Implementation of Facial Expression Detection as a CAPTCHA

[1]Anvesh Sinha, [2]Dr. Sandhya Tarar
[1]Dept. of Computer Engineering, Gautam Buddha University, Greater Noida, India
[2]Gautam Buddha University, Greater Noida, India

Abstract
CAPTCHA system protects against automated scripts (or bots) by generating tests that humans can pass but computer programs cannot. Recent advancements in optical character recognition (OCR) techniques have led to successful attacks on text-based CAPTCHAs. To counter it, the design of these CAPTCHAs has become more complex, making it more challenging for humans to solve them successfully. Hence a new CAPTCHA system is proposed which works on the principle of facial expression detection. It exploits a real-time interaction from the user through front camera of laptop to get an input image. If the desired expression matches with input expression of the user, the CAPCTHA system gets unlocked. This makes the system more robust to bot attacks, also creating the user experience better due to its click-only interface and language-independency as no text is involved.

Keywords
CAPTCHA, bots, OCR, robust, facial expression detection, front camera

I. Introduction

CAPTCHA is an abbreviation for "Completely Automated Public Turing Test to Tell Computers and Humans Apart" [1]. It was invented in 2000 by Luis Von Ahn, Nicholas Hooper, Manuel Blum and John Langord of Carnegie Mellon University. It is defined as a Human Interaction Proofs (HIP) system which aims to differentiate a human and a computer program automatically. It acts as a reverse Turing test wherein the CAPTCHA plays the role of a judge and user (which can be a human or machine) behaves like a participant. If the problem, given by the CAPTCHA system, is solved by the interacting user successfully, then the user is recognized as a human or else labelled as a machine.

The main objective of CAPTCHA is to act as a defensive system for security against automated scripts, called as bots. These are computer generated - artificially intelligent programs which acts like human deceiver to solve the CAPTCHA and then perform malicious activities like misusing free e-mail services, spamming in blogs via comments, creating worms and spams in e-mail, carrying dictionary attacks on password systems, etc. Hence the need of powerful, yet simple CAPTCHA systems exists.

As discussed by Sinha & Tarar [2], there are numerous CAPTCHA systems which are put to use for protection against the automated scripts. Some of them are image-based, text-based, video-based and audio-based CAPTCHAs. Their reliability against the attacks are thoroughly studied and analyzed. Due to high success rate of attacks against these already existing systems, need of a better system- both in terms of security as well as simplicity was felt.

The new CAPTCHA system proposed in this paper works on the principle of facial expression detection wherein the interacting user is required to select a particular expression and then produce the same to unlock CAPTCHA system. This act shall enable him/her to get access to required destination. This system exploits the need of a real-time human interaction required for differentiating human from computer program. Other major aspect of this system is to provide fast and simple interaction to the user for solving CAPTCHA.

Face plays an important role in social intercourse, conveying both identity and emotion. According to Luttin and Fasel, facial expression is defined as temporally deformed facial features such as eye brows, nose, eye lids, mouth, lips and skin texture generated by contractions of facial muscles [3]. In human-computer interactions study, the term "facial expression detection" means classifying facial features in six basic emotions, namely sadness, happiness, neutral or normal, surprise, disgust and anger. Although ability of humans to recognize and gather information from a face and the facial expression is remarkable, but for an intelligent computer system it is a suspect. A number of related research work exists in the area of facial expression detection, but higher accuracy remains problematic due to variety and complexity of faces and their particular expression.

In this paper, Daubechies wavelet transform is used to extract the features from a face for detecting facial expressions. It is performed by calculating value of approximate and detailed coefficients of decomposed image. These set of values are then used to train a 'normal' Naïve Bayes Classifier to predict the facial expression of any new user.

Rest of this paper is organized as follows: in the following section, related work in CAPTCHA and facial expression detection is documented and the studies are discussed. In section III, detailed explanation of the system is provided wherein the methodologies used are discussed and the database on which the system works is provided. Section IV, deals with the experiments conducted and a comprehensive analysis of generated results. Finally, concluding remarks and future direction is explored in section V.

