Detection of Good Quality Front Face from Inexpensive Low Resolution Surveillance Camera

Bhagyshree Manoj Karmarkar[1] Asst. Prof. Hirendra Hajare[2]
[1]PG Student [2]Assistant Professor
[1,2]Department of Computer Science & Engineering
[1,2]Ballarpur Institute of Technology, Ballarpur, India

Abstract— The face is important part of a body to identify any person. Every public place having the surveillance camera to monitor on public activities. From that video sequence if at any Circumstances we want to find the suspected person because of low resolution of surveillance camera it gets difficult to detect the face of person what we exactly wanted so that we need a mechanism which can take a input from that video sequence and convert it into face logs and applying some resolution techniques we can get the good quality and high resolution front faces from it. To deal with this approach we required an automated system which is working with LR and low-quality faces. Super-resolution (SR) is one of such technique for obtaining high resolution (HR) image from one or more LR input images. Super resolution faces two problems one which is related with images that it requires the similar image and second problem relates with improvement factor whose limit is of two. Because of this technique the improvement factor gets improved by four for whole system. The experimental results show that the proposed system can produce a front faces which have clear resolution and good quality from the video sequences.

Key words: Detection of Good Quality Front Face, Surveillance Camera

I. INTRODUCTION

In our daily life computer has lots of importance since it operates faster than human and it is easier to deal with the applications which are related to the human faces .With the rise of technology in 20th century cameras are found to be very useful tool to survey a large area with limited resources . Face recognition from the video sequences is important task today. Most of existing work focuses on identifying the faces from video sequences where input images are having high resolution and good quality. But identifying the exact face of person from surveillance camera which has poor quality and low resolution is quite difficult task. one of the most difficult task in identifying front face from surveillance camera is because of variable motions of person around the camera. Hence at public places the surveillance cameras are installed and human operators are monitoring on activities of people but it's not a proper solution we need a mechanism to bridge this gap. Our proposed approach exactly deal with this. Therefore we need an automated system which works on low resolution images and low quality faces .for this SR is one of the best technique to obtain the HR images from LR images .super resolution faces two problem one which is related with images that it requires the similar image and second problem relates with improvement factor whose limit is of two. two resolve above problems we require some approach that are use of three-step approach to resolve first problem and use of learning based super resolution .basically we use the RBSR and LBSR algorithms.

II. RELATED WORK

To detect a face is the first stage of any face recognition process. A lot of works has been done on this area. Since we are dealing with the video sequences coming from the surveillance camera there are chances of various orientation of human poses so to detect proper face from these variations is a research and quite tough task. Generally this process consists of three steps at first frame based detection is done .for this lots of traditional methods are applied which are applicable for still images like statistical modelling, neural network based method, SVM based methods. Ignoring the temporal information provided by video sequence is main drawback of this approach. Secondly, integrating detection and tracking, this indicates that identifying the face is the first step and then identify the person on the basis of it the whole process. Super resolution is the process of giving low resolution images and from that obtain the high quality images .Based on reasonable assumptions or prior knowledge about the observation model that converts the images having high resolution image to the images of low resolution ones. Super resolution algorithm is classified into two categories which are Reconstruction based super resolution (RBSR) and (LBSR) Learning based super resolution. Super resolution algorithm which based on reconstruction techniques works with more than one Low resolution input image. These LR inputs must have intra-image sub-pixel misalignments.

