On Video Base Face Recognition

Naimul Mefraz Khan, Xiaoming Nan, Azhar Quddus, Edward Rosales, and Ling Guan

Abstract

During this few decades, face recognition is a hot research topic and it also had a lot of high accurate algorithms exist in this world, but video-based face recognition had less explored. Usually for recognizing face, human will use the classic method, which is sparse representation-based face recognition method (SRC). However, this kind of classic method had some limitation, it is not suitable to use in video-based face recognition, because of the alignment changes and sensitivity towards pose. In the article, researchers were proposed a method that improved from SRC, they called it as video-based face recognition. They update the probe image into training matrix based on the video by using the Bayesian inference scheme and confidence criterion. This algorithms was successfully to solved the pose and alignment change problems.

Contents

1   Problem Solved

The illumination, occlusion are the challenges of video-based face recognition. To solve the problem on illumination and occlusion, research team were proposed the sparse representation-based face recognition method (SRC). Although SRC can solve the illumination, occlusion problem, but this method is sensitive on pose and alignment change. Based on the research, update the training matrix can improve the robustness against pose and alignment change in SRC method. So that, researchers had proposed Bayesian inference scheme and confidence criterion to update the training matrix.

2   Claimed Contributions

For analyze the update step and Confidence Criterion, researchers had created a video face database that had 11 subjects, and the subjects have to do the motions in few seconds, that is look straight, look right, look straight, look left and look straight. And the experiment had proven, when the subject look straight, the value of criterion P was stay close to 1, when the subject look left or right, the value of criterion P was quickly decrease ,but the values was keep above 0.5. When an impostor was appear in the video, the value of P is drop below 0.5 immediately.

Furthermore, researchers also tried their experiment on Youtube , and picked 100 subjects from the database by randomly. Besides, to test the effectiveness of inference scheme and confidence criterion that proposed by researchers, they used two kind of method to test, which is test with the update step and without the update step. The method without the update step is similar as SRC method. In the result of experiment, the accuracy of method that proposed by researchers but without update is 70.1%, and with the update strategy, the accuracy is higher ,which is 82.9%. The result shown, the method that proposed by researchers is better than the classic SRC method and other methods.

3   Related Work

Currently, there are having three categories of video-based face recognition , that is set-based, sequence-based and dictionary-based. In the first approach set-based, the usability of the images can be using in different conditions of training video. Based on the illumination or pose conditions, the images can be categorized into many different sets. Besides, some of the statistical measurements calculated are used to training and classification that from the sets. Image set distributions function are use to compared for classification and calculate for the experiment training videos. Furthermore, the image set can modeled as manifolds and linear subspaces. For comparing image sets, research team had presented a convex hull-based method, but those approaches are not using at temporal coherence in videos. For more effective to utilize the temporal information, most commend in the sequence-based approaches are using Hidden Markov Models. For the dictionary-based approach, researchers are using a specific sub- dictionaries that

built for illumination conditions and different pose. The probe videos will projected onto atoms in sub-dictionary, and the leftover residue will be use for recognition, but this method will cause a large computational effort. To get an ordered set of frame image calculating the video signature, researchers were used Hidden Markov Models.

4 Methodology

Researching in video-based face recognition, the challenging problem is changes in pose and alignment. In addition, the camera angle of catching videos, illumination, a moving scene and other imaging conditions that change quickly will affect the accuracy of face recognition. Usually when facing this kind of phenomena, that is not recommend to retrain the classifiers and reenroll the individual on a continuous basis. So that, dynamically incorporate the probe video into the training matrix intelligently is the key idea of the method. In original SRC method, training matrix is built once and it is never do any changing on the process. If the current frame have been identified in a test video and was succeed belongs to a particular class, then the next frame may belong to the same class. The most important part of such an approach determine the training matrix is always up to date. The update of training matrix is based on the failure in correctly classifying the current probe images, this detection step is important, so it had designed carefully. However, the system can be easily trick by an impostor, because the system can initiated to update the matrix by a nave way when the current probe image had failed to classified the same class as a previous image. To prevent those weaknesses, confidence criterion and Bayesian inference scheme had proposed by the researchers, confidence criterion is use to trigger the update step, and the Bayesian inference scheme is use to prevent impostors. For analysis the training videos frontal face image, there have some equation, to find the minimum residue,the equation is:

