International Journal of Review in Electronics and Communication Engineering
Volume 3 No 2 April 2015

FRONTAL FACE ESTIMATION FROM A LOW RESOLUTION VIDEO SEQUENCE

S.Shobhitha
IV B.Tech ECE
VITS, PDTR
sankepalli4u@gmail.com

S.Siddhendra Sai
IV B.Tech ECE
VITS, PDTR
siddhendra7262@gmail.com

S.Siddeswara Reddy
Asst. prof in ECE Dept.,
VITS, PDTR
siddeswara.reddy6@gmail.com

***Abstract*– Sustaining low-quality images from low budget surveillance cameras, to systems like, e.g., face recognition, produces inaccurate and unstable results. So, we need for a mechanism to bridge the gap between low-resolution, low-quality images and facial analysis systems. The face is one of the most important remote biometrics and is widely employed in many facial analysis systems, like face recognition, human–computer interaction, and so on. Real world challenges of these systems are that they have problem with low resolution images. So, our approach is to apply a reconstruction-based super-resolution algorithm. It has two main problems: First, it requires relatively similar images and second is that its improvement factor is limited by a factor close to two. To deal with the first problem, a three-step approach which produces a face-log containing images of similar frontal faces of the highest possible quality. To deal with the second problem, limited improvement factor, we use a learning based super-resolution algorithm applied to the result of the reconstruction-based part to improve the quality by another factor of two. The proposed system can indeed produce a high resolution and good quality frontal face image from low resolution video sequence.**

Keywords: **Face quality assessment, face log generation, super resolution, and surveillance video.**

I. INTRODUCTION

In military and civilian applications, high resolution images are always required. Therefore, having an automated system working with LR and low-quality face images is desirable. However, low-quality images do not have enough high-resolution (HR) details for facial analysis systems and using them directly in these systems is not reliable. Therefore, we need a mechanism for bridging this gap between LR images and facial analysis systems. Super-resolution (SR) is one of such mechanisms for obtaining a high resolution image from more low resolution input images. SR algorithms are classified into two classes: reconstruction-based SR (RBSR) [1]-[11] and learning-based SR (LBSR)[12]-[20].Reconstruction based SR algorithm work with more LR input images. The first RBSR system was a frequency domain algorithm proposed by Tsai and Huang [3]. Spatial domain solutions for RBSR were later developed [4]-[11]. These methods have better considerations for noise and blur than the frequency domain approaches. These methods mainly differ in three points: registration algorithm, the method for obtaining the final response of the system and the regularization method. Other RBSR methods include: no uniform interpolation (NUI) [6] based approaches, projection onto a convex set (POCS) [7] methods, maximum likelihood (ML) [8] methods, maximum a posteriori (MAP) [11] methods, to name but a few. Computational cost of NUI methods is low but they cannot use the degradation models if noise and blur are different in the LR images. POCS methods are easy to use but they have high computational costs, slow convergence and problems in producing a unique response. ML methods simply look for an HR image that maximizes the probability of having the LR inputs. ML methods are prone to be extremely ill-conditioned. MAP methods explicitly use the a priori information in the form of a prior probability on the HR image. These methods are also ill-posed and need some regularization term for stabilizing their final response. The improvement factor of the RBSR algorithms is close to two.

The relationship between some HR training images and their corresponding LR versions can be known by LBSR algorithm: resolution pyramids [12], neighborhood embedding [13], manifolds [14, 15], compressed sensing and sparse representation [16], neural networks[17,19], to name a few. Sustaining an LR input image to an LBSR, the system usually looks among its training data for the closest LR image and uses the learned relationship to predict the missing HR details. Usually, HR images are of double size of their corresponding LR images, i.e., they can improve the quality of LR inputs by a factor of two. However, unlike RBSR algorithms there is no guarantee that the hallucinated HR images produced by LBSR algorithms provide true HR details [20].

