Discrete Cosine Coefficients as Images Features for Fire Detection based on Computer Vision

Jean Paul Dukuzumuremyi[1], Beiji Zou, Carine Pierrette Mukamakuza
School of Information Science and Engineering, Central South University, Changsha, China
Email: dukuze@gmail.com, piecarine@yahoo.com

Damien Hanyurwimfura
College of Information Science and Engineering, Hunan University, Changsha, China
Email: hadamfr@gmail.com

Emmanuel Masabo
Dept of Computer Engineering and Information Technology, Kigali Institute of Science and Technology, Kigali, Rwanda
Email: masabem@gmail.com

***Abstract*— Fire hazards occurring recently in the world lead to the need of designing accurate fire detection systems in order to save human lives. The newest innovations continue to use cameras and computer algorithms to analyze the visible effects of fire and its motion in their applications like the adaboost classifier which is well known for its strength in rigid objects detection from images. This paper presents a Fire Detection System (FDS) with an algorithm that works side by side with the adaboost classifier to determine the presence of fire in an image taken by a normal web camera (webcam), in order to decrease the false alarms in an indoor scene. The images are first preprocessed and their selected discrete cosine coefficients are kept for analysis to get better coefficients that will be fed to a neural network for classification and results are compared to a statistical approach used in combination with binary background mask (BBM) and a wavelet-based model of fire's frequency signature(WMF) to test its accuracy.**

***Index Terms*— Computer vision, fire detection, neural network, Discrete Cosine Transform, adaboost classifier**

I. Introduction

Fire is useful to our life but nowadays it has shown to be one of threats to humanity; like on December 11th, 2005, fire and explosion happened in Buncefield oil depot. It caused 43 persons wound, and the direct economic loss is 2.5 billion pounds. On October 29th, 2009, the fire and explosion happened in a large oil tank farm at Rajasthan state of India. It caused the death of fifteen people and one hundred fifty people were injured [1]. As fire accidents frequently put in danger our life at the same time causing economical and ecological damages and as the scarcity of automatic fire detection systems continues to be a problem that really needs serious attention, to save human lives by preventing injuries and/or deaths, new approaches to detect fire can help to avoid those fire disasters. Many early fire-detection techniques have been explored and most of them are based on particle sampling, temperature sampling, relative humidity sampling, air transparency testing, smoke analysis, in addition to the traditional ultraviolet and infrared fire detectors. However, in large rooms and high buildings, it may take a long time for smoke particles and heat to reach a detector. Those detectors then must be set in the proximity of a fire or can't provide any additional information about the process of burning, such as fire location, size, growing rate, and so on and most of time their installation is very costive. Thus, they are not always reliable because energy emission of non-fires or by products of combustion, which can be yielded in other ways, may be detected by misadventure which usually results in false alarms. To provide more reliable information about fires, Computer Vision-based approaches are becoming more and more helpful. The newest innovations are continuing to use cameras and computer algorithms to analyze the visible effects of fire and its motion in their applications. Adaboost classifier is well known for its strength in rigid objects detection in images, but this approach presents some drawbacks when working on non-rigid objects like fire. The main difficulty is then to identify the fire if it does not occur at the expected position.

This paper proposes then an improvement due to the reprocessing of images and extraction of their features in frequency domain and their classification by a neural network.

II. Related works

There exists a video based fire detection (VFD) which is a newly developed technique as it can greatly complete the fire detection requirements in large rooms and high buildings, and even for outdoor environments [2]. Up to

Corresponding author: dukuze@gmail.com

doi:10.4304/jcp.9.2.295-300

now, most of methods make use of the visual features of fire including color, textures, geometry, flickering and motion, [3], [4].

Some researchers used RGB input as a simple and effective procedure for real-time applications. To improve the RGB systems, a fuzzy logic enhanced generic color model for fire pixel classification was proposed and it used YCbCr color space to separate the luminance from the chrominance more effectively than color spaces such as RGB[5]. An algorithm of early fire image detection and identification based on discrete fractal brownian incremental random field model is proposed in [6]. It is important to realize that most of fire detection systems used the heuristic fixed threshold values in their specific methods [7].

In addition to motion and color, another approach proposing smoke/fire detection by analyzing the video in wavelet domain has been treated in [8].

