A Hybrid Intelligence of Harmony Search Algorithm and Neural Network for Crowd Density Estimation

Dr. Saeed Mohammed Saeed Baneamoon
Department of Computer Engineering,
Faculty of Engineering & Petroleum,
Hadhramout University, Yemen
saeed_banaimoon@yahoo.com

Dr. Ali Salem Ali Bin Sama
Department of Computer Science and Information,
College of Sharia and Islamic Studies at Al-Ahsa, Imam
Muhammad bin Saud Islamic University, Saudi Arabia
alibinsama@gmail.com

***Abstract*— Crowd density estimation is an important task in crowd analysis of crowd places specially at popular events, this paper proposes a hybrid intelligence of harmony search algorithm and neural network to increase accuracy analysis of crowd density estimation, where Gabor wavelet is used at texture feature extraction stage for effectively extract crowd features of different density levels, and a hybrid intelligence of harmony search algorithm and neural network is used at pattern classification stage for recognition of patterns and regularities in data. Experimental results on a number of public benchmark images are used such as UCSD Dataset show that the proposed hybrid intelligence method provides better performance analysis of crowd places.**

Keywords- Crowd Density Estimation, Gabor filters, Harmony Search, Neural Network, Texture Image.

I. Introduction

In crowd phenomenon, crowd analysis of crowd density estimation is an important for monitoring, managing and safety control of public spaces, especially during the popular event that gather many people such as religious, festivals and sport events.

Therefore, a model is required to help those in charge of crowd management to make immediate decision for any sudden emergency in order to prevent disasters potential hazards.

Hence, this paper presents a hybrid harmony search algorithm and neural network for crowd density estimation, where this method is applied at pattern classification stage for recognition of patterns effectively.

II. Literature Review

Several techniques have been used to develop crowd density estimation problem based on crowd monitoring and analysis.

Choudri et al. in [1], have been used Human region detector to filter parts of human like heads based on counting only human-classified pixels rather than foreground pixels. This pixel-based crowd counting system can reduce the loss of people by using a more robust people counting based classifier when they get absorbed into the background after beings low or stationary. Conte et al. in [2], have been proposed density based clustering by applying shape technique which is more reliable to extract the shape of a set of points with the bounding box. They also in [3], have been developed a more robust real time method for counting moving pedestrians in a scene. Zhang et al. in [4], have been presented a method deal with the perspective distortion, this method has been subdivided the entire scene into smaller horizontal zones, each zone has a special size depending on its distance weights from the camera and the people counting is summed separately for each zone. Chen et al. in [5], have been adopted a method for counting crowds directly and local feature mining, in which features have been extracted from equally sized cells in a rectangular grid and the multiple output ridge regression have been used for capturing both global and local trends in the image. Ge et al. in [6], have been developed a mixture model of Bernoulli shapes for representing the foreground humans from a training dataset. Gall et al. in [7], have been proposed a Hough forest framework as a robust to partial occlusions and a typical part appearances. On the other hand this framework is a more general formulation which can be applied for tracking and action recognition beside object detection. Liang et al. in [8], have been developed a method based on feature points, a crowd flow tracking and counting for improving SURF point detecting process, detecting feature points that really belong to the moving crowd and also for reducing the time complexity. Ryan et al. in [9], have been presented an evaluation across multiple datasets for comparing the holistic, local and histogram based methods and to compare various image features and regression models. Neeta et al. in [10], have been introduced the survey compares different crowd datasets for crowd density estimation methods. Xiaohang et al. in [11], have been adopted an ensemble learning (EL) method using support vector regression (SVR) for se problems,for crowd density estimation in which SVR model is used to roughly predict the three features extracted from image blocks, including D-SIFT, ULBP, and GIST.

III. HARMONY SEARCH

Harmony Search (HS) was first proposed by Geem et al. in [12]. Harmony search is a music-based algorithm. HS is a meta-heuristic optimization algorithm inspired by the working principle of the musicians' improvisation of the harmony.

