A Meta-Analysis of Face Recognition Covariates

Yui Man Lui, David Bolme, Bruce A. Draper, J. Ross Beveridge, Geoff Givens, P. Jonathon Phillips

Abstract— **This paper presents a meta-analysis for covariates that affect performance of face recognition algorithms. Our review of the literature found six covariates for which multiple studies reported effects on face recognition performance. These are: age of the person, elapsed time between images, gender of the person, the person's expression, the resolution of the face images, and the race of the person. The results presented are drawn from 25 studies conducted over the past 12 years. There is near complete agreement between all of the studies that older people are easier to recognize than younger people, and recognition performance begins to degrade when images are taken more than a year apart. While individual studies find men or women easier to recognize, there is no consistent gender effect. There is universal agreement that changing expression hurts recognition performance. If forced to compare different expressions, there is still insufficient evidence to conclude that any particular expression is better than another. Higher resolution images improve performance for many modern algorithms. Finally, given the studies summarized here, no clear conclusions can be drawn about whether one racial group is harder or easier to recognize than another.**

I. Introduction

Since 2003, the authors have published a series of papers quantifying the impact of covariates on the performance of face recognition algorithms [21], [19], [20], [8], [7]. The motivation for these studies is both scientific and practical. The scientific reason is to better understand algorithms and phenomenology underlaying face image formation as expressed through recognition performance. The practical reason is to provide guidance to individuals responsible for fielding face recognition systems. The studies consistently show that changes in covariate values often affect the performance of a face recognition algorithm. The question taken up in this paper is, over multiple studies, which effects are consistent and which are not. For example, is there an emerging consensus about the way age, or gender, or image resolution influences face recognition performance.

Our approach is similar to that taken by Phillips and Newton in their iris recognition meta-analysis [32]. Here we present a quantitative summary of papers in the face recognition literature that report performance broken out by covariates. Face recognition covariates, in general, represent some aspect of the person, the image or some other aspect of the intended face recognition application. Covariates associated with people include such things as age, gender and race. They also include more transitive aspects of the person's appearance such as whether they are smiling or wearing glasses. Properties of the image includes such things as image resolution, the number of pixels between the eyes, or whether the image was compressed, and if so, by how much. Other aspects of the intended application may include such factors as whether or not images are taken outdoors, or how much time has elapsed between the time two images being compared were taken.

The rule for including a covariate in this analysis was that results appear in more than one study, and ideally from more than one group of researchers. After an extensive review of the literature, the decision was made to restrict our meta-analysis to results from experiments on frontal, still, visible light images of faces. The following covariates meet the above criteria for inclusion in our study: age, gender, and race of the subject; expression; elapsed time between the two images being compared; and size of the face in the image.

Some covariates not included in our analysis warrant special mention. Arguably, the two covariates with the greatest impact on face recognition performance are pose and illumination. As a result, these two covariates have received significant attention. The Yale [18] and PIE [42] databases were created to systematically study pose, illumination, and their interactions, and have been widely used by researchers. As a result, there is already a substantial literature on pose and illumination.

There are also a number of interesting covariates that are worth briefly mentioning, but are not included in our analysis because they are only reported in one study. Boult et al. [13] looked at how the time of day effects recognition performance in outdoor settings, and determined that recognition is better in the morning and afternoons than in either the daybreak or evening twilight in outdoor environment. The Foto-Fahndung report [3] noted that face recognition systems perform better in the middle of a day at the Mainz railway station. Image quality measures are also not analyzed, because virtually every study employs a unique quality measure. For example, Hsu and Martin [24] look at whether human subjective judgments of image quality correlate to machine recognition performance (they do), while Adler and Dembinsky [5] look at the correlation between a black-box vendor-supplied quality measure and performance. Abdel-Mottalib and Mahoor [4] study the effect of a machine quality measure based on kurtosis, while Weber [47] evaluates a quality measure based on "sharpness" and Beveridge et al. [8] measure quality via edge density. While there is clearly a relationship between image quality and recognition performance, structured comparisons are not possible without a common measure of image quality.

Yui Man Lui, David Bolme, Bruce A. Draper, and J. Ross Beveridge are with the Department of Computer Science, Geoff Givens is with the Department of Statistics, Colorado State University, Fort Collins, CO 80523, and P. Jonathon Phillips is with the National Institute of Standards and Technology, Gaithersburg, MD 20899.

A. Quick Summary of Findings

For the age covariate we found that virtually all studies report that older people are easier to recognize than younger ones. This is consistent with the intuition on the phenomenology of faces. Because faces change over time, it should not be surprising that older faces are either harder or easier to recognize than younger faces. Even though this effect is consistent over multiple studies, more research is needed to determine why older faces are easier to recognize.

Similarly, all studies agree that recognition performance degrades as the time between the query and target images increases. As expected, images taken on the same day consistently have a higher recognition rate than images taken on different days. Furthermore, it is expected that the appearance of the face changes as a person ages and therefore the time difference makes recognition more difficult. While this effect is clearly present in most studies, it is only significant when measured in years, not days or months.