$$\min_i r_i(y) = \parallel y - A\delta_i(x_1) \parallel_2$$

to get value of y, we have to use the equation below:

$$y = \alpha_{i,1}\, s_{i,1} + \alpha_{i,2}\, s_{i,2} + ... + \alpha_{i,n}\, s_{i,n}$$

$\alpha_{i,n}\, s_{i,n}$ are weight coefficients.
to get the value of A, the equation is:

$$A = [A_1, A_2, ..., A_M] = [s_{1,1}\,, s_{1,2}\,, ..., s_M, n_M]$$

A_n are the training classes. δ_i are denotes ith times of training.

$$x_1 = \arg\min \parallel x \parallel_1$$

x is sparse number and $\parallel x \parallel_1$ represents as l^1 norm. l^1 norm are the optimization problem for every frame. To get the confidence criterion value, researchers had given the equation as below:

$$P = e^{\dfrac{-\mid d - \max(d) \mid}{r}}$$

d are denotes the difference that between of minimum and the second-minimum residue, the max(d) is denotes the highest value of d, r denotes a tuning parameter, and the value of P is between 0 and 1. If the value of P is higher, the classified of probe image is more correctly.

For Bayesian inference scheme, it is count the posterior probability, so the equation is:

$$p(X_t \mid Y_{1:t}) = \lambda_t p(Y_t \mid X_t) p(X_t \mid Y_{1:t-1})$$

λ_t is the normalization factor, the value are same with X_t. And $p(Y_t \mid X_t)$ is image likelihood, $p(X_t \mid Y_{1:t-1})$ is the transition probability of process.

For calculate a probe image, first, have to find the value of confidence criterion P. Assume that the value of criterion is less than the predefined threshold value, we need to use equation:

$$p(X_t \mid Y_{1:t}) = \lambda_t p(Y_t \mid X_t) \sum_{X_{t-1}} p(X_t \mid X_{t-1}) p(X_{t-1} \mid Y_{1:t-1})$$

to find the current update is safe or unsafe. If the update is safe, it will use the current probe image to replace the training image that choose from the classified class. Most of the probe image $(p(Yt|Xt))$ if detected more than 0.5, than the current probe image is list as safe stage. Last, the test video will get assign to the class that classified by highest number of frames.

5 Conclusion

In the article, the sparse representation-based classification (SRC) method had be presented for face recognition and using in videos and with the update strategy. Researchers used Bayesian inference scheme and confidence criterion method to detect the update step. However, the method that proposed by researchers still having some space to improve, it was because the accuracy of proposed method were not good enough. Although the proposed methods still need some improvement, but the experimental result on Youtube shown that the update strategy had a noticeable improvement which compare with classic SRC method .

6 What did you learn(algorithms/experiments details?)
Possible extension/Future work

What I learnt from this research is the improvement for the technologies. Nowadays, we have the greatest technologies which may help us to easily solve the issues that we face. In this research show that the update strategy with the calculation, we can produce higher security technologies productions, and it may help to reduce some criminal cases in public areas. Let me provided some examples, polices may use this technologies to track on the suspect through the monitor or CCTV record, and the system can provide and detect the suspects face with a highest accuracy.

7   References

1. M.K. Naimul , X.M. Nan, Q.Azhar, R.Edward, & G. Ling ,July 2014. "On Video Based Face Recognition Through Adaptive Sparse Dictionary",*IEEExplore.ieee.org/Xplore*, pp 1-6,August 2015.

2. , T. Hassner, L. Wolf & I. Maoz, 2011. "Face recognition in unconstrained videos with matched background similarity", *IEEE Conference on Computer Vision and Pattern Recognition.*, pp. 529-534, August 2015