III. PROPOSED SYSTEM

The main motive of this system is to design an automatic front face extraction from low resolution surveillance camera. For this our total project being classified as various section .First Section includes the face detection and Second Section contains the reconstruction techniques on detected face logs to obtain the good quality front face. Proposed approach is consists of following steps

1) Step 1: Input the video and apply the skin detection Algorithm
2) Step 2: Face Detection and Generate the face log.
3) Step 3: After generating face log
4) Step 4: Obtain the refined face log by applying some techniques to remove non face part
5) Step 5: Apply the resolution algorithm to obtain the high resolution front face which includes RBSR and LBSR techniques.
6) Step 6: Finally collect the final extracted good quality front faces from the video sequence

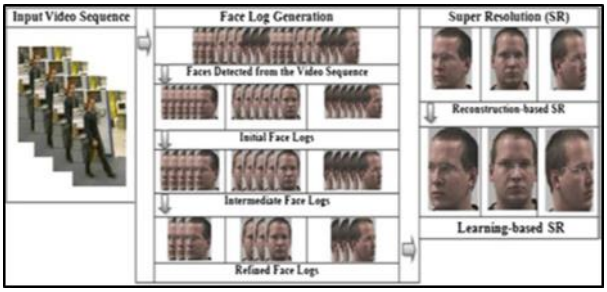

Fig. 1: Block Diagram of Architecture

Description of the Proposed Algorithm: Fig 1 describes about the block diagram of our proposed system architecture. Proposed approach is consists of following modules.

- Face Identification
- Face Log Generation
- SR image
- Final extracted image

A. Face Identification

To detect face first thing is to detect the skin part of body from video sequence after that we need to detect the face from it. This is the second step after detecting the skin. To detect the face and to separate it from skin region some parameters are taken into consideration. It works on CBIR (Content Based Image Retrieval) which operates on shape, size, colour and texture of skin.

B. Face Log Generation

In this section we explain about the face log generation techniques. These processes are: face detection, facial feature extraction, and face quality assessment (FQA).

1) Face Detection

Face-logs are considered as a concise and/or complete representation of a video sequence. For example if the face-logs are used for indexing video sequences, they may only contain the best face image of the sequence. If they are used for summarizing a video sequence, they should be complete. It means that they should contain the best side-view images (if any) as well as the best frontal image. If face-logs are used for SR, as in this paper, in addition to the best frontal (and side views) face(s) they need to have many more images that are closely similar to the best images. Every*m*th frame (3<m<15) of a video sequence and two different types of face-logs generated for different purposes. The proposed system in this paper uses one to three face-logs. The most important one is a face-log containing frontal and semi-frontal face images. The other two face-logs are associated with two side-view face images of the subject. For constructing these three face-logs, we first use a head-pose estimation method that is developed for LR images.

2) Facial Features Extraction

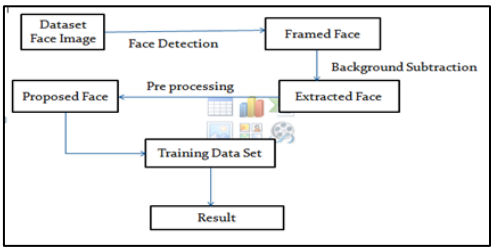

Fig. b: Flow of Block Diagram of Refined Face Generation

The following sub sections describes the details of extracting these facial features

a) Head-Pose Estimation

This feature is one of the most important features in assessing the usability of the face, because wide variation in pose can hide most of the useful features of the face. The previous face quality assessment systems [25, 26, and 27] have involved facial features like vertical position of the eyes, distance between the two eyes and vertical symmetry axis to estimate the pose of the face. It is obvious that most of these Features may be hidden in various conditions like having spectacles or different lightening condition or even in rotations more than 60° [27]. Hence using the facial features to estimate the pose of the face cannot be reliable. Furthermore in the quality assessment the exact rotation of the face is not important but choosing the least rotated face is. So, we deal with the face as a whole, and calculate the difference between the center of mass and the center of the detected face region. We calculate the center of mass using the following equation:

$$x_m = \frac{\sum_{j=1}^{n}\sum_{j=1}^{m} ib(i,j)}{A}, y_m = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m} jb(i,j)}{A} \quad \text{----------(1.1)}$$

Where, is the center of mass, *b* is the binary image containing the detected region as a face, *m* is the width, *n* is the height of the detected region and *A* is the area of this region.

