IOSR Journal of Electronics and Communication Engineering (IOSR-JECE)
e-ISSN: 2278-2834,p- ISSN: 2278-8735.Volume 11, Issue 1, Ver. IV (Jan. - Feb .2016), PP 14-19
www.iosrjournals.org

Accurate Face Recognition Using Stepwise Linear Discriminant Regression Classification and Hidden Conditional Random Fields (SWLDRC-HCRF)

T Syed Akheel[1], Dr. V.Usha shree[2] ,Dr.S.A.K Jilani[3]

[1]*Associate Professor, Dept. of Electronics and communication Engg, PVKK Institute of technology , Anantapur*

[2] *PhD, Professor, Dept. of Electronics and communication Engg JBIT Hyderabad,*

[3] *PhD, Professor, Dept. of Electronics and communication Engg MITS Madanapalle,*

Abstract*: This introduces an correct and potent face Recognition (FR) procedure. For function extraction, the proposed FR method employs stepwise linear discriminate analysis (SWLDA). SWLDA specializes in deciding on the localized elements from the expression frames utilising the partial F-test values, thereby lowering the inside type variance and growing the low between variance amongst distinct cognizance lessons. For attention, the hidden conditional random fields (HCRFs) model is utilized. HCRF is in a position of approximating a intricate distribution making use of a combo of Gaussian density functions. To acquire ideal outcome, the procedure employs a hierarchical awareness technique. Under during consciousness, on the first level, SWLDA and HCRF are employed to respect the faces category; whereas, on the 2d level, the label for the recognition inside the well-known category is set utilizing a separate set of SWLDA and HCRF, knowledgeable only for that class. So as to validate the process, four publicly on hand data sets have been used, and a complete of 4 experiments have been performed. The weighted common consciousness expense for the proposed FR approach was ninety six.37% throughout the four exclusive data units, which is a colossal development unlike the present FR ways.*

Index Terms*: Face recognition, Stepwise linear discriminant regression classification, and linear regression classification, hidden Markov models, hidden conditional random*

I. Introduction

Now a day's the identification to human machine interaction plays Associate in nursing important role in several applications like a industrial, Collages and public & personal corporations etc. identification has been receiving in depth attention over the past decade with increasing demands in automatic Face identification. Statistics is to spot people exploitation physiological or activity characteristics, like face, iris, retina, palm-print, fingerprint, etc. Among all the biometric techniques, Face recognition [1] is that the most well-liked technique and is with success employed in several applications. In previous several strategies are projected to attain booming face recognition in varied conditions. The well-known strategies embody principal element analysis (PCA) [2], freelance element analysis (ICA) [3], and linear discriminant analysis (LDA) [2], that decide to ask for a low-dimensional topological space for spatial property reduction to attain effective face recognition.

In 2013, Shih-Ming Huang and Jar-Ferr principle.. projected to additional embody the discriminant analysis thought into the regression classification, referred to as linear discriminant regression classification (LDRC), to spice up the lustiness of the LRC for pattern or face classification. Experimental results on FERET and AR face databases show the automated face recognition sometimes doesn't have enough coaching face pictures that build it a typical small-sample-size drawback [12]. The reconstruction error are often massive though the probe face image is delineate by the coaching face pictures of an equivalent category that makes the classification unstable.

In Xiaochao Qu, Suah Kim, Run Cui, Hyoung Joong Kim projected a completely particular linear cooperative discriminant regression classification (LCDRC) that makes use of the cooperative between-category reconstruction error (CBCRE) as an alternative than BCRE. The cooperative between-type illustration utilizes move-type coaching faces to symbolize a research face. The obtained CBCRE is smaller than every category-specific between-category reconstruction error (extra teaching faces symbolize the probe face higher). As a consequence, CBCRE are mainly viewed a sure of all of the classification-targeted between-type reconstruction mistakes and it can be for probably the most phase decided by means of these little classification-detailed between-type reconstruction errors so that the large type-exact between-type reconstruction error domination crisis are almost always eased.

