A Study of Non-frontal-view Facial Expressions Recognition

[1]Yuxiao Hu, [1]Zhihong Zeng, [2]Lijun Yin, [2]Xiaozhou Wei, [1]Jilin Tu and [1]Thomas S. Huang
[1]University of Illinois at Urbana-Champaign, [2]State University of New York at Binghamton
{hu3, zhzeng, jilintu, huang}@ifp.uiuc.edu, {lijun, xwei}@cs.binghamton.edu

Abstract

The existing methods of facial expression recognition are typically based on the near-frontal face data. The analysis of non-frontal-view facial expression is a largely unexplored research. The accessibility to a recent 3D facial expression database (BU-3DFE database) motivates us to explore an interesting question: whether non-frontal-view facial expression analysis can achieve the same as or better performance than the existing frontal-view facial expression method. Our extensive recognition experiments on data of 100 subjects with 5 yaw rotation view angles suggests that the non-frontal-view facial expression classification can outperform frontal-view facial expression recognition, given the manually labeled facial key points.

1. Introduction

Change in a speaker's affective state is a fundamental component in human communication [1]. It is expected that the next generation of human computer interaction should be proactive and human-centered, and have the ability to detect the change of the user's states, especially affective states, and initiate communications based on this information, rather than simply responding to user commands. A large number of studies in psychology confirm the correlation between some affective states (especially prototypical emotions such as happy, sad, anger, disgust, surprise, and fear) and facial expressions [9]. Due to the practical importance and the theoretical interest of facial expressions, automatic facial expression analysis has attracted the interests of many researchers in computer science, psychology, psychiatry, neuroscience, etc.

The potential applications of automatic facial expression recognition include affect-sensitive systems in computer-aided learning environment, customer service, intelligent driver assistance, and entertainment industry. These applications will change the vision of human-computer interaction in our daily life.

A recent survey [12] indicates a trend in machine understanding of facial expressions, which is moving from analysis of deliberately displayed affective expressions to analysis of spontaneous affective expressions. That is because increasing number of researchers have noticed the difference between posed affect expression and spontaneous affect expression. The existing algorithms based on posed affect expression are not expected to readily perform on real-life affect expression. This fact motivates researchers in this field to investigate the issues which have been largely unexplored. The recognition of non-frontal-view facial expression is one of these issues.

The public available and frequently used databases typically captured the frontal-view facial expressions [5], which result in the fact that most of existing studies of facial expressions can only handle frontal-view face displays. Recently, a 3D database named BU-3DFE has been collected by Yin et al. [11] and has been public available in the research community. Access to this database motivates us to explore some open issues about human non-frontal-view facial expressions.

One issue that attracts us most is: whether non-frontal-view facial expression recognition can achieve the same as or better performance than the current frontal-view facial expression methods, in the other words, whether the frontal view is the best perspective for a computer to recognize facial expressions?

We in this paper present investigation of this issue, based on geometric salient facial points and various classification methods. Different from our intuition, the extensive experiments on the data of 100 subjects of the BU-3DFE database suggest that the non-frontal view could be a better perspective than frontal view in automated facial expression recognition. Specifically, we got the better performance at non-frontal views

978-1-4244-2175-6/08/$25.00 ©2008 IEEE

than at the frontal view in recognition of prototypical emotions.

2. Related work

Most of existing efforts in this field, including studies on both posed expressions and on spontaneous facial expressions [3][4][6][8], focus on recognition of facial expressions in near-frontal-view recordings. The human behavior in less constraint environment, e.g., non-frontal-view face displays, challenges these existing methods.

An exemplar exception is the study of Pantic and Patras [7], who explored automatic analysis of facial expressions from the profile view of the face. Recently, Yin et al. [11], Wang et al. [10] and Chang et al. [2] used 3D expression data for facial expression recognition. In particular, the study of [10] analyzed the influence of viewing angle change on recognition performance of facial expression, based on the classifier trained on frontal-view faces. As shown by the experimental results in [10], the classifier performed poorly to recognize facial expression undergoing large view variation.

According to our best knowledge, this paper is the first attempt to explore the issue: whether non-frontal-view facial expression recognition can achieve the same or better performance than the current frontal-view facial expression methods? The answer to this question can provide insights toward the future research of non-frontal-view facial expression analysis.