The proposed system super resolution mechanism in this paper deals with the real world problems of SR system working with faces coming from a surveillance video sequence. Such a system has several problems: the slight-motions restriction of the objects, the ill-posed and ill-conditioned nature of the HR response, slow convergence of the system, and small magnification factors. Sustaining all frames of a surveillance video sequence to any SR system is impossible, because there are many face images in such a sequence that are not useful due to problems like not changing in the head-pose, blurriness, uneven illumination (due to changes in the lightening conditions), and so on. Including all such images in the process reduces the convergence speed of the system and makes the final response inaccurate and unreliable. Therefore, we need for a mechanism to assess the quality of the detected faces and reject the useless ones. Furthermore, faces given to an RBSR system should be very similar to each other. So, we introduce a three-step approach, which produces a face log containing images of similar faces of the highest possible quality.

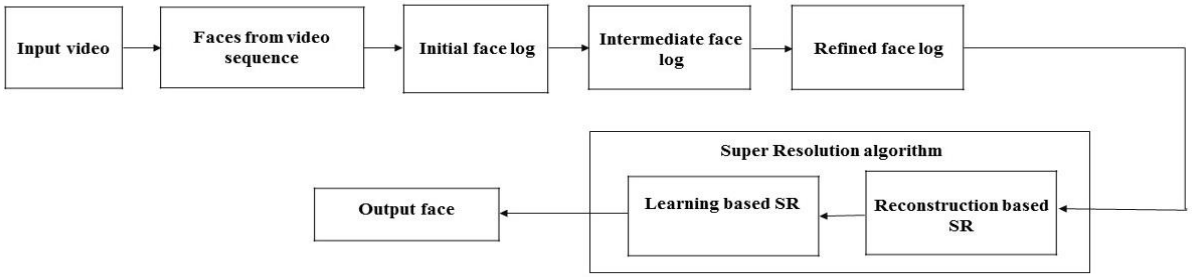

Figure 1: Block diagram of proposed system

In the first step, we classify face images in terms of head pose. In steps two and three, we consider this initial face log based on a number of quality and similarity measures, respectively, and end up with a refined face-log on which the SR algorithms can operate. The quality of the detected faces, we consider four facial features: sharpness, brightness, out of plan rotation and face size. For checking the similarity of the faces, two similarity measures are used: structural similarity and correlation coefficient. In the reconstruction-based part an MAP algorithm and in the recognition part a multilayer perception (MLP) is used. The reconstruction-based part improves the quality by a factor close to two and the learning-based part by a factor of two. It means that the final output of the system is improved by a factor close to four.

The block diagram of our system is shown in Fig. 1. In the first block of the system, face-log generation, the input video sequence is summarized to one [up to three face-log(s)]. The images inside the face-logs are very similar to each other and are of better quality compared to the other images of the sequence. In cascaded SR block, an RBSR is applied to the generated face-log(s) and produces an HR image. The quality of this HR image is improved by a factor of two compared to the LR images in the video sequence. This HR image is then fed to an LBSR to improve it even more.

II. REFINED FACE LOG GENERATION

In this section, we first explain the idea of face-log generation. Then, all the required processes for producing a face log(s) from a video sequence are described. These processes are: face logs, face feature extraction, face quality assessment (FQA), choosing the best face from the sequence.

A. Face-Log(s)

Face-logs are considered as a concise and/or complete representation of a video sequence. The content of face-logs depends on the application. 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 as well as the best frontal image.

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. Based on the value of this feature, we classify face images of the input video sequence into three face-logs containing frontal, left side-view, and right side-view face images. These face logs are denoted as initial face-logs. Each of these initial face logs goes into the same computations, separately.

B. Face Detection

Face detection is not the main focus of this paper but since some of the extracted features for the face(s) in this block are used in the assessment process too, we briefly describe it here.

