Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia Engineering 41 (2012) 1183 – 1189

Engineering Procedia

www.elsevier.com/locate/procedia

International Symposium on Robotics and Intelligent Sensors 2012 (IRIS 2012)

Performance Analysis of ANN based YCbCr Skin Detection Algorithm

A. M. Aibinu, A. A. Shafie and M. J. E. Salami

International Islamic University Malaysia (IIUM)
P. O. Box 53100 Gombak, Malaysia.
E-mail: maibinu@iium.edu.my

Abstract

Skin detection from acquired images has various areas of applications especially in automatic facial and human recognition system. The performance analysis of artificial neural network based –YcbCr skin recognition and three other techniques is evaluated in this work. Results obtained show that the use of YCbCr color model performs better than RGB colour model and the use of artificial neural network further improves the accuracy of the system.

Keywords: Acquired Image, Artificial Neural Network, Modeling Technique, Skin.

Nomenclature

ACC	Accuracy
ANN	Artificial Neural Network
CC	Correlation Coefficients
FN	False Negative
FP	False Positive
FPR	False Positive Rate
MSE	Mean Square Error
PR	Precision
RGB	Red-Green and Blue channel of an Image
YCbCr	luminance, Chromatic blue and Chromatic red channel of an Image
TN	True Negative
TP	True Positive
TPR	True Positive Rate
Subscripts	
T	transformed image
s	Skin pixel
ns	Non Skin pixel
1,2	Coefficient of Chromatic blue and red channel respectively

* Corresponding author. Tel.: +601-736-58-247; fax: +603-6196-4433.
E-mail address: maibinu@iium.edu.my

1877-7058 © 2012 Published by Elsevier Ltd. Open access under CC BY-NC-ND license.
doi:10.1016/j.proeng.2012.07.299

1. Introduction

Human skin surface and tone darkness depend on concentration of melanin, hemoglobin and some other pigments in the skin. As the concentration of these constituents increases, skin surface absorbs and reflects rays of light based on the skin roughness thus the difference in the perceived skin colouration. Figure 1 shows various forms of human skin. Several colours space techniques have been reported in the literature and among the known colour space models in digital image processing are: RGB (Red (R), Green (G) Blue (B) color mode); YCbCr (luminance (Y), chromatic blue (Cb), Chromatic red (Cr) color model); HSV (Hue(H), saturation(S), Value(V)) color model); CMY(Cyan(C), Magenta(M), Yellow (Y) color model) and CMYK (Cyan (C), Magenta (M), Yellow (Y), black(K)) color model [1-16].

An RGB image is an array of color pixels, where each color pixel is a triplet corresponding to red, green and blue components of the RGB image. In a typical M x N x 3 image, M represent the number of rows in the image while N represents the number of column present in that image and the last 3 represents the triplet R, G and B colour information in matrix form. However, in YCbCr (luminance(Y), hue (I), saturation (Q)) color model, the luminance information is represented by a single component Y while the color information is stored as two color difference components Cb, Cr. The difference between the blue component and reference value component is called Cb while the difference between the red component and reference value is referred to as component Cr [15]. Mathematically the relationship between RGB and YCbCr color scheme can be expressed as

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 65.481 & 128.553 & 24.966 \\ -37.797 & -74.203 & 112.000 \\ 112.000 & -93.786 & -18.214 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

Despite the availability of numerous colour space, the use of RGB and YCbCr have received favorable consideration in the literature [1,2]. Hence in this work, performance analysis of Skin Detection using YCbCr has been evaluated.

Figuer 1: Variation of Human Skin [3]

2.0 Review of Some Selected RGB and YCbCr Skin Detection Techniques

Lots of work have been reported regarding the use of RGB color space for skin-pixel identification [3, 5]. Kovac, Peer and Solina proposed the use of static threshold of the RGB channels of the acquired image in the determination of skin-color region [3]. Accuracy obtained was shown to depend on the illumination condition of the environment in static threshold condition . In addition, the author argued that a better approach to enhance the performance of this technique would have been the use of a color space that is able to separate luminance component from chromatic component with partial independence of the chromaticity and luminance [3]. Hence, the choice of YCbCr color space in this work.