In this propose during this analysis study, we tend to propose the employment of Stepwise Linear Discriminant Regression Classification for Face Recognition (SWLDRC). the aim of exploitation SWLDA as a feature extraction technique is to extract the localized options from faces that the previous feature extraction techniques were restricted in analyzing. As for the CBCRE, the present CBCRE models area unit restricted by their independence suppositions], which can scale back classification

In this research study, we propose the use of Stepwise Linear Discriminant Analysis (SWLDA) coupled with Hidden Conditional Random Fields (HCRF) for a sequence-based FR system named SH-FR. The block diagram of the SH-FR is shown in Fig. 1. Though SWLDA has been used in many different areas before [14], it is for the first time that it is being utilized as a feature extraction technique in an FR system. The purpose of using SWLDA as a feature extraction technique is to extract the localized features from faces that the previous feature extraction techniques were limited in analyzing. As for the HCRF, the existing HCRF models are limited by their independence suppositions [15], which may reduce classification accuracy. In this work, we have tried to overcome this limitation by approximating

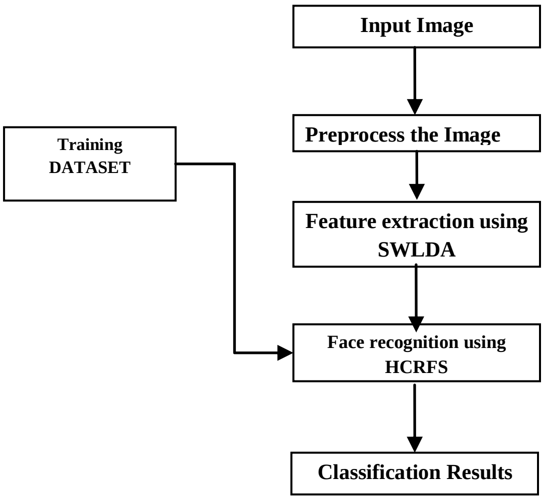

Figure: The block diagram of the SH-FR

II. Related Work

2.1 Linear Regression Classification (LRC)

In this the training face images of the ith class as $X_i \in R^{m\times n_i}$. Each column of X_i is an m dimensional face image of class i in which there are n_i training face images, and i =1, 2,. . . , c, where c is the total number of classes. All of the face images in this paper are assumed to be vectors by stacking the columns of original face images. Assume y is a probe face image that can be represented using X_i according to

$y = X_i\alpha_i,$ i=1,2,…..,c, (1)

Where $\alpha_i \in R^{m\times n_i}$ is the regression parameters; α_i can be calculated using the least-squares estimation as

$\widehat{\alpha_i} = (X_i^T X_i)^{-1} X_i^T y,\ i = 1,2, \dots .\, c,$ (2)

The reconstruction of y by each class can be obtained as

$\widehat{y_i} = X_i (X_i^T X_i)^{-1} X_i^T y = H_i y,\quad i = 1,2, \dots . .\, c,$ (3)

Where H_i is called a hat matrix that maps y into $\widehat{y_i}$. The reconstruction error of each class is calculated as

$e_i = \|y - \widehat{y_i}\|_2^2,\qquad i = 1,2, \dots \dots, c.$ (4)

LRC then assigns the y to the class that has the smallest reconstruction error.

2.2 Linear discriminant regression classification (LDRC)

Assume that all the training face images from the training matrix X =[x_1 . . . , x_i . . . x_n] $\in R^{m\times n_i}$, where n is the number of training face images and m is the dimensionality of each training face image. The class label of each x_i is denoted as l(x_i) $\in \{1,2, \dots c, \}$. Assume that the subspace rojection matrix is $U \in R^{m\times n_i}$. Each face image can be projected onto the subspace as

$y_{i=U^T x_i}$ (5)

The subspace projection matrix U is obtained by maximizing BCRE and minimizing WCRE simultaneously, where BSRE and WCRE are calculated as

