A Preliminary Investigation on the Sensitivity of COTS Face Recognition Systems to Forensic Analyst-style Face Processing for Occlusions

Felix Juefei-Xu*, Dipan K. Pal*, Karanhaar Singh*, and Marios Savvides
CyLab Biometrics Center, Electrical and Computer Engineering
Carnegie Mellon University, Pittsburgh, PA 15213, USA
{juefeix,dipanp,karanhaa}@andrew.cmu.edu, msavvid@ri.cmu.edu

Abstract

*Modern day law enforcement banks heavily on the use of commercial off-the-shelf (COTS) face recognition systems (FRS) as a tool for biometric evaluation and identification. However, in many real-world scenarios, when the face of an individual is occluded or degraded in some way, commercial recognition systems fail to accept the face for evaluation or simply return unusable matched faces. In these kinds of cases, forensic experts rely on image processing techniques and tools, to make the face fit to be processed by the commercial recognition systems (*e.g. *use partial face images from another subject to fill in the occluded parts of the face of interest, or have a tight crop around the face). In this study, we evaluate the sensitivity of commercial recognition systems to such forensic techniques. More specifically, we study the change in the rank-1 identification result that is caused by forensic processing of faces-of-interest that are unusable by the commercial recognition systems. Further, forensic processing of such faces is more of an art and it is extremely difficult to process faces consistently such that there is a predictable effect on the rank-n identification result. This study is meant to serve as an evaluation of the effect of a few forensic techniques intended to allow commercial recognition systems to process and match face images that were otherwise unusable. Our results indicate that COTS FRS can be sensitive to the subjectivity in facial part swapping and cropping, resulting in inconsistencies in the identification rankings and similarity scores.*

1. Introduction

With technological and infrastructure advances in surveillance, developing face recognition technology effective in-the-wild is quickly gaining importance. Nonetheless, today's commercial off-the-shelf (COTS) face recognition systems (FRS) offer themselves as powerful tools for law enforcement agencies for narrowing down suspects from humongous government databases. When combined with subjective and manual evaluations, forensic experts and law enforcement agencies are able to make important breakthroughs. Although there still exist many challenges in designing FRS for images in-the-wild with variations in pose, illumination, expression and aging, COTS FRS are starting to become more robust in dealing with such variations. Forensic face recognition in particular is more challenging given the effect of variations on the performance since there is a high likelihood that the results are useful to a legal procedure. However, in many real-world scenarios, the forensic experts only have access to a heavily occluded image of the person-of-interest Figure 1. In such cases, COTS FRS usually fail to even detect a face in the image. When this occurs, the forensic expert usually has no other option other than manually processing the image to make it "fit" for processing by the COTS FRS. Such processing usually involves replacing or swapping the occluded part of the part with a corresponding part from some other face so as to make the face visually seem complete or plausible. Then the processed "face" is presented to the COTS FRS which even though detects the face, goes on to provide the investigator with very questionable identification rankings as we show in our experiments. This method of "facial part swapping" is common in law enforcement and investigation agencies

Figure 1. Example images showing subjects wearing sunglasses and scarf. COTS FRS may have difficulties dealing with such occluded faces.

*These authors contribute equally to this work, and should be considered co-first authors.

owing to the failure of COTS FRS in handling face with heavy occlusion. The goal of this paper to study the effect of this method of dealing with occlusions on three different popular COTS FRS and their performance.

Face Recognition systems have come a long way since its early days [3, 28, 19, 2, 1]. Recent efforts to deal with unconstrained face recognition have also been made [10, 11, 27, 26, 17, 18, 9, 16, 15, 14, 13, 8, 12, 7, 6]. They have improved in robustness to various kinds of degradation and variations. Li *et al*. [22] have previously approached age tolerant face recognition using both discriminative and generative models. Park *et al*. [24] proposed a method for automated facial marks detection to facilitate traditional large scale face recognition. Klare *et al*. [21] proposed methodology to match forensic sketches to mugshots. Klare *et al*. [20] have developed a method to match near infrared (NIR) probe images to visible spectrum gallery images. Thus face recognition art is constantly developing to handle more challenging scenarios. In recent times, the impact of unconstrained face recognition in forensics is starting to gain more attention [5], [4]. However, one crucial aspect of forensic face recognition that has not been looked at previously is the processing of occluded or partial faces. Dealing with occlusions is very important in cases where the best image of the person-of-interest available has a partial or occluded face.

The aim of the paper is to highlight the inconsistencies in identification rankings returned by COTS FRS when dealing with partial faces or pre-processed faces (faces with occluded facial parts replaced). Indeed, in our experiments we find that in general COTS FRS find it hard to even detect partial or occluded faces pointing to the yet open nature of this problem. Further, we find large inconsistencies and unpredictability in ranking performance with pre-processed faces and also between COTS FRS, with a some handling the pre-processed faces better than the others. Our experiments try to simulate three major types of occlusions that occur in real-world scenarios. Also, they try to simulate the forensic processing procedure of swapping or replacing out occluded facial parts with parts from other faces. We describe the simple method we use to achieve that in a later section.

