Handwriting Digital Recognition via Modified Logistic Regression

Cungang Wang
School of Computer Science, Liaocheng University, Shandong, China
Email: c.g.wanglc@gmail.com

***Abstract*—Motivated by a wide range of real world applications of hand writing digital recognition, e.g., postal code recognition, the past decades have seen its great progress. The related approaches are generally composed of two components, feature extraction and identification methods. We note that the previous approaches are limited by the following two aspects: (1) the feature is not adaptive enough to cover the great variance within data; (2) the recognition methods are suffered from local minima solution. Inspired by these observations and to overcome these limitations, we in this paper propose an approach HMM-MLR by exploiting hidden Markov model (HMM) and modified logistic regression (MLR). In the proposed approach, HMM is employed to model the trace of handwriting digital, which is able to model the large variance within digitals and can adapt to the data distribution. Then the features are extracted based on HMM and then delivered into MLR for recognition. Benefitting from the global optimum solution of MLR, the proposed approach could reach highly stable results. To verify the effectiveness of the proposed approach, we experimentally compare our proposed approach with other state-of-the-art approaches over Semeion handwritten digit dataset. The experimental results show that, over both recognition accuracy and recall, for different rounds of experiments and different number of training samples, our HMM-MLR exhibits significant improvement over others.**

***Index Terms*—Handwriting Digital Recognition; Hidden Markov Model; Modified Logistic Regression**

I. INTRODUCTION

Researchers started the investigation of pattern recognition theory and applications since 1920s. Along with the rapid development of computer and artificial intelligence techniques, it has been a branch of science until early 1960s. The continuous development of recognition techniques [1] obtains a wide range of applications, for instance, character recognition, image processing, speech recognition, biosensor. Among them, handwriting digital recognition and image recognition [2] are important applications, on the base of image processing, image representation, pattern recognition techniques. The progress of these techniques brings more and more researchers to this field, working on both algorithms and applications [1, 2].On the other hand, according to the statistics, for human being, more than 70% information is obtained through visual signal, which indicates that visual signal is the most important way for communication. The large amount of images or visual information has been more and more important during the object recognition in recent years [1]. For this reason, image recognition becomes a popular topic [1-3].

In recent years, the fast development of artificial intelligence field provides a number of new approaches [1, 2] especially for image recognition and handwriting digital recognition [2]. As an important tool for understanding the real world, recognition techniques play an increasingly important role in both industry and daily life. With the development and applications of Internet and information science, the scale of image is increasing. Specifically, handwriting digital recognition techniques are used to handle more 10 millions of mails every data. Moreover, handwriting digital recognition techniques are demanded by character recognition is document analysis.

Researchers introduced a number of methods [1, 6-10] into image recognition [1, 5], such as neural network, support vector machine, Bayesian decision theory, hidden Markov models (HMM) [11]. These methods can be categorized into two classes, discriminative learning approaches and generative learning approaches. Typical discriminative approaches include support vector machine, neutral network, naïve Bayes, which typical generative approaches include HMM and so on.

Artificial neural network is based on empirical risk minimization and suffers from over-fitting, slow learning speed and other disadvantages during the pattern recognition. For support vector machine (SVM) [5], its model parameters exert a tremendous influence to the identification performance [6,7]. That is, appropriate parameters tend to have high performance of image recognition and handwriting digital recognition while inappropriate parameters cause degeneration in both computational efficiency and recognition accuracy. The naive Bayesian method could hardly satisfy the large scale identification problem. And it is difficult to realize the assumption of recognition unconditional independence. Additionally, it increases the learning and training complexity.

A stochastically process is named as Markov process if its future state at time $t+1$ only depends on its current state at time t, and is independent with its past states before time t. Hidden Markov Model (HMM) is a model on the base of the Markov assumption, with hidden states introduced. The observed events are not conducted deterministically. And they were connected

doi:10.4304/jmm.9.8.1048-1053

using a set of probability distribution. HMM was a dual random process, which consisted of two randomly processes, Markov chain as well as general randomly process. Markov chain described the state transition by means of transition probability, while general stochastically process described the relationship between state and observation series through observation probability. HMM [11] is a probability model which describes the statistic property of stochastically process by parameters. The model has been widely adopted in the areas of speech identification, character recognition, genetic analysis as well as bioinformatics engineering, etc.