$BCRE = \frac{1}{n(C-1)} \sum_{i=1\ j=1}^{n} \sum_{j\neq l(x_i)}^{c} \left\| y_i - \hat{y}_{ij}^{inter} \right\|_2^2$ (6)

$WCRE = \frac{1}{n} \sum_{i=1}^{n} \left\| y_i - \hat{y}_{ij}^{inter} \right\|_2^2$ (7)

Where $\hat{y}_{ij}^{inter}$ is the reconstruction of y_i by the jth class and $l(y_i) \neq j$. $\hat{y}_{ij}^{inter}$ is the reconstruction of y_i by the $l(y_i)$ class.

2.3 Linear collaborative discriminant regression classification

X =[x_1 . . . , x_i . . . x_n] ∈ $R^{m\times n_i}$ be the whole training faces image matrix, where X_i =[x_1 . . . , x_i . . . x_n] ∈ $R^{m\times n_i}$. Notice that both CBCRE and WCRE have the factor of $\frac{1}{n}$, therefore, it is safe to eliminate $\frac{1}{n}$ from CBCRE and WCRE simultaneously without affecting the value of the ratio of CBCRE over CRE. Under some algebraic deduction, CBCRE and WCRE can be written as:

In CBCRE and WCRE can be rewritten as:

$CBCRE = tr(U^T E_b U)$ (8)

$WCRE = tr(U^T E_w U)$ (9)

Where $\hat{y}_{ij}^{inter} = \hat{y}_{ij}^{inter}\ \hat{\propto}_{ij}^{inter}$ and $\hat{y}_{ij}^{intra} = \hat{y}_{ij}^{intra}\ \hat{\propto}_{ij}^{intra}.\hat{y}_{ij}^{inter}$ is the Y with Y_i eliminated and $\hat{y}_{ij}^{intra}$ is the Y_i with Y_{ij} eliminated .However, we can calculate ^a in the original space and use ^a as an approximation of a. From the definition of CBCRE in (7), we can see that the difference between CBCRE and BCRE is that CBCRE uses cross-class collaborative representation and BCRE uses class-specific representation .According to the relationships between X and Y, CBCRE and WCRE can be rewritten as follows:

To maximize CBCRE and minimize WCRE simultaneously, the maximum margin criterion (MMC) is adopted to maximize the following criterion

III. Classification Of HCRF

As for the classification module, a significant number of ways had been employed for correct expression classification. In authors exploited artificial neural networks (ANNs) in order to classify exclusive facial expressions and completed a 73% realization price. Nonetheless, ANN is a black box and has incomplete capacity to explicitly categorize feasible major relationships. Besides, ANNs may just take very long time to coach and may lure in dangerous local minima. Additionally, authors and employed help vector machines (SVMs) for their FER approach. But, in SVMs, the commentary probability is calculated using oblique approaches; in other phrases, there is not any direct estimation of the likelihood.

Furthermore, SVMs simply overlook temporal dependencies among video frames, and as a consequence each and every frame is expected to be statistically independent from the rest. Similarly, authors of utilized Gaussian combination items (GMMs) to admire exceptional types of facial expressions. However facial features would be very sensitive to noise; thus, quick versions in facial frames can't be modeled via GMMs and would cause misclassification. Lots of the aforementioned classifiers have been employed for frame-centered classification. Then again, probably the most normally used sequence-headquartered classification system is the Hidden Markov units (HMMs). HMMs have their own potential in handling sequential information when frame-degree points are used, whereas vector-headquartered classifiers, equivalent to GMMs, ANNs, and SVMs, fail to study the sequence of the characteristic vectors.

