

Super-Resolution Pipeline for Fast Adjudication in Watchlist Screening

Vitaliy Tayanov¹, Eric Granger¹, Miguel Bordallo² and Abdenour Hadid²

¹ École de technologie supérieure, Université du Québec, Montreal, Canada

² Center for Machine Vision Research, University of Oulu, Oulu, Finland

Abstract—Although still-to-video face recognition is an important function in watchlist screening, state-of-the-art systems often yield limited performance due to camera inter-operability and to variations in capture conditions. Therefore, the visual comparison of faces captured in unconstrained low-quality videos against a matching high-quality reference facial still image captured under controlled conditions is required in many surveillance applications to limit the number of costly false matches. To improve the visual appearance of faces captured in videos, this paper presents a new super-resolution (SR) pipeline that is suitable for fast adjudication of face-matches produced by an automated system. In this pipeline, face quality measures are used to rank and select face captures belonging to a facial trajectory, and multi-image SR iteratively enhances the appearance of a super-resolved face image. Face selection is optimized and registered using graphical models. Experiments with the Chokepoint dataset show that the proposed pipeline efficiently produces super-resolved face images by ranking best quality ROIs in a trajectory. To select the best face captures for SR, this pipeline exploits a strong correlation existing between pose and sharpness quality measurements.

Keywords—Video Surveillance, Face Recognition, Visualisation, Super-Resolution, Image Quality.

I. INTRODUCTION

In order to enhance security and situational awareness, many public security organizations have deployed video surveillance technologies at e.g., airport inspection lanes, portals and border crossings. Watchlist screening is a key application in video surveillance [1], [2]. In such applications, a target individual is enrolled to a still-to-video face recognition system by isolating facial regions of interest (ROIs) within reference still images captured a priori under controlled conditions. The facial model of an individual is designed using one or more discriminant reference patterns extracted from ROIs. During operations, video streams are captured over a network of video surveillance cameras. ROI patterns corresponding to faces detected in frames are matched against the facial models of each target individual in the watchlist. Finally, a tracker follows the movement of different persons appearing in the camera viewpoint, across consecutive frames, and the system combines matching scores over facial trajectories¹ for spatiotemporal recognition of target individuals.

Although several powerful techniques have been proposed to recognize frontal faces in still images, watchlist screening presents a challenging problem because these systems must provide accurate and timely face recognition in changing video

¹A trajectory is a set of ROIs that correspond to the same high-quality track of an individual, and that is linked to a face-match report.

surveillance environments [3], [4]. The appearance of faces in videos vary due to uncontrolled captured conditions (variations in pose, expression, scale-resolution, illumination, blur, etc.), and to camera inter-operability. In addition, the facial model of a target individual is often a poor representative of faces to be recognized during operations because they are designed during enrolment using few reference stills.

State-of-the-art commercial technologies have limited robustness to the challenges described above, and rarely leverage spatiotemporal information that is abundant in videos [5]. The majority of face recognition research has focused on improved matching performance, for instance, through multiple face representations, synthetic generation of reference stills, and adaptation of facial models [3], [6]. Some specialized academic systems have also been proposed in recent years for face recognition in video surveillance. Most notably, Li *et al.* proposed a Transduction Confidence Machine kNN for open-set face recognition [7], and Kamgar-Parsi *et al.* proposed a morphing approach that generates synthetic reference ROI patterns for robust matching [2]. Finally, several recent systems are modeled as individual-specific detectors, each one implemented using classifiers designed with reference ROIs from target and non-target individuals [8], [4]. A multi-classifier framework is proposed in [3] to implement such detectors based on multiple different patches and descriptors extracted from a single reference face image.

State-of-the-art technologies do not provide the levels of performance needed by public safety organizations for fully automated watchlist screening in real-world environments [3], [5]. Still-to-video face recognition systems may yield several false matches and non-matches errors. However, these systems can still be beneficial for decision support, especially in medium- to large-scale video surveillance networks, and in cluttered scenes, where many faces are matched every day to a large watchlist. Through an adjudication process², potential matches of the system may be confirmed by a human operator. The number of false matches may therefore be reduced while maintaining an appropriate level of security.