The paper is organized as follows. Section 2 presents our method to replace the occluded parts with the corresponding parts from some other face. Section 3 presents experiments the three COTS FRS and Section 4 provides a discussion on the same. Section 5 concludes the paper.

2. Our Methodology for Forensic Face Swapping

We now present our approach to simulating the forensic facial parts swapping process that is commonly undertaken at law enforcement and investigative agencies. The goal is to replace an occluded part of the face with the corresponding part from another face. This is done so that the COTS FRS is able to detect the face and that it can provide recognition rankings for the processed probe image. In the real-world, forensic experts manually do this swapping using sophisticated image and photo editing tools and arrive at a visually plausible result. However, to maintain consistency and to have explicit control over the process, we use facial landmarks of the images determined before the occlusion was introduced. Once the image is artificially occluded, we use the landmarks from a source image (the image whose corresponding part we swap into the occluded image) to crop out the needed part. The part is then morphed onto the corresponding landmarks of the occluded face by finding a simple non-reflective similarity transform between the source image eye coordinates to the image-of-interest eye coordinates. Figure 3 shows the three common types of occlusions explored in this paper for one subject and their corresponding “pre-processed" images with the occluded facial part swapped in from multiple sources. We also simulate subjectivity in the facial part swapping process by allowing for small perturbations in the fitted part onto the occluded image. This simulates the subjectivity in the facial part swapping process which is introduced due to the human factor in the forensic process. Figure 4 illustrates the perturbations that are introduced into the images for one subject for the same source image.

3. Experiments

3.1. Database

In all experiments, we use a part of the FRGC ver2.0 database [25]. We train on 466 subjects and freeze them as our gallery and test on 5 randomly chosen subjects. It should be noted that since we focus on exploring the sensitivity of various COTS FRS on forensic face processing it suffices to explore in depth a few probe images. The images present to the systems were the original uncropped versions of the images to allow the systems maximum flexibility in feature extraction. The size of all probe images was $2,000 \times 3,008$ and $2,272 \times 1,704$ for the target images. The probe images were manually landmarked for the automated face processing. The probe used in this study are presented in Figure 2.

3.2. Sensitivity to Facial Part Swapping

In this experiment, we focus on the sensitivity of multiple commercial off-the-shelf (COTS) face recognition system (FRS) to forensic facial part swapping. There exist at least two sources of error in the face part swapping method. The first type of error concerns with the size and the location of the facial part being swapped. For instance, one can imagine that swapping out the eyes from a subject would re-

Figure 2. Probe images randomly selected from the FRGC database for processing and matching in this study. Sensitivity of COTS FRS to such changes can lead to a degree of impracticality in the use of FRS.

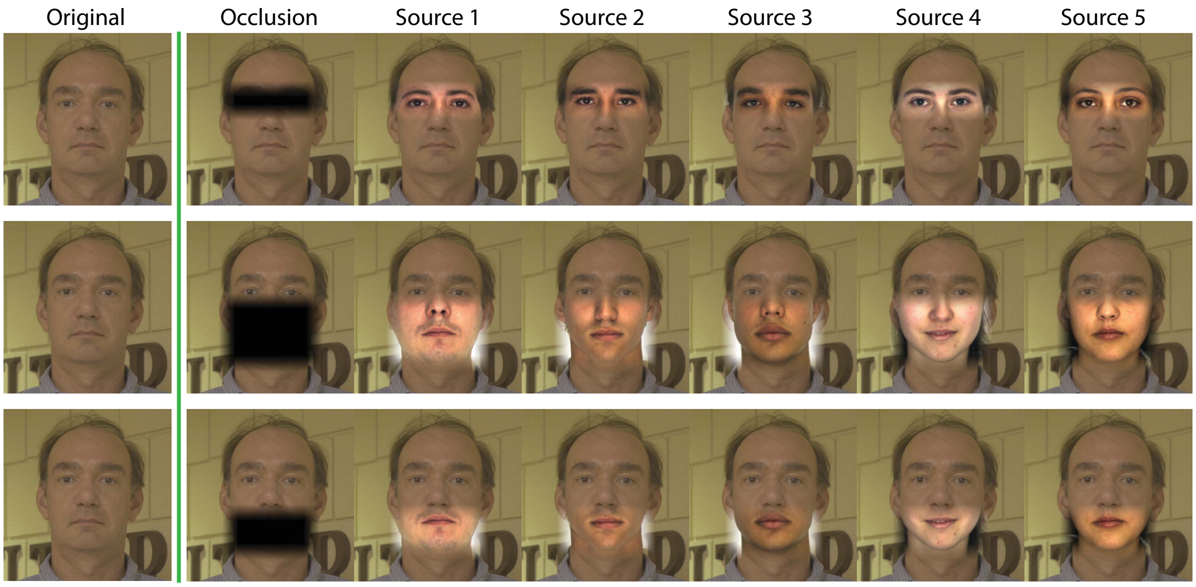

Figure 3. The figure illustrates the three types of occlusions used for a single subject. It also shows the processed image in which the occluded part is replaced by the 5 randomly chosen source images. Sensitivity of COTS FRS to such changes can lead to a degree of impracticality in use of the FRS.

