionary. For the given input teaching, it is firstly divided into teaching blocks for pretreatment. Next, the pre-processed teaching block in the dictionary is encoded, followed by the synthesis of all the teaching block coding coefficients to obtain the characteristic description of teaching. Experiments on the LIVE standard database shows that when the characteristic dimension is relatively low, the method proposed in this paper can

improve the accuracy of teaching quality by 8% compared with the method based on unsupervised feature learning. At the same time, the proposed method is also superior to the existing PE multimedia teaching evaluation methods.

## 2 Data assimilation

In this paper, in view of the problems existing in the unsupervised feature learning, its dictionary learning process is improved and a kind of data assimilation framework is put forward. The process of data assimilation framework is similar to that of unsupervised feature learning. However, in the dictionary learning stage, we introduce the active learning strategy. In combination with the characterization of the samples themselves and the difference between samples, the teaching blocks with maximum information volume is selected through iteration as the atoms of the dictionary so as to obtain a more complete dictionary and further learn to obtain better representation of the teaching features. The detailed flow of data assimilation framework is shown in Figure 1. Next, we will introduce the specific steps as the following.

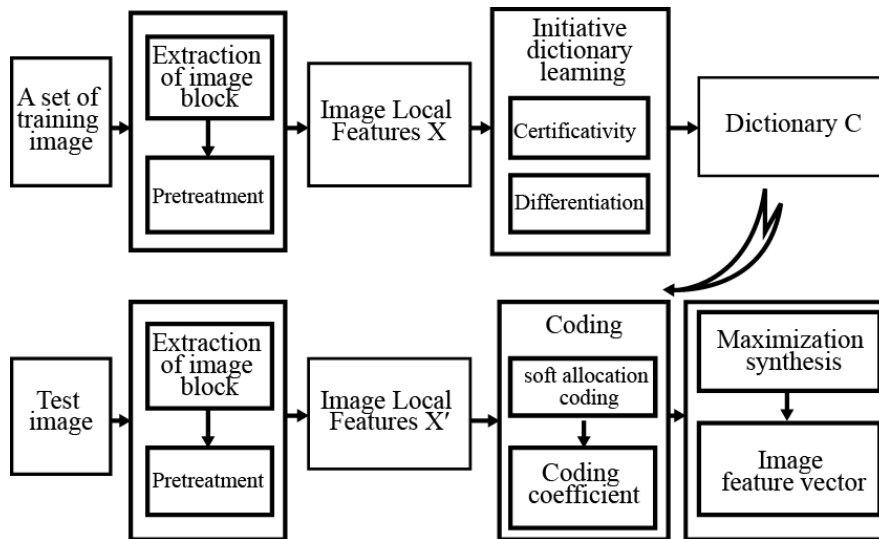


Fig. 1. Data Assimilation Framework

For a set of unlabeled teaching, we randomly select  $M$  teaching blocks of  $b \times b$  from these teaching and convert each teaching block into a vector form. We will extract the set of teaching blocks as follows:

$$P = \{p_1, p_2, \dots, p_M\},$$

$p_i (i = 1, \dots, M)$  is dimension vector of  $d$ , and  $d$  is the dimension of the teaching block vector. When the teaching is the gray scale teaching, it is  $d = b \times b$ . When

the teaching is the color teaching, it is  $d = b \times b \times 3$ . We can pre-process the extracted teaching block.

## 2.1 Pretreatment

In the pre-treatment phase, we firstly normalize the extracted teaching blocks, and the normalization method can be expressed as

$$\tilde{p}_i = \frac{p_i - m_i}{\sigma_i}, i = 1, \dots, M.$$

$m_i$  and  $\sigma_i$  are the mean and standard deviation of the pixel values  $p_i$ .  $\tilde{p}_i$  is the normalized teaching block. The normalized set of teaching blocks is expressed as

$$\tilde{P} = \{\tilde{p}_1, \tilde{p}_2, \dots, \tilde{p}_M\}.$$

Later, we used zero component analysis (Zero Components Analysis, ZCA) to whiten it, and the mathematical expression is as follows:

$$x_i = W_{ZCA} \tilde{p}_i, i = 1, \dots, M;$$

$$W_{ZCA} = \left( \frac{1}{M} \sum_{i=1}^M \tilde{p}_i \tilde{p}_i^T \right)^{-1/2}.$$

$x_i$  stands for a whitened vector of  $\tilde{p}_i$ ,  $W_{ZCA}$  is the transform matrix of ZCA, and  $X = \{x_1, x_2, \dots, x_M\}$  is a set of teaching blocks after whitening.