Given a color image, first of all, according to a Gaussian model of skin color, a probability image of the input image is produced. Then using an adaptive threshold, this probability image is segmented to a binary one which has the skin regions separated from the non-skin ones. Here after a cascading classifier using the extracted features for each region decides if this region is a face or not (See figure 3.).

The extracted features for each face include: face size, center of the mass and orientation of the face, number of holes inside the face, the holes area to its

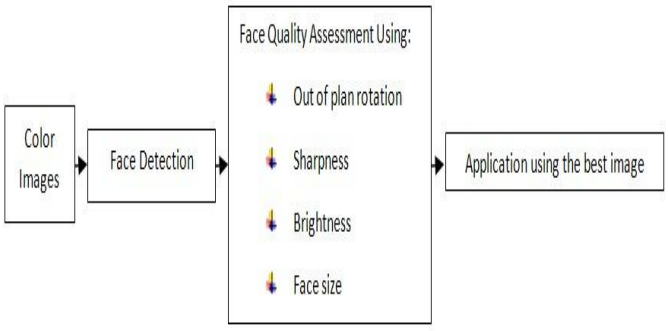

Figure 2: Block diagram of refined face log

Figure 3: Face detection process, from left to right: Input color image, its probability and segmented counterpart and detected faces

C. Facial feature extraction:

For each face region detected by the Face Detection, we use both some of the extracted features from the face detection block, and also new features to assess the quality of them. For each feature we assign a locally computed score so that we can decide which image is the best in terms of quality in the given sequence of images. The following subsections describe the details of these features and the scoring process.

Pose estimation: least out-of-plan rotated face(s):

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. Whenever the rotation of the face increases the difference between these two points increases too.

Given a face in a binary image as shown in figure 3, we calculate the center of mass using the following equation:

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

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.

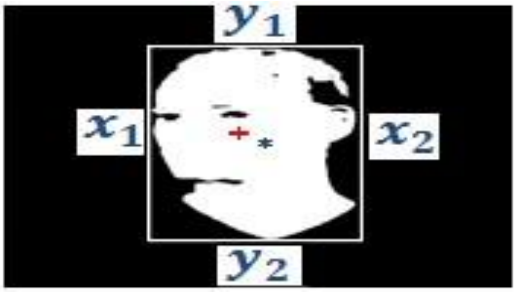

Figure 4: Center of mass (+) and center of the region (*)

Then we calculate the center of the region detected as a face using the equation 2

$$x_c = x_2 - x_1/2 \quad , y_c = y_2 - y_1/2 \qquad (2)$$

where are the right most and the left most pixel in the face region and are the lowest and top pixel, respectively, in this region as shown in figure 4.

						
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

Figure 4.1: A sequence of different head poses and the associated values for the distance and S_1

					
Distance	89.19	35.55	19.60	24.04	79.22
S_1	0.21	0.55	1	0.81	0.24

Figure 4.2: The introduced feature in presence of spectacles and the associated scores

Now we calculate the distance between these two centers as:

$$D=\text{sqrt}((x_c-x_m)^2+(y_c-y_m)^2) \quad (3)$$

The minimum value of this distance in a sequence of images gives us the least outof-plan rotated face as shown in figure 4. To convert this value to a local score in that sequence we use the following equation for each of the images in the sequence:

$$S_1=D_{min}/D \quad (4)$$

where D_{min} is the minimum value of the *D* in the given sequence.

Since the center of mass and the detected region are known from the face detection block the only computation for obtaining this feature is equation 3. The technique used by [27] in order to compute this feature, involves the analysis of gradients to locate the left and right sides of face as well as the vertical position of the eyes. From these values the approximate location of the eyes is estimated and the brightest point between the eyes is expected to lie on the face's axis of symmetry. Their method is not effective when subjects are wearing glasses, or when faces are not close to frontal. While our method is robust in these cases (see the following figure).