Gender, on the other hand, is a point of contention. Studies disagree about whether men or women are easier to recognize, although most agree that any effect of gender on recognition performance is small. Gender is also an example of a covariate that interacts with other covariates. The effect of gender, for example, decreases as the age of the subject increases.

Regarding expression, there is consensus that performance improves when a person's expression is the same in both the query and target images. Less clear is whether, if forced to enroll only a single image, it is better to request the person to smile or instead to adopt a neutral expression.

Image resolution effects also do not always agree, but in this case the chronology of the studies takes on particular significance. Older studies, often using algorithms that match whole face images in dimensionality reduced subspaces showed a sweet spot in terms of resolution and actually in some cases showed that higher resolution images could reduce recognition performance. However, more recent studies with more modern algorithms appear to remove this barrier, and instead largely suggest that increasing resolution increases recognition performance.

Multiple studies show race matters in so much as recognition performance improves for individuals belonging to one race versus another. However, based upon current studies it is difficult to tease apart how much this has to do with the specific race versus the relative proportion of different races in a particular data set. There is a suggestion in the current studies that recognition may be easier when a person belongs to a race that appears less frequently in the system's database of enrolled people.

II. Methodology

A three stage methodology is used here. The first stage consists of deciding what are the initial selection criteria for including papers in this study and a strategy for searching the literature for papers. For the first stage the selection criteria were papers that reported experimental results broken out by a subject, image, or quality covariate. The search utilized IEEExplore, Google Scholar, major computer vision conferences, biometric and face conferences and workshops, and references in papers.

The second stage consists of refining the selection criteria and deciding how to summarize findings. At this stage the selection criteria was restricted to papers reporting on frontal, still, visible light images of faces. Some papers had to be dismissed because they did not specifically relate a change in a commonly recognized measure of face recognition performance to a clearly identified covariate. It was also the case that for covariates such as age, gender and expression, the vast majority of papers we reviewed addressed how to classify face images by age, gender or expression, rather than relating these attributes to face recognition. At the conclusion of the second stage we had tabulated just under 100 distinct effects drawn from 46 studies.

In the third stage, each of these findings was further reviewed in order to form some reasonable quantitative summary of the finding. As the reader will see in the tabulated results presented below, ultimately we had to be able to relate a change in a covariate, for example gender, to a percent change in a commonly used measure of face recognition performance. Examples of such measures include rank 1 identification rate or verification rate at a fixed false accept rate. Of the nearly 100 candidates coming out of stage two, for only about half were we able to take this final step.

The tabulated results coming out of stage three are presented in two forms below. There are abbreviated tables that indicate succinctly the findings of different studies with little auxiliary information. These are included in the main body of the paper and are sufficient to see the major trends in our analysis. For each major table, there is an expanded version included in a supplemental document that provides supplemental details such as the data set used in the study, the algorithm used in the study, and what if any controls were used to account for other covariates.

The covariate results are summarized in six tables. Each row of a table represents a single research finding, and the rows are typically ordered chronologically so that the most recent, and therefore presumably most relevant, findings appear near the top. The columns to the left, filled with asterisks, show the reported effect of a covariate on recognition. These columns share a common header, indicating the direction of the effect.

The magnitude of the effect is reported as a percent change in a common performance measure. For example, reported effects for verification studies are measured as the change in verification rate at a fixed false reject rate. For rank based identification studies, the magnitude reflects the change in the rank one recognition rate. In both cases, magnitudes are coarsely quantized into ranges.

An asterisk may appear in more than one column, indicating an interaction with some other covariate leading to a range of observed performance. The right most column cites the paper in which the finding is reported.

TABLE I

AGE EFFECT ON RECOGNITION PERFORMANCE.

Older Easier								Younger Easier	
+17.5%	15%	10%	5%	0%	5%	10%	15%	+17.5%	Source
		*	*						[9]
	*	*							[9]
			*						[9]
		*							[7]
			*						[7]
*									[7]
*									[34]
*									[44]
				*					[23]
	*								[20]
			*						[19]
		*							[19]
		*							[19]
		*							[21]
	*								[35]
*									[35]
		*							[35]
			*						[35]
		*							[35]
		*							[35]
		*							[35]
				*					[35]

TABLE II

TIME BETWEEN IMAGES EFFECT ON RECOGNITION PERFORMANCE.

Less Time								More Time			
+17.5%	15%	10%	5%	0%	5%	10%	15%	+17.5%	Time Span	Same Day	Source
			*						120	No	[26]
*									120	No	[41]
*									108	No	[40]
*									37	Yes	[41]
*									37	Yes	[35]
*									37	Yes	[35]
*									37	Yes	[35]
		*							37	Yes	[35]
		*							37	Yes	[35]
		*							37	Yes	[35]
	*								37	Yes	[35]
*									37	Yes	[35]
*									24	Yes	[6]
*									24	Yes	[6]
		*							8	Yes	[7]
			*						8	Yes	[7]
				*					8	Yes	[7]
*									3	Yes	[28]
			*						2.5	No	[16]

III. AGE

Age is an example of a covariate where there is general, although not universal, consensus: older people are easier to recognize than younger ones. The evidence for this conclusion is summarized in Table I. While the magnitudes of the age effect are different for various algorithms, in 20 out of 22 results, older people are easier to recognize than younger people.