Adjudication is a challenging task because the appearance of faces captured in video can differ considerably from reference stills in the watchlist and between one another in a facial trajectory. Some visualisation tools are available commercially to assist an analyst with face matching, but they are designed

²Adjudication is the visual comparison of probe faces captured in videos against a matching reference still image.

primarily for forensics or other non-real time applications, or for matching of cooperative face captures [9]. Specialized visualization tools are required to support the fast adjudication of face-match reports (produced by an automated system) in real-world watchlist screening applications, where the faces are captures without cooperation.

The principal novelty presented in this paper is a new super-resolution (SR) pipeline to improve the visual appearance matched captured in videos. Given a facial trajectory produced by an automated Face recognition system, this pipeline first employs face quality assessment and optimization of a graphical model to rank its ROIs. During each iteration, the highest quality ROIs that remain in the trajectory is dewarped, registered, and upscaled. Then, a multi-image SR algorithm is employed to enhance the appearance of a face image. This iterative process ends when the quality of the super-resolved face image no longer improves.

A challenge of SR algorithms from multiple images is that they usually involve complex iteration processes. To compose a super resolved image, several facial ROI must be selected from a facial trajectory. In a Greedy search approach, each iteration should explore the benefit of composing the image with all the possible combinations of ROIs, in a costly iterative process. In contrast, our proposed pipeline selects and registers the ROIs a priori by using quality measurements.

Faces in video surveillance may be in motion and out of focus, and segmentation is only effective at a so called "working distance." A challenge of applying SR in watchlist screening is the subjective nature of quality enhancement. Face quality measurements can rely on full, reduced or no reference image. Since the characteristics and quality of ROIs isolated in a reference still images may differ considerably to that of ROIs in a trajectory, no-reference (NR) measures are employed for ranking and selection. Finally, a specialized target-based evaluation methodology was used to assess the benefit of proposed pipeline using the Chokeypoint dataset [10] in which videos are captured under semi- and uncontrolled conditions.

II. ADJUDICATION IN WATCHLIST SCREENING

This paper proposes the combination of efficient ROI visualization algorithms to support watchlist screening with the fast adjudication of automatically produced face matches. One assumes that the system captures ROIs along trajectories for matching with face models stored in the watchlist. When a person k is automatically matched to a watchlist individual, the operator may display and compare probe ROIs from trajectory T_k alongside a reference still ROI of that individual. To facilitate adjudication, different types of visualisation functions may be considered: the detection of facial components and zones of interest with high information content such as facial marks and areas with high contrast and brightness. in probe ROIs for local face comparison. This paper focuses on enhancing the appearance of probe ROIs captured in trajectory T_k through the employment of a novel super-resolution pipeline. In par-

ticular, face quality of ROIs in T_k is exploited in order to analyse, rank and select ROIs for the SR process. Following subsections will provide additional background on face quality assessment and image SR algorithms.

A. Quality Assessment:

Quantitative image quality measures (QMs) have been exploited to control the quality of biometric samples, as well as to select a biometric modality, algorithm, and/or system parameters [11]. In this paper, QMs are considered to analyse, rank and select ROIs in a T_k . For instance, the highest quality ROI in a T_k must be selected at the outset based on some measure. Quality assessment may be performed using image- and face-specific measures, and may rely on reference images.

Several face image quality standards have been proposed for spatial quality assessment [12], but they assume frontal views under controlled conditions. QMs can generally be divided into measures of image-specific qualities (contrast, sharpness, compression artefact, etc.) and face-specific qualities (geometry, pose angle, eye visibility, etc.). In addition, there are 3 categories for objective image QMs depending on the availability of reference ground truth information [13]. Spatial full-reference (FR) algorithms are provided with an original undistorted reference. With reduced-reference (RR) approaches, the algorithm is provided with distorted or poor quality images and some additional prior information on the original undistorted references for quality assessment. Finally, no-reference (NR) algorithms are used without a reference image.