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. Following [29], if , be a part of the image which contains the face, # , be the result of applying a low-pass filter to it, then the average value of the pixels of the following equation is the sharpness of the face:

$$E=\text{abs}(a(x,y)-la(x,y)) \quad (5)$$

Since it is difficult, at least computationally, to consider an upper limit for the best value of sharpness for all face images in order to have an acceptable normalization, we have used a local maximum. In this way, after calculating the sharpness for all of the chosen faces we assign the following score to the sharpness of each of them:

$$S_2=E/E_{max} \quad (6)$$

Where E_{max} 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.

Brightness:

Dark images of a face are in general not usable, so we need a way to measure the brightness of the face. Since the region of the face is usually a small region then we can consider the average value of the illumination component of all of the pixels in this region as the brightness of that region. So the following score determines the brightness of the image.

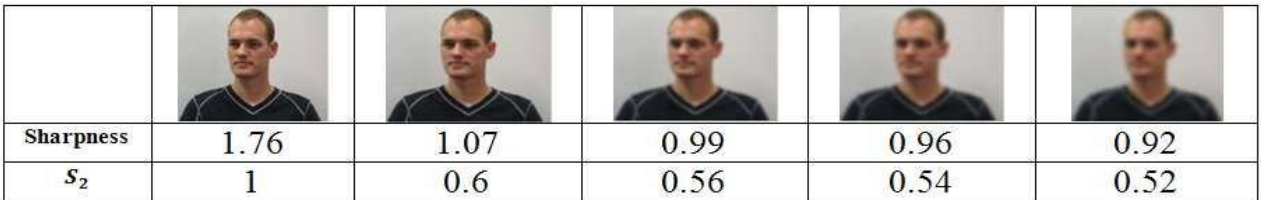

Figure 4.3: An image with different sharpness conditions and the associated scores

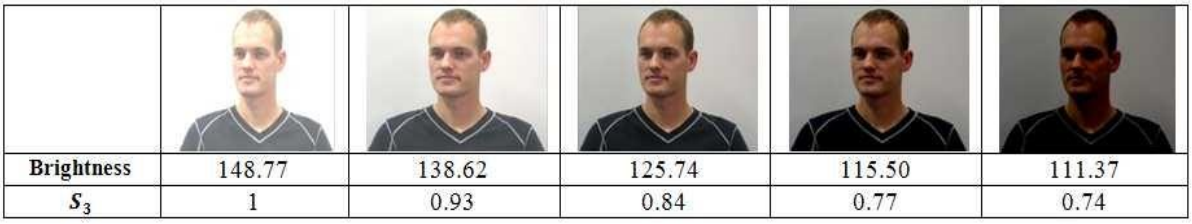

Figure 4.4: .An image with different brightness conditions and their associated score

$$S_3=B/B_{max} \quad (7)$$

Where B_{max} is the maximum value of the brightness in this sequence. Figure 4.4 shows some images of one person in different brightness conditions and their associated scores.

The brighter the image, the higher the score, yields the risk of favoring too bright images. In the real surveillance sequences too bright images are uncommon and in the case of a too bright image a face detector is highly to disregard the face anyway.

Image resolution*:*

Faces with higher resolution can yield better results than lower resolution ones. But it is only true up to a specific limit [27]. This limit depends on the application which is going to use the face after assessing its quality. But usually

considering 50 and 60, respectively, for the width and height of the face is suitable [27,28]. So we can define the score related to the image resolution as follows:

$$S_4=\min\left\{1,\frac{width}{50},\frac{height}{50}\right\} \quad (8)$$

Choosing the best face in a given sequence:

After calculating the four above mentioned features for each of the images in a given sequence, we combine the scores of these features into a general score for each image, as shown in the following equation:

$$S=\frac{\sum_{i=1}^{4} wisi}{\sum_{i=1}^{4} wi} \quad (9)$$

Where ***si*** are the score values for the above features and ***wi***the associated weights for each score. The images are sorted based on their combined scores and depending on the application, one or more images with the greatest values in S are considered as the highest quality image(s) in the given sequence