A fixed threshold technique for skin-pixel region identification in YCbCr color space was proposed in [2, 4]. Pixels satisfying the threshold values obtained offline were subsequently regarded as the likely skin-pixel. However, the major shortcoming of this approach lies in the use of fixed threshold which has been shown to depend on illumination and race. Yogarajah et. al [6] proposed an adaptive threshold and illumination invariant approach as solution to fixed threshold methodology in [2, 4]. Taking a cue from the argument proposed by Zheng et. al in [7] that, only little variation exists between the skin on human face and the other part of the body, the proposed technique firstly determines face in an image and subsequently use the color distribution in the face to classify other skin-pixel in the image. Though the dynamic threshold approach [6] outperforms the fixed threshold approach [2, 4], however this method is only applicable for images with faces as the method is unable to work in an image without face. Hence, the dynamic threshold approach is inappropriate for the current work as the acquired images does not involve the use of facial part.

In another work, skin-pixel detection using a very narrow band of values in only the Cr and Cb matrix of YCbCr model of an acquired image was reported in [8]. Unlike the reported method in [2, 4], two sets of static threshold ranges of $133 \le Cb \le 173$ and $77 \le Cr \le 127$ were proposed as skin-pixel band. The authors demonstrated that the band derived

from Cr and Cb remains effective in isolating skin region irrespective of skin-color variation. However, further analysis have shown that this method is illumination and race dependent.

The use of only the Cb and Cr information in a YCbCr color scheme was proposed in [1,14]. The approach involve the use of artificial neural network for coefficients determination while modification and the use of pseudo random number was also examined. Extracted skin-pixel and non skin-pixel from an image are fed to artificial neural network and from the adaptive weight and learning coefficients of the activation function, model coefficients required for skin extraction are obtained during offline training phase. Those coefficients are subsequently used in extracting likely skin-pixel during testing stage and this approach has been shown to be invariant to illumination variation, skin-color and image background.

Light compensating approach in RGB color scheme before color scheme transformation was proposed in [9]. The proposed technique involves accurate determination of average luminance of the input image, followed by the determination of compensation factor for R and G channel in the RGB acquired images. The compensated images were then transformed to YCbCr where Y and Cb channel are discarded and static threshold in the range of $10 \le Cr \le 45$ is applied to the Cr channel only. Lastly, it was demonstrated in [9] that the proposed technique is also illumination invariant.

Artificial Neural Network (ANN) based YCbCr skin detection technique was reported in [1]. An ANN is a global search technique which emulates the biological neurons of the human body. It has massive parallel distributed structures with high capability of learning and generalization upon convergence. ANN consists of interconnection of simple processing elements called nodes. Figure 2 shows a typical ANN with input and output nodes and single hidden layer.

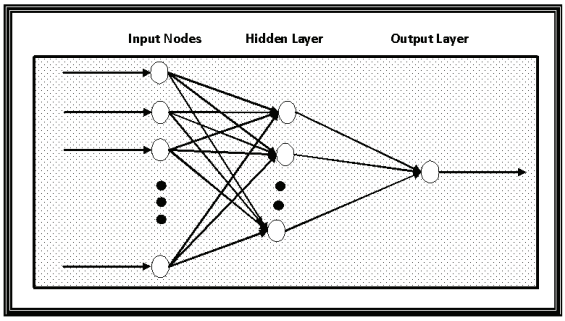

Figure 2: ANN Structure

A two layer network with two input nodes, 5 nodes in the hidden layer and 1 node at the output layer have been used for the extraction of likely skin pixel. Upon convergence, the parameters, i.e the synaptic weights and coefficients of the adaptive activation function of the network were extracted to compute $\mathbf{K_1}$ and $\mathbf{K_2}$. Skin regions and non-skin pixels are determined using

$$f_T(x,y) = \sum_{x=0}^{X-1}\sum_{y=0}^{Y-1} K_1 Cb(x,y) + \sum_{x=0}^{X-1}\sum_{y=0}^{Y-1} K_2 Cr(x,y) \quad (2)$$

where K_1 and K_2 are extracted coefficients and Cb and Cr are obtained from the input image using (1), x and y are row and column number respectively, X and Y are the total number of rows and columns in the input image and $f_T(x,y)$ is the segmented binary image. Steps involved in the extraction of the required coefficients are:

Step 1 Extract likely skin-region of interest from any acquired image from the database. Convert the acquired skin-region of interest to $\mathbf{Cb_s}$ and $\mathbf{Cr_s}$ data vector using (1).