## 2.2 Active dictionary learning

We think that the amount of information in the teaching block from  $X$  should contain two aspects: (1) the characterization of the teaching block-the choice of the teaching block should be the same as possible with the teaching block. Many of the teaching blocks are similar, so that we only choose a small number of samples, which can represent the entire set of teaching blocks; (2) the difference between teaching blocks - the newly selected teaching blocks should have a large difference from the selected teaching blocks. The dictionaries we construct can characterize a wider range of teaching distortions. In the section, we describe the process of characterization and difference, as well as the process of active dictionary learning.

**Representativeness.** We use the Radial Basic Function (RBF) to calculate the similarity between the samples. We measure the similarity between the samples and the

corresponding neighborhood samples. In this paper, we use the formula (1) to estimate characterization of teaching blocks:

$$R(x_i) = \frac{1}{M} \sum_{x_j \in X, j \neq i} \exp\left(-\|x_i - x_j\|_2^2 / 2\sigma_R^2\right) \quad (1)$$

$R(x_i)$  is the characterization of the teaching block  $x_i$ .  $\|x_i - x_j\|_2^2$  is for the Euclidean distance between  $x_i$  and  $x_j$ .  $\sigma_R$  is the distance for the radial basis function. In this paper, we choose the formula (2) for  $\sigma_R^2$ :

$$\sigma_R^2 = \rho \left( \text{median}_{x_i, x_j \in X, i \neq j} \|x_i - x_j\|_2^2 \right) \quad (2)$$

Median is the median value,  $\rho$  is the scalar to adjust the  $\sigma_R^2$  dimension. Experimental results show that when we take any value  $\rho$  between 0.01 and 1, the method can achieve better performance. In this paper, we choose  $\rho = 0.1$ .

**Diversity.** We use  $S$  to show the set of teaching blocks (i.e. dictionary). It has been chosen to represent the set of unselected teaching blocks  $U$ . For the sample  $x_i \in U$ , we calculate the difference  $x_i$  based on the minimum angle between  $x_i$  and  $S$  of all the teaching blocks.

$$D(x_i) = \min_{x_j \in S} \frac{1}{\pi} \cos^{-1} \left( \frac{x_i \cdot x_j}{\|x_i\| \|x_j\|} \right) \quad (3)$$

$x_i \cdot x_j$  shows the inner product between  $x_i$  and  $x_j$ ,  $\cos^{-1}$  represents the cosine function, and  $\|g\|$  represents the modulus of the vector.

**Active sample selection.** As the teaching blocks have the largest amount of information, both the characterization and the difference shall be met. We use the weighted average of the characterization and difference measure as the measure of the amount of information in the teaching block. Then, we select the teaching block containing the largest amount of information as the new. The criteria for sample selection can be expressed as

$$\arg \max_{x_i \in U} \left( \lambda R(x_i) + (1 - \lambda) D(x_i) \right),$$

$\lambda \in [0, 1]$  is the weighting factor used to weigh the impact of sample characterization and discrepancy on dictionary construction.

It is very difficult to choose all the teaching blocks from the information in  $U$ . Therefore, in the active dictionary learning, we adopt the method of stepwise iteration: in each iteration, the comprehensive characterization and difference selection include the most information teaching blocks to join the dictionary  $S$ . It can update the collection  $U$  as well as the difference measure of  $x_i \in U$ . This is followed by a sufficient number of teaching blocks in  $S$ . It is important to note that the representation of all the samples  $X$  does not change. There is no need for recalculation of the representational measure  $x_i \in U$  in the iterative process. The procedure is described in Procedure 1.

Process 1. Active dictionary learning.