In improvising, a musician has three possible choices are as follows:

1. play any note of music from the harmony memory (a series of pitches in harmony).
2. play a note similar to another one stored in memory (adjusting the pitch slightly);
3. play new or random notes.

These three options are formalized into quantitative optimization process, and the three corresponding components are as follows [12, 13, 14]:

1. usage of harmony memory,
2. pitch adjusting, and
3. randomization.

The usage of harmony memory is similar to the choice of the best-fit individuals in genetic algorithms (GA), this determined by a harmony memory accepting or considering rate parameter $\boldsymbol{r_{accept}} \in [0,1]$. If this rate is too low, only few best harmonies are selected and it may converge too slowly. If this rate is extremely high (near 1), generation of new solutions in the HS method makes full use of all the HM members.

The pitch adjustment determined by a pitch bandwidth $\boldsymbol{b_{range}}$ and a pitch adjusting rate $\boldsymbol{r_{pa}}$ that determines the probability of a candidate from the HM to be mutated. In practice, the pitch can be adjusted linearly as follows:

$$X_{new} = x_{old} + b_{range} * \varepsilon \qquad (1)$$

where

x_{old} : is the existing pitch or solution from the harmony memory.

X_{new}: is the new pitch after the pitch adjusting action.

ε : is a random number generator in the range of [-1,1].

The third component is the randomization, which is to increase the diversity of the solutions to explore various diverse solutions.

The basic steps of harmony search are as follows:

1) parameter initialization
2) initialize Harmony memory (HM)
3) improvise new harmony
4) update the HM
5) repeat step 2-3 until a termination criteria reached.

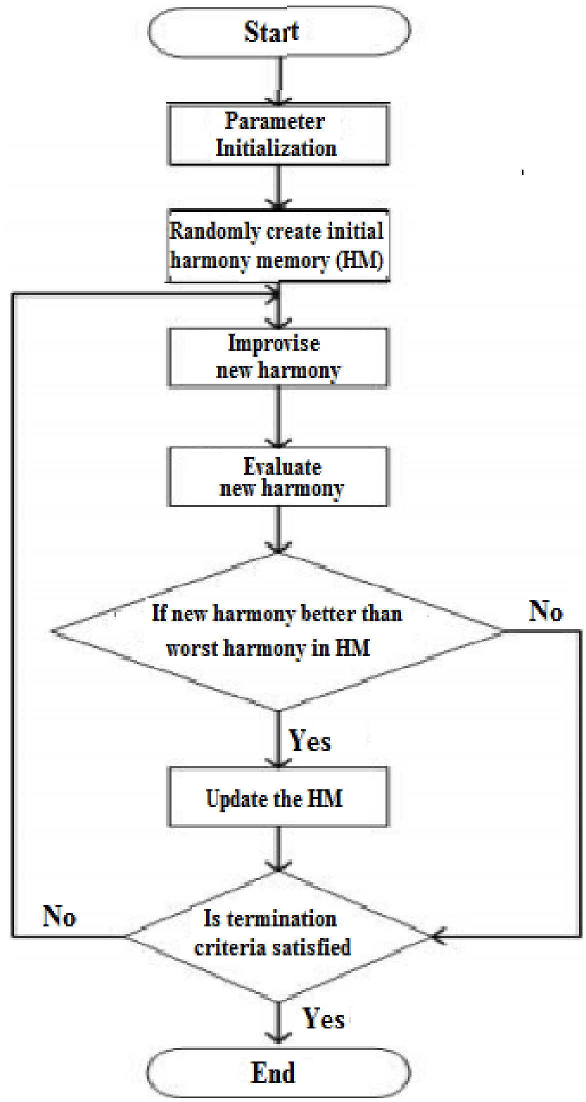

Figure 1: Harmony search flowchart

The Pseudo code of harmony search method are as follows:

