Automatic Track Creation and Deletion Framework for Face Tracking

Renimol T G, Anto Kumar R.P
Dept. of Information and Communication, St.Xavier's Catholic College of Engineering, Nagercoil, Tamilnadu, India.

Abstract
The proposed approach consists of improving the track management by the creation and deletion of the track when occlusion or failure occurs.. In this approach multiface tracking can be possible. Track creation and deletion will avoid errorness failure and improve track management. We improve the accuracy of face detection by using cascade classifiers. Also the face tracking is improved by using Haar Cascade algorithm. Haar cascade, very rarely addressed in the literature, is difficult due to object detector deficiencies or observation models that are insufficient to describe the full variability of tracked objects and deliver reliable likelihood (tracking) information. To achieve this, long-term observations from the image and the tracker itself are collected and processed in a principled way using decision tree algorithm, deciding on when to add and remove a target to the tracker. Proposed algorithm increases the performance considerably with respect to state-of-the-art tracking methods not using long-term observations and HMMs.

Keywords
Haar cascade classifier, track management, face detection, failure detection

1. Introduction

Face Tracking[1] is a computer technology that determines the locations and sizes of human faces in arbitrary (digital) images. It detects facial features and ignores anything else, such as buildings, trees and bodies. Face detection is used in biometrics, often as a part of a facial recognition system. It is also used in video surveillance, human computer interface and image database management. Face detection is gaining the interest of marketers. A webcam can be integrated into a television and detect any face that walks by. The system then calculates the race, gender, and age range of the face. Once the information is collected, a series of advertisements can be played that is specific toward the detected race/gender/age. Face Tracking is also being researched in the area of energy conservation. Face detection can be regarded as a specific case of object-class detection. In object-class detection, the task is to find the locations and sizes of all objects in an image that belong to a given class. Examples include upper torsos, pedestrians, and cars. Face detection can be regarded as a more general case of face localization[3]. In face localization, the task is to find the locations and sizes of a known number of faces (usually one). In face detection, one does not have this additional information. Early face-detection algorithms focused on the detection of frontal human faces, whereas newer algorithms attempt to solve the more general and difficult problem of multi-view face detection. That is, the detection of faces that are either rotated along the axis from the face to the observer (in-plane rotation), or rotated along the vertical or left-right axis (out-of-plane rotation), or both. The newer algorithms take into account variations in the image or video by factors such as face appearance, lighting, and pose.

Multi Object Tracking is an experimental technique used to study how our visual system tracks multiple moving objects. It was developed in 1988 in order to test a theoretical proposed mechanism called a Visual Index[6]. The theory postulates a small number indexes or pointers that pick out and stay attached to individual objects in the visual field independent of their changing properties, and thus allows them to be tracked. The theory addresses a problem of how conceptual descriptions can pick out individual visual objects despite the fact that descriptions themselves are insufficient in general to pick out tokens, as opposed to types. With the continuous development and progress of science and technology, facial biometric recognition technology has become more sophisticated and perfect. As the world's most cutting-edge biometric technology and image processing technology, facial biometric recognition technology will play a unique and irreplaceable role in today's public security prevent, arrest fugitives, network security, financial security and many other areas. It is a milestone of development and progress of human society science and technology. It will be widely used in public security, aviation, ports, customs, banks, large enterprises, large-scale conferences, high-end clubs, important streets, docks and other places of security[9]. It will bring revolutionary changes for safety defend and prevention of the current complicated domestic security situation. Fuzhou Haijing Science &Technology Development Co., Ltd. engaged in facial biometric recognition technology for many years, put into lots of manpower, material resources, overcome many problems with the

Manuscript received February 5, 2015
Manuscript revised February 20, 2015

cooperation of domestic and foreign famous universities and research institutes, and successfully developed the worlds most advanced and cutting-edge facial biometric recognition technology. This technology has the highest recognition accuracy, fastest recognition speed, least affected by the external environment.