sult in a larger drop in the identification rank than say swapping a part of the cheek or neck. This is a type of error that the forensic expert has no control over, since he has no control over the type of occlusion present in the face of interest. The second type of error is a kind that the expert has more control over. It concerns not with which part of the face of interest is being swapped, but rather whose corresponding part does the expert use to swap and fill in the occluded part. Even though the expert has more control over this error, it is almost impossible to predict the error or the effect of the source image on the identification rank.

To simulate these two kinds of errors, we present an array of face matching runs for each of the 5 probes. For each probe, we generate 3 other probes with varying occlusions. In the real world, the most common occlusions that occur are the ones that occlude the eyes (occlusion 1), the nose and the mouth (occlusion 2) and just the mouth (occlusion 3). We generate the occluded versions of the 5 probes and form the set of occluded probes. Following this, for each image in the set of occluded probes, we swap in the occluded face part with the corresponding part from 5 different and randomly chosen source images from the same database. Thus, overall for each of the 5 original probe images, we generate $3 \times 5 = 15$ probe images for testing with each of the 3 COTS FRS. We test three different COTS FRS for all experiments. Note that COTS FRS 1 required the user to provide the eye locations of the face whereas COTS FRS 2 and 3 use their internal face detector to find the face.

Figure 3 showcases the types of probes tested on for this experiment for one subject. Table 1 shows the identification rankings (out of 466) for the 5 test subjects for different occluded images. Table 2 shows the identification rankings for the three COTS FRS for each of the occlusions for the 5 source face images. The 5 source face images are taken from the AR face database [23] which significantly reduce the correlation between the source images and the probe/target images in the FRGC database where our experiments are based on.

3.3. Sensitivity to Subjectivity in Facial Part Swapping

In the previous experiment, we found that COTS FRS systems are very sensitive to certain types of occlusion and the source image used to fill in that occlusion. However, given a particular occluded image, it practically impossible to predict the "best fit" of facial parts to the face-of-interest.

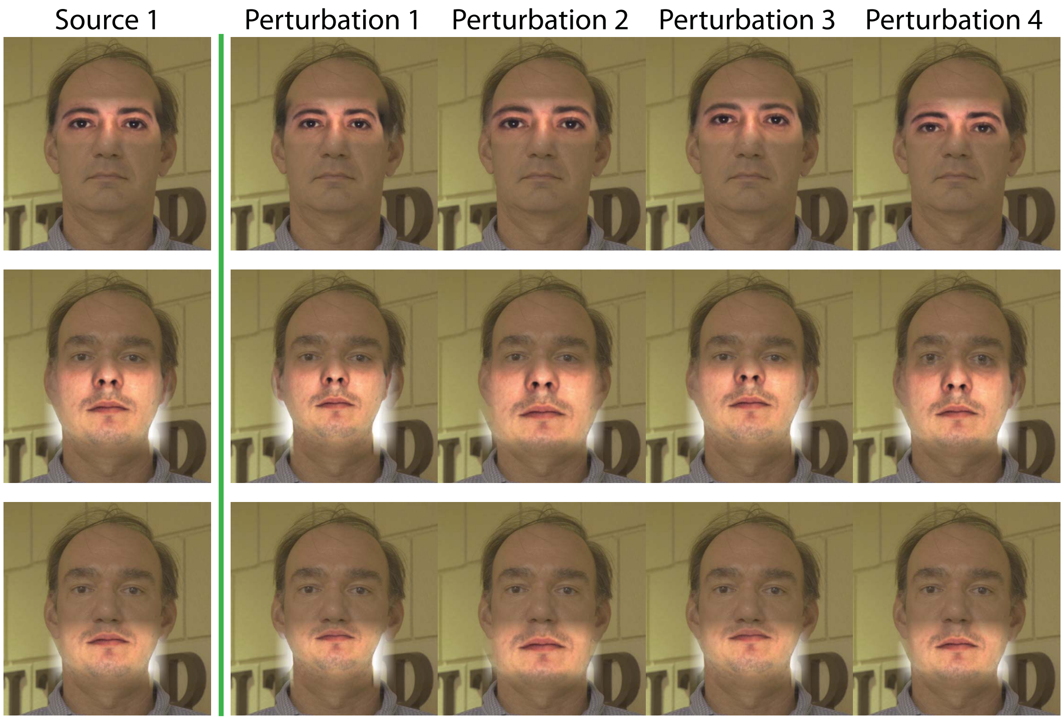

Figure 4. The figure illustrates the four types of perturbations that are introduced for a single subject. It also shows the processed image for the three types of occlusions explored in this image.

