International Journal of Hybrid Information Technology
Vol. 9, No.9 (2016), pp. 189-198
http://dx.doi.org/10.14257/ijhit.2016.9.9.18

Eye Fatigue State Recognition of Gabor Wavelet Optimization HMM Algorithm

Yu Xiang[1]

[1]*Department of Information Technology, Tianjin Chengjian University,*
Tianjin City, China
147940866@qq.com

Abstract

It is easy for the internet learners to generate learning fatigue because of the long-term lack of emotional interaction in the learning process, and learning fatigue often manifests through the eye condition, in order to do effective monitoring for remote intelligent tutoring system, the learning fatigue eye state recognition algorithm is put forward based on Gabor wavelet and HMM. The algorithm has certain distinguishing characteristics aiming at the degree of eye openness of network learner under 3 learning states: normal learning, fatigue and confusion, first, it does gray difference disposal for eye image by Laplace operator in $YCbCr$ *color space, then, it selects two-dimension Gabor kernel function to build 48 optimal filters, obtain 48 characteristic values, these 48 characteristic values generate 48 eigenvectors, at last, it use a set of observation sequence* O *formed by eigenvector of HMM for eye state image to do eye state recognition. Experimental results show that the recognition rate of this algorithm for network learning reaches 95.68%, and this algorithm has a good robustness.*

Keywords*: Learning fatigue; Wavelet transform; Hidden Markov Model; Facial recognition*

1. Introduction

Due to the lesser study on fatigue detection technology in the field of education, detection methods on network learning fatigue detection, and mature driver fatigue technology has certain reference significance on the early warning research on learning fatigue of network learners. The existing face recognition research shows that the state of eye and mouth of network learners in the learning fatigue is an important factor to show whether the learners have fatigue or not[1][2]. This paper presents a learning fatigue eye state recognition algorithm based on Gabor wavelet and HMM and this algorithm is nested in remote intelligent tutoring systems. The method does gray difference disposal to eye image by Laplace operator in $YCbCr$ color space; then it structures and selects two-dimension Gabor kernel function, structure 48 optimal filters, gets 48 eigenvalues, these 48 eigenvalues generate 48 eigenvectors, using HMM to do state recognition to observation sequence O, and this sequence is composed of eigenvectors of eye image.

2. Image Preprocessing

In our method, using the following equation to obtain color space of $YCbCr$ from RGB color space, among, R, G, and B respectively are red, green, and blue components of color image. Standard range of Y, Cr and Cb is [0, 255].

ISSN: 1738-9968 IJHIT

$$\begin{cases} Y = 0.299R + 0.587G + 0.114B \\ Cr = R - Y \\ Cb = B - Y \end{cases} \tag{1}$$

Compared with the facial forehead area, eye area has low intensity Y, low red degree (Cr) and high blue degree (Cb). By the fact, we can process the input image to grayscale images in advance. PDF of gray difference between neighboring pixels can be approximatively stimulated by generalized Laplace operator. See gray method in Figure 1.

$$f_L(\Delta i) = \frac{1}{Z_L} \exp(-\left|\frac{\Delta I}{\lambda}\right|^{\beta}) \tag{2}$$

Among ΔI is the gray difference, λ depends on the distance between the positions of two sample images, β is a parameter which is approximately equal to 0.5, ZL is normalization constant. In the following, we assume that $\beta = 0.5$, this means $ZL = 4\lambda$.

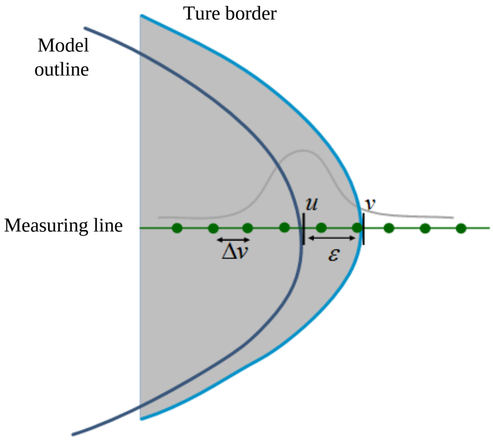


Figure 1. Definition of Edge Contour

Figure 2 shows the distribution of gray difference in eye image. The gray difference calculated on different skin scale is shown as the following figure. We can see from the figure, the distribution of gray difference of eye image can be approximatively stimulated by $YCbCr$. In addition, you can still see the width k of the distribution increases with scale.