From the above disadvantages, this paper proposes an approach for handwriting digital recognition that is on the basis of the hidden Markov model (HMM) and modified logistic regression (MLR). First, we enhance and thin the image of handwriting digit, so that we are able to extract the skeleton of digitals and then are able to separate them into individual digital. Second, we extract the location points from the skeleton and quantize the sequence of location points to a sequence of direction vectors. The reason is that, the direction vector is robust to rotation of digital which location sequence does not. Third, we use HMM to model the distribution of direction sequence of digitals. This can be implemented using standard learning method. Forth, the transition matrix conditioned on sample is employed as feature to represent the sample. Such feature exploits the data distribution and hidden states. Fifth, the extracted feature is delivered into MLR for recognition.

The contributions of the proposed method in this paper are threefold: (1) we propose to model the handwriting digit using hidden Markov model (HMM) and extract features based on HMM. To our best knowledge, the proposed approach is novel to this field. (2) We propose a modification for logistic regression so that it has better generalization ability than logistic regression. The modified logistic regression (MLR) is computationally efficient. (3) The proposed approach HMM-MLR combines the abilities of HMM and MLR for a comprehensive recognition [12].

The remainder of this work is organized as follows. Section 2 will present the proposed method, including hidden Markov model for feature extraction, modified logistic regression for recognition. Section 3 will verify the proposed method by a group of carefully designed experiments. We draw a conclusion in Section 4.

II. Our Proposed Schema

On the basis of the above discussion, in this section, we propose a method for handwritten digit recognition. Our approach is based on the hidden Markov model and modified logistic regression, and is able to surmount the limitation of previous approach. Our developed algorithm is graphically illustrated in Fig. 1. There are three main steps. First, collect data; second, feature extraction via HMM model in this Section; third, MLR model training and test [13].

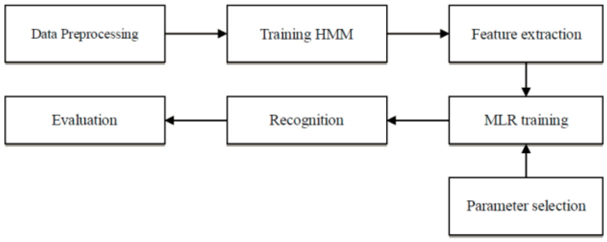

Figure 1. The flow chart for experiment

A. Modeling Handwritten Digit Using HMM

The hidden Markov model (HMM) is a probabilistic generative model developed to model the distribution of sequences. It is composed of a set of hidden states and a set of observed states which are controlled by the hidden states. Here we first introduce the mathematical notations and then present the formulation of hidden Markov model. The quintuple $\lambda=(N,O,A,B,\pi)$ is employed to describe HMM [11], where $O=\{V_1,V_2,\ldots,V_M\}$ is the set of M observed states; $S=\{S_1,S_2,\cdots,S_N\}$ is the set of N hidden states; o_t is the observed state at time t; q_t is the hidden state at time t.

Given the above notations, we now introduce the basic probabilities. Let $A=\{\alpha_{ij}\}$ be state transition probability matrix, where α_{ij} is transition probability from state S_i to state S_j. Let $B=\{b_{kj}\}$ be the output probability matrix where b_{jk} is the output probability from hidden state S_j to observed state V_k. Let $\pi=\{\pi_i\}$ be the initial probability where $\pi_i=P(q_1=S_i)$. When solving an estimation problem, for a given model and state transition series $q=(q_1,q_2,q_3...q_T)$, the probability of an observed series $O=(o_1,o_2,o_3...o_T)$ could be computed through the following equation:

$$P(o|q,\lambda)=P(o_1|q_1)\cdots P(o_T|q_T)=b_{q_1}(o_1)\cdots b_{q_T}(o_T) \qquad (1)$$

If λ is given, the probability of $q=(q_1,q_2,q_3...q_T)$ can be compute through

$$P(q|\lambda)=\pi_{q_1}\alpha_{q_1q_2}\cdots\alpha_{q_{T-1}q_T}$$

Therefore, the estimation probability of the observed sequence can be is written as,

$$P(o,q|\lambda)=P(o|q,\lambda)P(q\mid\lambda) \qquad (2)$$

An observation sequence may have more than one corresponding state transition sequences; thereby all the state transition sequences could be is expressed as

$$\begin{aligned}P(o\mid\lambda)&=\sum\nolimits_q P(o,q\mid\lambda)\\&=\sum\nolimits_{q1,q2,\ldots qT}\pi_{q1}b_{q1}(o_1)a_{q1q2}b_{q2}(o_2)\ldots a_{qT-1qT}b_{qT}(o_T)\end{aligned} \qquad (3)$$

The hidden Markov model can be efficiently solved using forward-backward algorithm [11].