Table 1. Identification Rankings (out of 466) for COTS FRS 1, 2 and 3 for the three kinds of occlusions. ⊘ indicates the FRS failed to detect a face.

Subject ID	***Original Probe***	***Occlusion in Eye***	***Occlusion in Nose & Mouth***	***Occlusion in Mouth***
COTS FRS 1				
1	1	47	32	1
2	1	2	255	5
3	1	3	86	1
4	1	1	8	1
5	1	1	231	1
COTS FRS 2				
1	1	⊘	⊘	1
2	2	⊘	⊘	3
3	1	⊘	⊘	1
4	1	⊘	⊘	⊘
5	1	⊘	⊘	⊘
COTS FRS 3				
1	1	⊘	⊘	1
2	1	⊘	⊘	1
3	1	⊘	⊘	233
4	1	⊘	⊘	⊘
5	1	⊘	⊘	⊘

For instance, the location of the eyes of a person might be slightly different than usual. The same might be the case for the nose, jaw line and other biometric features. Under these circumstances, when the forensic expert has no prior about these biometrics, he is forced to hallucinate or imagine the position and the relative orientation of the features. This brings in subjectivity into the forensic process which can have a dramatic effect on the performance of the COTS FRS. This experiment aims to study this precise effect.

We follow the same protocol as in our previous experiment. However, now instead of using 5 different sources for the facial parts, we fix one randomly chosen subject as the facial parts source image. To introduce subjectivity into the forensic facial part swapping, we perturb the location of swapping in the facial parts into the face-of-interest. Figure 4 illustrates an example of the perturbed facial part swapping method for one subject. Table 3 showcases the main result of this experiment. It shows the identification rankings for the three occlusion types for four kinds of perturbation.

3.4. Sensitivity to Image Cropping

In this experiment, we focus on another issue that faced by forensic investigators. In law enforcement, when the subject is uncooperative, in many cases, the best image available of the subject is a part of a larger crowd. In such cases, the investigator is forced to crop out the face from the image and present the cropped image to the COTS FRS. We simulate this situation by considering one random image from the database and obtaining 10 different crops of the face varying in tightness. We further place the cropped image in either a white and a black background. Figure 5 showcases the set of probes generated for this experiment.

Table 2. Identification Rankings (out of 466) for COTS FRS 1, 2 and 3 by swapping in from 5 different source face images. ⊘ denotes that the FRS failed to detect a face, = denotes the identification ranking did not change, and - denotes that no ranking change is reported due to FRS's failure to detect a face.