QMs that rely on a reference image (i.e., watchlist stills) are often considered as image distortion measures. One well-known example is the Structural Similarity index (SSIM) [14] that measures the distortion between probe and reference images as a combination of three factors: loss of correlation, luminance distortion and contrast distortion. Another more recent example is Visual Information Fidelity (VIF) index [15] that estimates the mutual information of two images in the spectral domain.

Difficulties in quality assessment lead to issues in image resolution enhancement. One challenge in watchlist screening is camera inter-operability and that fact that reference still ROIs or prior information may not be representative of ROIs captured in a camera viewpoint during operations. In addition, with full-reference measures, finding the relations between distortions relies on a training process and verification by involving subjective adjudications [13]. Reduced-reference measures therefore provide an interesting compromise. However, it is preferable for SR to consider quality measures that estimate pose, sharpness, brightness, etc., that do not rely on references information [16].

B. Super-Resolution:

SR algorithms allow to infer a high-resolution image from one or more lower-resolution ones. They have been used for

many different applications, such as satellite and aerial imaging, medical image processing, video enhancement and facial analysis. In this paper, SR algorithms are employed to produce enhanced higher-resolution face images (for comparison to a watchlist ROI) based on multiple lower-resolution ROIs in a T_k . Prior to comparison with watchlist ROIs, one or multiple lower-resolution ROIs from T_k are registered, and higher-resolution face images is inferred using a spatial or frequency domain algorithm.

Multi-image SR algorithms are mostly reconstruction-based. They address the aliasing artifacts that are present in low-resolution images (due to under-sampling process) by simulating the image formation model. In contrast, most single-image SR algorithms rely on the prior learning-optimization of a set of training images to estimate the mapping function between low- and high-resolution images [17]. In the spatial domain, multi-image SR algorithms map from low- to high-resolution using two main steps: registration (image alignment and correction) and reconstruction (image fusion) based on the estimation of the degradation function. Other probabilistic spatial domain algorithms learn data sets to map from low- to high-resolution images.

The ROIs captured in a T_k have lower resolution than watchlist image, and this face scale often varies over time as the person moves in the scene. Once upscaled to a common size, probe ROIs appear noisy against a watchlist ROI. In this paper, multiple-image SR algorithms in the spatial domain will be developed to enhance probe ROIs in a T_k (set of low-resolution images). There are however several challenges to implementing a cost-effective technique for SR in watchlist screening. The accuracy of reconstruction-based algorithms is highly dependent on the registration process (motion estimation, correction and alignment between the low-resolution images). This may pose challenges in real-world environments since ROIs in trajectories may vary considerably over time, and provide few good quality ROIs for SR. Ranking and selecting the most representative probes ROIs, facial components and zones is an important first step of multiple-image SR.

III. VISUALISATION PIPELINE

Fig.1 shows a block diagram of the proposed SR pipeline. Using images with low quality and miss-aligned ROIs for reconstruction (fusion) leads to a decline in quality of the resulting super-resolved face image. Therefore, ROIs captured in a trajectory $T_k = \{ROI_1, ROI_2, \dots, ROI_n\}$ are initially ranked into T'_k according to different face-specific no-reference QMs. The overall measurement, QM_{SR} , is given by a weighted sum of individual QMs relevant to video surveillance. Face pose, resolution and sharpness QMs from [16] are considered in this paper.

Depending on the number n of available ROIs in T'_k , either single- or multiple-image SR algorithm are applied to produce a final super-resolved face image. This image is used for visualisation purposes.

