The Application of Hidden Markov Model in the Eye Movement Data

Liwang Ma[1a],Mi Li[2b],Xiaodong Wang[c],Shengfu Lu[d]

International WIC Institute, Beijing University of Technology,Beijing 100124, China

Beijing International Collaboration Base on Brain Informatics and Wisdom Services,Beijing 100124, China

Beijing Key Laboratory of MRI and Brain Informatics,Beijing 100053, China

E-mail:[a]malw_happy@163.com,[b]limi@bjut.edu.cn,

[c]wangxiaodong@emails.bjut.edu.cn, [d]lusf@bjut.edu.cn

Ning Zhong

International WIC Institute, Beijing University of Technology,Beijing 100124, China
Beijing International Collaboration Base on Brain Informatics and Wisdom Services,Beijing 100124, China
Beijing Key Laboratory of MRI and Brain Informatics,Beijing 100053, China
Maebashi Institute of Technology Maebashi-City 371-0816, Japan
E-mail: zhong@maebashi-it.ac.jp

Hidden Markov model (HMM) is a statistical model that is widely applied in speech recognition, text information extraction, gait recognition. eye movement data is obtained from tracking eye movement which has the characters of temporality and spatiality, and Hidden Markov model can describe these two characters. Therefore, Hidden Markov model is applyied to the classification Internet user based on eye movement data. HMM can be used to establish the formation and evolution of the visual behavior related to different users online. Different user models are established. The classification principle is maximum generating probability, the model of biggest output probability is the type of user. The genetic algorithm is introduced to optimize the Initial parameters of Hidden Markov model to improve the accuracy of user classification. The experimental results show that the user classification method based on the eye movement data can be used to classify the old users and young users with high accuracy. It is proved that HMM is feasible and effective for classification.

---

[1]Speaker

[3]Corresponding Author

1. Introduction

Currently computer can only passively accept the user's information request through a keyboard, mouse, touch screen, but The users are able to quickly get a lot of information from the computer interface and audio, which caused an imbalance of interactive bandwidth. To solve this problem, computer is required to have the ability to simulate human senses of receiving information, to identify and understand the surrounding environment. The web Site is important interactive interface, and Vision is the main manner that users access information on the website. There are many research on the website by vision, Silva et al. (2012) used eye tracking technology to identify the target (plain text and multimedia) to improve the accuracy of the web site to find the key target[1]. Al-Wabil, A et al. (2008) recorded the users with dyslexia gaze time, gaze sequence and gaze direction by eye tracking technology to study the visual attention of the structure of web page navigation[2]. Li Qian, et al. studied the visual behavior of the user's browsing web pages to optimize web page design by using eye tracking technology[3].But these researches focused on the study of the same user on different types of information on the web, and there are few research that focused on the users classification of the same information.

There are differences between eye movement patterns of different ages online when users browse the Web[4-5].In this paper ,we propose a user classification of different age based on eye movement data. By analyzing eye movement patterns and pupil diameter, saccade distance, watching the frequency of eye movement, pattern recognition methods is used to the classification. The Hidden Markov Models(HMM) was selected as the Pattern Recognition Methods. Han Yu   chang has proved that the eye movement has the time sequence and spatial sequence features by studying the sequence when subjects observe different shapes and colors graphic, and   The Hidden Markov model can describe the temporal and spatial, so Hidden Markov models have been applied to the study of eye movement. Elhelw (2008) found that gaze data string can reveal the characteristics to attract the most observers note with discrete time Markov chain of fixation sequences[6]. Salvucci et al. (2002) described three kinds of methods based on sequence matching and Hidden Markov Model which was proved that the method and expert method can explain accurately the eye movement[7].Aziz F et al. extracted features from the electrooculography (EOG) signal traces of the user as commands to maneuver the wheelchair and used HMM to determine the state of the wheelchair system and generates commands for navigating the wheelchair accordingly[8].

These studies show that the Hidden Markov model can describe eye movements observed sequences of different users. Hidden Markov models is used to build different user models, and the classification is based on the principle of maximum output probability .Due to influence of the initial parameters for the hidden Markov model, genetic algorithm(GA) is selected to optimize the initial parameters to improve the recognition rate.