1. Initialization: for sample set  $X$ , the size  $K$  of the dictionary is built. The number  $k = 0$  has the selected samples and the dictionary  $S = \phi$ , and the sample set is selected;
2. The formula (1) is estimated the representation of the sample  $U$ ;
3. Select one of the most representative samples in  $U$  as the initial dictionary to remove it:

$$c_1 = \arg \max_{x_i \in U} R(x_i),$$

$$S = S \cup c_1, U = U - c_1, k = k + 1;$$

4. Use equation (3) to calculate the difference between  $x_i \in U$  and  $S$ ;
5. Select the sample with the largest amount of information to join the dictionary  $S$  and remove it  $U$ :

$$c_{k+1} = \arg \max_{x_i \in U} (\lambda R(x_i) + (1 - \lambda) D(x_i)),$$

$$S = S \cup c_{k+1}, U = U - c_{k+1}, k = k + 1;$$

6. If it is  $k < K$ , return to 4;
7. Output the dictionary  $S$ .

In order to facilitate the description, we will learn the dictionary  $S$  written in the form of a matrix  $C = [c_1, c_2, \dots, c_K]_{d \times K}$ . We use the first dictionary atom. Then, we

use the  $c_i$  as the  $i$ -th dictionary atomic normalization  $\tilde{c}_i = c_i / \|c_i\|$ ,  $i = 1, \dots, K$ , with the length of each vector 1. Normalized dictionary is expressed as

$$\tilde{C} = [\tilde{c}_1, \tilde{c}_2, \dots, \tilde{c}_K]_{d \times K}.$$

For simplicity, we refer to the normalized dictionary atom with "dictionary atom" below. Figure 2 shows some of the dictionaries of this paper based on active dictionary learning. It can be seen that the dummy atoms can represent edges, noise, and textures. In the local structure, the atoms have a greater difference, which makes the characteristics of the school, with the strong ability to distinguish.

### 2.3 Local feature coding

For a test teaching, we firstly extracted  $m$  teaching blocks and pre-process them to get their local feature set:

$$X^t = \{x_1^t, x_2^t, \dots, x_M^t\},$$

$x_i^t$  stands for the  $i$ -th local feature vector,  $i = 1, 2, \dots, M$ .

We will have the soft allocation coding in the dictionary  $\tilde{C}$  for  $x_i^t \in X^t$ . The coding coefficient  $x_i^t \in X^t$  is expressed

$$z_i = [z_{i,1}, \dots, z_{i,K}, z_{i,K+1}, \dots, z_{i,2K}]^T,$$

$$z_{i,j} = \max(x_i^t \cdot \tilde{c}_j, 0),$$

$$z_{i,j+k} = \max(-x_i^t \cdot \tilde{c}_j, 0), j = 1, \dots, K;$$

$x_i^t \cdot \tilde{c}_j$  shows the inner product of the  $x_i^t$  and  $j$  dictionary atoms  $\tilde{c}_j$ . We will test the corresponding code to teaching in all the expression as a matrix form

$$Z = [z_1, z_2, \dots, z_m]_{2K \times m}.$$

## 2.4 Local feature synthesis

Furthermore, we adopt the max-pooling strategy to synthesize the coding coefficients of all teaching blocks to obtain a fixed dimension of the teaching feature vector  $f$

$$f = \Psi_{max}(Z) = [f_1, f_2, \dots, f_{2K}]^T \in R^{2k},$$

$\Psi_{max}$  shows the maximum value of each row in each line of  $Z$ , and the  $j$ -th element that represents of  $f$  can be expressed as

$$f_j = \max\{z_{1,j}, z_{2,j}, \dots, z_{m,j}\}, j = 1, \dots, 2K;$$

$\max$  is taken for the maximal operation.

## 3 Evaluation of physical education in multimedia teaching quality

Because the support vector regression (SVR) has better generalization ability. We use SVR to construct the numerical mapping between teaching features and teaching quality. We realize the blind prediction of teaching quality. The method includes training and testing two processes:

1. In the training process, we first use the data assimilation method to extract the eigenvector of the training set. Then, we input the characteristics of the training set and the subjective quality score into the SVR model to train the optimal model parameters.
2. In the test process, for any given test teaching, we first use the data assimilation method to extract its eigenvector. Then, we input it into the trained SVR model. We can estimate the quality of the test teaching scores.