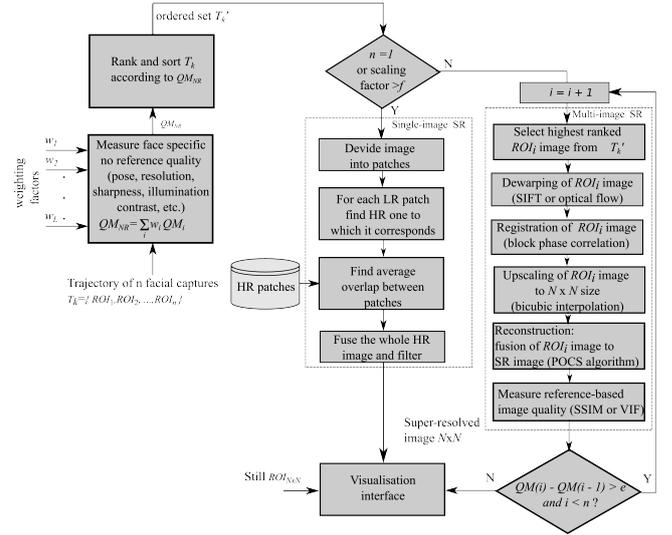


Fig. 1: Block diagram of visualisation algorithm.

Single-image algorithms can not use motion information from the trajectory of face images T'_k . That is why only approach is to use learning techniques to learn dependencies between low and high resolution patches (learning step). Then one should find the appropriate high-resolution patch corresponding to the low-resolved one in terms of some minimal distance (recognition step). The principle advantages of these algorithms are that they are not sensitive to the scaling factor and the problem is not ill-posed in this case. However such algorithms require training data with appropriate resolutions for each scaling factor and are computationally costly.

In the multi-image case, the highest face quality ranked ROI_i is initially selected among remaining ones in T'_k . Then, this ROI undergoes dewarping according to the still ROI, using a SIFT or optical flow algorithm. Dewarping is required to have a head orientation similar to that on the still image. For larger changes techniques based on 3D modelling or Active Appearance (Shape) models [21] are adequate. This should facilitate face comparison due to similarity in head orientation. The image registration algorithm is based on phase correlation (PC). It has a number of useful properties. Among them is robustness against blur and noise, and low computational complexity. It is suitable for smooth surfaces, lack of details, low number of feature points [18]. Image fusion is needed to enhance face resolution which can be measured by VIF index. Among all, VIF index shows how many high frequency components mutually shared between reference image and fused super-resolved one. That is why VIF can be used to control quality of super-resolved image measured against still image. A simple way to upscale ROI_i to the same $N \times N$ size as the still ROI is via interpolation (bi-cubic interpolation). However this does not provide real resolution enhancement because it does not add high frequency components to the image spectrum.

Due to computational complexity of single-image algo-

rithms and very rare case when only one image is available, one is going to apply only multi-image SR algorithm. For single-image case (if any) image is upsampled using some interpolation algorithm.

Projection onto Convex Sets (POCS) is employed as an iterative reference reconstruction (fusion) algorithm and is suitable for real time applications [17]. It is based on defined constraints which are convex sets to restrict the space of the super-resolved image. Restrictions include smoothness, positivity, bounded energy, reliability, etc. A permitted solution corresponds to the intersection of these particular sets. By iteratively applying this approach, a solution is obtained as the intersection of such constraints. POCS is compared with Papoulis-Gerchberg (PG) algorithm [19]. PG algorithm uses low-pass filter to fill up zero values of unknown pixels after creating of SR grid. This leads also to known values modification. Then known values are set to original values and that creates high-frequency components in image spectrum. After projection to the low-resolution grid and subtracting from the original low-resolved image, one obtains a residual. This residual is used to modify super-resolved image for the next iterations. As seen, both do not use any regularisation. The main difference is using constraints for residual and spatial filtering (in case of POCS) and approaching the predefined value of residual and low-pass filtering (in case of PG).

The reconstruction part of SR algorithm is optimised using bilateral or Wiener filtering [20] after each image fusion process to preserve contours or deblur an image. These filters are computationally costly. Using them after each image addition (but not during restoration) allows to significantly reduce the complexity of the image restoration process.