Nonetheless, traditional HMMs are situated on Markovian property, which presumes that the present state depends most effective on the previous state. Due to the fact that of this assumption, labels of two contiguous states need to hypothetically occur consecutively in the determined sequence. Lamentably, this presumption shouldn't be normally proper sincerely. Every other obstacles of HMMs include their generative nature and the independence assumption between states and observations [56]. A no generative model akin to highest entropy Markov model (MEMM) was developed with a view to unravel the barriers of HMM, and it produced higher results in comparison with HMM. However, MEMM has a commonly known predicament called the “label bias crisis”. Conditional Random Fields (CRF) and HCRF [58], the generalizations of MEMM, have been then proposed to take the whole skills of MEMM and to remedy the “label bias difficulty”. HCRF extends the capacity of CRF with hidden states making it in a position to be trained hidden structure of the sequential knowledge. Each of them use global normalization as an alternative of per-state normalization. For this reason, they enable weighted scores, making the parameter house greater than these of MEMM and HMM. The following dialogue provides the underlying concept of HCRF, and analyzes the barriers of their current implementations.

In this mapping from input X to labels $Y \in \Gamma$,for instance, computed as

$$p(Y|X;\Lambda) = \frac{\sum_{\bar{S}} \exp\{(\Lambda, \mathrm{f}(Y,\bar{S},X),\}}{Z(X;\Lambda)}$$

$$z(X,\Lambda) = \sum_{\bar{S}} \exp\{(\Lambda, \mathrm{f}(Y,\bar{S},X),\} \qquad (10)$$

Show the above equation that with some specific set of parameters (Λ),HCRF,s dependencies are similar to those of HMM. For example with above features vectore, the diagonal –covariance Gaussian distribution can be defined as

$$\Lambda_{y'}^{pr} = \log(u_{y'}), \forall y' \in Y,$$

$$\Lambda_{SS'}^{Tr} = \log(A_{ss'}), \forall ss' \in \bar{S},$$

$$\Lambda_S^{Occ} = -\frac{1}{2}(\log(2\pi\sigma_s^2) + \frac{\mu_s^2}{\sigma_s^2} \quad (11)$$

$$\Lambda_S^{M1} = \frac{\mu_s^2}{\sigma_s^2}$$

$$\Lambda_S^{M2} = -\frac{1}{2\sigma_s^2}, \quad (12)$$

Where u in (8) is the prior distribution of Gaussian –HMM and A is transition matrix, then the numerator of the condition probability can be written as

$$\sum_{\bar{s}} \exp\{(\Lambda, f(Y, \bar{S}, X),\} = \sum_S u(S1) \prod_{t=1}^{T} A(S_{t-1}, S_t) N(x_t^2, u\, s_t, \sigma S_t), \quad (13)$$

Where N denotes the Gaussian distribution. The conditional likelihood of X given Y is computed with a Gaussian-HMM that has a previous distribution u, and a transition matrix A. A extra generalized variant of the HCRF model has been proposed so as to handle extra complicated distributions utilizing a linear combo of Gaussian density features, and is given as

$$p(Y|X;\Lambda) = \frac{\sum_{\bar{S}} \sum_{m=1}^{M} \exp\{(\Lambda, f(Y, \bar{S}, X),\}}{Z(X;\Lambda)} \quad (14)$$

where *M* is the number of components in the Gaussian mixture. Although, there are some existing works that employed the above HCRF model and showed good result. They did not address and overcome the limitations of the model. As we can see in the above equation, that the model can only utilize diagonal-covariance Gaussian distribution. In other words, the variables (columns of *xi* , *i* = 1, 2, . . . , *N*) are assumed to be pair-wise independent. Hereafter, we call these model diagonal covariance Gaussian mixture hidden conditional random fields (DCGM-HCRF). In addition, equations (10), (11), and (12) imply that with a particular set of values, the observation density at each state will converge to Gaussian form. Unfortunately, there is algorithm that could guarantee this convergence. Therefore, these assumptions may result in a decrease of accuracy. In order to inherit the advantages of HCRF model and completely tackle the limitations of the existing work, we propose the use of HCRF algorithm that is able to explicitly utilize mixture of *full covariance Gaussian mixture hidden conditional random fields* (FCGM-HCRF).