Images containing faces are essential to intelligent vision-based human computer interaction, and research efforts in face processing include face recognition, face tracking, pose estimation, and expression recognition, etc. However, many reported methods assume that the faces in an image or an image sequence have been identified and localized. To build fully automated systems that analyze the information contained in face images, robust and efficient face detection algorithms are required. Given an image, the goal of face detection is to identify all image regions which contain a face regardless of its three-dimensional position, orientation, and lighting conditions. Such a problem is challenging because faces are not rigid and have a high degree of variability in size, shape, color, and texture. Target tracking has a number of applications such as human computer interface, security, Surveillance and video conferencing[10]. The human face poses even more problems than other objects since the human face is a dynamic object that comes in many forms and colors . However, facial detection and tracking provides many benefits. Facial recognition is not possible if the face is not isolated from the background. Human Computer Interaction (HCI) could greatly be improved by using emotion, pose, and gesture recognition, all of which require face and facial feature detection and tracking . Although many different algorithms exist to perform face detection, each has its own weaknesses and strengths. Some use flesh tones, some use contours, and other are even more complex involving templates, neural networks, or filters. These algorithms suffer from the same problem; they are computationally expensive .

An image is only a collection of color and/or light intensity values. Analyzing these pixels for face detection is time consuming and difficult to accomplish because of the wide variations of shape and pigmentation within a human face. Pixels often require reanalysis for scaling and precision. , Haar Classifiers, to rapidly detect any object, including human faces, using AdaBoost classifier cascades that are based on Haar-like features and not pixels .

Face Tracking" generally involves two stages:

Face Detection, where a photo is searched to find any face (shown here as a green rectangle), then image processing cleans up the facial image for easier recognition.

Face Tracking, where that detected and processed face

The previous papers use the algorithms as given below

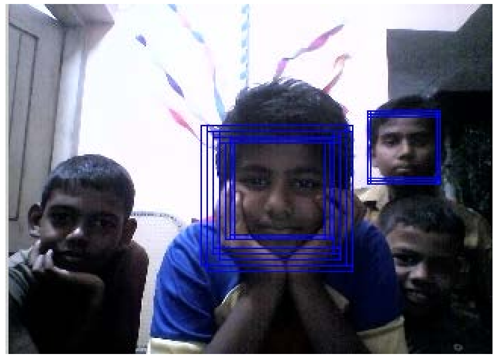


Figure no:1 Face Tracking

1.1 Markov chain Monte Carlo (MCMC)

Markov chain Monte Carlo (MCMC) methods are a class of algorithms for sampling from probability distributions based on constructing a Markov chain that has the desired distribution as its equilibrium distribution. The state of the chain after a large number of steps is then used as a sample of the desired distribution. The quality of the sample improves as a function of the number of steps. The most common application of these algorithms is numerically calculating multi-dimensional integrals. In these methods, an ensemble of "walkers" moves around randomly. At each point where the walker steps, the integrand value at that point is counted towards the integral. The walker then may make a number of tentative steps around the area, looking for a place with reasonably high contribution to the integral to move into next. Random walk methods are a kind of random simulation or Monte Carlo method. However, whereas the random samples of the integrand used in a conventional Monte Carlo integration are statistically independent, those used in MCMC are correlated. A Markov chain is constructed in such a way as to have the integrand as its equilibrium distribution.

1.2 Hidden Markov model (HMM)

A hidden Markov model (HMM) is a statistical Markov model in which the system being modelled is assumed to be a Markov process with unobserved (hidden) states.Hidden Markov models are especially known for their application in temporal pattern recognition such as speech, handwriting, recognition,part, musical score following, partial discharges and bioinformatics. A hidden Markov model can be considered a generalization of a mixture model where the hidden variables (or latent variables), which control the mixture component to be selected for each observation, are related through a Markov process rather than independent of each other. Hidden Markov models can model complex Markov processes where the states emit the observations according to some probability distribution. One such example of distribution is Gaussian distribution, in such a Hidden Markov Model the states output is represented by a Gaussian distribution.

Moreover it could represent even more complex behaviour when the output of the states is represented as mixture of two or more Gaussians, in which case the probability of generating an observation is the product of the probability of first selecting one of the Gaussians and the probability of generating that observation from that Gaussian. In the hidden Markov models considered above, the state space of the hidden variables is discrete, while the observations themselves can either be discrete (typically generated from a categorical distribution) or continuous (typically from a Gaussian distribution). Hidden Markov models can also be generalized to allow continuous state spaces[6]. Examples of such models are those where the Markov process over hidden variables is a linear dynamical system, with a linear relationship among related variables and where all hidden and observed variables follow a Gaussian distribution. In simple cases, such as the linear dynamical system just mentioned, exact inference is tractable (in this case, using the Kalman filter); however, in general, exact inference in HMMs with continuous latent variables is infeasible, and approximate methods must be used, such as the extended Kalman filter or the particle filter.