begin
Objective function f(x), $x=(x_1,x_2, \ldots,x_d)^T$
Generate initial harmonics (real number arrays)
Define pitch adjusting rate (r_{pa}), pitch limits and *bandwidth*
Define harmony memory accepting rate (r_{accept})
while *(t < Max number of iterations)*
Generate new harmonics by accepting best harmonics
Adjust pitch to get new harmonics (solutions)
If *(rand >* r_{accept}*), choose an existing harmonic randomly*
else if *(rand>*r_{pa}*), adjust the pitch randomly within limits*
else *generate new harmonics via randomization*
end if
Accept the new harmonics (solutions) if better
end while
Find the current best solutions
end

The probability of randomization is derive as follows:

$$P_{random} = 1 - r_{accept} \quad (2)$$

and the actual probability of adjusting pitches is

$$P_{pitch} = r_{accept} * r_{pa} \quad (3)$$

IV. Neural Networks

Artificial Neural Network (ANN) is a machine learning models that have been studied for many years with the hope of mimics the capabilities of thinking, remembering and problem solving of human brain. These models attempt to achieve "Human-like performance" via dense interconnections of computing cells (artificial neurons) through weighted [15]. ANN can classify depending of different properties are as follows:

1) Type of connection
   - *Fully interconnected* models for which the output of each neuron may be connected to the inputs of all neurons.
   - *Partially interconnected* models for which the output of a given neuron is allowed to connect only to certain of its neighbours.
2) Direction of feed signal
   - *Feed-Forward (FFNN)* in which signals propagate in only one direction from an input stage through intermediate neurons to an output stage.
   - *Feedback NN (FBNN)* in which signals propagate from output of any neuron to input of any neuron.
3) Number of Layers
   - *Single Layer (SL)* in which include only one layer.
   - *Multi Layer (ML)* in which must include at least one hidden layer, is one of the most widely used models among ANNs.

ANN algorithms are able to adapt continuously based on current results to improve performance. The performance of learning procedure depends on many factors such as choice of error function, net architecture and values of the weights. The three major learning paradigms each correspond to a particular learning task. These are supervised learning, unsupervised learning and reinforcement learning. [15, 16].

V. A Hybrid Intelligence of Harmony Search and Neural Network

In this paper a Generalized Delta Rule (back propagation) is adopted where the determination of the error is a recursive process which start with the O/P units and the error is back propagated to the I/P units. Therefore, the rule is called Error Back propagation (EBP) or simply back propagation (BP).

A hybrid intelligence of harmony search and neural network is proposed, where harmony search is used to optimize weighted connections, number of neurons in hidden layer, and to overcomes several disadvantages of NN such as long training time, local minimum trap and large number of parameters.

VI. Gabor wavelet features

A Gabor wavelets are usually called Gabor filters in the scope of applications. A Gabor filter is a linear filter used for texture analysis whether there are any specific frequency content in the image in specific directions in a localized region around the point or region of analysis [17].

In this paper, a Gabor filter is applied at texture feature extraction stage to generate a bank of filters at different frequencies and spatial orientations to obtain new Gabor-filtered image using the following formula:

$$G(x,y) = \left(\frac{1}{2\pi\sigma_x\sigma_y}\right) \exp\left(-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{x^2}{\sigma_y^2}\right)\right) \exp\left(2\pi jW(x\cos\theta + y\sin\theta)\right) \quad (4)$$

Where

W- is the radial frequency of the Gabor filter and its value lie in $W \in [0,0.5]$.

θ - orientation variable which controls the angle of the filter and its value lie in $\theta \in [0, \pi]$.

σ- the scale variable controls the shape of the Gaussian function and its value within $\sigma \in [0, 2\pi]$.

In this stage, a total of 40 Gabor filters are generated, at eight orientations $(0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}, \pi, \frac{5\pi}{4}, \frac{3\pi}{2}, \frac{7\pi}{4})$, and five scales $(0, \frac{\pi}{4}, \frac{\pi}{2}, \pi, \frac{3\pi}{4})$.