Stepwise Linear Discriminant Regression Classification For Face Recognition (Swldrc)

The expressions that end in high within-class variance and low between-class variance. Therefore, a technique is needed that not solely provides the dimension reduction, however additionally will increase the low between-class variance to extend category separation before the options square measure fed to the classifier. so as to resolve this downside, many ways are planned within the machine learning literature, like kernel discriminant analysis (KDA), generalized discriminant analysis (GDA), and linear discriminant analysis (LDA) [64]. Among these, LDA has been wide utilized in FER systems. However, LDA may be a linear technique that's restricted in flexibility once applied to a lot of advanced datasets. For a lot of details on LDA, please ask a previous study. Consequently, this work employs a strong feature extraction technique known as Stepwise Linear Discriminant Regression Classification for Face Recognition (SWLDRC). SWLDRC is straightforward to elucidate, has sensible prognostic ability, and computationally, it's more cost-effective than alternative existing ways. Some limitations of the prevailing works, like illumination amendment, don't have an effect on the performance of the SWLDRC. SWLDRC solely extracts alittle set of options by using forward and backward regression models. In forward regression, the foremost correlative options square measure hand-picked supported partial F-test values, whereas in backward regression, the smallest amount vital options square measure aloof from the regression model. In each cases, F-test values square measure calculated on the idea of outlined category labels. The advantage of this methodology is that it's terribly economical for seeking localized options. the particular variety of the extracted options used is two hundred. For a lot of details on SWLDRC.

Hidden Conditional Random Fields (HCRF)

As mentioned before that the present HCRF utilizes diagonal covariance Gaussian distributions in the function operate and does not assurance the convergence of its parameters to a few certain values at which the conditional likelihood is modeled as a mixture of typical density capabilities. When you consider that of this property, the existing HCRF losses quite a lot of expertise. This is among the foremost disadvantages of the existing HCRF model. With a view to remedy this drawback, we explicitly involve full covariance Gaussian distributions in the characteristic services at the observation degree. For the prior and transition possibilities, we used the same equations of [59] as described in (3) and (four). Mathematically, our contribution will also be defined as

$$\int_S^{0b}(Y, \bar{S}, X) = \sum_{t=1}^{T} \log\left(\sum_{m=1}^{M} \Gamma_{s,m}^{obs} N(x_t^2, \mu_s, m, \Sigma_s, m)\right)(\delta(S_t = S)) \quad (15)$$

The (15) presents the observation of the input at each state. Where *M* is the number of density functions,Γ is used in order to consider the contextual information of the whole observation, is the mixing weight of the *mth* component with mean μ_s,*m* and covariance matrix $\sum_{m=1}^{M}\Gamma_{s,m}^{obs}\mathrm{N}\left(x_t^2,\mu_s,m,\Sigma_s,m\right)$ in (15) can be computed as

$$\boldsymbol{N}(\boldsymbol{x_t^2},\boldsymbol{\mu_s},\boldsymbol{m},\boldsymbol{\Sigma_s},\mathbf{m})=\frac{\mathbf{1}}{(2\pi)^{\frac{D}{2}}|\Sigma_s,\mathbf{m}|^{\frac{1}{2}}}\mathbf{exp}\left(-\frac{1}{2}(\boldsymbol{x_t^2}-\boldsymbol{z},\boldsymbol{m})\,\boldsymbol{\Sigma_{s,m}^{-1}}{}'(\boldsymbol{X_t^2}-\boldsymbol{\mu_s},\boldsymbol{m})\right) \quad (16)$$

where D is the dimension of the observation, and Σ_s, m is the full covariance matrix. As we can see in (15), by changing Γ,μ and $\Sigma_s$we can create any mixture of the normal densities. So, the corresponding observation weight (Λ_s^{ob}) is not necessary to be updated during the training phase. Therefore,

$$\Lambda_s^{ob}=1,\forall\in\bar{S} \quad (17)$$

As a result, the conditional probability that is used to model the system can be rewritten as

$$p(Y|X;\Lambda,\Gamma,\mu,\Sigma)=\frac{\sum_{\bar{S}}\exp(P(\bar{S})+T(\bar{S})+o(\bar{S}))}{Z(X;\Lambda,\Gamma,\mu,\Sigma)} \quad (18)$$