In this paper, we use the  $\nu$ -SVR model and the linear kernel function. We adopt the LIBSVM toolbox to implement the method.

## 4 Experimental results and analysis

### 4.1 Data set

In the process of eigen learning, we use the teaching data in the CSIQ database as the marker data. As the teaching in the CSIQ database, there is no overlap with the teaching content in the LIVE database. We use the CSIQ database for feature learning, which can more accurately verify the effectiveness of the method. CSIQ database contains 30 original tutorials and their corresponding 866 distortion teachings. The



CSIQ database contains six kinds of distortions: JPEG compression, JPEG2000 compression, additive pink Gaussian noise, Gaussian Blur, global contrast decrements and white noise. In addition, we generated some new types of distortion teaching: speckle noise, Poisson noise, and salt-peppernoise. In this paper, we used the brightness components of the teaching. We randomly selected 105 gray scale teaching blocks with of  $8 \times 8$  from these unlabeled teaching. It can estimate the ZCA transform matrix and construct the dictionary.

In order to verify the performance of this method, we experimented with the LIVE standard data. The LIVE database contains 29 original teaching methods and 5 class distortion teaching methods (total of 808), which is generated by the 29 original teaching ways. JPEG2000 compression, JPEG compression, Gaussian noise, WN, Gaussian blurring, Gblur, and Fast fading in Raylia Channel (FF) are given in the database (DMANS). The lower the DMOS value of teaching, the better the quality of teaching.

We used the teaching quality evaluation to calculate the Spearman Rank-order Correlation Coefficient (SRCC) and the Pearson's Linear Correlation Coefficient (PLCC) between the predicted mass fraction and the teaching DMOS value. As an indicator of method performance, the higher the SRCC and PLCC values, the better the performance of the method. We used the median of PLCC and SRCC as the estimation of the method performance in 100 experiments. In each experiment, we randomly selected 23 original teaching from the LIVE database. Its corresponding distortion teaching is viewed as a training set in the remaining teaching on the test. We estimated the parameters of the  $\nu$ -SVR model by cross-validation on the training set. In sections 5.2 and 5.3, the weighting factor in the active dictionary learning is  $\lambda = 0.5$  for the results. In section 5.4, we analyze different  $\lambda$  to the impact on performance.

#### 4.2 Validity of data assimilation

In order to verify the validity of the data assimilation framework, we compared this method with the feature learning based on  $K$  mean clustering (CORNIA method [10]). In the quality prediction, we use  $\nu$ -SVR model and the linear kernel function. We tested the performance of the multimedia teaching evaluation method in the dictionary  $K$  at 200, 400, 800, 1200, 2500, 5000 and 10,000. Figure 2 shows the SRCC values of the two methods on the entire test set. The number of atoms changes in the dictionary.

It can be seen that both methods have achieved good performance. When the size of dictionary is high, the  $K$  value means clustering. The active dictionary learning can be used to obtain a complete dictionary. It has a good ability to characterize. However, when the size of the dictionary is small (i.e., the feature dimension is low), this method is improved by nearly 8%, which is compared with the CORNIA method. This is mainly because the  $K$ -means clustering mainly considers the representation of the sample. The differences between samples are resulting in the small dictionary size. The dictionary atom can not characterize a variety of distortion patterns. In the

data assimilation, we consider the characterization of the sample and the difference between the samples. When the dictionary size is small, the obtained teaching features still have high discriminative ability, which can effectively overcome this defect. Therefore, data assimilation can greatly reduce the consumption of computer storage and computing resources while maintaining better performance. With the mobile phone, Tablet PC and other electronic equipment storage and computing power are limited. The data assimilation of teaching quality evaluation of the promotion and application is of great importance.