VII. Proposed Method

This paper presents a hybrid intelligence algorithm for crowd density estimation based on harmony search method. The general structure of the proposed method is shown in Figure 2. As indicated that the proposed structure consists of three stages namely input stage, texture feature extraction stage, and classification stage. In the first stage, the input image is divided into a cell for feature extraction. The second stage is responsible about encoding the extracted cell using a texture descriptors Gabor wavelet [18]. These features are computed from the input cell and combined together as a single vector that will be fed to the last stage for recognition purposes. The last stage is concern about the classification of the input pattern. A hybrid intelligence of harmony search algorithm and neural network are used in this study to perform the classification task.

The output decision of classification stage will be one of four classes including empty region, low density region, medium density region, or high density region as shown in Figure 3.

VIII. Experimental Results

To evaluate the effectiveness of the proposed hybrid method on crowd density estimation, a public databases is employed which is the UCSD pedestrian (UCSD) [5]. The total numbers of 2000 frames, the 10 number of frames per second, the 11-46 range of crowd number in a frame, as well as the total number of 49885 labeled pedestrians for UCSD database.

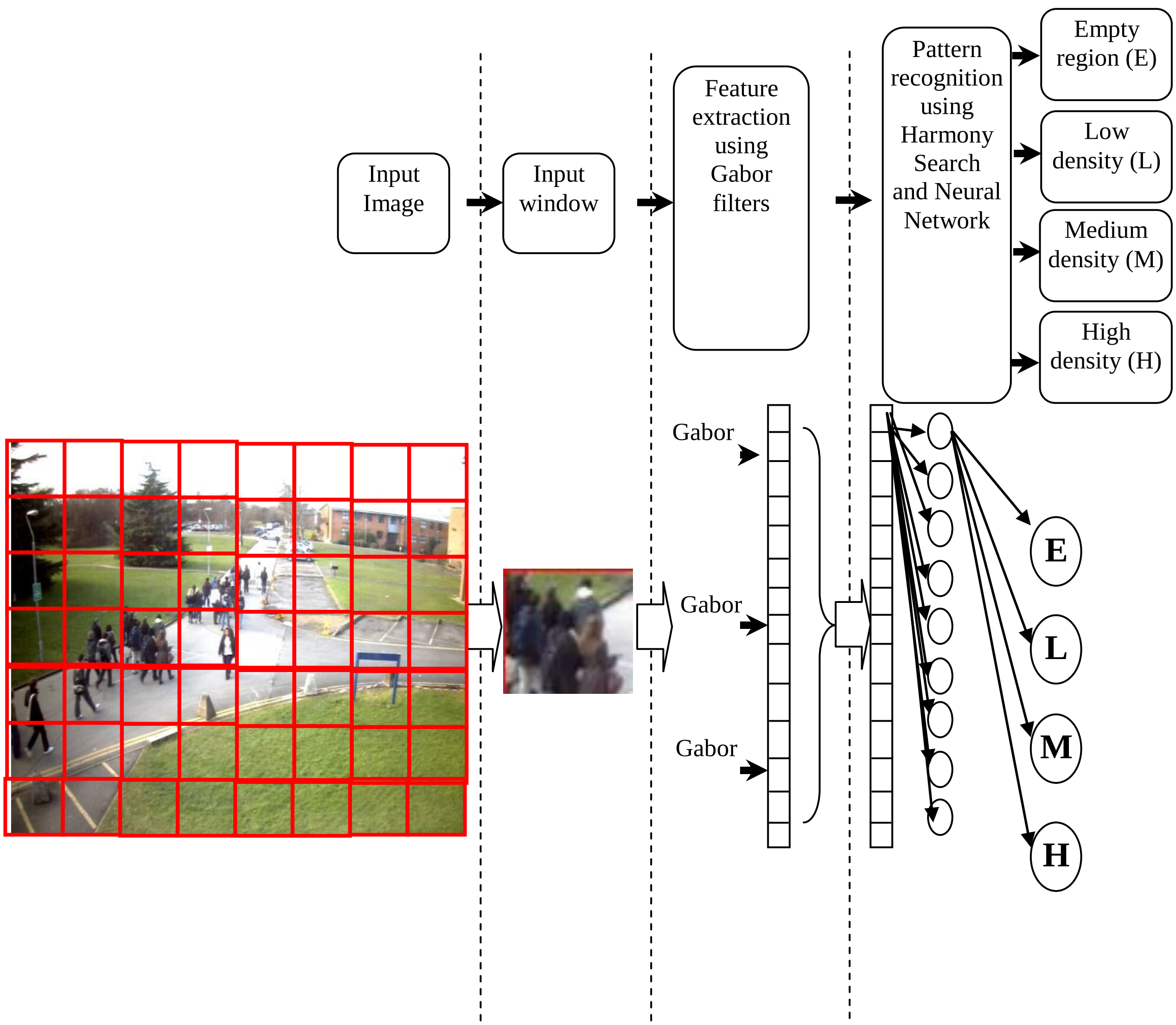