3. BU-3DFE database

Having enough labeled data of human affective expressions is a prerequisite in designing automatic affect recognizer. To our best knowledge, The BU-3DFE database of Yin and colleagues [11] is only public available emotion database that contains 3D range data of six prototypical facial expressions.

In BU_3DFE 3D facial expression database, there are 100 subjects who participated in face scans, including undergraduates, graduates and faculty from State University of New York at Binghamton. The resulting database consists of about 60% female and 40% male subjects with a variety of ethnic/racial ancestries.

Each subject in the database performed seven expressions (including neutral), captured by a 3D face scanner. With the exception of the neutral expression, each of the six prototypic expressions (happiness, disgust, fear, angry, surprise and sadness) includes four levels of intensity.

4. Features and classifiers

In our study, we generate multi-view images of facial expressions from the available 3D data. The data in our experiment includes images at 5 yaw angles (0, 30, 45, 60, and 90 degrees), illustrated in Figure 1. In order to reduce the noise from automatic feature extraction, we used the geometric points around mouth, eyes and eyebrows, which are manually labeled, as the features in our study. Figure 2 illustrates these features points, represented by white points, at different yaw rotation angles. The neutral face of each subject is used as the reference for person-independent test. We first calculate the geometric 2D displacement of the facial feature points between emotional and neutral expressions of the person at the corresponding angle, and then, normalize these distances to zero mean and unit variance. The resulted vectors are used as the inputs of classification.

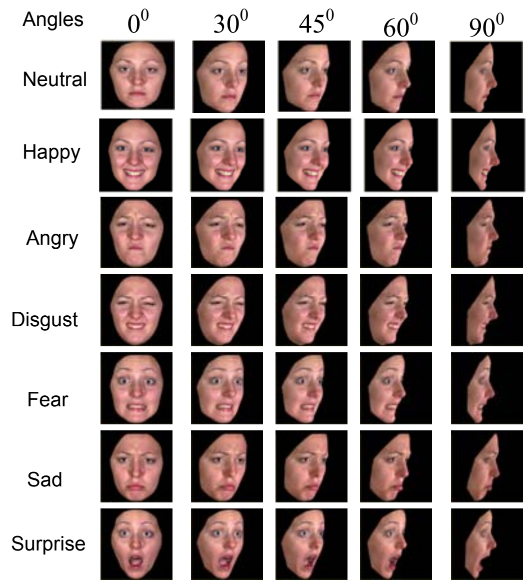
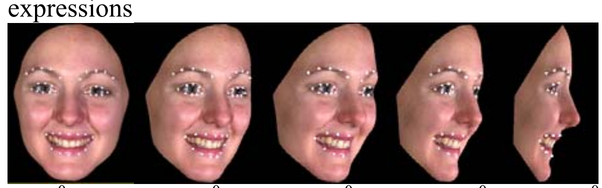


Figure 1. The examples of multi-view facial expressions



(a) 0^0 (b) 30^0 (c) 45^0 (d) 60^0 (e) 90^0

Figure 2. The white points represent manually labeled geometric facial features at different view angles.

We in this study apply various classifiers [13] including

ldc: Linear Bayes Normal Classifier,

qdc: Quadratic Bayes Normal Classifier,

parzen: Parzen classifier,

svm: Support Vector Classifier with linear kernel,

knnc: K-Nearest Neighbor Classifier,

For ***knnc***, we also investigate the influence of different dimension reduction methods including Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA) and Locality Preserving Projection (LPP) [14].

5. Experimental results

We apply five-fold cross validation to test person-independent emotion recognition. In details, we randomly divided 100 subjects of the BU-3DFE database into 5 groups without overlap. Each group includes 20 subjects. For the test, all the data in one group are used as the test data, and the data of the remaining groups are used as training samples. This experiment is repeated 5 times, each time using different group as the test data. The statistics of our experimental data are 100 subjects, 6 emotions with 4 intensity levels, 5 view angles (0^0, 30^0, 45^0, 60^0, or 90^0). So the total number of images is 12000.

Table 1 is the experimental results (error rates) of multi-view emotion recognition based on various classifiers. Each row in Table 1 shows the error rates of these classifiers with respect to yaw rotation angle (0, 30, 45, 60 and 90 degrees). The last row is the average performance of these classifiers in the multi-view emotion recognition experiment. We also applied the knnc classifier (k=5) on different dimension reduction algorithms, including original feature space, PCA, LDA and LPP.

