Research Article Impact Factor: 0.621 ISSN: 2319-507X
Vrushali Katote, IJPRET, 2014; Volume 2 (9): 290-298 IJPRET

INTERNATIONAL JOURNAL OF PURE AND APPLIED RESEARCH IN ENGINEERING AND TECHNOLOGY

A PATH FOR HORIZING YOUR INNOVATIVE WORK

SMILE DETECTION WITH IMPROVED MISDETECTION RATE AND REDUCED FALSE ALARM RATE

VRUSHALI KATOTE[1], PROF. LOKESH BIJOLE[2]

1. Computer Engineering department, ME student, Padmashri Dr. V. B. Kolte College of Engineering, Malkapur, Amravati University.
2. Computer Engineering department, Faculty of Engineering, Padmashri Dr. V. B. Kolte College of Engineering, Malkapur, Amravati University

Accepted Date: 27/02/2014 ; Published Date: 01/05/2014

Abstract: Smile detection in face images captured in unconstrained real world scenarios is an interesting problem with many potential applications. There are some of efficient approaches for accurate smile detection which are different from the traditional user interface, such as keyboard and mouse. In first approach the intensity differences between pixels in the grayscale face images are used as features. The method adopted is Ada Boost to choose and combine weak classifiers based on intensity differences to form a strong classifier. In the second approach, a real-time, accurate, and robust smile detection system is proposed which is compared with the smile shutter function of Sony DSC T300. In another approach a smile detector for the real world applications is developed. To develop a smile detector all the required characteristics of the training dataset, image registration, image representation, and machine learning algorithms are explored. Techniques from the psychophysics literature are presented for detailed diagnosis and refinement of the obtained smile detector.

Keywords: Histogram Equalization, Single Scale Retinex, Discrete Cosine Transform, LBP, Tan Trigs, Misdetection Rate, False Alarm Rate, Threshold.

Corresponding Author: MS. VRUSHALI KATOTE

Access Online On:

www.ijpret.com

How to Cite This Article:

Vrushali Katote, IJPRET, 2014; Volume 2 (9): 290-298

PAPER-QR CODE

INTRODUCTION

A smile is the most common facial expression that occurs in people daily life. It often indicates pleasure, happiness, appreciation, or satisfaction. Smile detection has many applications in practice, such as interactive systems (e.g., gaming), product rating, distance learning systems, video conferencing, and patient monitoring.

Here we explores whether current machine learning methods can be used to develop an expression recognition system that operates reliably in more realistic conditions. The necessary characteristics of the training data set, image registration, feature representation, and machine learning algorithms are also explored in the approaches which are discussed here. These approaches are used for the accurate smile detection which is different from the traditional user interface approaches which use the keyboard and mouse [1].

LITERATURE SURVEY

There is an amount of literature which focused specifically on smile detection. Recently, Whitehill et al. presented a comprehensive study on practical smile detection. They collected the GENKI database consisting of 63 000 real-life face images from the Web. They investigated different parameters, including size and type of data sets, image regist1ration, facial representation, and machine learning algorithms. Their study suggests that high detection accuracy is achievable in real-life situations. They argued that the order of 1000–10 000 images that have a wide range of imaging conditions and personal variables is required for training. There is also a Sensing component company Omron has recently released smile measurement software. It can automatically detect and identify faces of one or more people and assign each smile a factor from 0% to 100%. Omron uses 3D face mapping technology and claim its detection rate is more than 90%. But due to its less availability rate it was not used on a large scale. In 2007, Sony has released its first consumer camera Cyber-shot DSC T200 with smile shutter function. The smile shutter function can detect at most three human faces in the scene and automatically takes a photograph if smile is detected.