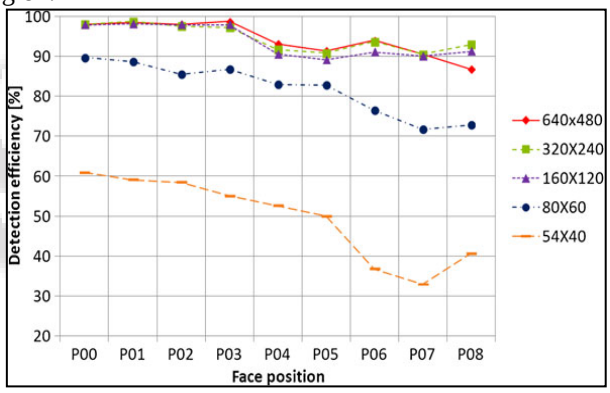

Fig. C: Influence of image resolution and various face positions on face detection process

Distance	62.26	57.26	46.28	24.82	18.97	15.04
S_1	0.24	0.26	0.32	0.6	0.79	1

Fig. 2: Sequence of Different Head Poses and the associated values for the distance and S1

b) Sharpness

Since in real world applications the objects are moving in front of the camera, it is possible that the captured image is affected by motion blur, so defining a sharpness feature can be useful for FQA. Well-focused images, which have a better sharpness compared to blurring images, should get a higher score for this feature. After calculating the sharpness for all of the chosen faces we assign the following score to the sharpness of each of them:

S2=E/Emax----------(1.2)

Where Emax is the maximum value of the sharpness in this sequence. Figure 6.shows some images of one person with different values in sharpness and their associated scores.

Fig. 3: An image with Different Sharpness Conditions and the Associated Scores

c) Brightness

Changes in the lightening conditions are very common in surveillance scenarios, e.g., like a camera mounted above an ATM machine. In such cases, brightness soft he faces can change very easily. Facial features in a bright face can be extracted much easier than in a dark face. Thus, it is important to assess the brightness of the face images of a given video sequence. We have defined the brightness as the mean of the illumination component of the face in *YCbCr*color space.

Fig. 4: An Image with Different Brightness Conditions and Their Associated Score

d) Resolution

For covering as much of the scene as possible, surveillance cameras are usually setup with wide fields of view. Therefore, there is a large distance between the camera and objects. This causes the faces to appear very small. However, due to the motion of the object the size of the faces is likely to change. To involve this feature in the quality assessment we assign the highest quality score of resolution to the biggest face. Accordingly, the quality score decreases for the smaller faces.

Feature	Head-Pose	Sharpness	Brightness	Resolution
Weight	1.7	0.9	0.6	0.8

Table 1: (Weights of the Facial Features Involved in the FQ)

3) Face Quality Assessment

In order to compare a face image of a specific person with the other face images of the same person from a video sequence, we need to assign a quality score to each face. To do so, we have combined the normalized value of the above-explained features into a quality score for each face. Following Equation is used to normalize the head-pose val

$$Q_{1X_i} = \frac{P_{\min}}{P_{X_i}} \quad \text{----------(1.4)}$$

Where Xi is the ith face image in the given video sequence (i is changing from one to the size of the initial face-log, m1), PXi is its head-pose value, and Q1Xi is its first quality score .Pmin is the minimum value of the head-pose feature in the face-log.

C. Super-Resolution

Super-resolution (SR) imaging aims to overcome or compensate the limitation or shortcomings of the image acquisition device/system and/or possibly ill-posed acquisition conditions to produce a higher-resolution image based on a set of images that were acquired from the same scene. With rapid development and deployment of image processing for visual communications and scene understanding, there is a strong demand for providing the viewer with high-resolution imaging not only for providing better visualization (fidelity issue) but also for extracting additional information details (recognition issue). For examples, a high resolution image is beneficial to achieve a better classification of regions in a multi-spectral remote sensing image or to assist radiologist for making diagnosis based on a medical imagery. In video surveillance systems, higher-resolution video frames are always welcomed for more accurately identifying the objects and persons of interest.