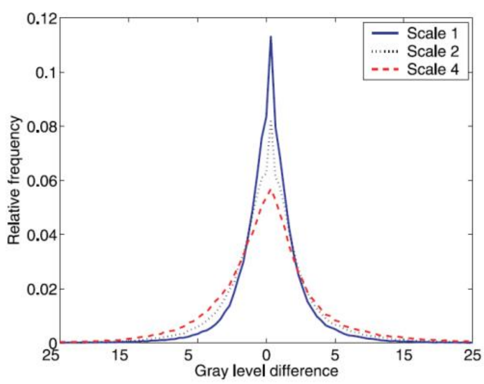


Figure 2. Gray Difference Distribution of Different Scales in Eye Image

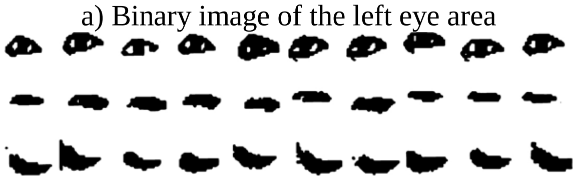
Figure 2 The distribution of gray difference on different scales in eye image. The scale parameters refer to the distance between gray measurements. "Measure 1" refers to the distance of a pixel, "measure 4" refers to the distance between the four pixels

Figure 2 shows the distribution of gray distribution in eye image. The gray difference calculated at different $\Delta\nu$ scales is shown as the figure. We can see from the figure, the distribution of gray distribution in eye image can be approximatively stimulated by generalized Laplace operator, see the definition in equation (2). In addition, we can still see the width of the distribution increases with scale. The reason for this phenomenon is that, with the increase of $\Delta\nu$, the correlation of the pixel values becomes weak.

Then, use threshold T to binarize gray level image to “binary image” by simple global threshold. After linearization, the next task is to get the 4 connected components, to label, and then to find the center of each chunk. We stick labels on two eyes, mouth, ears, *etc*. See detailed results of the connected components in [4].

The image after processing is shown as Figure 3.

a) Binary image of the left eye area



b) Binary image OF the right eye area

Figure 3. Image Processing of Eye Area

3. Feature Extraction of Gabor Filter

2D Gabor wavelet transform is an important tool for time-frequency domain to do signal analysis and processing, and its transform coefficients have good visual characteristics and biological backgrounds. Gabor filter with different parameters can capture the local feature information in an image, and correspond with different spatial frequency, spatial position and orientation. Since characteristics of Gabor filter, it is not sensitive for brightness and facial expression changes, therefore, Gabor filter is widely used for image coding, handwriting digital recognition, face recognition and edge detection.

We do bidimensional Gabor wavelet transform to gray level image after image preprocessing, thereby we can obtain a fatigue characteristics of the facial area of a driver. Two Gabor wavelet kernel function is:

$$\psi_j(\vec{k},\vec{x}) = \frac{\vec{k}_j^2}{\sigma^2}\exp\left(-\frac{\vec{k}_j^2 \bullet \vec{x}^2}{2\sigma^2}\right)\left[\exp(\vec{k}_j^2 \bullet \vec{x}) - \exp\left(-\frac{\sigma^2}{2}\right)\right] \tag{3}$$

Wherein, $\vec{k}_j^2$ can ensure that filters with the different frequency bandwidth have nearly the same energy. $\exp(-\frac{\sigma^2}{2})$ is the DC shunt excitation to compensate image, therefore, the filter is not sensitive to finishing lamination. The advantage is that it maintains the spatial relationship, while it can describe spatial frequency structure.

$$\vec{k}_j = \begin{pmatrix} k_v \cos\varphi_u \\ k_v \sin\varphi_u \end{pmatrix}, k_v = 2^{-\frac{v+2}{2}}\pi, \varphi_u = \mu\frac{\pi}{8} \tag{4}$$

$\vec{k}_j$ constitutes a different wavelet (values of its side are different), this paper has used changes of four dimensions and six directions.