The presence of two asterisks in the first two rows of Table I are the result of an interaction between age and environment, more specifically whether query images were taken indoors or outdoors. There is a smaller interaction between age and gender for the results summarized in the last eight rows of Table I, although the associated change in performance falls below the $\pm 2.5\%$ granularity of the table an consequently only single asterisks appear.

IV. TIME BETWEEN IMAGES

The previous section surveyed results about the effect of subject age on recognition performance. A related but distinct issue has to do with the time lapse between when the query and target images were taken. Here again we find consensus among studies (see Table II): recognition performance is better if the elapsed time between the query and target images is small, but the effect only becomes noticeable when the time lapse is measured in years, not days or weeks or even months.

Table II has two columns not found in Table I: "time span" and "same day". The "time span" column records the greatest time difference between a query and test image considered by a study, measured in months. Thus the shortest time span considered was 2.5 months, while the longest was ten years (120 months). The "same day" column reports whether any of the query and target image pairs were collected on the same day, or whether the images were always taken at least one day apart. Another difference in format between Table II and Table I is that the rows in Table II are ordered by the time span rather than chronologically. As shown in Table II, 14 of 15 findings [1] report that smaller time spans lead to better recognition, while only one finding suggests no effect either way.

We do not expect a perfect uniformity in results; the studies we are surveying employ different algorithms, different datasets, etc. Nonetheless, Table II leads us to certain conclusions. Every study that considered image pairs taken more than 24 months apart found that smaller time spans improve performance, and in many cases the effect was large, up to 17.5% or more. Studies with smaller time spans tend to show a smaller effect, albeit in the same direction, and one such study show no effect at all. This is the basis for our conclusion that time spans matter, but the effect is only large when the time span can be measured in years.

We also suspect that images taken on the same day are significantly easier to match than images taken just one day apart. Some of the studies in Table II use pairs of images taken on the same day as their smallest possible time span; others only compare images taken at least one day apart. Among the studies that look at time spans of 24 months or less, the one that does not consider same-day image pairs show no effect and a small effect (5%). The studies that do consider same-day pairs tend to show larger effects (up to 17.5% or more). We infer that same-day pairs are much easier to match than different-day pairs, for reasons that have more to do with other effects, such as changes in

[1] 14 studies had time span more than a year plus one interaction.

TABLE III

GENDER EFFECT ON RECOGNITION PERFORMANCE.

Male Easier								Female Easier		
+17.5%	15%	10%	5%	0%	5%	10%	15%	+17.5%	Interactions with other covariates	Source
				*						[9]
			*	*	*				Outdoor/Indoor	[9]
			*							[9]
				*	*				Outdoor/Indoor	[8]
					*					[7]
				*						[7]
			*							[7]
		*								[20]
				*						[19]
				*						[19]
				*						[19]
				*						[21]
	*	*	*	*					Age	[35]
		*	*	*	*	*			Age	[35]
	*	*	*	*					Age	[35]
					*	*			Expression, Lighting	[22]

TABLE IV

EXPRESSION EFFECT ON FACE RECOGNITION PERFORMANCE.

Same Expression								Different Expression	
+17.5%	15%	10%	5%	0%	5%	10%	15%	+17.5%	Source
*									[7]
		*							[7]
		*							[7]
			*						[19]
			*						[19]
				*					[19]
		*							[21]

Neutral vs. Smiling								Smiling vs. Neutral	
				*					[7]
				*					[7]
			*						[7]

Regular vs. Alternate								Alternate vs. Regular	
					*				[10]
					*				[10]
					*				[10]

illumination, camera position or clothing, than with aging. We believe that when same-day pairs are not considered, the effect of aging over a few months or even one or two years is small. This conclusion is also supported by a study by Bartlett [6] that suggests it is easier to match same-day images with different expressions than to match different-day images with the same expression (see Section VI).

V. GENDER

Many studies have looked at the effects of gender on face recognition as shown in Table III, but something less than a consensus emerges. There is a slight result in favor of men being easier to recognize; men are easier to recognize than women 7 to 5, with 6 reporting no effect. One plausible inference might be that gender has little or no effect on face recognition, since 11 studies have reported either no effect or a range of effects which includes no effect. Also, no study has reported as large an effect as reported in some of the studies of age and time.

Gender effects, when reported, often interact with other covariates. Note the addition in Table III of a column indicating interactions between gender and other covariates. For example, Phillips et al. [35] reported that for the Identix [1] system , the gender effect could range from men being 10% easier to women being 10% easier, depending on the age of the subject (see row 14 of Table III). Other studies have shown interactions between gender and setting (indoor vs. outdoor) and among gender, expression and illumination. In all but one finding with an interaction, the interaction not only changed the magnitude of the gender effect, it changes its direction in terms of whether men or women are easier to recognize.