2. Hidden Markov Model (HMM)

Hidden Markov model (HMM) is a stochastic process determined by the two interrelated mechanisms, an underlying Markov chain having a finite number of states and a set of random functions. The observer sees only the output of the random functions associated with each state

and cannot directly observe the states of the underlying Markov chain; hence prefixed 'hidden'. Hidden Markov model can be expressed by the next notion: $\lambda=\{S,O,A,B,\pi\}$.

S is the set of states $S=[S_1,S_2,\ldots,S_N]$,N is the number of states in the model,

O *is* the set of sequence of observations $O=[O_1O_2O_3..O_M]$,M is the number of sequence,

A the transition probability distribution matrix $A=\{a_{ij}\}$, where a_{ij} is the transition probability of the Markov chain transiting to state *j*, given the current state *i*, that is $a_{ij}=P(q_{t+1}=S_j| q_t=S_i)$,$1\leq i$, $j\leq N$，where qt is the state at time t and a_{ij} has the following properties: $a_{ij}\geq0$, $1\leq i$, $j\leq N$, $\sum_j a_{ij}=1$

B is the observation symbol probability distribution matrix $B=\{b_j(k)\}$, where $b_j(k)$is the random function associated with state *j*. $b_j(k)=P(O_t=w_k| q_t=S_j)$, which represent the Observation vector is w_k in the state S_j at time t. $b_j(k)$ has the following properties:$1\leq j\leq N$,$1\leq k\leq M$, $\sum_k b_j(k)=1$

π is the initial state distribution matrix $\pi=\{\pi_i\}$ in which $\pi_i=P(q_1=i)$,$1\leq i\leq N$.

Usually Hidden Markov model parameters can be abbreviated as follow: $\lambda=\{A,B,\pi\}$.

3. TheUser Identification Based on Eye Movement Data

3.1 The Process of Eye Movement Data

The eye movement data is acquired by eye tracker Through data processing and analysis to obtain eye movement indicators, the common eye movement indicators include fixation duration, fixation times, pupil diameter, saccade distance etc…

Fixation duration is also called gaze duration, which can reflect the time of user to extract information from the gaze target and can also reflect the he degree of interest in watching the objectives and the degree of difficulty of extracting useful information. Fixation duration must be greater than 100 ms.

Fixation times is that the number of fixation point in a certain area within a certain period of time. The fixation times can reflect the frequency when people gaze one area and the gaze target and the user's area of interest. By analyzing the fixation times in one area, getting the user's gaze target and regions of interest is helpful to carry out research of the user's intention.

The average pupil diameter of left and right eye is used as pupil diameter(PD).The change of pupil diameter can objectively reflect human mental activity and has a close relationship with the degree of mental effort during the process of information. There are difference between the degree of mental effort of different users when they obtain and process information, so Pupil diameter is also different.Suppose the i participant P_i has m points in the task j,which PD_{ijk} is the pupil diameter of the k point. $PD_{ijk}=(Left_{ijk}+Right_{ijk})/2$,which $Left_{ijk}$ is the pupil diamete of left eye and $Right_{ijk}$ is the pupil diamete of right eye.

Saccade distance(SD) is the Euclidean distance between the two fixations when gaze point coordinates change, which reflect how much the amount of information to be obtained. Saccade distance of different users' visual behavior in the Internet can reflect how much information to be obtained and the perceptual span.Suppose (x_i,y_i) is the coordinate of the point i and (x_{i+1},y_{i+1}) is the coordinate of point i+1. Saccade distance of point i is the Euclidean distance,

$$SD_i=\sqrt{(x_{i+1}-x_i)^2+(y_{i+1}-y_i)^2}$$

In biometric research, the classification with multiple features is significantly better than just using a biometric identification, thus, we use the single indicators and combination indicators as Classification feature set. Before building the Classification feature set, Eye movement indicators are normalized, then the Classification feature set is used as the Observation sequence of Hidden Markov models.