$k_v = 2^{-\frac{v+2}{2}}\pi \quad (v = 1,2,3,4)$, six directions $\varphi: 0-6/8\pi \quad (\mu = 1,2,...,6)$, $\pi/8$. The interval of σ is the length of the filter, we take $\sigma = \pi$, and assume that the frequency is 1. Input image $I(x, y)$ and convolution wavelet, then

$$g(\vec{k}_j, \vec{x}) = \iint I(\vec{x}, \vec{y})\varphi_j(\vec{k}, \vec{x})dxdy \tag{5}$$

Wherein, $g(\vec{k}_j, \vec{x})$ is the amplitude. Therefore, the value of the Gabor filter is 48, which constitute a group of optimum filtering representative of the target feature. These filters constitute the wavelet subspace, and project image to the wavelet subspace to obtain the wavelet coefficients and extract the mean and variance to represent the statistical characteristics of the driver's facial image.

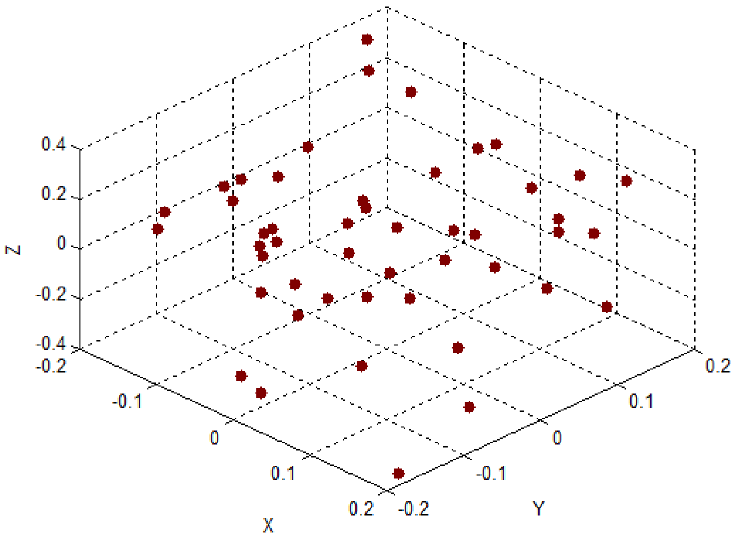


Figure 4. Distribution Diagram of 48 Feature Points

4. Eye State Recognition

Hidden Markov Models (HMM) can find state transfer characteristic hidden in eye phase diagram. This paper selects one-dimensional HMM, the eye state images after the extraction of eye state feature do subsequent recognition, which can improve eye state recognition effect.

4.1. Basic Definition

HMM is a relatively sophisticated stochastic process stochastic matching model, which is described by parameters. Hidden Markov Models (HMM) includes implied layer and observation layer, observation layer is the hidden markoff chain, observation layer is the actual observed quantity, and that is the living example that we want to identify.

A HMM can be denoted as formula (6):

$$\lambda = (N, M, \pi, A, B) \tag{6}$$

A represents the state transition probability matrix, B represents the observation value probability matrix, M represents the length of the observation sequence, N represents the number of implicit states, π represents the initial state probabilities.

4.2. HMM Training

To optimize HMM parameters, some use a single image for training, some use multiple images for training in accordance with specific conditions. Training steps:

(1) To do Gabor filter value calculation to classified eye state image, to find the eigenvalues to generate observation sequence Oi, and to regard Oi as an eigenvector of observation images.