The parameter learning task in HMMs is to find, given an output sequence or a set of such sequences, the best set of state transition and output probabilities. The task is usually to derive the maximum likelihood estimate of the parameters of the HMM given the set of output sequences. No tractable algorithm is known for solving this problem exactly, but a local maximum likelihood can be derived efficiently using the Baum–Welch algorithm or the Baldi–Chauvin algorithm. The Baum–Welch algorithm is a special case of the expectation-maximization algorithm.

1.3 Reversible-jump Markov chain Monte Carlo

Reversible-jump Markov chain Monte Carlo[5] is an extension to standard Markov chain Monte Carlo (MCMC) methodology that allows simulation of the posterior distribution on spaces of varying dimensions. Thus, the simulation is possible even if the number of parameters in the models is not known.

1.4 Bayesian Filter

A Bayes filter[8] is an algorithm used in computer science for calculating the probabilities of multiple beliefs to allow a robot to infer its position and orientation. Essentially, Bayes filters allow robots to continuously update their most likely position within a coordinate system, based on the most recently acquired sensor data. This is a recursive algorithm. It consists of two parts: prediction and innovation. If the variables are linear and normally distributed the Bayes filter becomes equal to the Kalman filter. In a simple example, a robot moving throughout a grid may have several different sensors that provide it with information about its surroundings. The robot may start out with certainty that it farther from its original position, the robot has continuously less certainty about its position; using a Bayes filter, a probability can be assigned to the robot's belief about its current position, and that probability can be continuously updated from additional sensor information.

2. PROPOSED MODEL

This section gives an overview on the major Face detection and recognition system module.

A. The pre-processing module

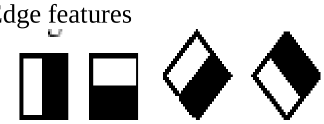
The aim of the pre-processing phase was to obtain images which have normalized intensity, uniform size and shape. In this module the images are normalized and enhanced to improve the recognition of the system. The preprocessing steps implemented are as follows:

- Conversion of grayscale image
- Image size normalization
- Histogram equalization

B. Face detection using HAAR Cascade Classifiers

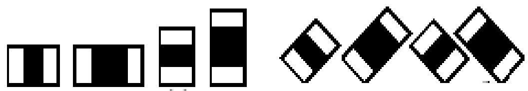
The human face poses even more problems than other objects since the human face is a dynamic object that comes in many forms and colors . However, facial detection and tracking provides many benefits. Facial recognition is not possible if the face is not isolated from the background. Human Computer Interaction (HCI) could greatly be improved by using emotion, pose, and gesture recognition, all of which require face and facial feature detection and tracking . Although many different algorithms exist to perform face detection, each has its own weaknesses and strengths. Some use flesh tones, some use contours, and other are even more complex involving templates, neural networks, or filters. These algorithms suffer from the same problem; they are computationally expensive . An image is only a collection of color and/or light intensity values. Analyzing these pixels for face detection is time consuming and difficult to accomplish because of the wide variations of shape and pigmentation within a human face. Pixels often require reanalysis for scaling and precision. , Haar Classifiers, to rapidly detect any object, including human faces, using AdaBoost classifier cascades that are based on Haar-like features and not pixels .

1. Edge features



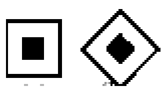
(a) (b) (c) (d)

2. Line features



(a) (b) (c) (d) (e) (f) (g) (h)

3. Center-surround features



(a) (b)

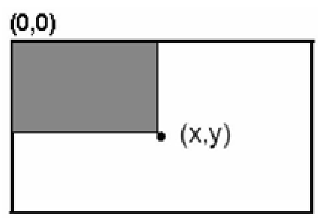


Figure no:2 Summed area of integral image

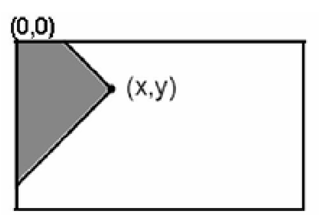


Figure no :3 Summed area of rotated integral image

The core basis for Haar classifier object detection is the Haar-like features. These features, rather than using the intensity values of a pixel, use the change in contrast values between adjacent rectangular groups of pixels. The contrast variances between the pixel groups are used to determine relative light and dark areas. Two or three adjacent groups with a relative contrast variance form a Haar-like feature. The simple rectangular features of an image are calculated using intermediate representation of an image, called the integral image . The integral image is an array containing the sums of the pixels' intensity values located directly to the left of a pixel and directly above the pixel at location (x, y) inclusive. So if A[x,y] is the original image and AI[x,y] is the integral image then the integral image .