Table 1 Experimental Result of Different Illumination Normalization Methods on Gabor

Illumination Normalization	Gabor (%)
HE	89.68 +- 0.62
SSR	86.50+-1.35
DCT	87.05+-0.96
LBP	88.70+-0.57
Tan-Triggs	85.75+-0.47

SYSTEM WORKING

Figure 1 some images used for smile detection

The experiment on the publicly available database where there are number of images out of which some are of smiling faces and some non smiling faces. For the experiment all the images are converted to a grayscale. After that using Gabor filters which gives high accuracy for the grayscale image dimensions is used. As shown in the above figure it is seen that there is varying illumination which is one of the difficulties for smile detection in real-life faces. To normalized the illumination some of the illumination normalization is used some of that methods are:

1) Histogram equalization (HE): HE is a simple and widely used technique for normalizing illumination effects.

2) Single-scale retinex (SSR): SSR is a basic photometric normalization technique, which smoothes the image with a Gaussian filter to estimate the luminance function.

3) Discrete cosine transform (DCT): the DCT-based normalization sets a number of DCT coefficients corresponding to low frequencies as zero to achieve illumination invariance.

4) LBP: one important property of LBP operators is their tolerance against monotonic illumination changes; therefore, LBP can be used as a preprocessing filter to remove illumination effects.

5)Tan–Triggs: Tan and Triggs proposed a series of steps to counter the effects of illumination variation, local shadowing, and highlights while still preserving the essential elements of visual appearance.

The results obtained are shown in Table 1 Surprisingly, for Gabor features only the HE technique achieves improved accuracy, whereas all other methods fail to improve. Hence the HE method is applied to the grayscale face image and then smile detection with the baseline approaches is performed. After extracting intensity difference features from face images preprocessed by HE, we run AdaBoost to choose the discriminative features and combine the selected weak classifiers as a strong classifier. AdaBoost provides a simple yet effective approach for stage wise learning of a nonlinear classification function. It combines the feature selection and classifier training steps in one process.

SMILING DETECTION SCHEME

There is a proposed fast and generally low misdetection and low false alarm video-based method of smile detector. It have 11.5% smile misdetection rate and 12.04% false alarm rate on the FGNET database. The smile detect algorithm is as follows:

1. Detect the first human face in the first image frame and locate the twenty standard facial features position.

2. In every image frame, use optical flow to track the position of left mouth corner and right mouth corner with accuracy of 0.01 pixels and update the standard facial feature position by face tracking and detection.

3. If x direction distance between the tracked left mouth corner and right mouth corner is larger than the standard distance plus a threshold T_{smile}, then we claim a smile detected.

4. Repeat from Step 2 to Step 3.

Table 2 Misdetection rate and false alarm rate with different thresholds.

Threshold	Misdetection Rate	False Alarm Rate
0.4**Dstd*	6.66%	19.73%
0.5**Dstd*	9.25%	14.04%
0.55**Dstd*	11.50%	12.04%
0.6**Dstd*	13.01%	8.71%
0.7**Dstd*	18.82%	4.24%
0.8**Dstd*	25.71%	2.30%

It is important to note that the feature tracking will accumulate errors as time goes by and that would lead to misdetection or false alarm results.. Here a method to automatically refine for real-time usage is proposed. Below the algorithm is shown. For each following image, the use of normalized cross correlation (NCC) is done for block matching method to calculate the best matching block to the pattern image around the new mouth region and calculate their cross correlation value. The NCC equation is:

$$C = \frac{\sum_{(x,y)\in R,(u,v)\in R'} (f(x,y)-\overline{f})(g(u,v)-\overline{g})}{\sqrt{\sum_{(x,y)\in R} (f(x,y)-\overline{f})^2}\sqrt{\sum_{(u,v)\in R'} (g(u,v)-\overline{g})^2}}$$

The equation shows the cross correlation between two blocks R and R'. If the correlation value is larger than some threshold, it means the mouth state is very close to the neutral one rather than an open mouse, a smile mouth or other state. Then we would relocate feature positions. To not take too much computation time on finding match block,

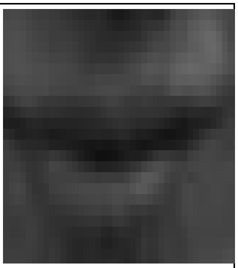
Cross correlation value 0.925

We set the search region center by initial position. To overcome the non-sub pixel block matching, we set the search range to a three by three block and find the largest correlation value as our results.