3.2 The Training of HMM

Before using HMM for the classfication, an HMM must be trained to describe the observation sequence of the word accurately. There are many criteria that can be used for this problem. The most common one is to maximize the probability of the observation sequence. The BaumWelch algorithm developed by Baum is one of the most successful optimisation methods for this problem. The BaumWelch algorithm is a hillclimbing algorithm that includes a set of re-estimation formulas and guarantees that the re-estimated HMM $\bar{\lambda}$ will equal to or better than the initial HMM model. The re-estimation formulas are shown as follows:

$$-\pi = \gamma_1(i)$$

(3.1)

$$-a_{ij} = \sum_{t=1}^{T-1} \xi_t(i,j) / \sum_{t=1}^{T-1} \gamma_t(i)$$

(3.2)

$$-b_j(k) = \sum_{t=1, O_{t=W_k}}^{T-1} \gamma_t(i) / \sum_{t=1}^{T-1} \gamma_t(i)$$

(3.3)

Where the $\bar{A}$, $\bar{B}$, $\bar{\pi}$ are the model parameters of $\bar{\lambda}$, $\xi_t(i, j)$ is the probability that the state is S_i at time t and the state is S_j at the next time t+1 under the known λ and observation sequence O , $\gamma_t(i)$ is the probability that the state is S_i at time t under the known λ and observation sequence O.

With the Baum Welch algorithm, we can obtain a new model parameters $\bar{\lambda} = \{ \bar{A}, \bar{B}, \bar{\pi} \}$,which is closer optimization goal than λ. However Baum Welch algorithm being a hill climbing algorithm which is sensitive to the initial value of parameters, any arbitrary chosen initial model λ will usually lead to a sub-optimal model[9], so we need optimization algorithm to change that.

3.3 The Training of Hidden Markov Model Based on Genetic Algorithm

The genetic algorithm (GA) is a searching process based on the laws of natural selection and genetics which is a method to search the optimal solution. It can evaluate each individual of the population from any of a random initial population by using selection, crossover, mutation, then the new individual is generated to achieve a stable state of optimization of the breeding and selection. GA use a random set rather than a point to search global optimal solution in the process of generating new population, so it can reduce the possible of falling into the local optimal solution. Kwong have successfully applied GA into HMM-based speaker recognition system, which the model parameters is better than obtained from the traditional training Baum-Welch algorithm[10]. Inspired by this we apply the GA to improve the online user identification training HMM model parameters. The flow chart based on GA-HMM is flowed in Figure 1:
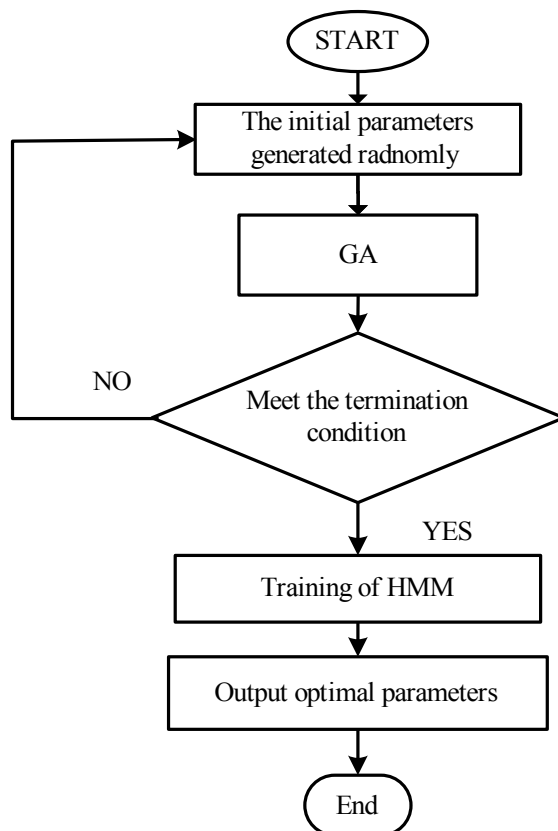
Figure1: The flow chart based on GA-HMM

When Hidden Markov Model parameters are optimized by genetic algorithm, the parameters must be encoded first. Due to A, B, π are three real matrix and the dimension may reach thousands, we use real matrix coding method which can direct reflect the change of parameters rather than binary encoding which is not real .After encoding A, B, π, the parameters sequence are combined into a gene cluster which constitute an individual(chromosome).