	ID	***Occlusion in Region***	***Source Image 1***	△	***Source Image 2***	△	***Source Image 3***	△	***Source Image 4***	△	***Source Image 5***	△
					Swapping the Eye Region							
COTS FRS 1	1	47	233	↓ 186	169	↓ 122	44	↑ 3	376	↓ 329	429	↓ 382
	2	2	8	↓ 6	73	↓ 71	9	↓ 7	159	↓ 157	183	↓ 181
	3	3	267	↓ 264	212	↓ 218	52	↓ 49	67	↓ 64	87	↓ 84
	4	1	99	↓ 98	430	↓ 429	89	↓ 88	47	↓ 46	19	↓ 18
	5	1	6	↓ 5	71	↓ 70	153	↓ 152	5	↓ 4	18	↓ 17
COTS FRS 2	1	⊘	24	-	16	-	27	-	20	-	20	-
	2	⊘	7	-	31	-	3	-	22	-	33	-
	3	⊘	3	-	⊘	-	2	-	8	-	5	-
	4	⊘	1	-	⊘	-	4	-	6	-	1	-
	5	⊘	1	-	8	-	1	-	1	-	1	-
COTS FRS 3	1	⊘	1	-	1	-	1	-	1	-	1	-
	2	⊘	1	-	2	-	3	-	2	-	1	-
	3	⊘	1	-	1	-	1	-	1	-	1	-
	4	⊘	1	-	1	-	1	-	1	-	1	-
	5	⊘	1	-	1	-	1	-	1	-	1	-
					Swapping the Nose and Mouth Region							
COTS FRS 1	1	32	1	↑ 31	1	↑ 31	1	↑ 31	2	↑ 30	2	↑ 30
	2	255	1	↑ 254	4	↑ 251	1	↑ 254	2	↑ 253	1	↑ 254
	3	86	2	↑ 84	5	↑ 81	17	↑ 69	235	↓ 149	227	↓ 141
	4	8	1	↑ 7	1	↑ 7	1	↑ 7	1	↑ 7	1	↑ 7
	5	231	1	↑ 230	1	↑ 230	2	↑ 229	1	↑ 230	1	↑ 230
COTS FRS 2	1	⊘	1	-	1	-	2	-	7	-	3	-
	2	⊘	1	-	1	-	1	-	1	-	2	-
	3	⊘	1	-	4	-	1	-	7	-	2	-
	4	⊘	1	-	1	-	1	-	1	-	1	-
	5	⊘	1	-	1	-	1	-	1	-	1	-
COTS FRS 3	1	⊘	1	-	1	-	1	-	1	-	1	-
	2	⊘	2	-	45	-	12	-	3	-	15	-
	3	⊘	1	-	130	-	1	-	3	-	1	-
	4	⊘	1	-	⊘	-	1	-	1	-	1	-
	5	⊘	1	-	⊘	-	1	-	1	-	1	-
					Swapping the Mouth Region							
COTS FRS 1	1	1	1	=	1	=	1	=	1	=	1	=
	2	5	1	↑ 4	1	↑ 4	1	↑ 4	1	↑ 4	1	↑ 4
	3	1	4	↓ 3	3	↓ 2	9	↓ 8	84	↓ 83	28	↓ 27
	4	1	1	=	1	=	1	=	1	=	1	=
	5	1	2	↓ 1	1	=	1	=	1	=	1	=
COTS FRS 2	1	1	1	=	1	=	1	=	1	=	1	=
	2	3	2	↑ 1	2	↑ 1	2	↑ 1	1	↑ 2	9	↓ 6
	3	1	1	=	1	=	1	=	1	=	1	=
	4	⊘	1	-	1	-	1	-	1	-	1	-
	5	⊘	1	-	1	-	1	-	1	-	1	-
COTS FRS 3	1	1	1	=	1	=	1	=	1	=	1	=
	2	1	2	↓ 1	1	=	2	↓ 1	1	=	2	↓ 1
	3	233	1	↑ 232	1	↑ 232	1	↑ 232	1	↑ 232	1	↑ 232
	4	⊘	1	-	1	-	1	-	1	-	1	-
	5	⊘	1	-	1	-	1	-	1	-	1	-

Table 3. Identification Rankings (out of 466) for COTS FRS 1, 2 and 3 swapping in from one source image but with 4 different perturbations. ⊘ denotes that the FRS failed to detect a face, = denotes the identification ranking did not change, and - denotes that no ranking change is reported due to FRS's failure to detect a face.

	ID	***Source Image 1***	***Perturbation 1***	$\triangle$	***Perturbation 2***	$\triangle$	***Perturbation 3***	$\triangle$	***Perturbation 4***	$\triangle$
					Swapping the Eye Region					
COTS FRS 1	1	233	206	↑ 27	180	↑ 53	314	↓ 81	206	↑ 27
	2	8	29	↓ 21	4	↑ 4	12	↓ 4	29	↓ 21
	3	267	291	↓ 24	192	↑ 75	132	↑ 135	291	↓ 24
	4	99	251	↓ 152	98	↑ 1	253	↓ 154	251	↓ 152
	5	6	13	↓ 7	17	↓ 11	15	↓ 9	13	↓ 7
COTS FRS 2	1	24	21	↑ 3	10	↑ 14	15	↑ 9	8	↑ 16
	2	7	5	↑ 2	9	↓ 2	17	↓ 10	2	↑ 5
	3	3	3	=	6	↓ 3	3	=	1	↑ 2
	4	1	⊘	-	1	=	⊘	-	2	↓ 1
	5	1	⊘	-	2	↓ 1	6	↓ 5	⊘	-
COTS FRS 3	1	1	1	=	1	=	1	=	1	=
	2	1	1	=	1	=	1	=	1	=
	3	1	1	=	1	=	1	=	1	=
	4	1	1	=	1	=	1	=	1	=
	5	1	1	=	1	=	1	=	1	=
					Swapping the Nose and Mouth Region					
COTS FRS 1	1	1	1	=	1	=	1	=	1	=
	2	1	1	=	4	↓ 3	1	=	1	=
	3	2	4	↓ 2	2	=	2	=	6	↓ 4
	4	1	1	=	1	=	1	=	1	=
	5	1	1	=	1	=	1	=	1	=
COTS FRS 2	1	1	1	=	3	↓ 2	2	↓ 1	3	↓ 2
	2	1	1	=	1	=	2	↓ 1	1	=
	3	1	1	=	1	=	1	=	1	=
	4	1	2	↓ 1	1	=	2	↓ 1	2	↓ 1
	5	1	3	↓ 2	5	↓ 4	9	↓ 8	1	=
COTS FRS 3	1	1	1	=	1	=	1	=	1	=
	2	2	3	↓ 1	110	↓ 108	1	↑ 1	108	↓ 106
	3	1	45	↓ 44	2	↓ 1	1	=	1	=
	4	1	⊘	-	1	=	242	↓ 241	1	=
	5	1	⊘	-	1	=	⊘	-	1	=
					Swapping the Mouth Region					
COTS FRS 1	1	1	1	=	1	=	1	=	1	=
	2	1	2	↓ 1	4	↓ 3	1	=	3	↓ 2
	3	4	8	↓ 4	2	↑ 2	9	↓ 5	2	↑ 2
	4	1	1	=	1	=	1	=	1	=
	5	2	1	↑ 1	1	↑ 1	1	↑ 1	1	↑ 1
COTS FRS 2	1	1	1	=	1	=	1	=	1	=
	2	2	3	↓ 1	2	=	1	↑ 1	1	↑ 1
	3	1	1	=	1	=	1	=	1	=
	4	1	1	=	1	=	1	=	1	=
	5	1	1	=	1	=	1	=	1	=
COTS FRS 3	1	1	1	=	1	=	1	=	1	=
	2	2	1	↑ 1	148	↓ 146	2	=	70	↓ 68
	3	1	298	↓ 297	1	=	355	↓ 354	1	=
	4	1	⊘	-	1	=	⊘	-	1	=
	5	1	⊘	-	1	=	⊘	-	1	=