B. Modified Logistic Regression for Identification

Both real and binary responses could adopt Logistic Regression, whose output posterior probabilities can be processed expediently and sent to other systems. It attempts to simulate the class label's conditional probability offer its observation:

$$\mathrm{p}(\mathrm{y}\mid\mathrm{x}) = \frac{1}{1+\exp\left(-y\left(w^T x+b\right)\right)} \tag{4}$$

where $x=(x_1,\cdots,x_m)^T$ is the example vector; m is the number of features; $\mathrm{y}\in\{+1,-1\}$ is the category label; $\mathrm{w}=(w_1,\cdots,w_m)^T$ is the weight vector; b is decision intercept. The weight can be estimated as,

$$\hat{w} = \arg\min_w \left\{\frac{1}{n}\sum_{i=1}^{n}\log\left[1+\exp\left(-y_i\left(w^T x_i+b\right)\right)\right]+\lambda w^T w\right\}$$

Please notice which the Hessian matrix of the objective function $O(\mathrm{w})$ is:

$$\mathrm{H} = \frac{d^2 O(w)}{dwdw^T} = \frac{1}{n}\sum_{i=1}^{n}\frac{\exp\left(-y_i w^T x_i\right)}{\left(1+\exp\left(-y_i w^T x_i\right)\right)^2} x_i x_i^T + 2\lambda\mathrm{I}$$

Here the identity matrix is I. Thanks to $\lambda > 0$, the Hessian matrix above mentioned is positive definite, that means the objective function of regularized LR with strict convexity, hence, we can see its solution, which is unique and global.

In this part, we first state that, by constructing a series of optimization problems, the solutions converge to the solution of SVM. Thus, SVM could be solved by simple unconstrained optimization techniques. Then we put forward our simple MLR-CG algorithm that uses CG as its inner loop.

To simplify our derivations, we take advantage of the augmented weight vector $\mathrm{w}=(b,w_1,w_2,\ldots,w_m)^T$ and the augmented sample vector $\mathrm{x}=(1,x_1,x_2,\ldots,x_m)^T$ from right now unless otherwise specified. To keep the SVM optimization problem unchanged, its form becomes,

$$\hat{w} = \arg\min_w \left\{\frac{1}{n}\sum_{i=1}^{n}\max\left\{0,1-y_i w^T x_i\right\}+\lambda\sum_{j=1}^{m} w_j^2\right\}$$

The intercept $\omega_0 = b$ is not in the regularization term. We also need not to penalize the intercept ω_0 in the regularized LR to approximate SVM:

$$\hat{w} = \arg\min_w \left\{\frac{1}{n}\sum_{i=1}^{n}\log\left(1+\exp\left(-y_i w^T x_i\right)\right)+\lambda\sum_{j=1}^{m} w_j^2\right\}$$

From former discussions we could see which loss functions play an important role in the SVM and LR. The SVM loss function can be approximated by the loss of the following modified LR:

$$g_\gamma(x,y,w) = \frac{1}{\gamma}\ln\left(1+\exp\left(-\gamma\left(yw^T x-1\right)\right)\right) \tag{5}$$

If we could approximate the SVM loss function, $g_{svm}(x,y,w)=\max\left\{0,1-yw^T x\right\}$

With the above sequence of functions $\{g_\gamma\}$, then the problem can be resolved with simple unconstrained optimization techniques.

The conjugate gradient method follows the above convergence proof. That is to say, we calculate the solution of SVM by solving the problems a sequence of sub-optimization problems. Especially, we solve each sub-optimization problem by means of the conjugate gradient. For solving large-scale nonlinear optimization problems, conjugate gradient is one of the most popular methods. More importantly, it is compared with other methods in fitting LR, which shows that it is more efficient. We could determine that the HS direction in the experiment is more efficient than the other two directions. We make a list of our conjugate gradient method below, which is an iterative method in the inner loop [14].