Figure 2: The structure of the proposed method

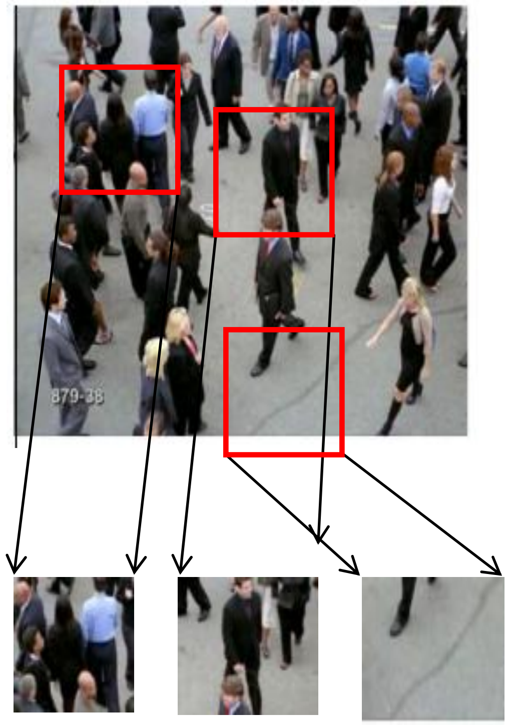

Figure 3:Example of crowd density classes (a) high density region, (b) low density region, and (c) empty region

A. The Estimation of Crowd Density

The density estimation experiments mainly include training and testing. The experimental images are taken from UCSD database. Then the sequence square window are obtained by extracting the square window from the crowd images. The square window size is 25×25. According to different crowd density, the experimental images are classified into low density, medium density and high density. Some crowd images of different density are shown in Figure 4 Totally 360 images are selected, including 120 images of low, medium and high density, respectively. We select 50 images for different density, totally 150 images. Having extracted the texture feature values of the foreground object, we train a hybrid intelligence of harmony search algorithm and neural network and obtain the training parameters, then begin to test. Numerical experiments have been conducted on MATLAB7.0 (the computer's CPU is Intel Core i7, 3.0 GHz, 8GB RAM).

In scenes of low, medium, and high density crowd, we extract means and standards deviations features generated by Gabor wavelet filter in the horizontal and vertical directions and constitute them into 11-dimensional, which is input into a hybrid intelligence of harmony search algorithm and neural network classifiers. We obtain artificial neural network model by training samples.