As mentioned, there is a need to know the threshold value to do the refinement. We have a real-time case in to show the correlation value changes with smile expression and off-line case on FGNET face database to decide the proper threshold Fig 4 shows a sequence of images and their correlation value corresponding to the initial mouth pattern. These images give us some level of confidence that using correlation to identify the neutral or smile expression is possible.

To show stronger evidence, we run a real-time case by doing seven smile activities with 244 frames and record their correlation value. Table 3. shows the image index and their correlation values. If we set 0.7 as our threshold, we would have mean correlation value 0.868 and standard deviation 0.0563 for neutral face and mean value 0.570 and standard deviation 0.0676 for smile face.

Cross correlation value 0.767



 Cross correlation value 0.502	

The difference of mean value 0.298 = 0.868-0.570 is greater than two times sum of standard deviation 0.2478 = 2 x (0.0563+0.0676). To have more persuasive evidence, we run on FGNET face database in the following.

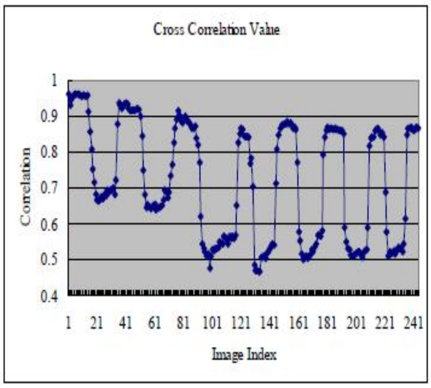


TABLE 3 Cross Correlation value of mouth pattern with seven smile activities

Above we have clear evidence that neutral expression and smile expression have a great difference on correlation value. We obtain more convincing threshold value by **Figure 5 Cross correlation value of mouth pattern for smile activity.**

Running cross correlation value's mean and standard deviation on FGNET face database. There are eighteen people, who have three sets of image sequences for each. Each set has 101 images or 151 images and roughly half of them are neutral face and others are smile face. We drop some false performing datasets. By setting threshold value 0.7, we have neutral face mean of mean and standard deviation correlation value 0.956 and 0.040. At the same time, smile face values are 0.558 and 0.097. It is not surprising that smile face has higher variance then neutral face since different user has different smile type. We set three standard deviation distances 0.12 = 3*0.04 as our threshold. If correlation value is beyond the original value minus 0.12, we can refine user's feature position automatically and correctly.

Table 3 illustrates the detection result with Sony T300 of Person 1 in FGNET face database. Figure 6 and Figure 7 show the total detection and false alarm rate results of the fifty video sequences in FGNET. We have a normalized detection rate 88.5% and false alarm rate 12% while Sony T300 has a normalized detection rate 72.7% and false alarm rate 0.5%.

Figure 6 Comparison of detection rate.

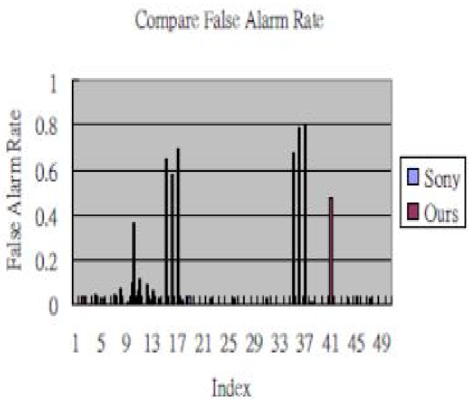


Figure 7 Comparison of false alarm rate.

CONCLUSION

Smile detection in face images captured in real-world scenarios is an interesting problem with many applications. In practice, because of the limited computational resource, it is desired that the features used can be computed easily and efficiently. In this paper, the intensity differences between pixels in the grayscale face images are used as simple features.