The select of fitness function of GA will affect the results, so we select $P(O|\lambda)$ as the fitness function to ensure the optimization results.

We use proportional selection method as a selection operator, the formula is as follows :

$$F_i / \sum_{i=1}^{M} ¿ F_i$$, i=1,2,…M, where M is the population size, F_i is the fitness of individual I, p_i is

$$p_i = ¿$$

the possible to be selected. Arithmetic crossover is used as crossover operation , which is two individuals produce two new individuals by a linear combination. The cross arithmetic operation target is normally individuals by real coding. The single point random mutation is used as mutation operator ,it refers that one or several genes are randomly assigned to carry out mutation operation a mutation probability p_m.

Genetic algorithm is an iterative process, during each iteration, fitness calculation, selection, crossover and mutation are performed until meeting the termination condition. There are two ways to terminate the Genetic Algorithm, the first one is the maximum number of iterations T, the second one is the minimum deviation. we select the first one .

There are four operating parameters(Population size、 Maximum iterations T、 Crossover probability p_c、 Mutation probability p_m) of GA to be set in advance . M ranges from 50 to 500,

T ranges from 10 to 100, p_c ranges from 0.1 to 0.9, p_m ranges from 0.01 to 0.1. p_c and p_m should not be too large or too small, or leading to premature convergence and falling into local optimum.

When Hidden Markov Model is optimized by Genetic Algorithms, we use GA optimize the initial parameters of HMM, and then the parameters is trained by Baum-Welch to building different users' models. The classification is based on the principle of maximum output probability.

3.4 Model Classification

In order to identify the type of Internet user, $P(O| \lambda_i)$ is calculated which represent the conditional probability of the Observation sequence in model λ_i. Based on the maximum output probability principle, the result is the model correspond to the maximum $P(O| \lambda_i)$, that is $\lambda^* = \arg \max P(O| \lambda_i)$.

$$P(O, \lambda) = \sum_{i=1}^{N} \alpha_t(i)\beta_t(i) = \sum_{i=1}^{N}\sum_{j=1}^{N} \alpha_t(i)a_{ij}b_i(o_{t+1})\beta_{t+1}(j), 1 \leq t \leq T-1 \qquad (3.4)$$

Where $\alpha_t(i) = P(O_1 O_2 \ldots O_t, q_t = S_i \mid \lambda)$ is the forward probability that the observation sequence is $O_1 O_2 \ldots O_t$ until time t and the state is S_i at time t. $\beta_t(i) = P(O_{t+1} O_{t+2} \ldots O_T, \mid q_t = S_i, \lambda)$ is the backward probability that the state is S_i at time t and after time t the observation sequence is $O_{t+1} O_{t+2} \ldots O_T$.

4. Experimental Verification

More and more elderly people have begun to use the network along with the popular of Internet. With increase of age ,the vision of elderly were decrease, the ability to regulate of the eye was decreased[11], the vision became narrow[12], Cognitive function became weak[13], the ability of information processing was decreased[14]. The physiological changes of the elderly lead to the change of visual behavior when they browse the Web, and their fixation duration, fixation times, pupil diameter, saccade distance would also change . This method will be applied to classfy the Internet old users and youth users. Pupil diameter, saccade distance and gaze frequency and its combinations are used as observation sequence.

Twenty-six young (aged eighteen to twenty-eight,be undergraduate and graduate students)and twenty-six elderly (aged fifty-one to seventy-nine, be able to skillfully Internet)people carried out the task of browsing the single-page web without scrollbar. Everyone browsed ten web pages. The equipment produced in Swedish was non-intrusive Tobii T120 Eye Tracker . The sampling frequency is 120Hz.

We collected two hundred and sixty groups of young people and two hundred and sixty groups of elderly people sample set. One hundred and eighty groups of young people and one hundred and eighty groups of elderly people sample set were used as training set. The left eighty groups of young people and eighty groups of elderly people sample set were used as testing data set.