(2) General model $\lambda = (N, M, \pi, A, B)$, determine the number, condition number and state transition of components of Gaussian mixture probability allowed by model.

(3) Calculate initial parameters of model, pay attention to corresponding with N_t (t moment) state, and then evenly split the training data. State transition matrix $A = (a_{ij})$, we take $a_{ij} = 0$, when $j < i$ or $j > j-1$. The initial probability distribution $\pi = (\pi_1, \pi_2, \cdots, \pi_N)$, assuming the first state $\pi_1 = 1$, if $\pi_i = 1(i \neq 1)$, $B = \{b_j(O_i)\}$ use a Gaussian probability density function, $B = \{b_j(O_i)\}$can be calculated based on the formula (8):

$$b_j(o_i) = \frac{1}{\sqrt{(2\pi)^n \left|\sum_j\right|}} \exp\left\{-\frac{1}{2}(o_i - \mu_j)^T \sum_j^{-1}(o_i - \mu_j)\right\} \tag{7}$$

Wherein, $\sum_j$ and μ_j are covariance matrix and mean value of Gaussian probability density function.

$$\mu_j = \frac{1}{T_j}\sum_{i=1}^{T_j} o_i \tag{8}$$

$$\sum_j = \frac{1}{T}\sum_{i=1}^{T_j}(o_i - \mu_j)(o_i - \mu_j)^T \tag{9}$$

(4) The determination of HMM's optimal state sequence. The parameters of Gaussian mixture model use Viterbi segmentation [14], combined with K-means clustering segmentation method.

(5) Using Baum-Welch algorithm [14] to estimate the parameters again. Confirming $\lambda = (N, M, \pi, A, B)$, and optimizing parameters on the basis of model, so that the $P(O \mid \lambda)$ gets the maximum value, $P(O \mid \lambda)$ is a category in eye fatigue state.

4.3. HMM Eye Fatigue State Recognition

First, we do gray difference disposal for $YCbCr$ color space by using Laplace operator, then we use Gabor filter to do disposal for eye image and train in group observation sequence O with the help of forward - backward algorithm, and the observation sequence is constituted by the eigenvectors, the model used by training is $\lambda_i (1 \leq i \leq 3)$, we strike probability $P(O \mid \lambda_i)$ in turn. The corresponding model of $\max_i P(O \mid \lambda_i)$ is the categories that the eye status which is to be recognized belongs to.

$$\varphi_n = \arg\max_i P(O \mid \lambda_i) \tag{10}$$

5. Experimental Results and Analysis

This study attempts to analyze the basic process and the principle of e-learning from a network learning view. The great majority of the network learners are adults, we sample face recognition database CAS-Peal-Rl, use research methods and means of cognitive psychology, and carry out the relevant network fatigue learning experiments. By analyzing learning state of network learners, we improve the quality of online course design, improve the level of service of distance education, provide supports to network learning support, so that e-learning platform can get more active use.

CAS-PEAL-R1 is a large face database in China, it includes 1040 people, 99 450 face images, and image size is 360 × 480 pixels, including facial expressions, gestures, lighting and accessories 4 kinds of changes. This algorithm of this paper is in CAS-Peal-R1, the number of samples is 20,000, 10,000 and 5000. The results are shown in Table 1.

Table 1. Comparison of Velocity of Gabor + HMM Algorithms and other + HMM's Eye State Algorithm (Unit: Seconds)

Experiment number of image	Gabor+HMM	Other+HMM
20000	2.46	4.5
10000	1.58	2.16
5000	1.34	1.1

As we can see from Table 1, when the number of test images is large (20000 and 10000), Gabor + HMM algorithm is more efficient than other + HMM algorithms (including PCA + HMM, ICA + HMM, DCT + HMM), however, when the number is small (5000), Gabor + HMM algorithm is slower than other + HMM algorithms, because the use of Gabor + HMM algorithm spends more time in recognizing the facial image classification of facial image advance line. Therefore, we can conclude that the algorithm is suitable for large-scale face database, and the face database is generally greater than 10000.