In Table 1, the experiments show that the svm has the best performance with the average error rate of 0.335, which is highlighted by the bold black font with the underline style.

The rank of classifiers with respect to average recognition performance from best to worst is: *svm, knn(lda), knn(lpp), ldc, qdc, knnc(pca), pazen, knnc(orig).*

The lowest error rate of each classifier with respect to the view angle is highlighted by the bold black font. For five classifiers (i.e., ldc, svm, knnc(lpp), knnc(lda) and knnc(pca)), 45 degree in yaw rotation is the best view to recognize facial expressions. The pazen and knnc(orig) has the best recognition performance at 30 degree and the qdc has the best performance at 60 degree.

Table 2 presents the more details of multi-view emotion recognition based on the svm classifier (the best classifier in our experiment setup, which lists the error rates of each emotion with respect to view angle. The emotions with recognition performance from best to worst are listed as follows:

surprise> happy>sad>disgust>angry>fear

Among these six emotions, surprise is easiest to be recognized with the average error rate of 0.205, which is highlighted by the bold blue font with underline style. And fear is the most difficult to be recognized.

Table 1: multi-view emotion recognition performance (error rate) based on different classifiers.

	ldc	*qdc*	*pazen*	*svm*	*knnc (k=5)*			
					orig	*pca*	*lda*	*lpp*
0^0	0.349	0.513	0.522	0.330	0.533	0.512	0.351	0.337
30	0.321	0.511	**0.487**	0.303	**0.499**	0.482	0.313	0.308
45	**0.299**	0.495	0.494	**0.285**	0.499	**0.467**	**0.291**	**0.296**
60	0.347	**0.455**	0.493	0.330	0.520	0.486	0.340	0.343
90	0.453	0.511	0.599	0.428	0.633	0.614	0.445	0.463
ave	0.354	0.497	0.519	**<u>0.335</u>**	0.537	0.512	0.348	0.350

Table 2: error rate of different emotions based on svm

	angry	*disgust*	*fear*	*happy*	*sad*	*surprise*	*ave*
0	0.365	0.3125	0.485	0.268	0.38	0.1725	0.3304
30	0.33	0.315	0.488	0.225	0.31	0.1525	0.3033
45	**0.3075**	**0.2875**	**0.475**	**0.218**	0.285	**0.14**	**<u>0.2854</u>**
60	0.3725	0.355	0.538	0.23	**0.28**	0.2025	0.3296
90	0.3775	0.4425	0.58	0.453	0.355	0.3575	0.4275
ave	0.3505	0.3425	0.513	0.279	0.322	**<u>0.205</u>**	

Table 3: confusion matrix of svm at the frontal view

%		Recognized					
		angry	disgust	fear	happy	sad	surprise
Ground truth	angry	**63.5**	6.25	6.75	2.75	20.25	0.5
	disgust	7.75	**68.75**	7	9.75	2.75	4
	fear	6	9	**51.5**	11	10.75	11.75
	happy	3.25	7	11.75	**73.25**	3.5	1.25
	sad	23.5	2.25	9.25	2.75	**62**	0.25
	surprise	0.5	3.5	12.25	0.75	0.25	**82.75**

Table 4: confusion matrix of svm at the best non-frontal view (45 degree)

%		Recognized					
		angry	disgust	fear	happy	sad	surprise
Ground truth	angry	**69.25**	6.75	4.75	1.75	17	0.5
	disgust	7.5	**71.25**	8	5	4	4.25
	fear	5.25	9.75	**52.5**	17.5	7	8
	happy	2	3.5	15	**78.25**	1	0.25
	sad	18	2.5	7.5	0.5	**71.5**	0
	surprise	0.5	3.5	8.75	1.25	0	**86**

The lowest error rates of each emotion with respect to view angle are highlighted by the underline style in Table 2. All of emotions have the lowest error rate at non-frontal view. The best performance of recognition is at 45 degree for the emotions except 'sad' which has the best performance at 60 degree.

Table 3 and Table 4 are the confusion matrices at the frontal view and best non-frontal view (45 degree), based on the svm classifier. All of the emotions have improved performance at 45 degree, compared with at frontal view. Specifically, sad has considerable improvement from 62% at the frontal view to 71.5% at 45 degree. Fear and disgust have little improvement, 1% and 2.5% increase respectively from the frontal view and 45-degree view.