As shown in Figure 2, the recognition rate by HMM using single feature is not high, the highest recognition rate of pupil diameter as a classification feature is 81.25 percent. The recognition rate is increased to 93.65 percent by using the combined features.
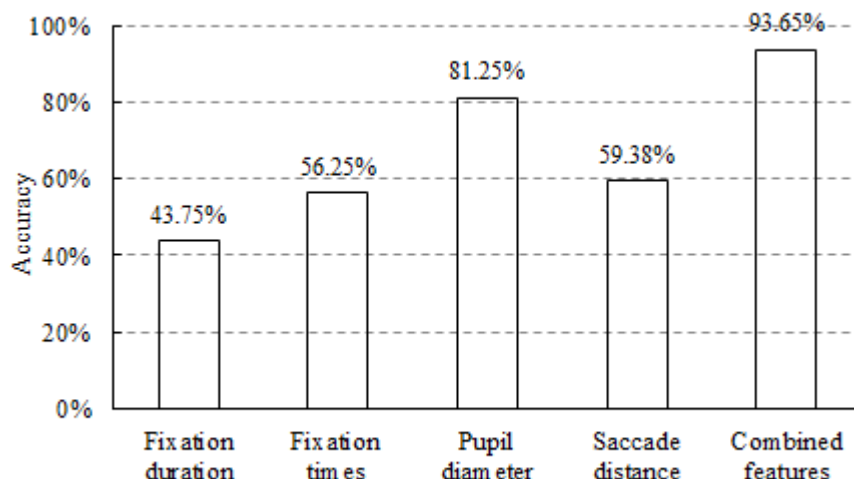
Figure2: Classification results of individual feature and combined features based on HMM

Since the initial parameters of Baum-Welch is random obtained, When the initial parameter randomly selected is better, the output probability value can be maximized, the resulting model parameters are relatively close to the global optimum, but the randomness of the parameters is too strong, so we can not determine the optimal parameters. We random five groups of parameters and calculate their P (O | λ) using forward-backward algorithm. The results are shown in Table 1.

	Maximum P（O|λ）	Minimum P（O|λ）	Average P（O|λ）
Random 1	1.1583	0.9730	1.0878
Random 2	1.3650	0.8904	1.0627
Random 3	1.5049	0.9912	1.2402
Random 4	1.0724	0.6499	0.8483
Random 5	1.1124	0.5452	0.7921

Table 1: The training under random initial parameters

As can be seen from Table 1, The average P （O| λ） of the first three groups are larger different with the latter two groups. This shows that there is a big randomness of parameters, which is not conducive to training. If the initial parameters we choose randomly is not the best ,the parameters of HMM is not optimal when the parameters is re-estimated using the equal (3.2.1)、(3.2.2) and (3.2.3)， which caused the model obtained inaccurate. To determine the best initial parameters of GA, we choose different populations, crossover and mutation rate. The results are shown in Table 2.

During the experiment, we set the maximum number of iterations of the algorithm 100. Several experiments show that the population tend to be stable, the best individual has been produced after normally 70 iterations. From the table 2,With increasing population, P (O | λ) of the best individuals in population also increased. This shows that when M become large , the

range of genetic algorithm searches increases, the opportunity to obtain the best individual increases. Especially when the M was increased to 200, p_c=0.8, p_m=0.1, P (O | λ) reaches to 1.8936 and the average of P (O | λ) reaches to 1.249946,which proved the a wide range of optimization. The main parameter of GA is set , M=200, p_c=0.8, p_m=0.1, T=100.

M	P_c	P_m	Maximum P（O|λ）	Minimum P（O|λ）	Average P（O|λ）
50	0.1	0.01	1.2729	0.9207	1.117829
50	0.6	0.05	1.6468	0.9557	1.249934
50	0.8	0.1	1.5901	0.9963	1.249931
100	0.1	0.01	1.4485	0.7192	1.249934
100	0.6	0.05	1.7814	0.9548	1.249951
100	0.8	0.1	1.7588	0.8302	1.249940
200	0.1	0.01	1.4380	0.7776	1.249923
200	0.6	0.05	1.6167	0.7847	1.249938
200	0.8	0.1	1.8936	0.9700	1.249946

Table2: The results of different parameters of GA

The recognition accuracy between HMM and GA-HMM is compared as shown in Table 3. From Table 3, the accuracy of the model has been improved by using GA to optimized the Initial parameters of HMM, which reflect the global optimization features of GA, and also prove that GA-HMM is valid for eye movement data.