	HR (90x120)	LR1 (23x30) (Zoomed to 90x120)	LR2 (23x30) (Interpolated to 90x120)	SR (90x120)
VA	100%	67%	67%	100%
KE				100%

Table 7a – Face recognition results of web camera captured video for 23x30 frames

	HR (90x120)	LR1 (45x60) (Zoomed to 90x120)	LR2 (45x60) (Interpolated to 90x120)	SR (90x120)
VA	100%	83%	83%	100%
KE				100%

Table 7b – Face recognition results of web camera captured video for 45x60 frames

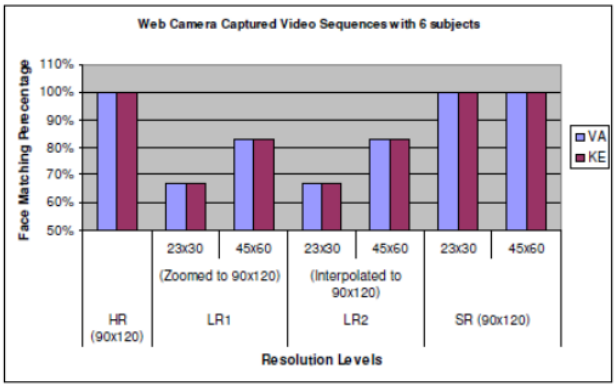

Fig 34: Face matching results of web camera captured video with 6 subjects (VA [8], KE [9]).

1) Reconstruction Based Super Resolution

In order to construct the higher quality images from the low resolution faces detected from the video sequences in surveillance camera we need to apply reconstruction technique. By warping Based on the imaging model each LR image has been created, blurring and down-sampling the HR image. It means that each Xi, i = {1, 2 . . . m3} where LR images in the refined face-log have been obtained by

$$X_i = DB_iW_iH + n_i \quad \text{----------(1.3)}$$

Where D, Bi, and Wiare the down-sampling is given by D, blurring is given by Bi, and warping matrix is given by Wi, H is the HR image and niis the introduced noise in the imaging process to produce the ith LR image from the HR image H.

2) Learning Based Super Resolution

In order to prepare the training data for any network we have first obtained the face areas of HR images and converted to gray scale. Then, an LR image is created for each of these HR images by down-sampling the images by a multiple of two. Then, all of these LR/HR pairs which are obtained from the above procedures which are applied in order to get high quality images are fed to the network as the training samples and the network learns the relationship between them.

IV. SIMULATION RESULTS

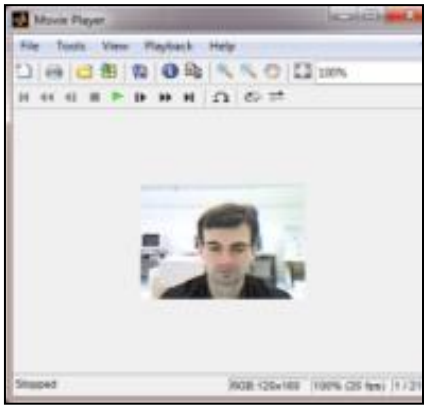

Fig. 1: Input the video

Fig. 1 Describes About the Inputting the Video to Generate Face Log.

Fig. 2: Face Log Generation

Fig. 2 This Figure Describes About the Face Log Generation and Face Rotation Depending on Head Pose Estimation

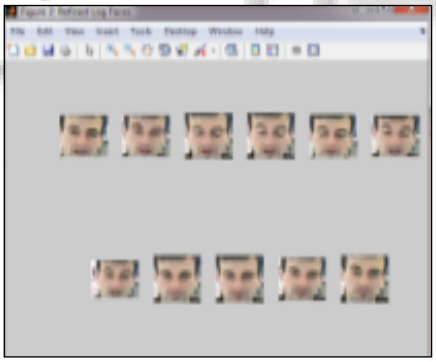

Fig. 3: Refined Face Log

Fig. 3 Describes About the Refined Face Log Generation Technique and Morphological Operation and Applying Facial Features Extraction Techniques