D. **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. Equation (10) is used to normalize the head-pose value

$$Q_{1X_i}=\frac{P_{\min}}{P_{X_i}} \quad (10)$$

where X_i is the *i*th face image in the given video sequence (*i* is changing from one to the size of the initial face-log, m_1), P_{Xi} is its head-pose value, and Q_{1Xi} is its first quality score. P_{min} is the minimum value of the head-pose feature in the face-log. For normalizing the other three features, we have used

$$Q_{jX_i}=\frac{F_{X_i}}{F_{\max}} \quad (11)$$

where X_i is the *i*th face image in the given video sequence, Q_{Jxi} is the jth quality score of X_i,j = {2,3,4}, F_{xi} is the value of the current feature of X_i, and F_{max} is the maximum value of the current feature in the initial face-log. F_{xi} can be sharpness, brightness or resolution of the ith face image (X_i). Having normalized the above feature values, we use a weighting system to combine them into a quality score for each face. Table I shows the values of the weights of the above features. Obtaining these values is discussed in the experimental results section. The combination is done using (12) in which N_{Xi} is the normalized quality score for the *i*th face image in the intermediate face-log (X_i), Q_{jXi} is its *j*th

quality score feature, and w_j is the weight of this score

$$N_{X_i}=\frac{\sum_{j=1}^{4} W_j Q_{jX_i}}{\sum_{j=1}^{4} W_j}. \quad (12)$$

TABLE 1: Weights of the Facial Features Involved in the FQA

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

III. SUPER-RESOLUTION

SR algorithm cascades both types of reconstruction-based and learning-based SR algorithms. The size of the smallest face that our face detector can detect is 24 × 24. Since the subjects are moving in the video sequence, we have faces of different sizes in the obtained face log. However, in order to be able to apply the RBSR to the images in the refined face-log we resize all of them to 46×56 pixels after face quality assessment. The RBSR produces a HR image of size 92×112 from the LR images of the refined face-log. Then, the proposed system feeds this image (of size 92×112) to the learning-based part to improve its quality even more.

Before applying the RBSR, the images in the refined face log need to be first registered to compensate for their misalignments.

A. Face Image Registration:

The used registration algorithm in this paper is inspired by [4]. This approach takes into account horizontal shift a, vertical shift b and rotation angle θ between the LR images in the refined face-log. Suppose Y is the reference image in the refined face-log and X_i is the ith face image in the log and it is going to register with Y

$$\mathbf{X_i(k,l)=Y(kcos\theta-lsin\theta+a,\ lcos\theta+ksin\theta+b)} \quad (13)$$

Using the first two terms of Taylor's series expansion of sin(θ) and cos(θ) gives

$$\mathbf{X_i(k,l)=Y(k+a-l\theta-k\theta^2/2,l=+b+k\theta-l\theta^2/2)} \quad (14)$$

Expanding X_i to its own Taylor's series gives

$$\mathbf{X_i(k,l)=Y(k,l)+(a-l\theta-k\theta^2/2,l+b+k\theta-l\theta^2/2)\ (\partial y/\partial l)} \quad (15)$$

Therefore, the error between X_i and Y can be written as

$$\mathbf{e(a,b,\theta)=\sum [Y(k,l)+(a-l\theta-(k\theta^2/2))(\partial Y/\partial k)+(b+k\theta-(l\theta^2/2))(\partial Y/\partial l)-X_i(k,l)} \quad (16)$$

The summation is over the overlapping areas of X_i and Y. The minimum of this error can be obtained by computing its derivatives with respect to a, b, and θ and setting them to zero

$$\begin{aligned}\sum Y_k^2 a+\sum Y_k Y_l b+\sum CY_k\theta&=Y_k D_t\\ \sum Y_k Y_l a+\sum Y_l^2 b+\sum CY_l\theta&=Y_l D_t\\ \sum CY_k a+\sum CYb+\sum C^2\theta&=CD_t\end{aligned} \quad (17)$$

where $Y_k=\partial Y/\partial k$, $Y_k=\partial Y/\partial l$, $D_t=X_i-Y$, and $C=kY_l-lY_k$. Solving these equations for a, b, and θ minimizes the difference between the reference image Y and image X_i (ith face image in the refined face-log) warped by (a,b,θ). As it is clearly known [4], the above equations can be obtained under assumptions that are valid only for small displacements. It means that direct use of this registration algorithm in any RBSR algorithm would be extremely inaccurate and unreliable due to the free motion of the objects.