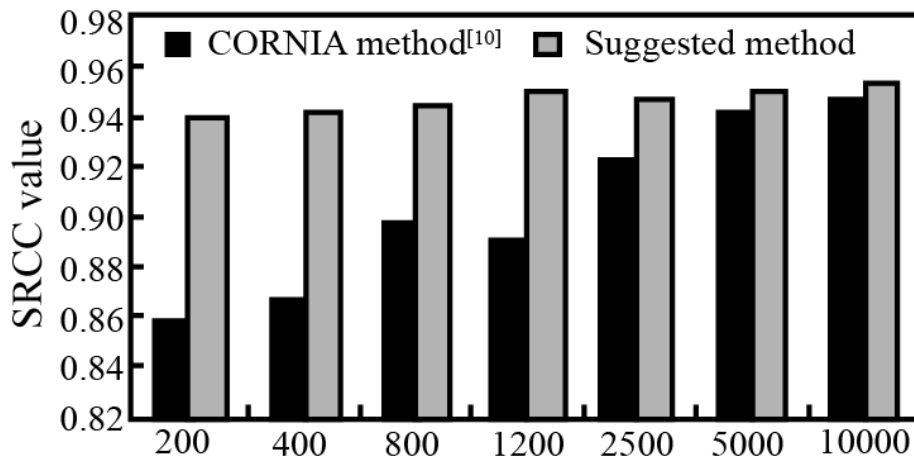


Fig. 2. SRCC values with the  $K$  number of atoms in the dictionary histogram

### 4.3 Consistent experiment

In this section, we compare the methods proposed in this paper with the existing methods of evaluation of multimedia multimedia teaching. There are BLIINDS-II [5], BRISQUE-L [6], NSS-TS [8], and CORNIA [10]. For BLIINDS-II, we use the support vector regression for quality prediction. CORNIA and the results of this method corresponding to the dictionary size are 10000.

Table 1 and Table 2 show the SRCC and PLCC values for each method in the LIVE database and for all the test sets. They show the best performance indicators in bold. Table 1 and Table 2 discover the method. The overall performance of this method is superior to that of other physical education multimedia teaching evaluation methods, which further validates the validity of the data assimilation method.

**Table 1.** SRCC values for different methods on the LIVE database

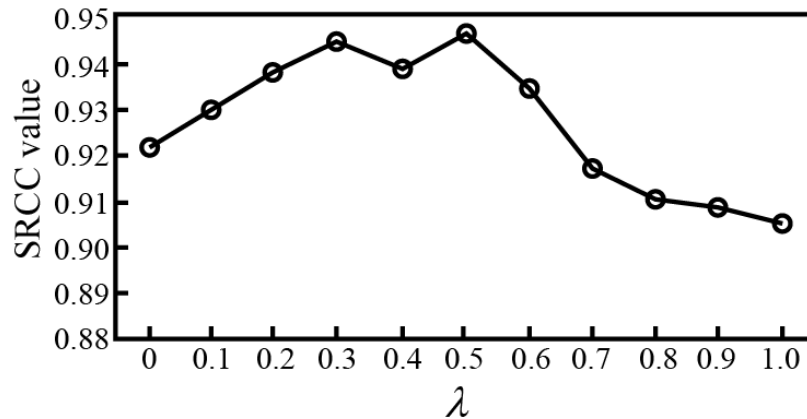
Method	JP2k	JPEG	WN	Gblur	FF	Total
BLIINDS-II	0.934	0.900	0.946	0.941	0.907	0.922
BRISQUE-L	0.928	0.913	0.979	0.961	0.921	0.936
NSS-TS	0.931	0.915	0.971	0.939	0.935	0.930
CORNIA	0.924	0.939	0.969	0.959	0.920	0.946
Suggested method	0.931	0.948	0.959	0.964	0.923	0.951

**Table 2.** The PLCC values for different methods on the LIVE database

Method	JP2k	JPEG	WN	Gblur	FF	Total
BLIINDS-II	0.934	0.920	0.936	0.927	0.900	0.909
BRISQUE-L	0.932	0.935	0.966	0.945	0.924	0.926
NSS-TS	0.947	0.933	0.963	0.950	0.942	0.926
CORNIA	0.949	0.964	0.977	0.954	0.940	0.948
Suggested method	0.956	0.970	0.971	0.961	0.944	0.952

#### 4.4 Effect of data assimilation parameters on performance

In order to prove the effect of the weighting factor  $\lambda \in [0,1]$  on the final performance, in the process of data assimilation, we selected different learning characteristics and made quality prediction. Figure 3 shows that the SRCC index  $\lambda$  of the method on the whole LIVE database. The weighting factor  $\lambda$  changes in the other dictionary size. There are similar experimental phenomena. At the same time, PLCC shows a similar change.