Ranking is also used to limit the number of SR iterations (ROIs). During image fusion one adds ROIs from the ranked set T'_k , and computes a super-resolved face image based on QMs (full-reference and no-reference based) in current and previous iterations. If the difference in quality exceeds a predefined threshold one continues fusing ROIs. Otherwise, the fusion process stops, and the final super-resolved face image is obtained. In comparison with classical image fusion where all ROIs are used at the same time, iterative fusion by adding one image after each previous fusion and use mentioned QMs as indicators of super-resolution quality allows to find the optimal solution.

Image selection algorithm has 3 options: no-reference, full-reference and time-based (TB) QM. No-reference QM based algorithms select first images from the ranked set T'_k . Ranking is done by total no-reference QM which is weighted sum (or product) of particular QMs. Full-reference based algorithm selects first best image from the ranked sequence and then one ranks one more time images against the full-reference-based QM (SSIM, VIF, etc. [14], [15]). Then first images should be selected to the subset used for image fusion. TB approaches select the first best image and their neighbours for fusion. Among all these image selection options, the less computationally demanding is no-reference and time-based

because they do not need to do one more ranking based on full-reference QMs (as in case of full-reference option).

To describe the potential efficiency of the proposed pipeline, consider the greedy search algorithm. In this context, one-level greedy search means that, at the very beginning, best image should be selected and then combined with $n - 1$ ROIs left in trajectory T'_k , giving $n - 1$ pairs to produce a reconstructed SR image. After that $n - 2$ images are combined with best pair of images and so on until there is no improvement by FR QM. The complexity of this greedy search is high because, at each iteration, all remaining ROIs in T'_k are tentatively combined with the SR image before selecting the best outcome. Given the difficulties with image differences, image alignment must be performed prior to image reconstruction. To address this issue, the proposed pipeline considers three options of image ranking and then select the optimal one to give a good approximation of the best solution. Within the pipeline, selection is based on the quality of low-resolution images. This reduces algorithmic complexity compared to combinatorial way used to achieve of optimal search solution. The idea of using such kind of image selection options is based on the fact that adding images with lower quality to the previous fusion set means that super-resolved image will be also of a lower quality (the higher quality of ROIs used for fusion the better quality of SR image will be).

IV. OPTIMISATION ROUTINE

Using NR and FR measures, image selection and registration techniques should be conjointly optimised. Such optimisation is desirable because there are three options for selection and two registration algorithms. During optimisation routine the images to be fused should both provide the best quality and appropriate displacements (rotations). No-reference QMs are normalised for each person in the scene by the maximum value during SR iterations. This allows improvements during iterations. Usually the quality of captures is not high and due to that it is more important to see changes in total no-reference QMs. VIF index is not normalised to see the coherence between SR image and still one. The principal idea of optimisation routine is to find the most probable values of image quality measures during image selection that maximise the probability of given or optimal value of VIF (SSIM) using definite registration and fusion techniques. This can be performed using probabilistic graphical model (GM) [22]. However this optimisation process needs wide learning for finding such dependencies. That is why this problem should be solved in several steps. First of them is considered here where one estimates dependencies between image selection and registration techniques based on PC and 2 image fusion techniques (POCS and PG). Also influence of different image quality measures on image selection quality is considered. Thus, GM can be presented as follows:

$$p(QM, IR, IS) = p(QM|IR, IS) \cdot p(IR, IS),$$

where $p(QM, R, IS)$ is a joint probability of FR (VIF index [15]) and total NR QMs, image registration and image selection. Here $QM \in [0, 1]$, $IR = \{1, 2\}$ and $IS = \{1, 2, 3\}$ are the numbers of image registration algorithms and image selection options. The total number of combinations is $|IR \times IS| = 6$. Because $p(IR, IS)$ has a constant value of $1/6$, only $p(QM|R, IS)$ is optimised. The optimisation routine is given as follows

$$\{IR_0, IS_0\} = \arg \max_{IR, IS} \{p(QM|IR, IS)\}.$$

It is important to establish the combination of which QMs (pose, sharpness and resolution) has the most significant impact on the total no-reference and full-reference based QM. As in the previous case, the joint probability distribution of all measures could be presented as follows