B. Reconstruction-Based Super-Resolution:

In order to reconstruct the HR image from the LR images of the refined face-log we assume that these images have been produced from the HR image by following an imaging model. Based on the imaging model each LR image has been created by warping, blurring and down-sampling the HR image [11]. It means that each X_i,i = {1,2,...,m_3} LR images in the refined face-log have been obtained by

$$X_i = DB_iW_iH + n_i \qquad (18)$$

where D, B_i, and W_i are the down-sampling, blurring, and warping matrix, respectively, H is the HR image and n_i is the introduced noise to the imaging process for producing the ith LR image from the HR image H.

Obtaining the HR image H from (18) is an inverse problem that is extremely ill-posed and ill-conditioned. Following [11] , a MAP method has been employed to obtain the HR response image and a Markov regularization term to convert the problem to a well-posed one. Staking all the m_3 LR image of the refined face-log into a vector X = [X_1,X_2,...,Xm_3], the HR result can be estimated by

$$\mathbf{H^{\wedge} = argmaxp(H|X).} \qquad (19)$$

Using the Bayesian rule, it can be written as

$$\mathbf{H^{\wedge} = argmax}\frac{p(X|H)p(H)}{p(X)}. \qquad (20)$$

Using the fact that LR images X are known (their probability is constant) and they are independent of each other, (20) becomes

Figure 5(a): Result of face log images

Figure 5(b): Result of the refined face log generation

Figure 5(c): Result after applying RBSR algorithm

Figure 5(d): Final extracted image

$$\mathbf{H^{\wedge} = argmax.} \prod_{i=1}^{m_3} p(X_i|H)p(H) \quad (21)$$

Then, using the monotonic logarithm function, we have

$$\mathbf{H^{\wedge} = argmax} \sum_{i=1}^{m_3} [\log p(X_i|H) + \log p(H)]. \quad (22)$$

Now we expand both terms of the above sigma. The image prior p(H) has a Gibbs form

$$p(H) = \frac{1}{C_3} \exp(-\Gamma(H)) \quad (23)$$

where C_3 is a constant and$\Gamma(H)$ is the prior energy function which works as the regularization term. Assuming the noise is i.i.d in (18), the likelihood distribution $p(X_i|H)$ in (22) can be expressed as

$$\log p(X_i|H) = \frac{1}{C_4} \exp\left(-\frac{||X_i - DB_iW_iH||^2}{2\sigma_i^2}\right) \quad (24)$$

where C_4 is a constant and σ_i^2 is the error variance. The MAP response of the system can be obtained by substituting (23) and (24) into (22) which gives

$$H^{\wedge} = \text{argmax} \sum_i [||X_i - DB_iW_iH||^2 + \lambda\Gamma(H)] \quad (25)$$

Where λ is the regularization parameter. This method improves the quality of the LR images in the refined face-log by a factor close to two. For further improvement of this HR image, it is fed to the next step of the system that is an LBSR algorithm.

C. Learning-Based Super-Resolution:

LBSR algorithm used in this paper is a MLP. The benefit of using neural networks is their auto-ability in learning the complicated space of the human faces. For training this MLP we have used around 400 frontal and semi-frontal face images.

To prepare the training data for the network, first, we have extracted the face areas of the HR images and converted them to grayscale. Then, an LR image is created for each of these HR images by down-sampling the images by a factor of two.