The results are shown in Figure 5 and Figure 6.

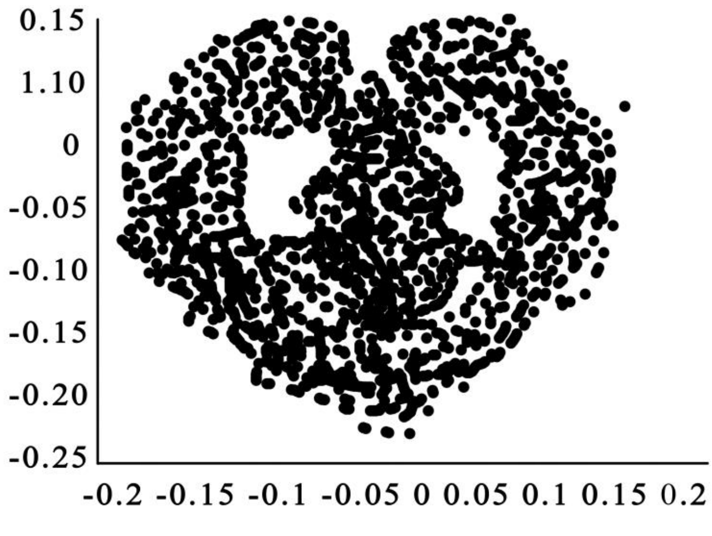


Figure 5. Recognition Effect of Gabor + HMM Algorithm

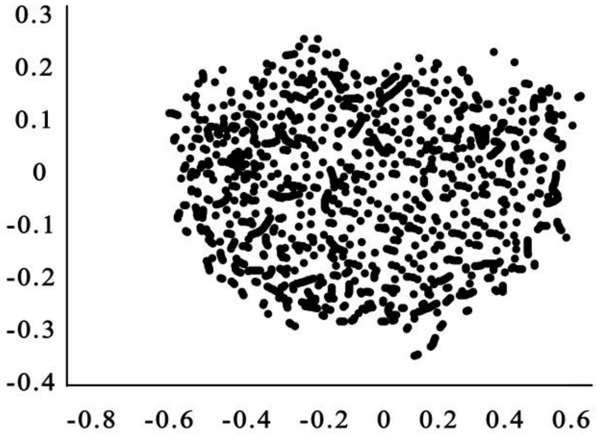


Figure 6. Recognition Performance of Other + HMM Algorithm

Figure 7 compares the recognition rate of GABOR+HMM 、 PCA+HMM 、 ICA+HMM and DCT+HMM of this paper, and draws the following conclusions: the dimension of GABOR+HMM is lower, the recognition rate is lower, when the dimension is 1, all of the recognition rates do not exceed 75%; the recognition rate significantly increases with the increase of dimension (when dimension is equal to 48, it reaches a maximum 97.68%, and this is consistent with the set of the front Gabor filter values), for PCA + HMM, ICA + HMM, DCT + HMM three methods, when the corresponding feature dimensions respectively reach 20,7,22, the corresponding recognition rates basically stabilize at 93.29%, 90.89%, 92.8%.