$$p(VIF, p, s, r) = p(VIF|p, s, r)p(p, s, r),$$

where p , s , and r are pose, sharpness and resolution NR-based QM, respectively. Let assume $\{p, s, r\} = \{0, 1\}$ depending if measure is included (1) or not (0) in total QM. Thus $p(p, s, r)$ is equal to $1/7$ (case when $\{p, s, r\} = 0$ is skipped) for every combination of values p, s and r . So just value of $p(VIF|p, s, r)$ is maximised. The optimal values of p, s and r are found as follows:

$$\{p_0, s_0, r_0\} = \arg \max_{p, s, r} \{p(VIF|p, s, r)\}.$$

The maximum value of $p(VIF|p, s, r)$ is equivalent to the maximum value of one of seven combinations.

In a typical adjudication process, an operator would upscale and compare the best quality video ROI from a trajectory to the reference still ROI.

V. EXPERIMENTAL RESULTS AND ANALYSIS

Experimental results are obtained using Chokeypoint dataset, where 29 subjects are captured while walking through portals, and also pipeline algorithms shown in Fig. 1. In real-world applications capturing conditions might be much different and challenging. Notwithstanding experiments show some dependencies between image selection options, dewarping, registration and fusion.

The optimal result is obtained for NR image selection combined with separable version of PC. Into account one took number of images needed for image fusion and also the subjective and objective coherence of the image quality. There is almost no difference between NR and FR-based image selection options. Also NR-based image selection option requires only one ranking instead of FR-based that requires two of them.

Table 1 shows 5 different combinations with no-reference QMs. Based on the average total no-reference QM and VIF, the best results have been achieved for $p = 0, s = 1, r = 1$. During experiments two more cases $p = 1, s = 0, r = 1$ and

$p = 0, s = 0, r = 1$ were eliminated due to very poor quality of the SR image and insolvency of pose detection for these cases.

Results in Table 1 show the significant improvement of using the proposed SR pipeline in terms of both no-reference and full-reference QMs. It should be noticed that resolution is not coherent with sharpness and pose on videos with translational motion (toward the camera). Also, resolution does not penalise the yaw and pitch as well as motion and optical blur, thus visualization result may be poor. On the other hand pose is not a real image quality measure. Based on obtained results, the best strategy is to select frontal views from appropriate cameras and then to take captures with optimal values of s and r . Thus dimensionality of QM space is reduced.

Table 1: Average no-reference , full-reference and time based QMs for each the three first iterations during SR image fusion process on the Chokeypoint data.

		NR-Separable PC		p (NR-Separable PC)	
i	Total NR QM	VIF	Total NR QM	VIF	
1	0.340 ± 0.091	0.204 ± 0.019	0.327 ± 0.089	0.198 ± 0.022	
2	0.971 ± 0.059	0.304 ± 0.038	0.965 ± 0.059	0.296 ± 0.051	
3	0.974 ± 0.052	0.306 ± 0.039	0.966 ± 0.051	0.297 ± 0.053	
		NR-Subpixel PC		s (NR-Separable PC)	
i	Total NR QM	VIF	Total NR QM	VIF	
1	0.236 ± 0.076	0.204 ± 0.019	0.340 ± 0.0900	0.204 ± 0.021	
2	0.763 ± 0.149	0.292 ± 0.062	0.964 ± 0.0604	0.307 ± 0.044	
3	0.964 ± 0.046	0.278 ± 0.070	0.978 ± 0.0469	0.309 ± 0.045	
		FR-Separable PC		p and s (NR-Separable PC)	
i	Total NR QM	VIF	Total NR QM	VIF	
1	0.329 ± 0.088	0.204 ± 0.019	0.330 ± 0.088	0.201 ± 0.018	
2	0.968 ± 0.058	0.305 ± 0.038	0.971 ± 0.059	0.282 ± 0.088	
3	0.975 ± 0.052	0.306 ± 0.039	0.970 ± 0.051	0.303 ± 0.044	
		FR-Subpixel PC		s and r (NR-Separable PC)	
i	Total NR QM	VIF	Total NR QM	VIF	
1	0.236 ± 0.080	0.204 ± 0.019	0.345 ± 0.077	0.207 ± 0.018	
2	0.697 ± 0.145	0.292 ± 0.058	0.970 ± 0.049	0.317 ± 0.038	
3	0.923 ± 0.136	0.283 ± 0.071	0.991 ± 0.017	0.327 ± 0.038	
		TB-Separable PC		p, s and r (NR-Separable PC)	
i	Total NR QM	VIF	Total NR QM	VIF	
1	0.344 ± 0.091	0.204 ± 0.019	0.340 ± 0.091	0.203 ± 0.019	
2	0.968 ± 0.056	0.298 ± 0.048	0.971 ± 0.059	0.304 ± 0.038	
3	0.967 ± 0.054	0.300 ± 0.049	0.974 ± 0.052	0.306 ± 0.039	