VI. EXPRESSION

Most work on expression is focused on recognizing a person's expression, with fewer studies relating expression to recognition performance. Our summary of these studies appears in Table IV. These studies have focused on two subtly different questions: (1) does a change in expression matter, and (2) if it does matter, what expression should be used when a single image of a person is to be enrolled in a database?

Studies on whether changes in expression matter are reported in the top part of Table IV. Not surprisingly, there is relative consensus on this question: 6 out of 7 findings report that it is easier to match subjects across images when the expressions are the same. For 5 of the 7 studies same versus different expressions improved recognition performance from between 5% to 10%.

It should be noted that half of these studies were conducted on the FERET data, in which people were asked to adopt a neutral expression for one image and an "alternate" expression for the other. The other half of the studies were conducted on the FRGC dataset, in which people were asked to adopt a neutral expression for one image and a smiling expression for the other. Nonetheless, when both parts of an image pair share a common expression, whatever it may be, recognition is easier.

The bottom half of Table IV is divided into two parts: one for the study by Blackburn et al. [10], the other for the study by Beveridge et al. [7]. Both these studies included results for cross expression recognition. Thus, both studies considered the situation where a new, query, image is compared to an enrolled, target, image and the expressions do not match.

In the study by Blackburn et al. [10], carried out in 2000, recognition performance was compared between two cases: 1) a query image with a regular expression was compared to a target image with an alternate expression, and 2) a query image with an alternate expression was compared to a target image with a regular expression. Blackburn et al. found that performance improved for cross-expression image pairs if the query image had the alternate expression and the enrolled image the regular expression. This is true across three algorithms, although the effect is small (about 5%).

In the study by Beveridge et al. [7], carried out on data and algorithms from FRGC 2006, the two cases compared were: 1) a query image with a neutral expression was compared to a target image with an smiling expression, and 2) a query image with smiling expression was compared to a target image with a neutral expression. The authors found that when the choice was restricted to smiling versus neutral, enrolling a smiling face is slightly better. However, the effect is small and algorithm dependent. In another study , Yacoob and Davis [50] [2] report that extreme expressions, whether smiling or angry, produce better recognition performance than neutral expressions.

Overall it is better to compare matching expressions than non-matching expressions, but from a performance standpoint, what expression should be enrolled when the query cannot be controlled is unknown. This finding is of interest in relation to the common practice of requiring a neutral expression on face images being used in an official capacity.

For example, the Canadian government has decreed that smiles are not allowed in passport photos [2]. Many U.S. states are adopting or considering similar regulations for driver's license photos. This bias toward neutral expressions makes sense if one assumes that neutral expressions are more likely to be present in query images against which these enrolled images are to be compared. It is not supported by evidence of any intrinsic superiority of one expression over another.

VII. Image Resolution

Image resolution is an important covariate that can greatly affect an algorithm's performance. It is also a covariate over which system engineers typically have a great deal of control. Modern cameras make comparatively high resolution images available with essentially no overhead. However, there are still computational incentives to downsample images in order to make storage and particularly comparison of images more efficient. Thus understanding the performance trade offs related to resolution is critical.

Table V presents results grouped into three sections. The top section of the table describes the effect of image resolution while both query and target images are varied based on the distance between the eyes. The middle section of the table illustrates the effect of image resolution while the resolution of the target images is fixed. The bottom section of the table depicts the effect of image resolution in terms of image size rather than pixels between the eyes.

There is general agreement that recognition performance degrades when image resolution is below 32×32 [12], [46], [14], [48], [25] as evident by the results shown in the bottom section of Table V. This statement suggests an operational lower bound on image resolution. Note that enhancement techniques, like super-resolution [48], may be applied prior to image matching when the resolution drops below this lower bound.

Still focusing on the bottom section of Table V, the optimal recognition performance can be found in image size between 32×32 and 64×64. Note the results for this section are derived for variants of PCA and LDA (see supplemental materials for details). Low resolution imagery may be adequate for these older techniques. In addition, Keil et al. [25] reported that a 37×37 face chip has the highest class separability for Fisher linear discriminant analysis, nonparametric discriminant analysis, and mutual information algorithms on the FRGC subset. Beveridge et al. [7] also reported that PCA benefits from using lower resolution shown in the fourth row on the top section of Table V.

What about more advanced algorithms? The FaceIt system [1] is one of the earliest commercial face recognition products. Like PCA and LDA algorithms, the FaceIt system derived no benefit from higher resolution imagery when tested in the FRVT 2000 evaluation [10] shown in the first row of the middle section of Table V. More specifically, according to the FRVT 2006 evaluation report [39], the FaceIt system performs gracefully with lower resolution imagery when the eye distance is varied between 75 and 400 pixels in both controlled and uncontrolled settings. This result is shown in the top section of Table V, row 14. Similar findings are also supported by Wheeler et al. [48] where the optimal recognition result is found in 37 pixels between the eyes using the FaceIt system on the GE video and the FERET dataset. Finally, O'Toole et al. [33] noted a similar finding; as shown in the top section of Table V, row 12, the FaceIt system does not favor very high resolution imagery.