To improve the solutions to the shortcomings enumerated above, many researchers opted computer vision based fire/smoke detection. Some techniques such as combination of color and stereo vision used in [8] have been developed to compensate small or gradual changes in the scene or the lighting. Although stereo vision and edge detection proposed in [9] can be used to detect independently moving targets in the presence of camera motion, it is not feasible for non-rigid object extraction since the movements of the fire/smoke are different. To increase reliability, some systems like the one used in [10][11] integrate multiple cues such as stereo, color and wavelet pattern to detect fire /smoke. However, color is very sensitive to illumination changes[11-12] which proves that frequency domain features can be more reliable cues than color in general situations, the reason why a new colour-based model of fire's appearance as well as a new Wavelet-based model of fire's frequency signature are proposed in [13] and this paper proposes an improvement by extracting images features in frequency domain (Discrete Cosine coefficients) used with adaboost method for object detection and their classification by a neural network.

III. Proposed approach

Our new fire warning system is designed to be equipped with a normal web camera detection module that detects fire in real-time based on frequency features of frames, and its main function is to improve the previous work in reducing false fire alarms by distinguishing fire information from other objects present in the scene. The common process of in detection methods consist of two steps: the first step is extracting features; the second is using learning method or classifying method to build classifiers based on those features[14]. This is achieved through frames reprocessing, discrete cosine for features extraction and recognition based on neural networks. The neural network modeling system is developed to model a system with some descriptive rules that are designed to deal with complex, ill-defined and unconstrained problems. The two main purposes of neural networks are: (i) to make an accurate and quick procedure to approach the desired result, (ii) to help in building a general and flexible mechanism which is applicable in solving various fields of modeling problems shown in[14] and [15]. As accuracy is very important in classifiers used for emergency applications, a high percentage of false negatives in screening systems increases the risk of real fire images by not receiving the attention they need, while a high false alarm rate causes unwarranted worries and increases the load on system resources.

IV. The algorithm

Our algorithm consists of four steps: adaboost classification, preprocessing, features extraction and classification. The adaboost classification is helped by the next steps in case it failed to detect an actually present fire in the video frame. Preprocessing deals with segmentation in order to isolate the new object, image enhancement, conversion, resizing, border normalization and histogram equalization after the object has been isolated as being within the indoor scene with the help of a predetermined disparity map. The features extraction is accomplished using discrete cosine transforms. Lastly we recognize the presence of fire in the frame based on neural networks classification. In our experiment, we took 1000 positives (images with fire) and 1000 negatives (images without fire) indoor images to use for the production of an adaboost xml description file. Every image has been resized to 256×192. The data set contains images taken in different circumstances of light and time to ensure variety[13]. We used several different environments such as candles, electrical tubes, paper fire and plastic fire in night shots or daytime shots to gain the variability. Fig. 1 depicts the general flowchart of our fire classification system. The system input is a web camera video images, the system output is the class label (1 for positive and 0 for negative) to determine whether there is fire or not in the frame,

V. System Setup and detection Algorithm

The system is set in a way that if the adaboost classifier manages to find a fire in the frame, so there is no need for further steps as shown in Fig. 2.

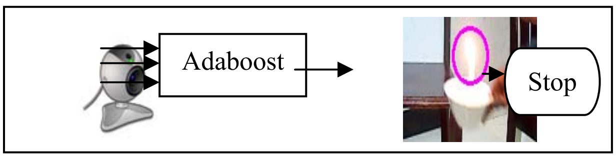


Figure.2 Flowchart for an adaboost classification.

If the adaboost failed to find the fire then an improvement of the system is done by the next steps: reprocessing, discrete cosine transform, and the neural networks classifier to determine whether the system was right or wrong as shown in Fig 3.

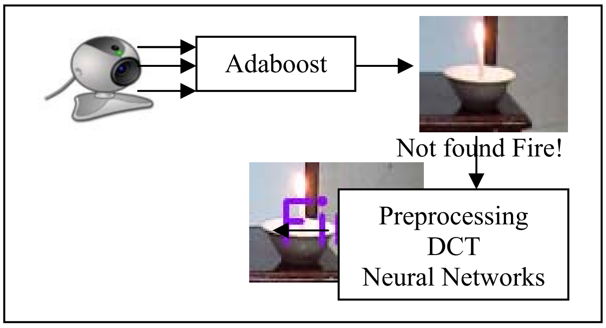


Figure.3 The Flowchart of the improved system after adaboost failed to find the existing fire

As the fire has to be detected in an indoor scene, there is need to determine the interior of the room so that outside scene is not taken into account. Given a 3D information recovered from the calculation of the disparity map, an issue to solve is the determination of the lowest boundary of the house interior in the disparity space. Then, the object of interest will be defined as all connected pixels with disparity values larger than the predetermined disparity threshold map. This stage mainly consists of three steps[16]:

1. Off-line determination of a unique disparity threshold map: Establish a pre-determined disparity threshold map (disparity template) that defines the background surfaces for suppression. This disparity template is built off-line.