Figure 5. The figure showcases the probe images generated using different kinds of occlusion with a white and black background. Sensitivity of COTS FRS to such changes can lead to a degree of impracticality in use of the FRS.

We examine the similarity score returned by the three different COTS FRS between the 20 cropped probe images and the original target image. Table 4 presents the similarity scores returned by the three COTS FRS for this experiment.

From Table 4, we find that for COTS FRS 1 and 2, the similarity scores varies for the same probe image owing to changes in cropping and a seemingly unrelated change in the color of the background. However, COTS FRS 3 returns near perfect results with a consistent similarity score for all different crops and the two color of the backgrounds.

4. Discussion

In this section we take a deeper look into the results for each COTS FRS.

4.1. COTS FRS 1:

From Table 1 we see that since COTS FRS 1 requires eye coordinates of the face, it is able to find the face for recognition. However, we find that occlusions in the nose and mouth causes the most degradation in performance followed by occlusion in the periocular region. We also notice that COTS FRS 1 does not seem to capitalize on only mouth features since mouth occlusions does not affect the performance very much. From Table 2 we find that COTS FRS 1 introduces a significant variation in the identification rankings when multiple source images are used to fill in the occluded periocular and nose and mouth region. However, there seems to be some merit in using the facial part swapping method since in the case of nose and mouth occlusions, rankings improve significantly after swapping. This result at the same time questions the credibility of the FRS since the swapped in part contains biometric information of another subject. It is questionable and surprising that the FRS uses some other biometric information to its benefit. In filling the mouth region with multiple source images, the identification performance is not affected a lot. Table 3 shows that COTS FRS 1 suffers similar significant variations in identification rankings when the same single source image was swapped into the periocular region with slight perturbation. In the case of swapping into the nose and mouth and just the mouth region, the performance degradation is far less severe. Thus, subjective differences in the forensic facial part swapping process leads to unpredicted changes and effects in the identification rankings making practical use of the COTS FRS 1 algorithm using facial part swapping difficult. Further, the decision of which source face to use to salvage facial parts from especially for the eyes and nose region seems to be a vital factor.

4.2. COTS FRS 2:

COTS FRS 2, as seen in Table 1, almost entirely failed to process the occluded faces which is precisely the problem faced by forensic experts. However, from Table 2 we find that COTS FRS 2 performs much more robustly than COTS FRS 1 in handling different source images. Nonetheless, it also introduces significant variability in rankings when the periocular region is swapped (a 30 rank drop equals a $30/466 = 6\%$ drop in ranking, amounting to a much larger number when a larger database is used). We note that the performance of COTS FRS 2 is much better and again would seem to provide some merit to the facial part swapping process used by forensic investigators. Table 3 shows that COTS FRS 2 is rather sensitive to the slight perturbations in the swapping method even when using the same source image especially in the periocular region. At

Table 4. Matching Scores for COTS FRS 1, 2 and 3 for different crop images on black and white background. B1-B10 correspond to crops from small to large with black background, and W1-W10 correspond to the crops from small to large with white background.