Then an LR image is created for each of these HR images by down-sampling the images by a factor of two. Then, all of these LR/HR pairs are fed to the network as the training samples and the network learns the relationship between them.

IV. EXPERIMENTAL RESULTS

The experiment have been carried out on these video sequence, it validates performance of cascaded SR algorithms.Fig.5 shows the result of applying the proposed system to low resolution video sequence.

HR images produced by the proposed system (and sometimes images of even the size of the LR inputs) are more than enough for recognition purposes. However, the recognition is not the main goal of this system. Otherwise, our system could have been biased by the concerns of the employed recognition algorithm. By the way, in order to show that our system is practical, in the fourth experiment, we have employed a face recognition system which is a linear associative memory. The recognizer is trained using the (manually extracted) bestface images of all the 122 video sequences of both databases. To do the experiment, the best images of the sequences are first found by the FQA algorithm. Then, they are resized to the required size of the inputs of the recognition algorithm by bilinear, bicubic, and Schultz *et al.* [5] algorithms. Then, the recognition rates of the face recognizer in these cases are monitored and compared against the case that the input of the recognizer is produced by the proposed system. The face recognizer performs better when it is fed using the HR images produced by the proposed system.

The best image chosen by the FQA algorithm is not necessarily the best one for the recognition. However, it is very likely that best image for the recognition is among the best few images chosen by the FQA algorithm. Therefore, in the second part of the fourth experiment, we have applied the recognition algorithm to the first five best images of the video sequence (which are resized to the required size of the recognition algorithm again by bilinear, bicubic and Schultz *et al.* [5] algorithms). Then, the max rule is applied to the recognition rates and the results are compared against the case in which the input of the recognition algorithm is obtained using the proposed system.

V. CONCLUSION AND FUTURE TRENDS

High quality frontal face images are required for the face analysis systems, but typical surveillance videos are of poor quality. Therefore, we need for a mechanism to bridge between low quality, low-resolution face images from video sequences and their applications in facial analysis system. Super-resolution is one of these mechanisms. The proposed system in this paper deals with the real-world problems of super-resolution systems working with surveillance video sequences. Free movement in such video sequences increases the registration errors and consequently is the main source of error for typical super-resolution systems. To deal with this problem we employed a face-log generation technique that

uses a face quality assessment algorithm to discard useless images in the input video sequence and classify images with similar motions to the same class. It has been shown that it reduces the registration errors and improves the results of the system.

Furthermore, typical reconstruction-based super-resolution systems have an improvement factor close to two. Therefore, for having improvement factors bigger than two, we accompanied the used reconstruction-based algorithm with a recognition-based super-resolution. This increases the improvement factor of the system to almost four.

REFERENCES

[1] V. Bannore, Iterative Interpolation Super Resolution Image Reconstruction. Berlin/Heidelberg, Germany: Springer-Verlag, 2009.

[2] S. Chaudhuri, Super Resolution Imaging, 2nd ed. New York: Kluwer Academic Publishers, 2002.

[3] T. S. Huang and R. Tsai, "Multi-frame image restoration and registration," Adv. Comput. Vis. Image Process., vol. 1, no. 2, pp. 317–339 , 1984.

[4] M. Irani and S. Peleg, "Improving resolution by image registration," Graph. Models Image Process., vol. 53, no. 3, pp. 231–239, 1991.

[5] R. R. Schultz and R. L. Stevenson, "Extraction of high-resolution frames from video sequences," IEEE Trans. Image Process., vol. 5, no. 6, pp. 996–1010, Jun. 1996.

[6] S. Lertrattanapanich and N. K. Bose, "High resolution image formation from low resolution frames using Delaunay triangulation," IEEE Trans. Image Process., vol. 11, no. 12, pp. 1427–1441, Dec. 2002.