FUTURE WORK

In this paper we focused on detecting smiles in poses within approximately ±20◦ from frontal. Developing expression recognition systems that are robust to pose variations will be an important challenge for the near future. Another important future challenge will be to develop comprehensive expression recognition systems capable of decoding the entire gamut of human facial expressions, not just smiles. One promising approach that we and others have been pursuing [5], [7], [13] is automating the Facial Action Coding System. This framework allows coding all possible facial expressions as combinations of 53 elementary expressions (Action Units). Our experience developing a smile detector suggests that robust automation of the Facial Action Coding system may require on the order of 1,000 to 10,000 examples images per target Action Unit. Datasets of this size are likely to be needed to capture the variability in illumination and personal characteristics likely to be encountered in practical applications.

REFERENCES

1. Smile Detection by Boosting Pixel Differences Caifeng Shan, Member, Ieee Transactions On Image Processing, Vol. 21, No. 1, January 2012

2. Face Detection and Smile Detection 1Yu-Hao Huang, Dept. of Computer Science and Information Engineering, National Taiwan University E-mail:r94013@csie.ntu.edu.tw

3. Chiou-Shann Fuh, Dept. of Computer Science and Information Engineering, National Taiwan University E-mail:fuh@csie.ntu.edu.tw

4. Toward Practical Smile Detection Jacob Whitehill, Gwen Littlewort, Ian Fasel, Marian Bartlett, Member, IEEE, and Javier Movellan Computer Vision and Image Understanding, vol. 98, 2005.

5. Towards Practical Smile Detection Jacob Whitehill, Gwen Littlewort, Ian Fasel, Marian Bartlett, and Javier Movellan, Computer Vision and Image Understanding, vol. 98, 2005.

6. Developing a Practical Smile Detector Jacob Whitehill, Gwen Littlewort, Ian Fasel, Marian Bartlett, and Javier Movellan, M. Pantic and L. Rothkrantz, "Automatic analysis of facial expressions: The state of art," IEEE Trans. Pattern Anal. Mach. Intell., vol. 22, no. 12, pp. 1424–1445, Dec. 2000.

7. B. Fasel and J. Luettin, "Automatic facial expression analysis: A Survey," Pattern Recognit, vol. 36, no. 1, pp. 259–275, Jan. 2003.

8. Z. Zeng, M. Pantic, G. Roisman, and T. Huang, "A survey of affect Recognition methods: Audio, visual, and spontaneous expressions," IEEE Trans. Pattern Anal. Mach. Intell., vol. 31, no. 1, pp. 39–58, Jan. 2009.

9. T. Kanade, J. Cohn, and Y. Tian, "Comprehensive database for facial expression analysis," in Proc. IEEE Int. Conf. Autom. Face Gesture Recog., 2000, pp. 46–53.

10. J. F. Cohn and K. L. Schmidt, "The timing of facial motion in posed And spontaneous smiles," Int. J.Wavelets, Multiresolution Inf. Process., vol. 2, pp. 1–12, 2004.

11. Z. Zeng, Y. Fu, G. I. Roisman, Z.Wen, Y. Hu, and T. S. Huang, "Spontaneous emotional facial expression detection," J. Multimedia, vol. 1, no. 5, pp. 1–8, Aug. 2006.

12. P. Ekman, W.V. Friesen, and P. Ellsworth, Emotion in the human face: Guidelines for research and an integration of findings, Pergamon Press, New York, 1972.

13. P. Ekman and W.V. Friesen, Unmasking the face. A guide to recognizing emotions from facial clues, Prentice-Hall, Englewood Cliffs, New Jersey, 1975.

14. P. Ekman, Emotion in the Human Face, Cambridge University Press, New York, 2 edition, 1982.

15. P. Ekman, "Facial expression of emotion," Psychologist, vol. 48, pp. 384–392, 1993.

16. P. Ekman, Telling Lies: Clues to Deceit in the Marketplace, Politics, and Marriage, W.W. Norton, New York, 3 editions, 2001.

17. L. Parr and B. M. Waller, "Understanding chimpancee facial expression: insights into the evolution of communication," Social Cognitive and Affective Neuroscience, vol. 1, no. 3, pp. 221–228, 2006.