$$AI[x,y] = \sum_{x' \le x,\, y' \le y} A(x',y')$$

The rotated integral image is calculated by finding the sum of the pixels' intensity values that are located at a forty five degree angle to the left and above for the x value and below for the y value. So if A[x,y] is the original image and AR[x,y] is the rotated integral image

$$AR[x,y] = \sum_{x' \le x,\, x' \le x-|y-y'|} A(x',y')$$

C. Decision tree algorithm

Decision trees are powerful, top-down, hierarchical search methodology to extract easily interpretable classification rules . Fuzzy decision trees are composed of a set of internal nodes representing variables used in the solution of a classification problem, a set of branches representing fuzzy sets of corresponding node variables, and a set of leaf nodes representing the degree of certainty with which each class has been approximated. We have used our own implementation of fuzzy ID3 algorithm for learning a fuzzy classifier on the training data. ID3 utilizes decision tree algorithm classification entropy of a possibilistic distribution for decision tree generation. decision tree using fuzzy ID3 algorithm for the skin-nonskin classification problem. We have taken leaf selection threshold 0.75[4].

The product-product-sum reasoning mechanism consists of the following three steps:

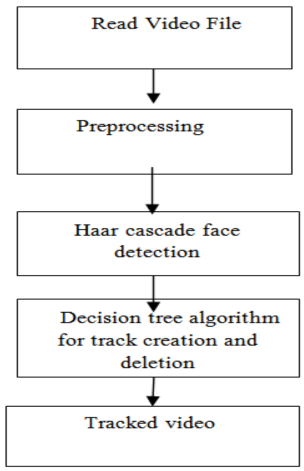
STEP 1. For the operation to aggregate membership values of fuzzy sets of node genres along the paths, the product is adopted.

STEP 2. For the operation of the total membership value of the path of fuzzy evidences and the certainty of the class attached to leaf-nodes, also the product is adopted.

STEP 3. For the operation to aggregate certainties of the same class from different paths, the sum is adopted.

Decision tree learning is one of the most successful techniques for supervised classification learning. For this section, assume that all of the features have finite discrete domains, and there is a single target feature called the classification. Each element of the domain of the classification is called a class. A decision tree or a classification tree is a tree in which each internal (non-leaf) node is labeled with an input feature. The arcs coming from a node labeled with a feature are labeled with each of the possible values of the feature. Each leaf of the tree is labeled with a class or a probability distribution over the classes.A tree can be "learned" by splitting the source set into subsets based on an attribute value test. This process is repeated on each derived subset in a recursive manner called recursive partitioning.

Proposed Block Diagram



3. Performance Analysis

The performance of face recognition is commonly evaluated using real time database. Sixteen different poses of 150 individuals are present in the real face database. A database of 150 images of different subjects of size 128 is taken for experimentation. The train image is limited in size (only 128)[1]. The log Gabor filter based PCA has a face recognition accuracy of 79% for the same test database. feature is much smaller than that of the whole image, detection of all three facial features takes less time on average than detecting the face itself. Using a 1.2GHz AMD processor to analyze a 320 by 240 image, a frame rate of 3 frames per second was achieved. Since a frame rate of 5 frames per second was achieved in facial detection only by using a much faster processor, regionalization provides a tremendous increase in efficiency in facial feature detection[7]. Using a 1.2GHz AMD processor to analyze a 320 by 240 image, a frame rate of 3 frames per second was achieved. Since a frame rate of 5 frames per second was achieved in facial detection only by [5] using a much faster processor, regionalization provides a tremendous increase in efficiency in facial feature detection.