	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	Original
FRS 1	2.4839	3.0353	3.0433	3.0786	3.0117	3.0032	3.0096	3.0096	3.0096	3.0096	3.0348
FRS 2	2.7100	3.0869	2.9209	3.0439	3.2881	0.6230	1.5840	2.8623	2.8662	2.4326	2.6602
FRS 3	0.9995	0.9995	0.9995	0.9995	0.9995	0.9995	0.9995	0.9995	0.9995	0.9995	0.9995
	W1	W2	W3	W4	W5	W6	W7	W8	W9	W10	Original
FRS 1	2.8613	2.9357	3.0711	3.0132	3.0032	3.0096	3.0096	3.0096	3.0096	3.0096	3.0348
FRS 2	3.3076	2.7275	3.1982	3.1904	3.2227	2.1816	1.8047	2.1172	2.3818	2.7324	2.6602
FRS 3	0.9995	0.9995	0.9995	0.9995	0.9995	0.9995	0.9995	0.9995	0.9995	0.9995	0.9995

times however, such as swapping into the nose and mouth occluded region, the FRS does introduce variation in independent rankings which would be exaggerated in a larger database.

4.3. COTS FRS 3:

Examining COTS FRS 3, we immediately see from Table 1 that it also almost completely fails in occluded images. Table 2 shows that COTS FRS 3 seems to perform similarly to COTS FRS 2 in response to different sources for the occluded regions of the face. COTS FRS 3 however, seems to be more robust to he periocular region swapping but less so for the nose and mouth. However, Table 3 shows that COTS FRS 3 introduces unpredictable variation in identification rankings owing to slight perturbations in the swapped in facial parts especially in the nose and mouth and also the mouth region.

Overall, we find that swapping in facial parts from other sources could help in some cases. However, note that we are able to make that observation since we have the ground truth labels. In a real world setting, it would be nearly impossible to directly accept the identification rankings as all three COTS FRS present them. We also note that forensic experts usually combine the automated facial matching effort with the manual counterpart wherein the investigator rejects many faces from the top ranks based on other information. However, manual processing is practical only up to a small number of images, say about a 100. Thus unpredictability and clear inconsistency in identification rankings and sensitivity to the source images and the slight orientation changes due to subjectivity in forensic face processing can make it impractical for the investigator to match the probe image in the database. These results highlights the need for a system which is consistent across the forensic techniques the investigators have to use in order to get an identification ranking. Further, a long term goal would be to have more robust FRS which work on partial models of faces. A cooperation between law enforcement agencies and the pattern recognition community can help the community develop technologies which directly address the biggest challenges faced by the agencies.

5. Conclusions

This work presents, to the best of our knowledge, the first preliminary study into the sensitivity of COTS FRS to forensic face processing techniques designed to deal with occlusion and cropping. Our experiments find that although there seems to be some merit in the techniques used by forensic experts, there are clear inconsistencies in the identification rankings returned by the COTS FRS. Factors which are out of the control of experts and investigators such as source images for facial parts and the subjectivity in the processing framework have a dramatic effect on the performance on the COTS FRS, in many cases for the worse.

The work clearly highlights and motivates more serious efforts needed in systems which are able to deal with occlusions and which are designed with the needs of forensic experts in mind. Variations in identification rankings owing to sensitivity to minute changes in the processed image can be magnified when dealing with the large scale datasets used by the law enforcement agencies. Although, there have been considerable advancements in commercial face recognition art, much work is needed in dealing with real-world scenarios which have the potential to have a huge impact in law enforcement and other related applications.

References

[1] T. Ahonen, A. Hadid, and M. Pietikainen. Face description with local binary patterns: Application to face recognition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 28(12):2037–2041, Dec 2006. 2

[2] P. N. Belhumeur, J. P. Hespanha, J. ao P. Hespanha, and D. J. Kriegman. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19:711–720, 1996. 2

[3] A. Goldstein and L. D. Harmon. Identification of human faces. *Proceedings of the IEEE*, 59(5):748–760, 1971. 2

[4] A. Jain, B. Klare, and U. Park. Face matching and retrieval in forensics applications. *MultiMedia, IEEE*, 19(1):20–20, Jan 2012. 2

[5] A. K. Jain, B. Klare, and U. Park. Face recognition: Some challenges in forensics. In *In Proc. Int. Conference on Automatic Face and Gesture Recognition*, 2011. 2

[6] F. Juefei-Xu, M. Cha, J. L. Heyman, S. Venugopalan, R. Abiantun, and M. Savvides. Robust Local Binary Pattern Feature Sets for Periocular Biometric Identification. In *Biometrics: Theory Applications and Systems (BTAS), 4th IEEE Int'l Conf. on*, pages 1–8, sep 2010. 2

[7] F. Juefei-Xu, M. Cha, M. Savvides, S. Bedros, and J. Trojanova. Robust Periocular Biometric Recognition Using Multi-level Fusion of Various Local Feature Extraction Techniques. In *IEEE 17th International Conference on Digital Signal Processing (DSP)*, 2011. 2