Where

$$P(\bar{S})=\sum_{S\in\bar{S}}\Lambda_{y'}^{pr}f_{y'}^{pr}(Y,\bar{S},X), \quad (19)$$

$$T(\bar{S})=\sum_{\{S'S\}\in\bar{S}}\Lambda_{S'S'}^{Tr}f_{S'S'}^{Tr}(Y,\bar{S},X), \quad (20)$$

$$o(\bar{S})=\sum_{S\in\bar{S}}f_o^{ob}(Y,\bar{S},X) \quad (21)$$

By putting the values of $(P(\bar{S})\ T(\bar{S})\ and\ o(\bar{S})$from (19), (20), and (21) respectively in (18), the updated conditional probability can be rewritten in (22), as shown at the bottom of this page. As mentioned before, our contribution is at the observation level; therefore, by putting the value of $(P(\bar{S})+T(\bar{S})+o(\bar{S})$from (15), the updated conditional probability for the system can be rewritten, as shown at the bottom of this page. The simple form of the conditional probability is defined in (22).

$$p(Y|X;\Lambda,\Gamma,\mu,\Sigma)=\frac{Score\ (Y|X;\Lambda,\Gamma,\mu,\Sigma)}{Z(X;\Lambda,\Gamma,\mu,\Sigma)} \quad (22)$$

The procedure of the proposed HCRF follows exactly the procedure of the [59]. Based on equations (22) further update the conditional probability using the well known forward and backward algorithms (as the algorithms used in HMM), which are defined in equations respectively, as shown at the bottom of this page.

$$\boldsymbol{p}(\boldsymbol{Y}|\boldsymbol{X};\boldsymbol{\Lambda},\boldsymbol{\Gamma},\boldsymbol{\mu},\boldsymbol{\Sigma})=\frac{\sum_{\bar{S}}\mathbf{exp}\left(\sum_{S\in\bar{S}}\boldsymbol{\Lambda_{y'}^{pr}}\boldsymbol{f_{y'}^{pr}}(\boldsymbol{Y},\bar{\boldsymbol{S}},\boldsymbol{X})+\sum_{\{S'S\}\in\bar{S}}\boldsymbol{\Lambda_{S'S'}^{Tr}}\boldsymbol{f_{S'S'}^{Tr}}(\boldsymbol{Y},\bar{\boldsymbol{S}},\boldsymbol{X})+\sum_{\{S'S\}\in\bar{S}}\boldsymbol{\Lambda_{S'S'}^{Tr}}\boldsymbol{f_{S'S'}^{Tr}}(\boldsymbol{Y},\bar{\boldsymbol{S}},\boldsymbol{X})\right)}{\boldsymbol{Z}(\boldsymbol{X};\boldsymbol{\Lambda},\boldsymbol{\Gamma},\boldsymbol{\mu},\boldsymbol{\Sigma})} \quad (23)$$

Therefore, the *Score* $Z(X;\Lambda,\Gamma,\mu,\Sigma)$of (24) is equal to the forward algorithm (α) and backward algorithm (β).

$$\boldsymbol{Score}(\boldsymbol{Y}|\boldsymbol{X};\boldsymbol{\Lambda},\boldsymbol{\Gamma},\boldsymbol{\mu},\boldsymbol{\Sigma})=\sum_{S\in\bar{S}}\boldsymbol{\alpha T}(\boldsymbol{s})=\sum_{S\in\bar{S}}\boldsymbol{\beta_1}(\boldsymbol{s}) \quad (24)$$

In the training phase, our goal was to find the parameters to maximize the conditional probability of the training data. In SH-FER, we utilize (Limited-memory Broyden-Fletcher-Goldfarb-Shanno) L-BGFS method to search the optimal point. However, instead of repeating the forward and backward algorithms to compute the gradients as others did [59], we run the forward and backward algorithms only when calculating the conditional probability, then we reuse the results to compute the gradients.