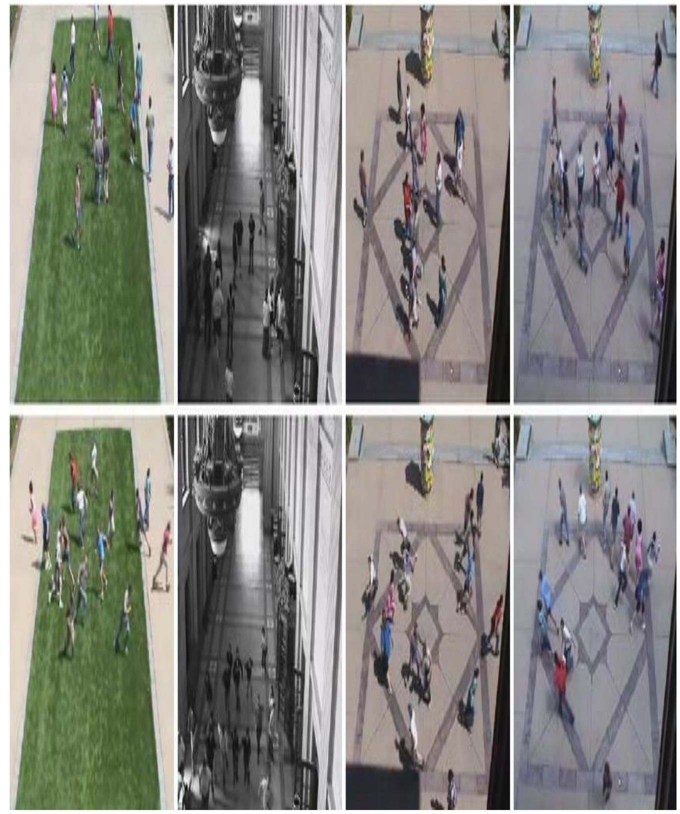

Figure 4: Sample images from UCSD

IX. CONCLUSION

In summary, we have presented a hybrid intelligence of harmony search algorithm and neural network for crowd density estimation in the sequences images from one databases. In the initial stage, the input images are divided into small square windows and then the texture feature by using Gabor wavelet is extracted. Furthermore, the feature vector computed as the mean and standard deviation of the magnitude value for each filter have been computed. After that, the feature vectors are fed into the training stage for solving the crowd density estimation. Besides this, the a hybrid intelligence of harmony search algorithm and neural network have been used for the classification in UCSD databases. The high accuracy with high speed for processing and low computational complexity enables our hybrid method possible for estimating crowd density. Experimental results demonstrate the effectiveness of our method. Nevertheless, there is a lot of scope for improvement in our hybrid method. In our future work we would like to reduce the false estimations, incorporate additional methods to reason out occlusions in a crowded scene and make the hybrid method more stable.

REFERENCES

[1] Choudri, S., Ferryman, J. M., Badii, A., Robust background model for pixel based people counting using a single uncalibrated camera. In: Twelfth IEEE International Workshop on Performance Evaluation of Tracking and Surveillance (PETS-Winter). IEEE, pp. 1–8, 2009.

[2] Conte, D., Foggia, P., Percannella, G., Tufano, F., Vento, M., A method for counting people in crowded scenes. In: Seventh IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS). IEEE, pp. 225–232, 2010.

[3] Conte, D., Foggia, P., Percannella, G., Vento, M., Counting moving persons in crowded scenes. Mach.Vis. Appl. 24(5), 1029–1042. 2013.

[4] Zhang, Z., Li, M., Crowd density estimation based on statistical analysis of local intra-crowd motions for public area surveillance. Opt. Eng., 51 (4) 047204-1, 2012.

[5] Chen Ke, Chen Change Loy, Shaogang Gong, Tony Xiang. Feature mining for localised crowd counting, British Machine Vision Association, pp. 21.121.11., 2012.

[6] Ge Weina, Robert T. Collins, R. Barry Ruback, Vision-based analysis of small groups in pedestrian crowds, IEEE Trans. Pattern Anal. Mach. Intell. 34 (5), 1003–1016, 2012.