II. Related Work

As reviewed and discussed by Sinha and Tarar [2], CAPTCHA techniques used presently are vulnerable to computer generated programs and also have issue of being less-user friendly. The high success rate of attacks on the CAPTCHA systems is a major source of concern. Hence a new CAPTCHA system on the line of facial expression detection was thought of and implemented to tackle the security problems as well as to ease the interface of CAPTCHA system and hence, its usage.

A number of papers exist on automatic facial expression detection. Extensive survey in the field of facial expression analysis is done in overview papers of Rothkrantz, et al. [4-5], by Sebe, et al. [6] and by Cowie, et al. [7]. The work on recognizing emotion form facial expression was started from a psychological perspective pioneered by Ekman in his extensive studies [8]. Studies related to computer – assisted detection of facial expression started from early 1990's. Lantis, et al. [9], used appearance model and flexible shape architecture to identify a person's gender and facial expressions. Yacoob, et al. [10] tracked non – rigid facial gesture, using local parameterized prototypes of image motion, and provided it to a rule based classifier of face expression. Huang, et al. [11], introduced Hidden Markov Model (HMM) architecture

for automatic segmentation and recognition of facial expression through video sequences. Cohen, et al. [12], proposed the use of Bayesian network classifier for training and classification along with a multi-level HMM for segmenting arbitrary long sequences automatically to their corresponding facial expression.

III. System Overview

The proposed system for CAPTCHA implementing facial expression detection works on three basic emotions, which are neutral (or normal), happy and sad. The system provides an option to the user to select any expression of his liking. On selecting an expression, the user has to mimic the same expression in-front of laptop's front camera. If the acquired picture's expression matches with the data, associated with that expression, the CAPTCHA will get unlocked, or else user needs to again repeat the process. The architecture of the system is shown below in fig. 1.

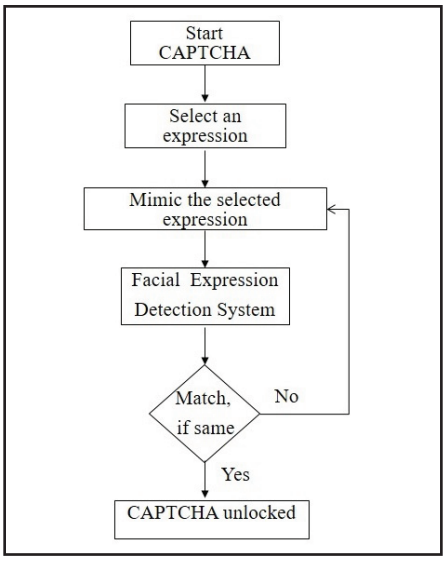

Fig. 1: Architecture of the Proposed System

The facial expression detection system used in the proposed CAPTCHA system uses three major techniques, namely Viola Jones algorithm, Daubechies wavelet transformation and Naïve Bayes classifier as shown in fig. 2.

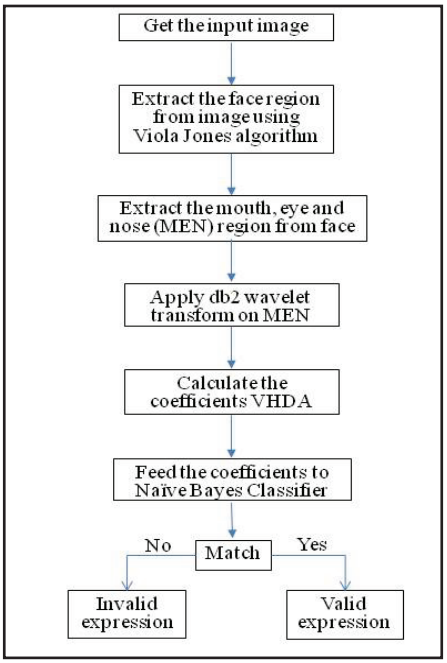

Fig. 2: Facial Expression Detection System

Due to the high detection rate and fast implementation of the Viola Jones algorithm [13], it has been used in the system as the first step. When an input image is captured, to remove the redundant area, Viola Jones algorithm is used to detect and extract the face region from input image. After cropping the face, three regions of interests, namely – mouth (M), eye (E) and nose (N) region, are then further detected and extracted using the same algorithm.