IV. Conclusion

In this propose a Step wise linear discriminant regression classification for face recognition and Hidden Conditional Random Fields. SWLDRC-HCRF finds a discriminate mathematical space by increasing cooperative between-class reconstruction error and minimizing the within-class reconstruction error at the same time. in depth experiments convey that LRC features a a lot of higher recognition accuracy on the mathematical space learned by SWLDRC than that of LCDRC.

Over the past two decades, FR systems have received a great deal of attention from the research community due to their application in many areas of pattern recognition and computer vision. However, recognizing human faces accurately is still a major concern. This lack of accuracy can be attributed to various causes, such as the failure to extract prominent features, and the high similarity among different face that results due to the presence of low between-class variance in the feature space.

References

[1]. W. Zhao, R. Chellappa, A. Rosenfeld, and P. J. Phillips, "Face recognition: A literature survey," ACM Comput. Surv., vol. 35, no. 4, pp. 399–458, 2003.
[2]. P. Belhumeur, J. Hespanha, and D. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," IEEE Trans. Patt. Anal. Mach. Intel., vol. 19, no. 7, 1997.
[3]. M. S. Bartlett, J. R. Movellan, and T. J. Sejnowski, "Face recognition by independent component analysis," IEEE Trans Neural Netw., vol. 13, no. 6, pp. 1450–1464, 2002.
[4]. I. Naseem, R. Togneri, and M. Bennamoun, "Linear regression for face recognition," IEEE Trans. Patt. Anal. Mach. Intel., vol. 32, no. 11, pp. 2106–2112, Jul. 2010.
[5]. I. Naseem, R. Togneri, and M. Bennamoun, "Robust regression for face recognition," Patt. Recognit., vol. 45, no, 1, pp. 104–118, Jan. 2012.