Figure 2 shows the visual results of the image fusion process. The image selection technique utilises no-reference quality measurements and the registration is performed using a separable version of phase correlation. The image shows how the quality of the image improves with the first two iterations.

Figure 3 shows the visual results of SIFT flow based dewarping compared with the reference image.

As seen from the obtained results in Table 2 SIFT dewarping produces images with better similarity (even after some degradation due to image alignment) to the still image (based on VIF). This is due to similar head orientation.



(a) first best ROI (b) two best ROIs (c) three best ROIs

Fig. 2: Results of the fusion of several ROIs in a trajectory. The quality measurements are improved with the first two iterations.



(a) dewarped ROI (b) Still ROI

Fig. 3: A dewarped image to match with the visualization reference.

Table 2: SIFT dewarping with PC vs PC (NR-based image selection) for each the three first iterations during SR image fusion process on one example from Chokeypoint data.

i	NR-SIFT-Separable PC		NR-Separable PC	
	Total NR	QM VIF	Total NR	QM VIF
1	0.508	0.298	0.601	0.180
2	1.000	0.332	0.701	0.191
3	1.000	0.333	0.987	0.191

VI. CONCLUSION

In this paper, a super-resolution pipeline has been proposed for fast adjudication in watchlists screening. Proposed image selection options approximate the best results obtained by greedy search. Correlation between sharpness and pose was established for translational motion. This gives a possibility to perform the dimensionality reduction in QM space using optimisation routine based on graphical models. Appropriate bilateral filtering allows to preserve image contours in a computationally easy way. In experimental results with Chokeypoint dataset, for POCS the optimal solution (in terms of visualisation) is obtained after 3 iterations during image fusion and for PG—after 2 iterations. With some limitations SIFT alignment enhance visual similarity due to similar head orientation of compared still and aligned images. Iterative fusion allows to obtain shorter computation and visualization time in comparison with classical scheme.

ACKNOWLEDGMENT

The financial supports of the Natural Sciences and Engineering Research Council of Canada and the Academy of Finland are fully acknowledged.