Step 2 Extract likely non-skin region of interest from any acquired image from the database. Convert the acquired non-skin region of interest to $\mathbf{Cb_{ns}}$ and $\mathbf{Cr_{ns}}$ data vector using (1)

Step 3 Concatenate the input vector obtained in step 1-2 to form ANN Input data for a network configure with adaptive coefficients of the activation and model synaptic weights of the network.

$$\text{Input Training Data} = \begin{bmatrix} Cb_{ns} & Cb_s \\ Cr_{ns} & Cr_s \end{bmatrix}$$

Step 4 Generate the target data of the same length with that of Input Data. 1 is assigned to skin ROI while 0 is assigned for Non skin ROI.

Step 5 Train the ANN and upon convergence, extracts model synaptic weights and coefficients of the adaptive activation function. Compute $\mathbf{K_1}$ and $\mathbf{K_2}$ as stated in [17]

In the testing phase, the skin and non-skin pixels are partitioned into two non-overlapping regions using

$$f_{T(x,y)} = \begin{cases} 1 & skin \quad pixel \\ 0 & non-skinpixel \end{cases}$$

3.0 Performance Measures, Results and Discussion

The definition of terms used for the performance measures are:

• True Positive (TP): If a skin-pixel (positive (P)) is classified as skin-pixel (positive (P)), then it is counted as a true positive.

• True Negative (TN): If a non skin-pixel (negative (N)) is classified as non skin-pixel (negative (N)), such is regarded as true negative.

• False Negative (FN): If a positive (P) pixel is classified as negative (N) pixel, such is regarded as false negative.

• False Positive (FP): If a negative (N) pixel is classified as positive (P) pixel, such is regarded as false positive.

The performance metrics used in this work are given as :

• Accuracy (ACC): Accuracy is the degree of closeness of a skin-pixel detection algorithm and the ground truth skin image. ACC is mathematically expressed as:

$$ACC = \frac{TP + TN}{\sum P + \sum N} \tag{3}$$

• The Precision (PR): The PR of a skin-pixel detection algorithm is calculated using

$$TPR = \frac{TP}{TP + FP} \tag{4}$$

• The True positive rate (TPR): The TPR of a skin-pixel detection algorithm is estimated as

$$TPR = \frac{TP}{\sum P} \tag{5}$$

• False Positive Rate (FPR): The FPR which is also known as the false alarm is defined as

$$FPR = \frac{FN}{FN + TN} \tag{6}$$

• Mean Square Error (MSE): The MSE is computed as

$$MSE = \frac{1}{NM} \sum_{x=1}^{M} \sum_{y=1}^{N} | S_R(x,y) - \hat{S}_R(x,y) |^2 \tag{7}$$

and $S_R(x,y)$ is the binarized ground truth image, $\hat{S}_R(x,y)$ is the binarized output of the method under consideration, N is the number of rows and M is the number of columns.

• Correlation coefficient (CC): Correlation coefficient (CC) quantifies the closeness between the binarized ground truth image and the binarized output of the method under consideration. CC varies between -1 to $+1$, in which $+1$ indicates that the data points are identical and -1 implies that the points are highly dissimilar. CC is computed from

$$CC = \frac{\sum_{x=1}^{M} \sum_{y=1}^{N} (S_R(x,y) - \bar{S}_R(x,y))(\hat{S}_R(x,y) - \bar{\hat{S}}_R(x,y))}{\sqrt{\sum_{x=1}^{M} \sum_{y=1}^{N} (S_R(x,y) - \bar{S}_R(x,y))^2 \sum_{n=1}^{M} (\hat{S}_R(x,y) - \bar{\hat{S}}_R(x,y))^2}} \tag{8}$$

where $\bar{S}_R(x,y)$ and $\bar{\hat{S}}_R(x,y)$ are the mean of the ground truth and output image respectively, $S_R(x,y)$ and $\hat{S}_R(x,y)$ as previously defined.

3.1 Results Obtained and Discussion

Performance analysis measures used in this work include ACC, FPR, TPR, PR , MSE and CC using skin-pixel detection algorithm proposed in this work and the proposed techniques and methods reported in [4, 3, 8, 9] are shown in Table 1 - Table 6. The database consists of 46 images of human face captured at different posture and illumination conditions [2, 4]. Ground truth images were obtained using manual segmentation technique and the output of each of the skin-pixel detection algorithm were compared with the ground truth in evaluating its performance as reported subsequently.