2. Current creation of a binary suppression mask using the disparity map of the present situation and the pre-determined threshold map. The disparity map of a new pair stereo is generated during the operation and compared with the disparity threshold map. The output is a binary mask and the background pixels are then assigned to zeros.

3. Employ morphological closing operator to remove the noise and smooth the foreground regions: due to noise and other artifacts coming from the calculation of the disparity map, the extracted object from the previous step might appear as containing holes or even being sliced into several parts. This refinement stage of the algorithm handles these problems through binary image processing techniques including dilation, singularity suppression and erosion.

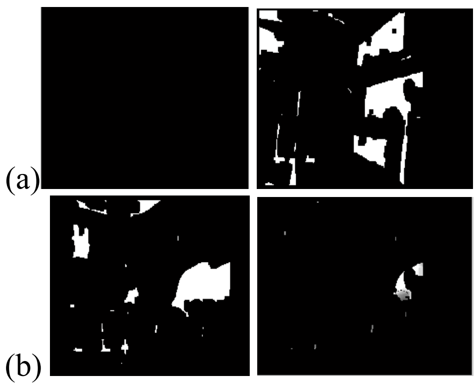


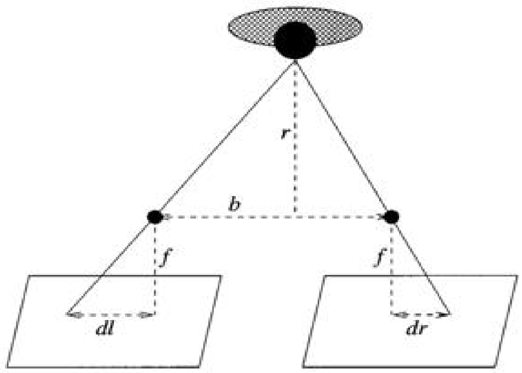
Figure.4 Results of the disparity-based object extraction on images in indoor scenes. The rows (a) and (b) show the disparity map of an original image (right column) and the extracted object image (left column) of a kitchen background, high flame, low flame and very low flame, respectively

A. Disparity Threshold Map Determination

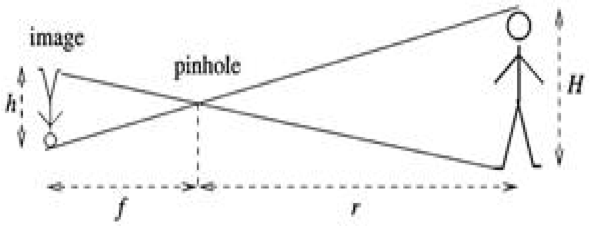
The disparity threshold map is generated off-line using a set of representative images of the interiors of a kitchen without fire presence. Both physical and virtual boundary surfaces with added artificial textures to the physical boundary are used. A physical boundary is defined as the interior of the kitchen or a structure that limits the view depth of the cameras. A virtual boundary surface can be established from the stereo camera configuration and the desired maximum view depth of the camera. Obviously, the selected boundaries should contain the maximum volume within which a fire will be classified. For better range accuracy, artificial textures are added to the physical boundary surfaces in creating the disparity threshold map in order to obtain the perceptual benefits of texture gradients. The goal of this approach is not to produce realistic textured images, but rather to aid the visualization of rendered surfaces. For the same reason, averaged disparity values of a number of image frames are used in the process. Experimentally, an average of over 100 frames is good enough to establish the disparity threshold map[11].