[6]. S.-M. Huang and J.-F. Yang, “Improved principal component regression for face recognition under illumination variations,” IEEE Signal Process. Lett., vol. 19, no. 4, pp. 179–182, Apr. 2012.
[7]. A. Martinez and R. Benavente, The AR Face Database, CVC Tech. Rep. 24 1998.
[8]. P. J. Phillips, H.Moon, S. Rizvi, and P. Rauss, “The FERET evaluation methodology for face recognition algorithms,” IEEE Trans. Patt. Anal. Mach. Intell., vol. 22, no. 10, pp. 1090–1104, Oct. 2000.
[9]. A. Mehrabian, “Communication without words,” Psychol. Today, vol. 2, no. 9, pp. 52–55, 1968.
[10]. V. Bettadapura. (2012). “Face expression recognition and analysis: The state of the art.” [Online]. Available: http://arxiv.org/abs/1203.6722
[11]. C. Lisetti and C. LeRouge, “Affective computing in tele-home health: Design science possibilities in recognition of adoption and diffusion issues,” in Proc. 37th IEEE Hawaii Int. Conf. Syst. Sci., Hawaii, USA, Jan. 2004, pp. 348–363.
[12]. X. Wu and J. Zhao, “Curvelet feature extraction for face recognition and facial expression recognition,” in Proc. 6th Int. Conf. Natural Comput. (ICNC), vol. 3. Aug. 2010, pp. 1212–1216.
[13]. S. Moore and R. Bowden, “The effects of pose on facial expression recognition,” in Proc. Brit. Mach. Vis. Conf., 2009, pp. 1–11.
[14]. Z. Zhu and Q. Ji, “Robust real-time face pose and facial expression recovery,” in Proc. Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 1. Jun. 2006, pp. 681–688.
[15]. M. S. Bartlett, G. Littlewort, M. Frank, C. Lainscsek, I. Fasel, and J. Movellan, “Recognizing facial expression: Machine learning and application to spontaneous behavior,” in Proc. IEEE Comput. Soc.Conf. Comput. Vis. Pattern Recognit. (CVPR), vol. 2. Jun. 2005,pp. 568–573.
[16]. Y.-L. Tian, T. Kanade, and J. F. Cohn, “Facial expression analysis,” in Handbook of Face Recognition. New York, NY, USA: Springer-Verlag,2005, pp. 247–275.
[17]. C. Shan, S. Gong, and P. W. McOwan, “Facial expression recognition based on local binary patterns: A comprehensive study,” Image Vis. Comput., vol. 27, no. 6, pp. 803–816, 2009.
[18]. C. Shan and R. Braspenning, “Recognizing facial expressions automatically from video,” in Handbook of Ambient Intelligence and Smart Environments. New York, NY, USA: Springer-Verlag, 2010, pp. 479–509.
[19]. D. J. Krusienski, E. W. Sellers, D. J. McFarland, T. M. Vaughan, and J. R.Wolpaw, “Toward enhanced P300 speller performance,” J. Neurosci. Methods, vol. 167, no. 1, pp. 15–21, 2008.
[20]. M. Ager, Z. Cvetkovic, and P. Sollich. (2013). “Phoneme classification in high-dimensional linear feature domains.” [Online]. Available: http://http://arxiv.org/abs/1312.6849
[21]. M. Nusseck, D. W. Cunningham, C. Wallraven, and H. H. Bülthoff, “The contribution of different facial regions to the recognition of conversational expressions,” J. Vis., vol. 8, no. 8, 2008,Art. ID 1.
[22]. K. Kaulard, D. W. Cunningham, H. H. Bülthoff, and C. Wallraven, “The MPI facial expression database—A validated database of emotional and conversational facial expressions,” PloS One, vol. 7, no. 3, p. e32321, 2012.
[23]. K. L. Schmidt and J. F. Cohn, “Human facial expressions as adaptations: Evolutionary questions in facial expression research,” Amer. J. Phys. Anthropol., vol. 116, no. S33, pp. 3–24, 2001.
[24]. X. Wang and K. K. Paliwal, “Feature extraction and dimensionality reduction algorithms and their applications in vowel re ognition,” Pattern Recognit., vol. 36, no. 10, pp. 2429–2439, 2003.
[25]. Y. Pang, Y. Yuan, and X. Li, “Iterative subspace analysis based on feature line distance,” IEEE Trans. Image Process., vol. 18, no. 4, pp. 903–907, Apr. 2009.
[26]. G. Aguilar-Torres, K. Toscano-Medina, G. Sanchez-Perez,M. Nakano-Miyatake, and H. Perez-Meana, “Eigenface–Gabor algorithm for feature extraction in face recognition,” Int. J. Comput., vol. 3, no. 1, pp. 20–30, 2009 .
[27]. S. R. V. Kittusamy and V. Chakrapani, “Facial expression recognition using eigenspaces,” J. Comput. Sci., vol. 8, no. 10, pp. 1674–1679, 2012.
[28]. J. Kalita and K. Das, “Recognition of facial expression using eigenvector based distributed features and Euclidean distance based decision making technique,” Int. J. Adv. Comput. Sci. Appl., vol. 4, no. 2, pp. 196–202, 2013.
[29]. Z. Abidin and A. Harjoko, “A neural network based facial expression recognition using Fisherface,” Int. J. Comput. Appl., vol. 59, no. 3, pp. 30–34, 2012.
[30]. V. J. Mistry and M. M. Goyani, “A literature survey on facial expression recognition using global features,” Int. J. Eng. Adv. Technol., vol. 2, no. 4, pp. 653–657, 2013.
[31]. F. Long, T. Wu, J. R. Movellan, M. S. Bartlett, and G. Littlewort, “Learning spatiotemporal features by using independent component analysis with application to facial expression.