[8] F. Juefei-Xu, K. Luu, M. Savvides, T. Bui, and C. Suen. Investigating Age Invariant Face Recognition Based on Periocular Biometrics. In *Biometrics (IJCB), 2011 International Joint Conference on*, pages 1–7, Oct 2011. 2

[9] F. Juefei-Xu, D. K. Pal, and M. Savvides. Hallucinating the Full Face from the Periocular Region via Dimensionally Weighted K-SVD. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2014 IEEE Conference on*, June 2014. 2

[10] F. Juefei-Xu, D. K. Pal, and M. Savvides. NIR-VIS Heterogeneous Face Recognition via Cross-Spectral Joint Dictionary Learning and Reconstruction. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2015 IEEE Conference on*, June 2015. 2

[11] F. Juefei-Xu, D. K. Pal, K. Singh, and M. Savvides. A Preliminary Investigation on the Sensitivity of COTS Face Recognition Systems to Forensic Analyst-style Face Processing for Occlusions. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2015 IEEE Conference on*, June 2015. 2

[12] F. Juefei-Xu and M. Savvides. Can Your Eyebrows Tell Me Who You Are? In *Signal Processing and Communication Systems (ICSPCS), 2011 5th International Conference on*, pages 1–8, Dec 2011. 2

[13] F. Juefei-Xu and M. Savvides. Unconstrained Periocular Biometric Acquisition and Recognition Using COTS PTZ Camera for Uncooperative and Non-cooperative Subjects. In *Applications of Computer Vision (WACV), 2012 IEEE Workshop on*, pages 201–208, Jan 2012. 2

[14] F. Juefei-Xu and M. Savvides. An Augmented Linear Discriminant Analysis Approach for Identifying Identical Twins with the Aid of Facial Asymmetry Features. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2013 IEEE Conference on*, pages 56–63, June 2013. 2

[15] F. Juefei-Xu and M. Savvides. An Image Statistics Approach towards Efficient and Robust Refinement for Landmarks on Facial Boundary. In *Biometrics: Theory, Applications and Systems (BTAS), 2013 IEEE Sixth International Conference on*, pages 1–8, Sept 2013. 2

[16] F. Juefei-Xu and M. Savvides. Subspace Based Discrete Transform Encoded Local Binary Patterns Representations for Robust Periocular Matching on NIST's Face Recognition Grand Challenge. *IEEE Trans. on Image Processing*, 23(8):3490–3505, aug 2014. 2

[17] F. Juefei-Xu and M. Savvides. Facial Ethnic Appearance Synthesis. In *Computer Vision - ECCV 2014 Workshops*, volume 8926 of *Lecture Notes in Computer Science*, pages 825–840. Springer International Publishing, 2015. 2

[18] F. Juefei-Xu and M. Savvides. Weight-Optimal Local Binary Patterns. In *Computer Vision - ECCV 2014 Workshops*, volume 8926 of *Lecture Notes in Computer Science*, pages 148–159. Springer International Publishing, 2015. 2

[19] T. Kanade. Picture processing system by computer complex and recognition of human faces. In *doctoral dissertation, Kyoto University*. November 1973. 2

[20] B. Klare and A. Jain. Heterogeneous face recognition: Matching nir to visible light images. In *Pattern Recognition (ICPR), 2010 20th International Conference on*, pages 1513–1516, Aug 2010. 2

[21] B. Klare, Z. Li, and A. Jain. Matching forensic sketches to mug shot photos. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 33(3):639–646, March 2011. 2

[22] Z. Li, U. Park, and A. K. Jain. A discriminative model for age invariant face recognition. *Trans. Info. For. Sec.*, 6(3):1028–1037, Sept. 2011. 2

[23] A. M. Martinez and R. Benavente. The ar face database. *CVC Technical Report*, 24, 1998. 3

[24] U. Park and A. Jain. Face matching and retrieval using soft biometrics. *Information Forensics and Security, IEEE Transactions on*, 5(3):406–415, Sept 2010. 2

[25] P. Phillips, P. Flynn, T. Scruggs, K. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek. Overview of the face recognition grand challenge. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, pages 947–954, June 2005. 2

[26] M. Savvides and F. Juefei-Xu. Image Matching Using Subspace-Based Discrete Transform Encoded Local Binary Patterns, Sept. 2013. US Patent US 2014/0212044 A1. 2

[27] K. Seshadri, F. Juefei-Xu, D. K. Pal, and M. Savvides. Driver Cell Phone Usage Detection on Strategic Highway Research Program (SHRP2) Face View Videos. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2015 IEEE Conference on*, June 2015. 2

[28] M. Turk and A. Pentland. Eigenfaces for recognition. *J. Cognitive Neuroscience*, 3(1):71–86, Jan. 1991. 2