**Fig. 3.** Variation curve of SRCC value with weighting factor  $\lambda$

It can be seen that the best performance can be obtained when it is  $\lambda=0.5$ . Its performance is generally decreasing when it is close to 0 or close to 1. Among them, when  $\lambda$  is close to 1, the performance decreases rapidly. At this point, the decrease in the weight loss in the active dictionary learning indicates that the difference is more important than the data assimilation. It is consistent with the analysis in section 5.2. The introduction of differences makes the data assimilation framework. This paper further validates the rationality of active dictionary learning, by considering the representativeness and difference of the sample comprehensively.

## 5 Conclusion

In this paper, the author introduces the strategy of active learning into unsupervised feature learning. This paper puts forward a data assimilation framework and applies it to the evaluation of physical education multimedia teaching. The experimental results show that the evaluation method of multimedia teaching in physical education is highly consistent with human eye perception. But, it also verifies that data assimilation is superior to the traditional unsupervised feature in learning methods. Because data assimilation can maintain a better performance, a significant reduction can be reduced in the computer storage and computing resources consumption. It is related to the promotion and application of learning of great significance. In the next work, we believe that the following issues are needed to be further studied and explored: (1) research data assimilation framework, different initialization methods, different coding methods and integrated approach to the impact of performance; (2) we study the data assimilation framework based on batch, to select multiple samples in each iteration. It is to improve the efficiency of active dictionary learning; (3) we can extend the data assimilation framework to target classification and can verify its effectiveness and exploit its possible improvement.

## 6 References

- [1] Miksch S, Aigner W. A matter of time: Applying a data–users–tasks design triangle to visual analytics of time-oriented data[J]. *Computers & Graphics*, 2014, 38(1):286-290. <https://doi.org/10.1016/j.cag.2013.11.002>
- [2] Rind A, Lammarsch T, Aigner W, et al. TimeBench: a data model and software library for visual analytics of time-oriented data.[J]. *IEEE Transactions on Visualization & Computer Graphics*, 2013, 19(12):2247. <https://doi.org/10.1109/TVCG.2013.206>
- [3] Tiede D. A new geospatial overlay method for the analysis and visualization of spatial change patterns using object-oriented data modeling concepts[J]. *Cartography & Geographic Information Science*, 2014, 41(3):227. <https://doi.org/10.1080/15230406.2014.901900>
- [4] Liu K. Data-Oriented Characterization of Application-Level Energy Optimization[C]// *Fundamental Approaches to Software Engineering*. 2015:316-331. [https://doi.org/10.1007/978-3-662-46675-9\\_21](https://doi.org/10.1007/978-3-662-46675-9_21)

- [5] Gschwandtner T, Aigner W, Miksch S, et al. TimeCleanser: a visual analytics approach for data cleansing of time-oriented data[J]. *Driven Business. ACM*, 2014:1-8.
- [6] Chen Q. Research for influence of physical education multimedia teaching on physical education motivation of students[J]. *Advances in Information Sciences & Service Sciences*, 2012, 9(6):1241-1255.
- [7] Xu J. A Study of Extension Strategies of Multimedia Online Teaching Platform in Physical education Teaching of Universities[J]. *Journal of Computational & Theoretical Nanoscience*, 2017, 20(1):421-430.
- [8] Lin X Y. Research of Information Physical education Teaching System based on Multimedia CAI[J]. *International Journal of Advancements in Computing Technology*, 2013, 12(6):429-435(7).
- [9] Che Y. Multimedia application in physical education teaching drawbacks[J]. *Electronic Test*, 2013, 30(1):147-152.
- [10] Zhang T, Wen H, Hu S. The impact of using the visual multimedia presentation platform on the teaching effectiveness of public physical education[J]. *International Journal of Information Technology & Management*, 2014, 13(1):64-71. <https://doi.org/10.1504/IJITM.2014.059150>

## 7 Author

**Lifu Xing** is with Ningbo Dahongying University, Ningbo, Zhejiang, China (188872807@qq.com).

Article submitted 07 February 2018. Resubmitted 15 March 2018. Final acceptance 31 March 2018. Final version published as submitted by the author.