On getting the mouth, eye and nose of input image, Daubechies discrete wavelet transform decomposition is applied on each region of interest individually. It helps us to decompose a multi-dimensional image into a variety of channels which represent the features of the image in form of different frequency bands. This removes the noise present in images, which works as a plus point for our system. After performing decomposition, the approximation and detailed components of the image are calculated in form of coefficients. These are a set of 4 different coefficients namely, A (approximate coefficient), V (vertical coefficient), H (horizontal coefficient) and D (diagonal coefficient). The coefficients V, H and D are sub-parts of the detailed coefficients. The Daubechies wavelet decomposed up to level one with 2 vanishing moments (i.e., db2) and is used here because of its continuous and real nature and for its root-mean-square (RMS) error compared to other wavelets [14].

Getting the VHDA coefficients of mouth, eye and nose regions of an input image, the data values are then fed to a Naïve Bayes classifier to classify the image into one of the three already existing classes (neutral, happy and sad) which have been created by training the classifier with training images. Naïve Bayes classifier used here, scales well to high-dimensional problems and is fast to compute [15]. Also because of variant nature of input images, use of Bayesian probability theory based classifier helps to classify images with good accuracy.

IV. Experiments and Results

The CAPTCHA system designed in this paper uses poor resolution camera which is required to check the authenticity of a human interacting with the system to allow his/her access. It is thought of using the image database of human's expression from poor resolution camera rather than using better quality images from net database. Hence, 39 images of different individuals with three different facial expressions – neutral, happy and sad were captured using laptop's front camera to create the database for proposed system. Some of images are shown in fig. 3.

Fig. 3: Images Used to Create the Database of Neutral, Happy and Sad Face

In this system, 20 images per expression are used to train Naïve Bayes classifier. Rest 14 images are used for testing the system.

Training and testing accuracy is well tabulated in Table 1.

Table 1: Training and Testing Accuracy for Neutral, Happy and Sad Expression

	Training accuracy (in %)	Testing accuracy (in %)
neutral	75	57
happy	70	64
sad	90	79

To understand the mismatching of required expression and expression given by the user, a confusion matrix is generated as shown in Table 2.

Table 2(a): Confusion Matrix for Training Images [20 Images]

	neutral	Happy	Sad
neutral	15	1	4
happy	5	14	1
sad	1	1	18

Table 2(b): Confusion matrix for testing images [14 images]

	neutral	happy	sad
neutral	8	1	5
happy	3	9	2
sad	2	1	11

With decent detection percentage accuracy for all the three gestures, this new CAPTCHA system is proposed. However, due to variability in facial gesture, illumination, distance from camera and quality of image these variations are encountered and have to be further looked upon.

V. Conclusion

In this paper, a new CAPTCHA system is presented which works on principal of facial expression detection. User is required to select and subsequently impersonate the selected expression to the system using front camera of laptop. Using techniques like face detection, Daubechies wavelet transform and Naïve Bayes classifier the system decides whether the expression is valid or not. This system requires a more real-time human interaction which helps it to distinguish a human being from a computer generated program. Also with ease of using this system as a CAPTCHA it is a better alternative for those CAPTCHAs which are not easy to comprehend. In future, a more robust but fast facial expression detection approach can be developed and used to further improve the proposed CAPTCHA system.

References

[1] L. Ahn, M. Blum, J. Langford,"Telling Humans and Computers Apart Automatically", Communications of the ACM, 47(2), pp. 57–60, 2004.