On the other hand, some modern algorithms take considerable advantage of high resolution imagery. As the top section of Table V reveals, most of the results summarized indicate considerable improvement in performance is possible when using higher resolution images. Recently, O'Toole et al. [33] have found that NevenVision [3] performs exceptionally well on very high resolution images as shown in the eleventh row on the top section of Table V. The fact that NevenVision performs better in higher resolution can be further supported by the FRVT 2006 evaluation [39] where NevenVision consistently yields higher recognition performance in both controlled and uncontrolled settings using high resolution imagery, see the fifteenth row on the top section of Table V. Furthermore, Beveridge et al. [7] also noted that algorithms provided by CMU [49] and NJIT [27] achieve better recognition results when high resolution imagery is used as shown in the fifth and sixth rows on the top section of Table V, respectively.

The choice of image resolution is clearly algorithm dependent which is demonstrated through a wide range of algorithms like LDA [12], PCA [46], FaceIt [1], CMU [49], NJIT [27], and NevenVision. When there is an interaction

[2] This study is noteworthy, but was not tabulated in our results because it reported results in terms of discrimination power that we could not easily map to explicit changes in recognition performance.

[3] Neven Vision, Inc was acquired by Google, Inc in 2006.

between indoor and outdoor, image resolution is also interacted with these environment settings. Beveridge et al. [8], [9] reported that higher resolution imagery is favored when the query image is taken outdoor whereas lower resolution may be adequate when the query is acquired in indoor. As the first three rows on the top section of Table V depicts, the environment effect could range from 17.5% or more easier when high resolution imagery is used in an outdoor setting to 10% easier when low resolution imagery is used on indoor images.

Face recognition is advancing quickly and will find more applications in uncontrolled environments. Therefore, we can safely surmise that the use of high resolution imagery will increase. As Table V shows, some modern algorithms have already taken advantages of the additional detail present in high resolution imagery.

VIII. Race

As shown in Table VI, there have been several studies which attempt to determine if members of one racial group are easier or harder for machines to recognize than members of another. Unfortunately, the findings from these studies are difficult to interpret because all of the studies confound race with sampling effects. To be specific, all three of the studies in Table VI draw their subjects from racially mixed populations in which the groups are unevenly represented. In the FERET, FRGC and FRVT2006 datasets, Caucasians are always the most heavily sampled group, followed by East Asians, by approximately 3 to 1. All other groups are insufficiently represented to report.

So, while Table VI shows that for these studies, Caucasians are harder to recognize than East Asians, this could be explained either as an intrinsic property of the two races, or alternatively, as a by product of the relative proportion of the data set constituted by each. If that latter explanation is the correct one, the presumably reversing the relative proportion of Caucasians and East Asians in the study would correspondingly suggest Caucasians are easier to recognize than East Asians. What is needed to resolve this ambiguity is a study comparing performance across two uniform populations of different racial groups, but to the best of our knowledge this has not yet been done.

TABLE V
Image resolution effect on recognition performance.

High Resolution								Low Resolution		
+17.5%	15%	10%	5%	0%	5%	10%	15%	+17.5%	Pixels between the eyes	Source
*	*	*	*	*	*				202 ∼ 134	[9]
*	*	*	*	*	*				202 ∼ 134	[9]
*	*	*	*	*	*	*			202 ∼ 134	[9]
								*	180 ∼ 120	[7]
	*								180 ∼ 120	[7]
*									180 ∼ 120	[7]
*									210 ∼ 137	[8]
				*					190 ∼ 110	[33]
								*	190 ∼ 110	[33]
				*					190 ∼ 110	[33]
			*						190 ∼ 110	[33]
					*				190 ∼ 110	[33]
*	*	*	*	*					400 ∼ 75	[39]
					*	*	*	*	400 ∼ 75	[39]
			*						400 ∼ 75	[39]
*	*	*	*	*					400 ∼ 75	[39]
		*	*						400 ∼ 75	[39]
			*	*					400 ∼ 75	[39]
			*	*					400 ∼ 75	[39]
*									48 ∼ 17	[48]

High Resolution								Low Resolution		
+17.5%	15%	10%	5%	0%	5%	10%	15%	+17.5%	Pixels between the eyes	Source
				*					60 ∼ 15	[10]
			*						60 ∼ 15	[10]
	*								60 ∼ 15	[10]

High Resolution								Low Resolution		
+17.5%	15%	10%	5%	0%	5%	10%	15%	+17.5%	Image size	Source
				*					128x128 vs 32x32	[12]
			*						32x32 vs 16x16	[12]
				*					144x108 vs 64x48	[46]
			*						64x48 vs 32x24	[46]
				*					144x108 vs 64x48	[46]
		*							64x48 vs 32x24	[46]
				*					64x64 vs 32x32	[14]
			*	*					64x64 vs 16x16	[14]
*	*	*	*						64x64 vs 8x8	[14]