B. Background Suppression

A comparison between the measured disparity value and the disparity threshold for all the pixels in the image creates a binary map having the same matrix correlation as the original image. For a given pixel location (i, j), assume the measured disparity value is $Db(i, j)$ and the disparity threshold is $Dt(i, j)$. The new binary map $Q(i, j)$ with the same matrix correlation as the original image is first created with the following rules: If $Db(i, j) \leq Dt(i, j) + \mu(i, j)$, then $Q(i, j) = 0$ otherwise $Q(i, j) = 1$, where the parameter $\mu(i, j)$ is used to control the degree of background suppression which can be a function of the disparity threshold $Dt(i, j)$, the location of the pixel or simply a constant. In our current implementation, it was chosen as a constant for simplicity. The disparity value is defined in such a way that it increases as the object distance decreases. In this case, the areas indicated by 0*s* in $Q(i, j)$ represent objects that are beyond the desired detection range. Similarly, the areas indicated by 1*s* in $Q(i, j)$ represent objects that are within the desired detection range. We have set up the stereo cameras so that their image planes are embedded within the same plane and separated with base length *(b)* of 10 cm as shown in Fig. 5. Under this condition, the difference between *dl* and *dr* is called the disparity, and is directly related to the distance of the object normal to the image plane[1].



(a)



(b)

Figure.5. (a) Stereo geometry, (b) image-forming geometry

C. On-line Morphological Operator Application for Object Enhancement

In order to reduce the noise effects, fill out the holes and/or connect several parts of the object of interest in the binary image, Image processing techniques of type opening and closing operations are used to achieve a smoother and reliable background suppression mask [16].

- A 3 × 3 8−way flat structuring element for morphological operators is applied to the initial binary mask $Q(i, j)$. Two different levels of erosion and other singularity suppression operators are introduced here and applied as follows:
- Erosion Level 1: If the summation of its 8−way neighbors is greater than or equal to 1, set $Q(i, j)$ to 1; otherwise set $Q(i, j)$ to 0[11].
- Erosion Level 2: If the product of its 8−way neighbors is 0, then set $Q(i, j)$ to 0, otherwise set $Q(i, j)$ to 1.
- Suppression: If the summation of its 8−way neighbors is less than or equal to 3, set $Q(i, j)$ to 0; otherwise no change to $Q(i, j)$.
- Since the segmentation is based on a low-resolution disparity calculation, the disparity calculation is applied block by block instead of pixel by pixel. The block size is 10×16 pixels. Multiple binary image processing steps should be applied to the new binary map to reject noise.

VI. FEATURES EXTRACTION

DCT (Discrete Cosine Transform) has been the features generator which has been applied to the image block and only the first 10 of the total coefficients are retained after they showed to be more discriminative than others after analysis and along with the eleventh attribute which determines the class label(0 or 1 given that a fire is present or not in the frame) they have contributed to the feed of the first layer of the neural network.

VII. NEURAL NETWORKS BASED FIRE RECOGNITION

A back-propagation neural network has been used as it can be trained for different kinds of scenes and can even deal with noisy data robustly. The design of the input data to the neural network (NN) is important as it directly affects the performance of the network. The goal is to make the input data maintain the shape information for recognition while being reduced to a manageable amount. The network is set to be of 3 layers (11-10-2):

- one input node per attribute for input layer
- 10 hidden nodes
- one output node per label

At the prediction stage - the highest probability can be accepted as the "winning" class label output by the network. The network uses a sigmoid function with alpha and beta parameters 0.6 and 1 specified respectively. The network will terminate the training after either 1000 iterations or a very small change in the network weights is below the specified values. The network gives out a vector of probabilities for each class, taking the class with the highest probability for simplicity and it should also check the separation of different probabilities in this vector in case two classes have very similar values.

VIII. EXPERIMENT RESULTS

We implemented the FDS by using Visual C++ 6.0 and OpenCV computer vision library. Windows XP, intel 1.86GHz and RAM 1GB are used as our computing environment.

For machine learning, 1000 images of fire and 1000 images without fire were used for testing the proposed method. Every image has been resized to 250×250. Several different environments were used; such as candles light, electrical tubes light, paper fire and plastic fire in night shots and day time shots to gain the variability. Another source of frames has been taken from the indoor video clips from Signal and Image Processing Group at Bilkent University in Turkey (http://signal.ee.bilkent.edu.tr/ VisiFire/Demo/FireClips/) and we compared the results with a Background Binary Mask denoted as BBM used with statistical color model as it has been developed on the same environments[7] and then we compared our results with a wavelet-based model of fire's frequency signature denoted as WMF[13].