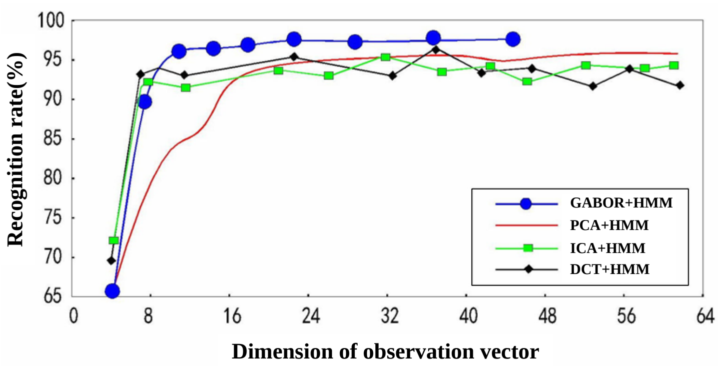
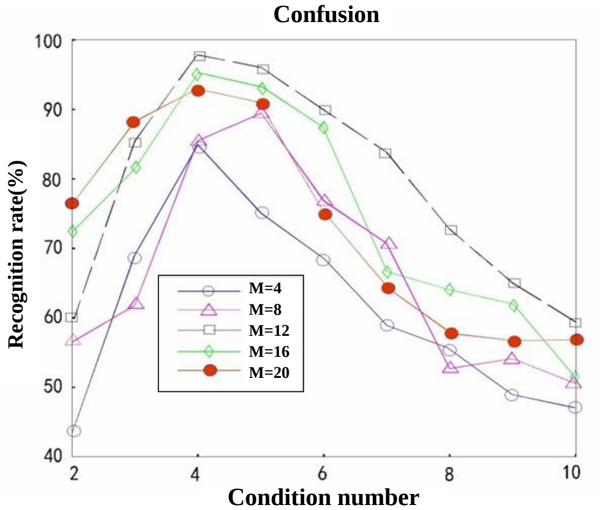
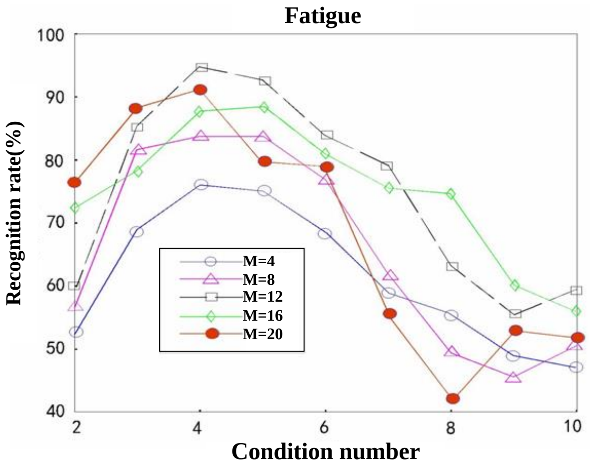


Figure 7. Recognition Rate of Gabor\, PCA, ICA and DCT under Different Observation Vector Dimension

In this paper, the selection of N and M is determined in accordance with experimental results, the value domain of N is 2~8, the value domain of M is 5~21. As it can be seen from Figure 8, when M = 12, N = 4, the recognition rate of confusion and fatigue is at a maximum. Under the promise of keeping the same number of mixed component of Gaussian probability, when N <4, the recognition rate increases as the increase of N; After N≥4, the recognition rate gradually decreases. The formula computing of observation probability matrix $B = \{b_j(O_i)\}$ is the formula (7), μ_j in formula (7) can be calculated by formula (8), $\sum_j$ represents formula (9) is used to calculate.



a) Confusion



b) Fatigue

Figure 8. The Influence of the Number of Mixing Element and Condition Number of Gaussian Probability for Recognition Rate

6. Conclusion

In these states, the recognition rate of normal learning is relatively low, and the reason is that concentration and fatigue are more obvious than expression features of normal learning. In remote intelligent educational systems, we improve the recognition rate in two ways: first, we require the camera to shoot color images, and update the Gaussian model parameters in time according to changes of Y component; second, we refine classification of learning fatigue, demarcation of eye open, eye close is clearer, which can enhance the accuracy and stability of the identification and tracking. Recognition results of this paper are transmitted to a remote intelligent tutoring system as feedback information, the teachers can adjust teaching progress in time, re-arrange teaching content, correct teaching methods and provide basis, which can provide individualized learning environment for network learners to compensate emotion absence of network learners.