REFERENCES

- [1] Jeremiah R. Barr, Kevin W. Bowyer, Patrick J. Flynn, and Soma Biswas, "Face recognition from video: a review," *Int'l Journal of Pattern Recognition and Artificial Intelligence*, vol. 26, no. 5, 2012.
- [2] Behrooz Kamgar-Parsi and Wallace Lawson, "Toward development of a face recognition system for watchlist surveillance," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 33, no. 10, pp. 1925–1937, 2011.
- [3] Saman Bashbaghi, Eric Granger, Robert Sabourin, and Guillaume-Alexandre Bilodeau, "Watchlist screening using ensembles based on multiple face representations," 2014.
- [4] Christophe Pagano, Eric Granger, Robert Sabourin, Gian-Luca Marcialis, and Fabio Roli, "Adaptive ensembles for face recognition in changing video surveillance environments," *Information Sciences*, vol. 286, pp. 75–101, 2014.
- [5] Lacey Best-Rowden, Brendan Klare, Joshua Klontz, and Anil K Jain, "Video-to-video face matching: Establishing a baseline for unconstrained face recognition," in *Biometrics: Theory, Applications and Systems (BTAS), 2013 IEEE Sixth International Conference on*. IEEE, 2013, pp. 1–8.
- [6] Xiaoyang Tan, Songcan Chen, Zhi-Hua Zhou, and Fuyan Zhang, "Face recognition from a single image per person: A survey," *Pattern Recognition*, vol. 39, no. 9, pp. 1725–1745, 2006.
- [7] Fayin Li and Harry Wechsler, "Open set face recognition using transduction," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 27, no. 11, pp. 1686–1697, 2005.
- [8] Miguel De-la Torre, Eric Granger, Paulo Radtke, Robert Sabourin, and Dmitry O Gorodnichy, "Self-updating with facial trajectories for video-to-video face recognition," 2014.
- [9] Anil K Jain, Brendan Klare, and Unsang Park, "Face matching and retrieval in forensics applications," *IEEE MultiMedia*, vol. 19, no. 1, pp. 20, 2012.
- [10] Yongkang Wong, Shaokang Chen, Sandra Mau, Conrad Sanderson, and Brian C Lovell, "Patch-based probabilistic image quality assessment for face selection and improved video-based face recognition," in *Computer Vision and Pattern Recognition Workshops (CVPRW), 2011 IEEE Computer Society Conference on*. IEEE, 2011, pp. 74–81.
- [11] Harin Sellaheewa and Sabah A Jassim, "Image-quality-based adaptive face recognition," *Instrumentation and Measurement, IEEE Transactions on*, vol. 59, no. 4, pp. 805–813, 2010.
- [12] Jitao Sang, Zhen Lei, and Stan Z Li, "Face image quality evaluation for iso/iec standards 19794-5 and 29794-5," in *Advances in Biometrics*, pp. 229–238. Springer, 2009.
- [13] Anush Krishna Moorthy and Alan Conrad Bovik, "Blind image quality assessment: From natural scene statistics to perceptual quality," *Image Processing, IEEE Transactions on*, vol. 20, no. 12, pp. 3350–3364, 2011.
- [14] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli, "Image quality assessment: from error visibility to structural similarity," *Image Processing, IEEE Transactions on*, vol. 13, no. 4, pp. 600–612, 2004.
- [15] Hamid R Sheikh and Alan C Bovik, "Image information and visual quality," *Image Processing, IEEE Transactions on*, vol. 15, no. 2, pp. 430–444, 2006.
- [16] Ali J. Abboud and Sabah A. Jassim, "Biometric templates selection and update using quality measures," 2012, vol. 8406, pp. 1–9.
- [17] Kamal Nasrollahi and Thomas B Moeslund, "Super-resolution: A comprehensive survey," *Machine Vision & Applications*, 2014.
- [18] Patrick Vandewalle, Sabine Sü, Martin Vetterli, et al., "A frequency domain approach to registration of aliased images with application to super-resolution," *EURASIP Journal on Advances in Signal Processing*, vol. 2006, 2006.
- [19] Priyam Chatterjee, Sujata Mukherjee, Subhasis Chaudhuri, and Guna Seetharaman, "Application of papoulis–gerchberg method in image super-resolution and inpainting," *The computer journal*, vol. 52, no. 1, pp. 80–89, 2009.
- [20] John C Russ and J Christian Russ, *Introduction to image processing and analysis*, CRC Press, Inc., 2007.
- [21] Frederick W Wheeler, Xiaoming Liu, and Peter H Tu, "Multi-frame super-resolution for face recognition," in *Biometrics: Theory, Applications, and Systems, 2007. BTAS 2007. First IEEE International Conference on*. IEEE, 2007, pp. 1–6.
- [22] Daphne Koller and Nir Friedman, *Probabilistic graphical models: principles and techniques*, MIT press, 2009.