From the results presented in Table I, out of 1000 images containing fire, the adaboost classifier counted up to only 720 images, the BBM counted up to 820, the WMF counted up to 850 whereas the improved FDS counted up to 890 images; all the systems getting 72%, 82% , 85%and 89% respectively.

From the results presented in Table II, out of 1000 images without fire, the adaboost classifier got 810 images, the BBM counted up to 840, the WMF counted

TABLE I.
DETAILED COMPARATIVE ANALYSIS ON POSITIVE IMAGES

Method	Overall positive images	False negative images
Adaboost	72%	28%
BBM[7]	82%	18%
WMF[13]	85%	15%
FDS	89%	11%

TABLE II.
TABLE II DETAILED COMPARATIVE ANALYSIS ON NEGATIVE IMAGES

Method	Overall negative images	False positive images
Adaboost	81%	19%
BBM[7]	84%	16%
WMF[13]	86%	14%
FDS	87%	13%

TABLE III.
PERFORMANCE COMPARATIVE ANALYSIS FOR BOTH APPROACHES

Methods	Overall positive images		
	Accu-racy(%)	Sensi-tivity	Speci-ficity
Adaboost	76.5	0.72	0.81
BBM[7]	83	0.82	0.84
WMF[13]	85	0.85	0.86
FDS	88	0.89	0.87

up to 860 whereas the FDS got 870; all the systems getting 81%, 84%,86% and 89% of the total images respectively. From the results, it is clear that the FDS based fire detection presented an advantage of getting more positive images as shown by the three different judging methods: accuracy, sensitivity and specificity, which were used as defined in equation (1), (2) and (3).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} x100\% \quad (1)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (2)$$

$$Specificity = \frac{TN}{TN + FP} \quad (3)$$

Where *TP* is the true positive, *TN* the true negative, *FP* the false positive, and *FN* the false negative.

IX. CONCLUSION

In this paper, an algorithm based on computer-vision has been proposed for fire detection using frequency features of the video frames. The input images were first transformed to grey scale then image processing techniques were used to reject noise introduced into the images through low resolution disparity calculations. To check the robustness of the used technique, we compared our results to those from adabost classifier and spacial statistical color model with Binary Background Mask (BBM) as well as the wavelet-based model of fire's frequency signature (WMF) and the averaged overall accuracy is about 11.5% better than that of adaboost classifier, 5% better than BBM and 2.5% better than that of WMF. The sensitivity of 0.72 and specificity of 0.81 showed that this approach can be used along with the adaboost classifier to get smarter fire detection.

ACKNOWLEDGMENT

This work is Supported by the National Science Foundation of China under Grant No. 60970098, Modeling for Human Visual System and its application; The National Science Foundation of China under Grant No. 61173122.

REFERENCES

[1] J. Tan, Y. Xie,T. Wang. "Fire and Explosion Hazard Prediction Base on Virtual Reality in Tank Farm",JOURNAL OF SOFTWARE, 2012, 7(3), pp. 678-682

[2] R. Lucile, A. Moulay and Tison Yves, "Dynamic Fire 3D Modeling using a Real-time Stereo Vision System", Journal of Communication and Computer ,2009, 6(10) pp 34

[3] T. Celik, "Fast and Efficiency Method for Fire Detection Using Image Processing", ETRI Journal, 2010, 32(6) pp.881-890.

[4] C. Jun, D. Yang, and W. Dong, "An Early Fire Image Detection and Identification Algorithm Based on DFBIR Model", Computer Science and Information Engineering, 2009 WRI World Congress, Los Angeles, CA USA, pp. 229 - 232.

[5] B. Cho, J. Bae, and S. Jung, "Image Processing-Based Fire Detection System using Statistic Color Model", Advanced Language Processing and Web Information Technology, ALPIT '08. In Proceeding of International Conference on Advanced Language Processing and Web Information Technology, Dalian Liaoning , 2008 pp: 245 – 250.

[6] A. Akhloufi, L. Rossi1, L. Abdelhadi and Y. Tison. "Dynamic Fire Modeling in Three-dimensional Space", First IEEE International workshop on Image Processing Theory, Tools & Applications, 2008. pp. 1-7

[7] B. Ko, K. HoCheong, and J. YealNam. "Early Fire Detection Algorithm based on Irregular Patterns of Flames and Hierarchical Bayesian Networks", Elsevier, Fire Safety Journal, 2010, 45(4) pp. 262-270.