	Young	Elderly	Average
HMM	92.65%	94.65%	93.65%
GA-HMM	98.75%	97.5%	98.125%

Table 3: Comparison of recognition rate between HMM and GA-HMM

5. Conclusion

In this paper, we propose a method of online user identification based on eye movement data, by which can identify the online old users and young users with higher accuracy. As the interface with the web page, using eye tracking technology to study the user, the eye movement data can be used to identify the user's real-time information, also can be used to further explore the user's interest and browsing habits. In this paper, the identification of online users is now in the primary stage, just a division of the user's age level, the next work need to narrow the age interval , making the user types achieve more accurate and detailed classification. In addition, the classification algorithm is based on maximum likelihood criterion, as the kinds of users increase, the recognition rate will be reduced. In the future, we can consider comparing the learning algorithm of other criterias (maximum mutual information, minimum classification error) with maximum likelihood criteria to select a better classification algorithm, which is the

next work.In addition , the time complexity should be considered because of the introduce of GA,which is also the next work.

References

[1] Silva, J.D.V. *Improvement of a Methodology for Website Keyobject Identification through the Application of Eye-Tracking Technologies.* IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent, Agent Technology. 59-63(2012)

[2] A. Al-Wabil, P. Zaphiris, Wilson, S. *Examining visual attention of dyslexics on web navigation structures with eye tracking* . Innovations in Information Technology. 717-721(2008)

[3] Q. Li, L.Y. Sun, J. Y. Duan. *Web page viewing behavior of users: an eye-tracking study* . Services Systems and Services Management. (1):244-249(2005)

[4] X.Y. Su, M. Li, S. F. Lu, *An eye tracking study on visual search and browse strategies of elderly people on web pages.* International Conference on Mechatronics Engineering and Computing Technology, ICMECT. 556-562：6154-6158(2014)

[5] Y. C. Han. *The order of eye movement observation of different shapes and colours* . Psychological Science. 20(1)：40-43(1997) (In Chinese)

[6] M. Elhelw, M. Nicolaou, A. Chung, et al. *A gaze-based study for investigating the perception of visual realism in simulated scenes*. ACM Transactions on Applied Perception (TAP). 5(1): 3(2008)

[7] D. D. Salvucci, J. R. Anderson. *Automated eye-movement protocol analysis*. Human-Computer Interaction. 16(1): 39-86(2001)

[8] F. Aziz, H. Arof, N. Mokhtar. *HMM based automated wheelchair navigation using EOG traces in EEG.* JOURNAL OF NEURAL ENGINEERING. 11(5):79-96 (2014)

[9] D. Bouchaffra, J. Tan. *Structural hidden Markov models using a relation of equivalence application to automotive designs*. Data Mining and Knowledge Discovery. 12():79- 96(2006)

[10] S. Kwong, C. W. Chan, K. F. Man, et al. *Optimization of HMM topology and its model parameters by genetic algorithms* .Pattern Recognition, 34:(2)509- 522(2001)

[11] S. L. Xu. *Changes of visual, auditory and psychological reactions and coping with the elderly in the aged*. Chinese Journal of mental health. 3: 136-138 (1998) (In Chinese)

[12] W. Wang , Z. Y. Tang, S. Y. Yin. *Analysis of 70 cases of elderly patients with static quantitative perimetry* . Practical Department of Ophthalmology. 11(2): 125-127(1993) (In Chinese)

[13]D. M. Li, T. Y. Chen. *Cognitive aging and elderly mental health*. Progress in Psychological Science. 14(4): 560-564(2006) (In Chinese)

[14] D. M. Li, C. Liu. *Study on the aging of cognitive speed*. Chinese Journal of Gerontology. 20(6): 383-385(2000) (In Chinese)