Algorithm 1: Conjugate Gradient
1. Initialize $w=0$, $\gamma=1.0$, $l=10$, and $\delta=10$
2. Repeat until convergence:
(a) Initialize conjugate gradient by collocation its search direction to minus gradient
(b) Minimize O_γ with l steps
(c) Increase $\gamma\leftarrow\gamma+\delta$

In fact, we should start from small γ as well as run not increase γ to infinity for two reasons. The first reason is which when γ is big the Hessian matrix is ill-conditioned, that will result in the unsuitability of our algorithm. Starting from small γ and gradually increase it will lead us obtain the stable solution.

For example, when $\gamma=200$, it is at most 0.003, that is already not influential for our problems. As well as later on, in our experiments we will give a summarize that this approximation will not degrade the performance of our trained classifier. We run not need increase γ after each conjugate gradient procedure since we should let conjugate gradient do at least several steps to fully adopt its power in finding conjugate directions; and also we do not require to wait until it converged before we change γ. We set both δ as well as l to be 10 in our experiments. And each round when γ is changed, conjugate gradient should be re-initialized. We use 200 conjugate gradient steps in our experiments as the stopping criteria. And we also make use of other criteria like the change of weight vector or objective value.

III. Experimental Results

In this section, we will validate our proposed method HMM-MLR for handwritten digit recognition. The experimental steps contain the following steps: (1) data collection; (2) model the data and extract the feature; (3) train the model and perform test. The experimental procedure are summarize in the above section. This part will sequentially report the dataset, verification criterion and experimental results [15].

A. Experimental Database

The Simenon Handwritten Digit instance used in our experiments were collected using Tactile Sal, Brescia, Italy and donated in 1994 to Simenon Research Center of Sciences of Communication, Rome, Italy, for machine learning research. 1593 handwritten digits from around 80 persons were scanned, stretched in a rectangular box 16×16 in a gray scale of 256 values. Then each pixel of each image was scaled into a binary value by a fixed threshold. Each person wrote on a paper all the digits from 0 to 9, twice. The commitment was to write the digit the first trial in the normal way (trying to write each digit accurately) as well as the second time in a fast way (with no accuracy). The best validation protocol for this database seems to be a 5×2CV, 50% Tune (Train +Test) and completely blind 50% Validation. This dataset consists of 1593 records and 256 attributes. Each record represents a handwritten digit, originally scanned with a resolution of 256 grays scale (28). Each pixel of the each original scanned image was first stretched, as well as after scaled between 0 and 1 (collocation to 0 every pixel whose value was under the value 127 of the grey scale as well as setting to 1 each pixel whose value in the grey scale was over 127). Each binary image was scaled again into a 16×16 square box (the final 256 binary attributes). The categories described as shown in Table I.

TABLE I. CLASS DISTRIBUTION OF SAMPLES

Categories	Number of samples
Man	873
Woman	720
Total number of samples	1593

B. Verification Criterion

To validate the ability of our proposed HMM-MLR method for handwritten digit recognition, as well as compare with different approach, we in this paper use the following assessment criterion to assess the HMM-MLR approach for handwritten digit recognition. Specifically, we define the recognition accuracy, recognition precision, identification recall and classification true negative rate. The definitions are report in Table II. In this table, TN represents true negative; TP denotes true positive; FN represents false negative; FP denotes false positive. These evaluation standards can be directly made use of to measure two class recognition problem of handwritten digit recognition, and multiple class recognition problem of handwritten digit recognition

TABLE II. THE EVALUATION CRITERION THE HMM-MLR APPROACH FOR HANDWRITTEN DIGIT RECOGNITION

Evaluation criterion	Definition
Accuracy	$(TP+TN)/(TP+TN+FP+FN)$
Precision	$TP/(TP+FP)$
Recall	$TP/(FN+TP)$
True negative rate	$TN/(TN+TP)$

C. Main Results

In the first experiment, we validate our proposed HMM-MLR approach for handwritten digit recognition, over the Semeion Handwritten Digit dataset. We employ two comprehensive criterions, accuracy and recall, for experimental verification. Identification accuracy and recall are two typical and popular measures for the correctness of the identification model. The experimental procedure is show in the experiment section. The preprocessing step as well as feature extraction step is important due to their encoding discriminant information. Our proposed algorithm HMM-MLR is trained by the above approach, and some parameters of HMM-MLR are obtained using cross-validation strategy. We conduct the test for multiple rounds, where in each round we stochastically divide the database to training configure as well as test configure.