IX. Conclusion

The goal of this review was to quantitatively summarize what is known from the literature about the effects of six covariates on face recognition: age, time, gender, expression, resolution and race. In so doing, we collect findings reported over different face recognition algorithms and data sets using different statistical methods. The goal was to look for consensus; if a result holds over different algorithms, data sets and studies, it is probably a universal property of face recognition. Conflicting findings, on the other hand, suggest that an effect is algorithm or data specific, or interacts with another unknown factor. Using this methodology, we identify four universal covariates: (1) older people are easier to recognize than younger ones; (2) recognition gets harder as the time between the query and target images grows, but this is only significant if the time difference is measured in years; (3) recognition is easier if the subject adopts the same expression in the query and target images; and (4) high resolution imagery yields better recognition performance for many modern algorithms. None of these results are particularly surprising.

Of perhaps greater interest are the effects we did not find. There is no universal gender effect. Although individual studies report that men are easier to recognize than women or vice-versa, there is no consensus. Gender effects appear to be algorithm and/or dataset dependent rather than universal. Expression is another interesting case. While it is universally easier to match images with the same expression, there is

TABLE VI

RACE EFFECT ON RECOGNITION PERFORMANCE.

Caucasian Easier								East Asian Easier		
+17.5%	15%	10%	5%	0%	5%	10%	15%	+17.5%	Interactions with other covariates	Source
							*			[9]
							*	*	Outdoor/Indoor	[9]
								*		[9]
					*	*			Outdoor/Indoor	[8]
					*					[7]
						*				[7]
						*				[7]
					*					[19]
					*					[19]
					*					[19]
					*					[21]
					*	*	*	*	Dist. Measure	[17]
						*				[17]

no empirical support for the accepted dictum that subjects should be enrolled with neutral expressions. Finally, although there are several studies suggesting a race effect, all of these studies are confounded with sampling imbalances between racial groups. As a result, there is no strong evidence to believe that one race group is easier to recognize than another; the more probable explanation is that whichever group is most heavily sampled in a population becomes the most difficult racial group to recognize.

Not surprisingly, this review suggests the need for additional studies. Studies that look at the effect of expressions during enrollment are needed, particularly in light of the current efforts to regulate passport and other identification photos. Studies that separate sampling effects from ethnic effects would also be useful. The most important studies, however, may be about other covariates. Even taking age, time and expression into account, it is still difficult to predict which images will fail to match. More research is needed.

ACKNOWLEDGMENTS

The work was funded in part by the Technical Support Working Group (TSWG) under Task T-1840C. Jonathon Phillips was supported by the Department of Homeland Security, Director of National Intelligence, and Federal Bureau of Investigation. The identification of any commercial product or trade name does not imply endorsement or recommendation by Colorado State University or the National Institute of Standards and Technology.