Table 1 shows the accuracy results obtained for some of the images in the database. Firstly, it is observed that the accuracy of the RGB based technique [3] decreases as the distinction between skin-pixel and background becomes more complex. Though high accuracy was obtained in some cases, in which computation of the TPR shows very low TPR as compared to YCbCr techniques [4, 8, 9]. Overall, the use of YCbCr skin pixel techniques show better performance as compared to the use of RGB method for skin-pixel detection. This shows the need for explicit colour components separation so as to remove the luminance component (Y) in the image. The de-correlation of luminance was ascribed to the dependency of skin colour on chrominance rather than luminance. Therefore, the Cb and Cr of the two regions are sufficients for skin detection.

Table 1: Accuracy of Skin-Pixel Detection Algorithms

	[1]	**[4]**	**[3]**	**[8]**	**[9]**
Image-1	0.9741	0.9732	0.9826	0.9587	0.9653
Image-2	0.9719	0.9718	0.9832	0.9564	0.9675
Image-3	0.9685	0.9667	0.9755	0.9390	0.9636
Image-4	0.9678	0.9651	0.9829	0.9462	0.9670
Image-5	0.9645	0.9622	0.9672	0.9047	0.9447
Image-11	0.9592	0.9611	0.9712	0.9425	0.9579
Image-12	0.9483	0.9510	0.9626	0.9302	0.9405
Image-13	0.9405	0.9580	0.9564	0.9234	0.9388
Image-14	0.9421	0.9616	0.9532	0.9087	0.9406
Image-15	0.9448	0.9608	0.9648	0.9342	0.9350
Image-21	0.9720	0.9028	0.9723	0.9677	0.9518
Image-22	0.9698	0.9157	0.9596	0.9645	0.9596
Image-23	0.9501	0.8911	0.9575	0.9425	0.9351
Image-24	0.9618	0.8693	0.9770	0.9497	0.9439
Image-25	0.9625	0.9127	0.9691	0.9540	0.9495
Image-31	0.9674	0.9653	0.9714	0.9473	0.9572
Image-32	0.9666	0.9652	0.9804	0.9506	0.9560
Image-33	0.9679	0.9650	0.9759	0.9396	0.9503
Image-34	0.9708	0.9669	0.9772	0.9496	0.9558
Image-35	0.9689	0.9696	0.9757	0.9588	0.9633
Image-41	0.9585	0.9415	0.8767	0.9054	0.9179
Image-42	0.9354	0.9291	0.7995	0.8691	0.9010
Image-43	0.9710	0.9524	0.9228	0.9466	0.9629
Image-44	0.9686	0.9646	0.8349	0.9305	0.9521
Image-45	0.9689	0.9653	0.8012	0.9163	0.9369
Image-46	0.9692	0.9482	0.7190	0.9023	0.9073

Other results obtained from the performance analysis of all the skin-pixel detection algorithms are stated in Table 2 to Table 6. In Table 2, the maximum MSE obtained using ANN-based YCbCr is 0.0783 as compared to other techniques. Further comparison of the proposed technique with other YCbCr based skin-pixel detection technique [4, 8, 9] was also evaluated in this work. This justify the use of offline training using ANN approach as oppose to the singular use of segmentation techniques proposed in other related work. The consistency in high accuracy and TPR values with low FPR value are noticeable from Table 2 for the proposed YCbCr technique with an average accuracy value of 0.9610 as compared to lower values which were obtained from other algorithms as shown in Table 1. MSE results obtained over the entire database is shown in Figure 2.

Training using ANN approach has allowed the network to learn likely skin region in the offline mode thus subsequent usage in online mode on unseen images from different database shows that upon convergence, the system is able to genaralize accurately.