We extensively compare our proposed HMM-MLR method for handwritten digit recognition with three algorithm, HMM-KNN, HMM as well as SVM. The classification results are show in Table III as well as Fig. 2. These experimental results indicate which: (1) our proposed method HMM-MLR outperforms all three compared methods significantly, under the distinct experimental configurations, distinct number of training example, and distinct evaluation standard. (2) Our developed method exhibit robustness against the round of experiments, and the verification criterion, which no wonder mean which our proposed algorithm could be used to a lot of tasks. The reasons are three folds. (1) The HMM-MLR has the ability to map the nonlinear data in the low dimensional space to the high dimensional space by the feature extraction, which makes the classification problem simple. (2) In comparison with empirical parameter selection approach, the selection method for parameters can adapt to the dataset. (3) The experimental procedure of our method could provide informative features and could maximize the discrimination ability.

TABLE III. CLASSIFICATION PERFORMANCE COMPARISION OF FOUR METHOD FOR HANDWRITTEN DIGIT RECOGNITION

Experiment round	approach	Evaluation Criterion	
		Accuracy	Recall
Round 1	HMM-KNN	85.65	84.57
	HMM	80.62	83.23
	SVM	83.72	87.27
	HMM-MLR(ours)	87.88	85.93
Round 2	HMM-KNN	85.05	83.55
	HMM	84.55	81.88
	SVM	84.32	87.71
	HMM-MLR(ours)	88.62	87.35
Round 3	HMM-KNN	85.56	83.69
	HMM	84.57	80.81
	SVM	82.73	83.25
	HMM-MLR(ours)	88.70	86.93

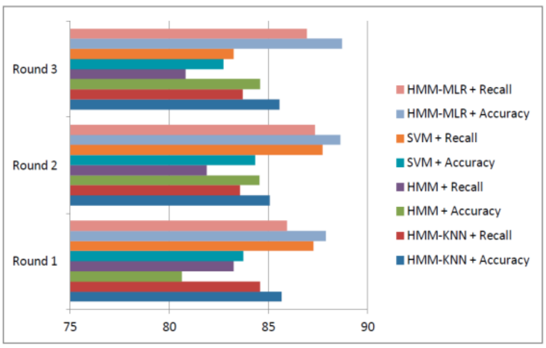

Figure 2. The performance comparision of experimented algorithm for handwritten digit recognition

In the second experiment, we validate the ability of the proposed HMM-MLR algorithm in handwritten digit recognition, by means of comparison experiment. Two popular criterions accuracy as well as recallis adopt for evaluation. Identification accuracy and recall are two typical and popular measures for the correctness of the identification model. The used database is Semeion Handwritten Digit dataset. The experimental step is show in above part. Our algorithm HMM-MLR is learnt by means of the method in above part, where some parameters of HMM-MLR are configure to defaults. The tests are run for several rounds over stochastically separate dataset.

The accuracy and recall is use as the assessment standard for the handwritten digit recognition. We do the experiment for 20 rounds as well as summarize the experimental results of partial round are in Table IV as well as Fig. 3. As report in Table IV, using by our method to determine the parameters, HMM-MLR for handwritten digit recognition reaches the highest performance of 91.29% under the standard of accuracy, while HMM-MLR achieve the highest performance of 88.45% under the standard of recall. Moreover, the average accuracy of HMM-MLR is 88.70% that outperforms which of HMM-KNN (85.43%). The potential reasons for these results are mainly threefold. Firstly, HMM-MLR has the ability to map the nonlinear data points in the low dimensional space to the high dimensional space by HMM based feature extraction, which simplifies the classification problem. Secondly, the parameter selection can be adaptive to different dataset, in comparison with the empirical parameter selection method. Thirdly, the framework of the proposed approach is composed of a group of comprehensive procedures which sequentially maximize the identification ability.