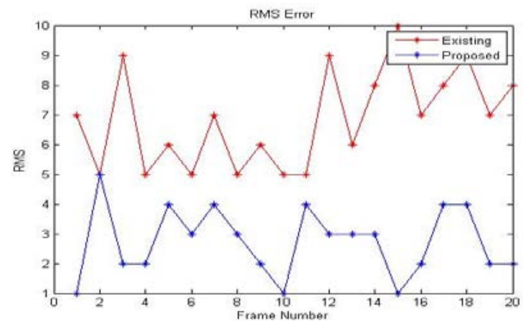


Fig no 2:Track creation and deletion

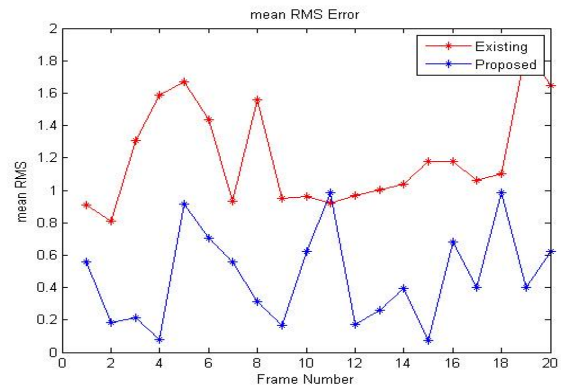


Fig no 3:Improved performance

4. Conclusion

This approach will reduce the computation time and improve the detection accuracy. The approach was used to construct a face detection system which is approximately 15 times faster than any previous approach. This paper brings together new algorithms, representations, and insights which are quite generic and may well have broader application in computer vision and image processing. Finally this paper presents a set of detailed experiments on a difficult face detection dataset which has been widely studied. This dataset includes faces under a very wide range of conditions including: illumination, scale, pose, and camera variation. Experiments on such a large and complex dataset are difficult and time consuming. We presented an on-line multi-face tracking algorithm that effectively deals with situations where detections are rare or uncertain. To achieve this, long-term observations from the image and the tracker itself are collected and processed in a principled way using haar cascade face detection and to *add* and *remove* a target to the tracker using decision tree algorithm. We evaluated our approach on more than 9 hours of videos with extensive annotation, and the results show that the proposed algorithm increases the performance considerably with respect to state-of-the-art tracking methods not using long-term observations and HMMs.

REFERENCES

[1] J. Fan, W. Xu, Y. Wu, and Y. Gong, “Human tracking using convolutional neural networks,” IEEE Trans. Neural Netw., vol. 21, no. 10, pp. 1610–1623, Oct. 2010.

[2] Sarala A. Dabhade, Prof. Mrunal S. Bewoor,” Real Time Face Detection and Recognition using Haar - Based Cascade Classifier and Principal Component Analysis International Journal of Computer Science and Management Research Vol 1 Issue 1 Aug 2012

[3] Z. Khan, T. Balch, and F. Dellaert, “MCMC-based particle filtering for tracking a variable number of interacting targets,” IEEE Trans. PatternAnal. Mach. Intell., vol. 27, no. 11, pp. 1805–1918, Nov. 2005.

[4] T. Darrell, G. Gordon, M. Harville, and J. Woodfill, “Integrated person tracking using stereo, color, and pattern detection,” Int. J. Comput. Vis., vol. 37, no. 2, pp. 175–185, 2000.

[5] C.-H. Kuo and R. Nevatia, “How does person identity recognition help multi-person tracking?” in Proc. Comput. Vis. Pattern Recognit., 2011, pp. 1217–1224.

[6] B. Benfold and I. Reid, “Stable multi-target tracking in real-time surveillance video,” in Proc. Comput. Vis. Pattern Recognit., Jun. 2011, pp. 3457–3464.

[7] M. Yang, Y. Wu, and G. Hua, “Context-aware visual tracking,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 31, no. 7, pp. 1195–1209, Jul.2009.

[8] E. Ricci and J.-M.Odobez, “Learning large margin likelihoods for realtime head pose tracking,” in Proc. Int. Conf. Image Process., Oct. 2009, pp. 2593–2596.

[9] Stefan Duffner and Jean-Marc Odobez ”Track Creation and Deletion Framework for Long-Term Online Multiface Tracking” IEEE Transactions on image processing, vol. 22, no. 1, january 2013

[10] Y. Wu, T. Yu, and G. Hua, “Tracking appearances with occlusions,” in Proc. Comput.Vis. Pattern Recognit., vol. 1. 2003, pp. 789–795.