Table 2: Accuracy of Skin-Pixel Detection Algorithms

	[1]	[4]	[3]	[8]	[9]
Minimum	0.9217	0.8195	0.7190	0.8691	0.9010
Maximum	0.9763	0.9754	0.9856	0.9677	0.9675
Mean	0.9610	0.9449	0.9447	0.9336	0.9471
Std. Dev.	0.0133	0.0340	0.0599	0.0236	0.0166

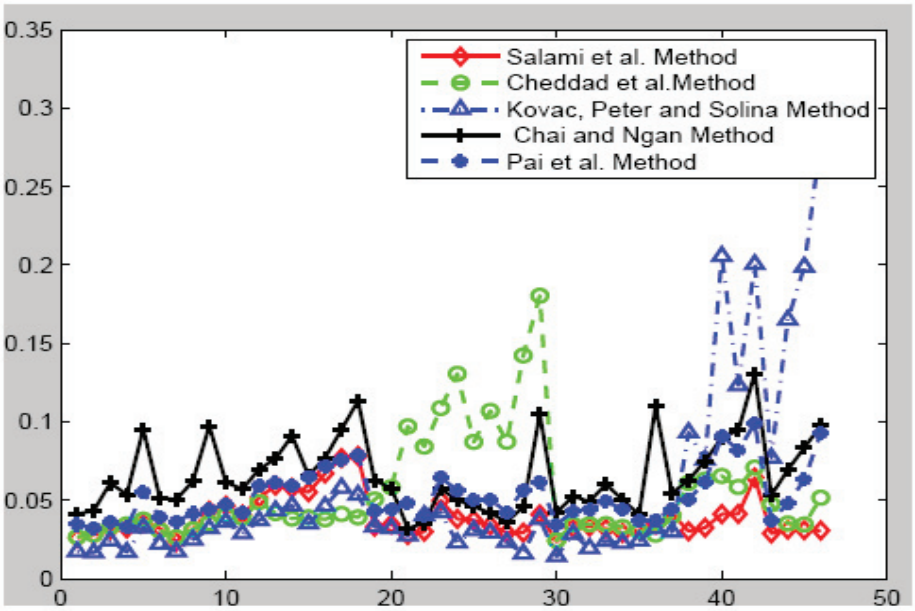

Figure 2: MSE of Skin-Pixel Detection Algorithms

Similarly, in Table 3, the correlation coefficients between the detected skin-pixel region and the ground truth for ANN-based YCbCr technique show highest mean correlation over the whole database, however the mean precision and mean FPR values obtained shown in Table 4 and Table 5 trail behind values obtained in the method reported in [4].

Table 3: MSE of Skin-Pixel Detection Algorithms

	M1	M2	M3	M4	M5
Min	0.0237	0.0246	0.0144	0.0323	0.0325
Max	0.0783	0.1805	0.2810	0.1309	0.0990
Mean	0.0390	0.0551	0.0553	0.0664	0.0529
SD	0.0133	0.0340	0.0599	0.0236	0.0166

Table 4: CC of Skin-Pixel Detection Algorithms

	M1	M2	M3	M4	M5
Min	0.7449	0.4919	0.4667	0.6247	0.6918
Max	0.9346	0.9325	0.9590	0.9029	0.9092
Mean	0.8818	0.8189	0.8541	0.8174	0.8438
SD	0.0437	0.1134	0.1183	0.0696	0.0503

Table 5: Precision of Skin-Pixel Detection Algorithms

	M1	[4]	[3]	[8]	[9]
Min	0.6092	0.6646	0.3227	0.4453	0.5446
Max	0.9376	0.9525	0.9732	0.8968	0.9093
Mean	0.8369	0.8601	0.8186	0.7433	0.7978
SD	0.0767	0.0577	0.1617	0.1082	0.0872

Table 6: FPR of Skin-Pixel Detection Algorithms

	[1]	[4]	[3]	[8]	[9]
Min	0.0130	0.0058	0.0062	0.0278	0.0207
Max	0.0922	0.0589	0.3244	0.1473	0.1113
Mean	0.0428	0.0314	0.0587	0.0772	0.0555
SD	0.0182	0.0128	0.0728	0.0280	0.0232

Similarly, the average TPR value obtained using [8] as shown in Table 7 shows better results compared to ANN-YCbCr technique, however detailed analysis showed that the technique reported in [8] was unable to detect difference between skin-region pixel and similar looking skin-region pixel in an image, hence the reason for very high FPR and TPR values in Table FPR1 and Table 6 respectively.

Table 7: TPR of Skin-Pixel Detection Algorithms

	[1]	[4]	[3]	[8]	[9]
Min	0.8888	0.3332	0.8018	0.8777	0.8311
Max	1.0000	0.9993	1.0000	1.0000	0.9988
Mean	0.9795	0.8532	0.9649	0.9852	0.9618
SD	0.0321	0.1888	0.0388	0.0280	0.0508

4.0 Conclusion

In this work, the performance analysis of the technique proposed in [1] has been evaluated. The proposed technique have shown better performance when compared with other techniques reported in the literature. Furthermore, ANN-based YCbCr shows better performance in detecting likely skin-pixels as compared to other techniques evaluated in this paper .