[8] N. Do, and M. Vetterli, "Wavelet-Based Texture Retrieval Using Generalized Gaussian Density and Kullback–Leibler Distance". IEEE Transactions On Image Processing,2002, 11(2), pp.146-158.

[9] T. Truong and J. Kim, "Fire Detection with Video using Fuzzy k-means and Back-propagation Neural Network", In the Proceedings of the 8th international conference on Advances in neural networks ,Part II, 2011, pp. 373–380.

[10] S. Ham, B. Ko, and J. Nam, "Fire-Flame Detection based on Fuzzy Finite Automation", In the proceeding of 20th International Conference on Pattern Recognition,2010, pp. 3919-3922.

[11] J. Gubbi, S. Marusic, and M. Palaniswami, "Smoke Detection in Video using Wavelets and Support Vector Machines", Fire Safety Journal, Elsevier, September 2009, 44(8) pp. 1110-1115.

[12] W. Li,B. Fu,L. Xiao , Y. Wang , P. Liu "A Block-based Video Smoke Detection Algorithm".Journal of software, 2013, 8(1)pp.63-70

[13] H. Kong et al. “Disparity Based Image Segmentation For Occupant Classification”, In proceeding of International conference on computer vision and pattern recognition workshop , 2004, pp. 126-133.

[14] P. Santana, P. Gomes, and J. Barata "A vision-based system for early fire detection", 2012 IEEE International Conference on Systems, Man and Cybernetics, 2012, pp. 739-744.

[15] Y. Ouyang, N. Sang. Robust Automatic Facial Expression Detection Method, Journal of Software, 2013 , 8(7) pp .1759-1764.

[16] J. Dukuzumuremyi, P. Umugwaneza, Z. Beiji, W. Lei, L. Yixiong. “Computer vision-based Image segmentation for Fire/Smoke Detection”, World Academy of Science, Engineering and Technology, Paris-France, June 24-26, 2011.

[17] H. Feng “A Self-Tuning Fuzzy Control System Design”. In Proceedings of 9th IFSA World Congress and 20th AFIPS International Conference. Vancouver, 2001,pp. 209-214.

Jean Paul Dukuzumuremyi has got his Master of Engineering Degree in computer application technology from the school of information science and engineering at Central South University, China in 2009. He is currently a Ph.D. candidate at the same university. His research interests include Digital Image Processing, Computer Graphics, Computer Vision, and their applications.

Beiji Zou holds a Ph.D in Computer science and technology from Hunan University, China.

Beiji is currently a Professor in School of Information Science and Engineering of Central South University, China. Since his graduation, Mr. Zou has been engaged in research of Digital Image Processing and Graphics. His papers were published in different Journal including Research and Development of Computer, Journal of Computer, Journal of Electronics these was presented at the International Conference sponsored by IEEE. In 1986, he participated in a project “3D Free-Form Surface Modeling”, which was supported by the youth award research funds from Chinese Academy of Science.

Beiji worked in Zhejiang University, China as a visiting scholar from Jul.1999 to Jun. 2000 and in Tsinghua University, China as a post-doctor from Jan. 2002 to Nov. 2003 respectively, engaged in research about Human Facial Expressional Recognition and Animation. He worked in the research field of Multimedia Technology in Griffith University of Australia from Dec. of 2003 to Dec. of 2004.

In recent years, he has focused on computer graphics, computer animation, multimedia technology and virtual reality technique. He has published sixties of papers and accomplished more than 10 projects.

Carine Pierrette Mukamakuza has got her Bsc Degree in computer application technology in 2012 from school of information science and engineering at Central South University, Changsha, China. She is currently a Master’s student at the same University. Her research interests include Digital Image Processing, Rough Sets and their applications.

Damien Hanyurwimfura is curently a Ph.D student at the college of information science and engineering, Hunan University, Changsha, China. He is a lecturer at Kigali Institute of Science and Technology (KIST), Rwanda. He received his Masters degree of engineering in computer science and technology from Hunan University, Changsha, China in 2010. His research interests include wireless network, security and data mining.

Emmanuel Masabo has got his Master of Engineering Degree in computer application technology from the school of information science and engineering at Central South University, China in 2009. He is a lecturer at Kigali Institute of Science and Technology (KIST), Rwanda.. His research interests include image processing and web security.