[7] Gall, J., Lempitsky, V., Class-specific hough forests for object detection. In: Decision Forests for Computer Vision and Medical Image Analysis. Springer, pp. 143–157, 2013.

[8] Liang, R., Zhu, Y., Wang, H., Counting crowd flow based on feature points. Neuro computing 133, 377–384, 2014.

[9] Ryan David, Simon Denman, Sridha Sridharan, Clinton Fookes, An evaluation of crowd counting methods, features and regression models, Elsevier, Computer Vision and Image Understanding 130 1–17, 2015.

[10] Neeta, A. Nemade., Gohokar, V. V, A survey of video datasets for crowd density estimation. In International Conference on Global Trends in Signal Processing, Information Computing and Communication (ICGTSPICC), (Jalgaon, India). IEEE, 2016.

[11] Xiaohang Xu, Dongming Zhang, Hong Zheng, Crowd Density Estimation of Scenic SpotsBased on Multifeature Ensemble Learning. In Journal of Electrical and Computer Engineering, Volume 2017.

[12] Geem ZW, Kim JH and Loganathan GV, A new heuristic optimization al-gorithm: Harmony search. Simulation, 76:60-68, 2001.

[13] Lee KS and Geem ZW, A new meta-heuristic algorithm for continuous engineering optimization: harmony search theory and practice. Comput. Methods Appl. Mech. Engrg., 194:3902-3933, 2005.

[14] X.-S. Yang, "Harmony Search as a Metaheuristic Algorithm", in: Music-Inspired Harmony Search Algorithm: Theory and Applications (Editor Z. W. Geem), Studies in Computational Intelligence, Springer Berlin, vol. 191, pp. 1-14, 2009.

[15] Ghosh, J., Bovik, A.C., Neural networks for textured image processing. In: I.K. Sethi, A.K. Jain (Eds.), Progress in Artificial Neural Networks and Statistical Pattern Recognition. North-Holland, pp. 133–154, 1991.

[16] Kim, G. A., TaekiKim, Moonhyun, Estimation of crowd density in public areas based on neural network. KSII Trans. Internet Inf. Syst. 6(9), 2 170–2190, 2012.

[17] Naghdy, G., Wang, J. & Ogunbona, P. Texture analysis using Gabor wavelets. Proceedings of SPIE -Human Vision and Electronic. The International Society for Optical Engineering. Imaging, pp. 74-85, 1996.

[18] C. Liu, H. Wechsler, Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition, IEEE Transactions on Image processing, 11, 467-476, 2002.

AUTHORS PROFILE

Dr. Saeed Mohammed Baneamoon received the B.Sc. and the M.Sc. degrees in computer engineering from Technical University, Sofia, Bulgaria in 1996, the M.Sc. degree in computer science from University of Technology, Baghdad, Iraq in 2003 and the Ph.D. degree in Artificial Intelligence from Universiti Sains Malaysia, USM, Penang, Malaysia in 2011. He is Assistant Professor in Department of Computer Engineering, Faculty of Engineering & Petroleum, Hadhramout University, Hadhramout, Yemen From 2011 till now. He is a Head of Computer Engineering Department, Faculty of Engineering & Petroleum, Hadhramout University, Hadhramout, Yemen From 2014 till now. His interest focuses on the field of Artificial Intelligence and Robotics. He has published a number of papers in journals and conferences.

Dr. Ali Salem Ali Bin Sama: Received the B.E degree in Computer Engineering, Balqa University, Faculty of Applied Engineering, Jordan, 1997, M.Sc. degree in Computer science from University of Science, Malaysia, 2006, PhD degree in computer Science, Malaysia, 2011, He has 11 international papers,5 international conferences ,He is assist. professor of department of engineering geology in Minerals & Oil Faculty, Aden University, Yemen, Assistant Professor in a Department of Computer Science and Information, College of Sharia and Islamic Studies at Al-Ahsa, Imam Muhammad bin Saud Islamic University, in Saudi Arabia, From 25/08/2014 until now.