References

[1] M.J.E Salami, A.M Aibinu, Safinaz Bt OAP Kader Mohideen and Siti Aisha Bt Mansor, ``Design Of An Intelligent Robotic Donationbox A Case Study" in Proceedings of the 4th IEEE International Conference on Mechatronics (ICOM'11), Kuala Lumpur, Malaysia, May 17-19, 2011.

[2] A. Cheddad, J. Condell, K. Curran, P. Kevitt,`` A new colour space for skin tone detection", *ICIP, IEEE* pp. 497-500, 2009.

[3] P. Peer, J. Kovac, and F. Solina, `` Human skin colour clustering for face detection" In *Proc. of EUROCON 2003 - International Conference on Computer as a Tool*, pp. 144-148, 2003.

[4] A. Cheddad, J.V. Condell, K.J. Curran, and P. McKevitt, ``A skin tone detection algorithm for an adaptive approach to steganography," *in Elsevier - Journal of Signal Processing*, pp. 2465-2478, 2009.

[5] V. Vezhnevets, V. Sazonov and A. Andreeva, ``A survey on pixel-based skin color detection techniques", *Graphicon-2003*, Moscow, Russia, 2003.

[6] P. Yogarajah, J. Condell, K. Curran, and P. Mc Kevitt, `` A Dynamic threshold approach for Skin Segmentation in Color Images" *Proceedings of 2010 IEEE 17th International Conference on Image Processing* September 26-29, Hong Kong, 2010.

[7] Qing-Fang Zheng, Ming-Ji Zhang, and Wei-Qiang Wang, ``A hybrid approach to detect adult web images,", in *PCM* (2), pp. 609–616, 2004.

[8] D. Chai and K. N. Ngan, ``Face segmentation using skin-color map in videophone applications", *IEEE Trans. Circuits Syst. Video Technol*. Vol 9, pp. 551–564 1999.

[9] Y. T. Pai, S. J. Ruan, M. C. Shie and Y.C. Liu, ``A Simple and Accurate Color Face Detection Algorithm in Complex Background", *IEEE 2006 Int. Conf. on Multimedia and Expo (ICME 2006)*, pp.1545-1548, 2006.

[10] Y. Feng, J. Li, L. Huang, C. Liu, ``Real-time ROI Acquisition for Unsupervised and Touch-less Palmprint", *World Academy of science, Engineering and Technology* Vol. 78. pp 823-827, 2011.

[11] J. Terrillon, M. N. Shirazi, H. Fukamachi, S. Akamatsu, `` Comparative Performance of Different Skin Chrominance Models and Chrominance Spaces for the Automatic Detection of Human Faces in Color Images, Automatic Face and Gesture Recognition", *IEEE*. pp 54-61, 2000.

[12] J. Yun, H. Lee, A. K. Paul,J. Baek, ``Robust face detection for video summary using illumination-compensation and morphological processing", *International Conference on Natural Computing*, pp. 710-714, 2000.

[13] C. Poon, D. Wong, H. Shen, `` A new method in locating and segmenting palmprint into region of interest", *International Conference on Pattern Recognition, IEEE* , Vol (4). pp 533-536, 2004.

[14] M. O. Rotinwa-Akinbile, A.M. Aibinu and M.J.E. Salami, ``Palmprint Extraction Using Model Based Approach", In Proc. of First International Conference on Informatics and Computational Intelligence, Indonesia, December, 2011.

[15] Y. Wang and B. Yuan, ``A novel approach for human face detection from color images under complex background," in *Pattern Recognition*, pp. 1983–1992, 2001.

[16] Rafael C. Gonzalez and Richard E. Woods. *'Digital Image Processing using MATLAB'*, 2nd edition. Prentice Hall, 2002. ISBN 0-201-18075-8.

[17] A. M. Aibinu, M. J. E. Salami, and A. A. Shafie "Artificial Neural Network Based Autoregressive (AR) modeling Technique with Application in Voice Activity Detection", Elsevier, Engineering Applications of Artificial Intelligence, 2012.