TABLE IV. THE PERFORMANCE COMPARISON OF DIFFERENT METHOD

Experiment	Algorithm	Evaluation Criterion	
		Accuracy	Recall
Round1	HMM-KNN	85.65	84.57
	HMM-MLR(ours)	87.88	85.93
Round 2	HMM-KNN	85.05	83.55
	HMM-MLR(ours)	88.62	87.35
Round 3	HMM-KNN	85.56	83.69
	HMM-MLR(ours)	86.78	87.96
Round 4	HMM-KNN	84.47	83.29
	HMM-MLR(ours)	88.19	86.10
Round 5	HMM-KNN	84.61	83.75
	HMM-MLR(ours)	91.29	86.83
Average	HMM-KNN	85.43	84.83
	HMM-MLR(ours)	88.70	86.93

In the third experiment, we conduct experiments over Semeion Handwritten Digit dataset. The dataset includes have 1583 samples with 256 attributes, and the dataset is divided into two part, the training data set with 1400 samples and the test data set with 183 samples. This experiment verifies the ability of the proposed HMM-MLR for handwritten digit recognition, and the capability of the optimization method. The experimental scheme is summarize in the previous part of this work. It adopts a criterion approach to learn the HMM-MLR based on the database Semeion Handwritten Digit dataset. In this experiment, HMM-MLR adopts the default parameters configure by means of authors. The assessment criterions are accuracy as well as recall where Identification accuracy and recall are two typical and popular measures for the correctness of the identification model.

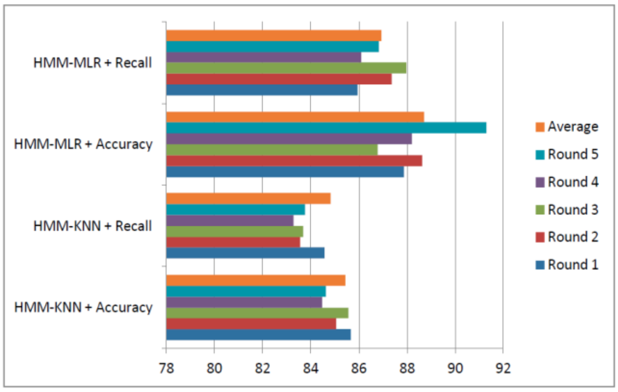

Figure 3. The comparison of experimented method over two criterion

TABLE V. THE PERFORMANCE COMPARISON OF HMM-MLR FOR HANDWRITTEN DIGIT RECOGNITION

Training Samples	Method	Evaluation Criterion	
		Accuracy	Recall
30%	SVM	72.85	73.83
	HMM-MLR (ours)	74.57	75.07
40%	SVM	78.44	81.49
	HMM-MLR (ours)	83.08	82.77
50%	SVM	84.44	86.66
	HMM-MLR (ours)	90.26	87.87
60%	SVM	87.87	91.04
	HMM-MLR (ours)	90.26	90.36
70%	SVM	90.12	91.71
	HMM-MLR (ours)	95.41	92.59

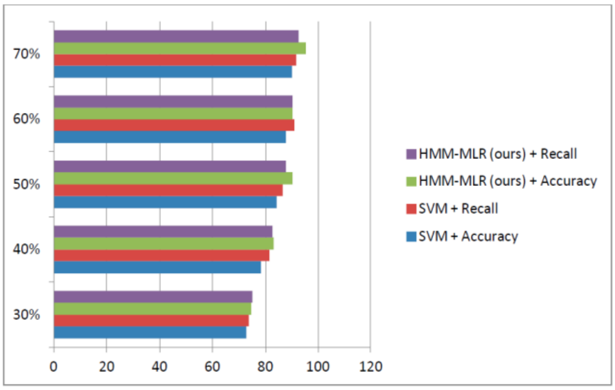

Figure 4. The performance comparison of HMM-MLR for handwritten digit recognition

Here we verify our developed algorithm by comparing it with HMM over two criterions, accuracy as well as recall. The experimental results for handwritten digit recognition are summarized in Table V as well as Fig. 4. As summarize in this table, we can see that, the values of accuracy as well as recall of our developed algorithm are much higher than that in HMM. That is to say, about 90.26-84.44% for accuracy and 87.87-86.66% for recall. We also found that, HMM-MLR is higher than HMM for distinct number of training data, over two criterions. In all experiments with different collocation, the results of our approach exhibit advantage over the HMM approach. These results can be intuitively explained. First, comparing with the traditional machine learning based methods the HMM-MLR can be well applied to the condition that the sample data is large scale, high dimension and contains a large number of heterogeneous information. Second, in comparison with the empirical parameter selection method, our method can adapt to the dataset. Third, the framework of the proposed algorithm contains a group of comprehensive procedures which sequentially maximize the performance.