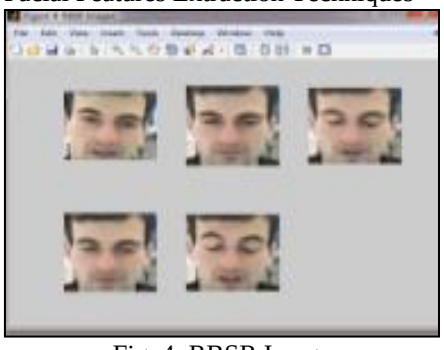

Fig. 4: RBSR Image

Fig. 4 Describes about the RBSR that is Reconstruction Based Super Resolution Algorithm

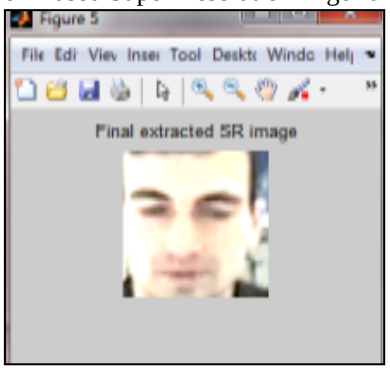

Fig. 5: Final Extracted Image

Fig. 5 Describes about the Final Extracted Image after RBSR and LBSR Technique

V. CONCLUSION AND FUTURE WORK

This approach is best solution to obtain the good quality of front face from low resolution surveillance cameras .Since we not getting the clear and high resolution front faces from the video sequences in surveillance camera for that we need a proper approach this is the best approach to deal with this problem.

REFERENCES

[1] Jian Sun, ZongbenXu, and Heung-Yeung Shum, "Image super resolutionusing gradient profile prior," in Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on, june 2008, pp. 1–8.

[2] Xiangjun Zhang and Xiaolin Wu, "Image interpolation by adaptive2-d autoregressive modeling and soft-decision estimation," Image Processing, IEEE Transactions on, vol. 17, no. 6, pp. 887–896, June2008.

[3] S. Baker and T. Kanade, "Limits on super-resolution and how to breakthem," Pattern Analysis and Machine Intelligence, IEEE Transactionson, vol. 24, no. 9, pp. 1167 – 1183, sep 2002.

[4] Shengyang Dai, Mei Han, Wei Xu, Ying Wu, Yihong Gong, and A. K. Katsaggelos, "Softcuts: A soft edge smoothness prior for color image super-resolution," Image Processing, IEEE Transactions on, vol. 18, no.5, pp. 969–981, May 2009.

[5] S. Mallat and Guoshen Yu, "Super-resolution with sparse mixing estimators," Image Processing, IEEE Transactions on, vol. 19, no. 11,pp. 2889–2900, Nov 2010.

[6] D. Glasner, S. Bagon, and M. Irani, "Super-resolution from a single image," in Computer Vision, 2009 IEEE 12th International Conferenceon, 29 2009-oct. 2 2009,

[7] M. Zontak and M. Irani, "Internal statistics of a single natural image," in Computer Vision and Pattern Recognition (CVPR), 2011IEEE Conference on, june 2011, pp. 977 –984.

[8] A. Sethi, N. Kumar, and N.K. Rai, "A spatial neighborhood based learning set-up for super resolution," in India Conference (INDICON),2012 Annual IEEE, 2012, pp. 790–793.

[9] W.T. Freeman, T.R. Jones, and E.C. Pasztor, "Example-based super resolution, "Computer Graphics and Application IEEE, vol. 22, no. 2,pp. 56 –65, mar/apr 2002.
[10] Jianchao Yang, J. Wright, T.S. Huang, and Yi Ma, "Image super resolutionvia sparse representation," Image Processing, IEEE Transactionson, vol. 19, no. 11, pp. 2861 –2873, nov. 2010.