Acknowledgement

The research is supported by Tianjin Xinghai plan (KJXH2014-13)

References

[1] Z. Lv, T. Yin and Y. Han, "WebVR——web virtual reality engine based on P2P network", Journal of Networks, vol. 6, no. 7, **(2011)**, pp. 990-998.
[2] J. Yang, J. Zhou and Z. Lv, "A Real-Time Monitoring System of Industry Carbon Monoxide Based on Wireless Sensor Networks", Sensors, vo. 15, no. 11, **(2015)**, pp. 29535-29546.
[3] D. Jiang, Z. Xu and Z .Lv, "A multicast delivery approach with minimum energy consumption for wireless multi-hop networks", Telecommunication Systems, **(2015)**, pp. 1-12.
[4] Y. Lin, J. Yang and Z. Lv, "A self-assessment stereo capture model applicable to the internet of things", Sensors, vol. 15, no. 8, **(2015)**, pp. 20925-20944.
[5] Ying Liang*. Satisfaction With Economic and Social Rights and Quality of Life in a Post-Disaster Zone in China: Evidence From Earthquake-Prone Sichuan, Disaster Medicine and Public Health Preparedness, 9(2):111-118.
[6] Y. Liang*, "Correlations Between Health-Related Quality of Life and Interpersonal Trust: Comparisons Between Two Generations of Chinese Rural-to-Urban Migrants", Social Indicators Research, vol. 123, no. 3, 677-700.
[7] Y. Liang* and P. Lu, "Medical insurance policy organized by Chinese government and the health inequity of the elderly: longitudinal comparison based on effect of New Cooperative Medical Scheme on health of rural elderly in 22 provinces and cities", International Journal for Equity in Health, vol. 13, **(2014)**, pp. 37, 1-15.
[8] D. Jiang, X. Ying and Y. Han, "Collaborative multi-hop routing in cognitive wireless networks", Wireless Personal Communications, vol. 86, no. 2, **(2016)**, pp. 901-923.
[9] Z. Lv, A. Halawani and S. Fen, " Multimodal hand and foot gesture interaction for handheld devices", ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), vol. 11, no. 1s, **(2014)**.
[10] J. Hu and Z. Gao, "Distinction immune genes of hepatitis-induced heptatocellular carcinoma", Bioinformatics, vol. 28, no. 24, **(2012)**, pp. 3191-3194.
[11] J. Hu, Z. Gao and W. Pan, "Multiangle Social Network Recommendation Algorithms and Similarity Network Evaluation", Journal of Applied Mathematics, 2013, **(2013)**.
[12]J. Hu and Z. Gao, "Modules identification in gene positive networks of hepatocellular carcinoma using Pearson agglomerative method and Pearson cohesion coupling modularity", Journal of Applied Mathematics, **(2012)**.
[13] Y.Geng, J. Chen, R. Fu, G. Bao and K. Pahlavan, "Enlighten Wearable Physiological Monitoring systems: On-Body RF Characteristics Based Human Motion Classification Using a Support Vector Machine", vol. PP, no. 99, pp. 1-16.
[14] X Song,., and Y.Geng, Distributed Community Detection Optimization Algorithm for Complex Networks. Journal of Networks, 9(10), 2758-2765.
[15] K. Pahlavan, P. Krishnamurthy and Y. Geng, "Localization Challenges for the Emergence of the Smart World", Access, IEEE, vol. 3, no. 1, pp. 1-11.
[16] J. He, Y. Geng, Y. Wan, S. Li and K. Pahlavan, "A cyber physical test-bed for virtualization of RF access environment for body sensor network", Sensors Journal, IEEE, vol. 13, no. 10, **(2013)**, pp. 3826-3836.

Authors

Yu Xiang, he received his M.S. degree in software engineering from Estonian Business School of Information and Technology in Tallinn, Estonia. He is currently a lecturer in the College of Computer Technology at Tianjin Chengjian University. His research interest is mainly in the area of Computer Software, Image Processing and Cloud Computing. He has published several research papers in scholarly journals in the above research areas and has participated in several books.