IV. SUMMARY

This paper proposed a handwriting digital recognition method based on hidden Markov model and modified logistic regression. Specifically, hidden Markov model models the distribution of sequence of digital, where the digital is preprocessed and represented as a sequence of direction vector. Then, the features of handwriting digital are extracted from hidden Markov model, i.e. use the matrix of transfer state as the feature. The extracted features are then delivered to modified logistic regression for recognition. The modified logistic regression method is an extension of logistic regression by modifying its loss function which approaches to the hinge-loss of support vector machine, but keeps the global optimum solution. The proposed approach HMM-MLR has the follows three advantages: (1) the sequence of direction vector is robust to the rotation of images; (2) the extract feature is adaptive to data distribution and exploits hidden states; (3) the modified logistic regression has global optimum solution which produces robust results. The experiment results over real world database, by comparing with other popular methods validate the advantages of our proposed approach.

REFERENCES

[1] Duda, Richard O and Hart, Peter E and Stork, David G. Pattern classification. *John Wiley & Sons*, 2012.

[2] Haijun Li, Application Research of Food Quality Detection Based on Computer Vision, *Journal of Computers*, Vol 8, No. 7, pp. 1758-1762, 2013

[3] Xu, Lei and Krzyzak, Adam and Suen, Ching Y. Methods of combining multiple classifiers and their applications to handwriting recognition, *IEEE Transactions on Systems, Man and Cybernetics*, vol 22, no 3, pp. 418-435, 1992.

[4] Huang Weiyong, Shao Xiaogen, Chen Kui, Coal-and-Gas Outburst Forecast Using CCPSO and SVM, *Computer Science*, Vol. 39, No. 11, pp. 216-220, 225, 2012

[5] Haralick R M. Statistical and Structural Approaches to Texture, *Proceeding of IEEE*, Vol. 67, No. 5, 1957

[6] Tang Jingtian, Hu Dan. SAR image classification research based on SVM. *Remote Sensing Technology and Application*, Vol. 23, No. 3, pp. 56-58, 2008

[7] Li Yi, Huachun Zhou, Fei Ren, Hongke Zhang, "Analysis of Route Optimization Mechanism for Distributed Mobility Management", *Journal of Networks*, Vol. 7, No. 10, pp. 1662-1669, 2012

[8] Yi Yan, Zhao Tao, Qi lili. Evaluation and research of knowledge management maturity based onneural network of particle swarm. *Journal of Information*, Vol. 29, No. 5, pp. 124-128, 2010

[9] Jian Wu, Zhi-ming Cui, Jian-ming Chen, Guang-ming Zhang, A Survey on Video-based Vehicle Behavior Analysis Algorithms *Journal of Multimedia*, Vol. 7, No. 3, pp. 223-230, 2012

[10] Jing Li, Xiaowen Zhang Geometrically Invariant Watermarking Scheme Based on Local Feature Points, *Journal of Multimedia*, Vol. 7, No. 3, pp. 231-238, 2012

[11] Guangming Zhang, Zhiming Cui, Pengpeng Zhao, JianWu, A Novel De-noising Model Based on Independent Component Analysis and Beamlet Transform, *Journal of Multimedia*, Vol. 7, No. 3, 247-253, 2012

[12] Eddy, Sean R, Hidden markov models, *Current opinion in structural biology*, vol. 6, no. 3, pp. 361-365, 1996.

[13] Yueqiu Jiang, Xiangwen Fu, Hongwei Gao, A New Circular Region Detection Algorithm based on the Geometric Characteristics, *Journal of Software*, Vol. 8, No. 11, pp. 2899-2907, 2013

[14] Wei Zhan, Zhi-Qing Luo, Research of Vehicle Type Recognition System Based on Audio Video Interleaved Flow for Toll Station, *Journal of Software*, Vol. 7, No. 4, pp. 741-744, 2012

[15] Zhenbing Liu, Shujie Jiang, Huihua Yang, Jianguo Liu, Arbitrary-length Fast Hartley Transform without Multiplications, *Journal of Computers*, Vol. 8, No. 10, pp. 2656-2663, 2013

Cungang Wang received his B.E. degree in educational technology from Liaocheng University, Liaocheng, China, in 2000, and his M.S. degree in computer science and technology from Ocean University of China, Qingdao, China, in 2006. He worked as a lecturer at the School of Computer Science at Liaocheng University since 2008. His research interests include machine learning, image processing and pattern recognition.