[2] A. Sinha, S. Tarar,"Review Paper on Different CAPTCHA Techniques", International Journal of Computer Science and Technology, Vol. 7, Issue 1, Jan – March 2016

[3] B. Fasel, J. Luettin,"Automatic facial expression analysis: A survey,Pattern Recognition", Vol. 36(1), pp. 259–275, 2003.

[4] Pantic, M., Rothkrantz, L.,"Automatic Analysis of Facial Expressions: the State of the Art", IEEE Trans. On Pattern Analysis and Machine Intelligence, Vol. 22(12), pp. 1424-1445, 2000.

[5] Pantic M., Rothkrantz, L.J.M.,"Toward an affect-sensitive multimodal human-computer interaction", Proceedings of the IEEE, Vol. 91, No. 9, pp. 1370- 1390, Sept. 2003.

[6] Sebe, N., Cohen, I., Gevers, T., Huang, T.S.,"Multimodal Approaches for Emotion Recognition: A Survey", In Proceedings of SPIE-IS & T Electronic Imaging, SPIE, Vol. 5670: pp. 56-67, 2005.

[7] Cowie, R., Douglas-Cowie, E., Tsapatsoulis, N., Votsis, G., Kollias, S., Fellenz, W., Taylor, J.G.,"Emotion Recognition in Human-Computer Interaction", IEEE Signal Processing Magazine, pp. 32-80, January 2001.

[8] P. Ekman, W. Friesen,"Facial Action Coding System (FACS): Investigator's Guide". Conculting Psychol- ogists Press, 1978.

[9] Lanitis, A., Taylor, C., Cootes, T.,"A Unified Approach to Coding and Interpreting Face Images", In Proc. International Conf. on Computer Vision, pp. 368-373, 1995.

[10] Black, M., Yacoob, Y.,"Tracking and Recognizing Rigid and Non-rigid Facial Motions Using Local Parametric Models of Image Motion", In Proc. Int. Conf. on Computer Vision, pp. 374-381, 1995.

[11] I. Cohen, A. Garg, T. S. Huang,"Emotion Recognition from Facial Expressions using Multilevel HMM", Proc. of the 2000 NIPS Workshop on Affective Computing, Denver, Colorado, USA, December 2000.

[12] Cohen, L., Sebe, N., Garg, A., Chen, L., Huang, T., "Facial expression recognition from video sequences: Temporal and static modeling", Computer Vision and Image Understanding, Vol. 91(1-2), pp. 160-187, 2003.

[13] P.Viola, M. Jones,"Rapid Object Detection using a Boosted Cascade of Simple Features", Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 1, pp. 511-518, 2001.

[14] V. Saravanyah, S. Saranya, R. Balamurugan,"Facial Expression Recognition using Daubechies Wavelet PCA for Authentication", International Journal of Engineering Research & Technology (IJERT), Vol. 3 Issue 3, March – 2014.

[15] H. Jasso, P. Shin, T. Fountain, D. Pennington,"Using wavelets for the classification of hyperspectral images," In Fourth European Conference on Ecological Modelling and Fourth International Workshop on Environmental Applications of Machine Learning, 2004.

Anvesh Sinha is presently a final year student (2016 batch) of Integrated Dual Degree B.Tech and M.Tech from Gautam Buddha University, Greater Noida in Computer Engineering, with a specialization in Software Engineering.

ISSN : 0976-8491 (Online) | ISSN : 2229-4333 (Print)

Dr. Sandhya Tarar received her B.Tech in Computer Science and Engineering from U.P.T. University, in 2005, M.Tech in Computer Science from Aligarh Muslim University and Ph.D. in Biometrics from Gautam Buddha University, in 2014. Her research interest is in Pattern Recognition. She has a teaching experience of more than 10 years and have a number of research paper publications in both national and international journals. Presently, she is a Research Associate in Gautam Buddha University, Greater Noida.