REFERENCES

[1] Faceit. http://www.l1id.com.
[2] Passport canada. http://www.ppt.gc.ca/cdn/photos.aspx?lang=eng.
[3] Face recognition as a search tool "foto-fahndung", 2007.
[4] M. Abdel-Mottaleb and M. Mahoor. Application notes-Algorithms for Assessing the Quality of Facial Images. *IEEE Computational Intelligence Magazine*, 2(2):10–17, 2007.
[5] A. Adler and T. Dembinsky. Human vs. automatic measurement of biometric sample quality. In *CCECE*, 2006.
[6] M. Bartlett. *Face image analysis by unsupervised learning and redundancy reduction*. PhD thesis, University of California, San Diego, 1998.
[7] J. Beveridge, G. Givens, P. Phillips, and B. Draper. Factors that influence algorithm performance in the Face Recognition Grand Challenge. *Computer Vision and Image Understanding*, 2009.
[8] J. Beveridge, G. Givens, P. Phillips, B. Draper, and Y. M. Lui. Focus on quality, predicting frvt 2006 performance. In *FGR, Amsterdam, The Netherlands*, 2008.
[9] J. R. Beveridge, G. Givens, P. J. Phillips, B. Draper, D. S. Bolme, and Y. M. Lui. Frvt 2006: Quo vidas face quality. *Image and Vision Computing*, (to appear).
[10] D. Blackburn and M. B. J. Phillips. Frvt 2000 - evaluation report. *Technical Report*, 2001. http://www.frvt.org.
[11] D. S. Bolme, J. R. Beveridge, M. L. Teixeira, and B. A. Draper. The csu face identification evaluation system: Its purpose, features and structure. In *Proc. 3rd International Conf. on Computer Vision Systems*, Graz, Austria, Apr. 2003.
[12] B. Boom, G. Beumer, L. Spreeuwers, and R. Veldhuis. The effect of image resolution on the performance of a face recognition system. In *Proceedings of the Ninth International Conference on Control, Automation, Robotics and Vision*, pages 409–414, 2006.
[13] T. E. Boult and R. Scheirer, Walter J.and Woodworth. Faad: face at a distance. In *SPIE Defense and Security Symposium : Biometric Technology for Human Identification*, volume 6944, 2008.
[14] J. Czyz and L. Vandendorpe. Evaluation of LDA-based face verification with respect to available computational resources. In *Int'l Workshop on Pattern Recognition in Information Systems*, 2002.
[15] FG-NET consortium. FG-NET Aging Database. http://www.fgnet.rsunit.com/index.php.
[16] P. J. Flynn, K. W. Bowyer, and P. J. Phillips. Assessment of time dependency in face recognition: An initial study. In *Audio- and Video-Based Biometric Person Authentication*, 2003.
[17] N. Furl, P. Phillips, and A. O'Toole. Face recognition algorithms and the other-race effect: computational mechanisms for a developmental contact hypothesis. *Cognitive Science*, 26:797–815, 2002.
[18] A. Georghiades, P. Belhumeur, and D. Kriegman. From few to many: Illumination cone models for face recognition under variable lighting and pose. *IEEE Trans. Pattern Anal. Mach. Intelligence*, 23(6):643–660, 2001.
[19] G. Givens, J. Beveridge, B. Draper, P. Grother, and P. Phillips. How features of the human face affect recognition : A statistical comparison of three face recognition algorithms. In *CVPR*, 2004.
[20] G. Givens, J. Beveridge, B. Draper, and P. Phillips. Repeated Measures GLMM Estimation of Subject-Related and False Positive Threshold Effects on Human Face Verification Performance. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 40–40, 2005.
[21] G. Givens, J. R. Beveridge, B. A. Draper, and D. Bolme. A statistical assessment of subject factors in the pca recognition of human faces. In *CVPR Workshop on Statistical Analysis in Computer Vision*, 2003.
[22] R. Gross, J. Shi, and J. Cohn. Quo vadis face recognition? In *CVPR Workshop on Empirical Evaluation Methods in Computer Vision*, 2001.
[23] W. H. Ho, P. A. Watters, and D. Verity. Are younger people more difficult to identify or just a peer-to-peer effect. In W. G. Kropatsch, M. Kampel, and A. Hanbury, editors, *CAIP*, volume 4673, pages 351–359. Springer, 2007.
[24] R.-L. Hsu, J. Shah, and B. Martin. Quality assessment of facial images. In *Biometric Consortium Conference, 2006 Biometrics Symposium*, 2006.
[25] M. S. Keil, A. Lapedriza, D. Masip, and J. Vitria. Preferred spatial frequencies for human face processing are associated with optimal class discrimination in the machine. *PLoS ONE*, 3(7), 2008.
[26] H. Ling, S. Soatto, N. Ramanathan, and D. Jacobs. A study of face recognition as people age. In *IEEE Conf. ICCV*, 2007.
[27] C. Liu. Capitalize on dimensionality increasing techniques for improving face recognition grand challenge performance. *IEEE transactions on pattern analysis and machine intelligence*, 28(5):725–737, 2006.
[28] Z. Liu and S. Sakar. Outdoor recognition at a distance by fusing gait and face. *Image and Vision Computing*, 25(6), 2007.
[29] K. Messer, J. Matas, J. Kittler, J. Luettin, and G. Maitre. XM2VTSDB: The extended M2VTS database. In *Second International Conference on Audio and Video-based Biometric Person Authentication*, volume 964, pages 965–966, 1999.
[30] B. Moghaddam and A. Pentland. Probabilistic visual learning for object representation. *PAMI*, 19(7), 1997.
[31] H. Moon and J. Phillips. Computational and performance aspects of pca-based face recognition algorithms,. *Perception*, 30, 2001.
[32] E. Newton and P. Phillips. Meta-analysis of third-party evaluations of iris recognition. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, 39(1):4–11, Jan. 2009.