6. Discussion

Our experiment suggests that the frontal view is not the best perspective for a computer to recognize facial expressions. This phenomenon strikes us enormously because it conflicts with our intuition.

In realistic face-face interactive environment, humans always display frontal-view face to a companion. Non-frontal-view-face communication is always regarded as being not polite. This ritual habit actually causes non-uniform distribution of training samples for human learning process, i.e., the number of nearly-frontal-view face images is far larger than the number of non-frontal-view face images. This fact could result in human perception bias, i.e., humans seem more sensitive to subtle movement of frontal-view face and facial expressions than non-frontal-view ones.

Machine understanding of facial expressions is also dependent on the training data we provide. In our experiment, the classifiers for different angles were trained with the same number of training samples. That enables us to evaluate the influence of view angle on facial expression recognition on a fair platform.

Although the frontal-view face has most visible facial features than the non-frontal-view face, the displays of facial features of expressions are always symmetry at the frontal view. The facial symmetry results in the redundancy of information to some degree. However, an image of non-frontal-view face provides in some ways face depth information, which could contribute to the improvement of facial expression recognition. Specifically, 45 degree is a good view to get the depth information of face. The above explanation could be reasons behind our results.

7. Conclusion

The analysis of non-frontal-view facial expressions is a largely unexplored research. Building a facial expression system robust to non-frontal-view face is very important to advance the research of affective behavior in less constraint interactive environment.

In this paper, we focus on an issue about whether non-frontal-view facial expression recognition can achieve equal or better performance than frontal-view facial expression recognition. Our extensive experiments show that non-frontal view is better than the frontal view for a computer to recognize facial expressions, based on manually labeled facial point features. Actually automatic accurate extraction of these facial points at non-frontal face view is a big challenge which we are working on.

References

[1] Cohn, J.F. (2006), Foundations of Human Computing: Facial Expression and Emotion, Int. Conf. on Multimodal Interfaces (ICMI2006), 233-238

[2] Chang Y, Vieira M, Turk M, and Velho L (2005). Automatic 3D facial expression analysis in videos. Analysis and Modeling of Faces and Gestures, 3723, 293-307.

[3] Bartlett, M.S., Littlewort, G., Frank, M., Lainscsek, C., Fasel, I., and Movellan, J.(2005), Recognizing Facial Expression: Machine Learning and Application to Spontaneous Behavior, IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), 568-573

[4] Gu H, Ji Q (2004). An Automated Face Reader for Fatigue Detection. Int. Conf. Automatic Face and Gesture Recognition (FG2004). 111-116

[5] Kanade, T., Cohn, J., and Tian, Y. (2000), Comprehensive Database for Facial Expression Analysis, FG2000, 46-53

[6] Kapoor, A., Burleson, W., and Picard, R. W. (2007), Automatic prediction of frustration. Int. Journal of Human-Computer Studies. Vol. 65(8), 724-736.

[7] Pantic, M., and Patras, I. (2006). Dynamics of facial expression: recognition of facial actions and their temporal segments form face profile image sequences. IEEE Trans. Systems, Man and Cybernetics–Part B, Vol. 36, No.2, 433-449

[8] Sebe, N., Lew, M.S., Cohen, I., Sun, Y., Gevers, T., Huang, T.S.(2004), Authentic Facial Expression Analysis, FG2004

[9] Russell J.A., Bachorowski J. and Fernandez-Dols J. (2003). Facial and vocal expressions of emotion. Ann. Rev. Psychol. 54:329-349

[10] Wang, J., Yin, L., Wei, X., and Sun, Y. (2006). 3D Facial Expression Recognition Based on Primitive Surface Feature Distribution. CVPR2006, 1399-1406

[11] Yin, L., Wei, X., Sun, Y., Wang, J., Rosato, M. J. (2006). A 3D facial expression database for facial behavior research. FG2006, 211- 216

[12] Zeng, Z., Pantic, M., Roisman, G.I. and Huang, T.S. (2007). A Survey of Affect Recognition Methods: Audio, Visual, and Spontaneous Expressions. ICMI2007, 126-133.

[13] PRTools: http://www.prtools.org/

[14] He X.F., Yan S.C., Hu Y.X, Zhang H.J.(2003), Learning a Locality Preserving Subspace for Visual Recognition, Int. Conf. on Computer Vision(ICCV2003), 385-393