[7] A. J. Patti and Y. Altunbasak, "Artifact reduction for set theoretic super resolution image reconstruction with edge adaptive constraints and higher-order interpolants," IEEE Trans. Image Process., vol. 10, no. 1, pp. 179–186, Jan. 2001.\

[8] M. E. Tipping and C. M. Bishop, "Bayesian image super-resolution," Adv. Neural Inform. Process. Syst., vol. 15, pp. 1303–1310, 2002.

[9] X. Li, Y. Hu, X. Gao, D. Tao, and B. Ning, "A multi-frame image super-resolution method," Signal Process., vol. 90, no. 2, pp. 405–414, 2010.

[10] L. Zhang, H. Zhang, H. Shen, and P. Li, "A super-resolution reconstruction algorithm for surveillance images," Signal Process., vol. 90, no. 3, pp. 848–859, 2010.

[11] S. Baker and T. Kanade, "Hallucinating faces," in Proc. IEEE Int. Conf. AFGR, Mar. 2000, pp. 83–88.

[12] H. Chang, D.-Y. Yeung, and Y. Xiong, "Super-resolution through neighbor embedding," in Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 1. Jun.–Jul. 2004, pp. 275–282.

[13] O. Arandjelovic and R. Cipolla, "A manifold approach to face recognition from low quality video across illumination and pose using implicit super-resolution," in Proc. ICCV, 2007.

[14] B. Li, H. Chang, S. Shan, and X. Chen, "Aligning coupled manifolds for face hallucination," IEEE Signal Process. Lett., vol. 16, no. 11, pp. 957–960, Nov. 2009.

[15] J. Yang, J. Wright, T. Huang, and Y. Ma, "Image super-resolution as sparse representation of raw image patches," in Proc. Int. Conf. CVPR, 2008.

[16] Y. Huang and Y. Long, "Super-resolution using neural networks based on the optimal recovery theory," J. Comput. Electron., vol. 5, no. 4, pp. 275–281, 2007.

[17] C. Miravet and F. B. Rodriguez, "A two-step neural-network based algorithm for fast image super-resolution," Image Vis. Comput., vol. 25, no. 9, pp. 1449–1473, 2007.

[18] V. H. Patil, D. S. Bormane, and V. S. Pawar, "Super resolution using neural network," in Proc. IEEE 2nd Asia Int. Conf. Modeling Simul., May 2008, pp. 492–496.

[19] D. Glasner, S. Bagon, and M. Irani, "Super-resolution from a single image," in Proc. ICCV, 2009.

[20] A. Fourney and R. Laganiere, "Constructing face image logs that are both complete and concise," in Proc. 4th IEEE Can. Conf. Comput. Vis. Robot Vis., May 2007, pp. 488–494.

[21] K. Nasrollahi and T. B. Moeslund, "Complete face logs for video sequences using quality face measures," IET Int. J. Signal Process., vol.3, no. 4, pp. 289–300, Jul. 2009.

[22] P.Viola and M. Jones, "Robust real time face detection," Int. J. Comput. Vis., vol. 57, no. 2, pp. 137–154, 2004.

[23] Facegen Modeler [Online]. Available: http://www.facegen.com/ modeller.htm

[24] Image processing (ICIP),pp 2973—2976, IEEE Press,(2004)

[25] H. Fronthaler, K. Kollreider, J. Bigun, Automatic Image Quality Assessment with Application in Biometrics, In: International Conference on Computer Vision and Pattern Recognition, IEEE Press, (2006)

[26] G. Xiufeng, Z. Stan, L. Rong, P. Zhang, Standardization of Face Image Sample Quality, In: ICB, pp. 242--251, Springer-Verlag, Berlin (2007)

[27] A. Fourney, R. Laganiere, Constructing Face Image Logs that are Both Complete and Concise, In: 4th Canadian Conference on Computer Vision and Robot Vision, IEEE Press, Canada (2007)

[28] F. Weber, "Some quality measures for face images and their relationship to recognition performance", In: Biometric Quality Workshop. National Institute of Standards and Technology, Maryland, USA, March 8-9 (2006).