[33] A. O'Toole, P. Phillips, and A. Narvekar. Humans versus algorithms: Comparisons from the face recognition vendor test 2006. In *FGR, Amsterdam, The Netherlands*, 2008.
[34] U. Park, Y. Tong, and A. K. Jain. Face recognition with temporal invariance: A 3d aging model. In *FG*, pages 1–7. IEEE, 2008.
[35] J. Phillips, P. Grother, R. Micheals, D. Blackburn, E. Tabassi, and M. Bone. Frvt 2002 - evaluation report. *Technical Report*, 2002. http://www.frvt.org.
[36] 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 *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005. CVPR 2005*, volume 1, 2005.
[37] P. Phillips, H. Moon, S. Rizvi, and P. Rauss. The FERET Evaluation Methodology for Face-Recognition Algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(10):1090–1104, October 2000.
[38] P. Phillips, S. Sarkar, I. Robledo, P. Grother, and K. Bowyer. Baseline results for the challenge problem of HumanID using gaitanalysis. In *Fifth IEEE International Conference on Automatic Face and Gesture Recognition, 2002. Proceedings*, pages 130–135, 2002.
[39] P. Phillips, W. Scruggs, A. O'Toole, P. Flynn, K. Bowyer, C. Schott, and M. Sharpe. FRVT 2006 and ICE 2006 large-scale results. *National Institute of Standards and Technology, NISTIR*, 7408, 2007.
[40] N. Ramanathan and R. Chellappa. Face verification across age progression. *IEEE Transactions on Image Processing*, 15(11):3349, 2006.
[41] K. Ricanek Jr and E. Boone. The effect of normal adult aging on standard PCA face recognition accuracy rates. In *2005 IEEE International Joint Conference on Neural Networks, 2005. IJCNN'05. Proceedings*, volume 4, 2005.
[42] T. Sim, S. Baker, and M. Bsat. The cmu pose, illumination, and expression database. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(1):1615 – 1618, December 2003.
[43] R. Singh, M. Vatsa, and A. Noore. Face recognition with disguise and single gallery images. *Image and Vision Computing*, 2007.
[44] R. Singh, M. Vatsa, A. Noore, and S. K. Singh. Age transformation for improving face recognition performance. In A. Ghosh, R. K. De, and S. K. Pal, editors, *Pattern Recognition and Machine Intelligence, Second International Conference*, volume 4815 of *Lecture Notes in Computer Science*, pages 576–583. Springer, 2007.
[45] R. Veldhuis, A. Bazen, W. Booij, and A. Hendrikse. Hand-geometry recognition based on contour parameters. In *Proceedings of SPIE*, volume 5779, page 344, 2005.
[46] J. Wang, C. Zhang, and H. Shum. Face image resolution versus face recognition performance based on two global methods. In *Proceedings of Asia Conference on Computer Vision (ACCV'04)*, 2004.
[47] 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*, 2006.
[48] F. W. Wheeler, X. Liu, and P. Tu. Multi-frame super-resolution for face recognition. In *BTAS*, 2007.
[49] C. Xie, M. Savvides, and B. Kumar. Redundant class-dependence feature analysis based on correlation filters using frgc2.0 data. *Computer Vision and Pattern Recognition, 2005 IEEE Computer Society Conference on*, 3:153–153, June 2005.
[50] Y. Yacoob and L. Davis. Smiling faces are better for face recognition. In *In Proc. International Conference Face Recognition and Gesture Analysis*, pages 59–64, 2002.

A Meta-Analysis of Face Recognition Covariates: Supplemental Material

Yui Man Lui, David Bolme, Bruce A. Draper, J. Ross Beveridge, Geoff Givens, and P. Jonathon Phillips

I. Introduction to Supplemental Material

This supplemental materials document contains additional detail that has been omitted from the main paper. Primarily, this supplemental material consists of larger more detailed versions of the covariate effect summary tables presented in the main paper. For every table in the main paper, there's a matching table in this document which includes additional columns that indicate such things as the algorithm and the data set used in the study. There is also supplemental description primarily aimed at clarifying the additional information presented in tables.

There are 49 distinct algorithm implementations associated with results tabulated in this paper. These are enumerated in the Table VII. The first column provides a short name or key that is used to reference the algorithm in subsequent tables. The second and third columns provides a more detailed description along with a general type or category. The last two columns indicate the institution responsible for developing the algorithm and a citation, when available, describing the algorithm. There are 13 data sets represented in the findings summarized here, and these are enumerated in Table VIII.

In the full covariate tables, the column under the heading "controls" address how many other covariates were controlled for in the study. Other covariates were controlled for either by using a statistical model such as a Generalized Linear Model or by balancing the sample population with respect to another covariate. The controls are further enumerated in Table IX. The listing of covariates in Table IX divides covariates into three groups: properties of people (e.g. age), properties of images (e.g. resolution), and properties to how face recognition is being applied (e.g. indoors versus outdoors). The four remaining columns break the findings out according to algorithm and dataset, and identify the study by lead author and date. The columns labeled "Algorithm" and "Dataset" are short keys enumerated and defined in Tables VII and VIII respectively.

A. Age

Table X is a more complete version of the summary Table I that appears in the main paper. The effects shown on the left match between the two tables. The additional columns contain information about interactions, controls, algorithm, dataset, author, and year.

B. Time Between Images

Table XI is a more complete version of the summary Table II that appears in the main paper. The effects shown on the left match between the two tables as do the entries for Time Span and Same Day. The additional columns contain information about interactions, controls, algorithm, dataset, author, and year.

C. Gender

Table XII is a more complete version of the summary Table III that appears in the main paper. The effects shown on the left match between the two tables. The additional columns contain information about interactions, controls, algorithm, dataset, author, and year.

D. Expression

Table XIII is a more complete version of the summary Table IV that appears in the main paper. The effects shown on the left match between the two tables. The additional columns contain information about interactions, controls, algorithm, dataset, author, and year.

E. Resolution

Table XIV is a more complete version of the summary Table V that appears in the main paper. The effects shown on the left match between the two tables as do the entries for pixels between the eyes and image size. The additional covariates are interactions, controls, algorithms, dataset, author, and year.

F. Race

Table XV is a more complete version of the summary Table VI that appears in the main paper. The effects shown on the left match between the two tables. The additional covariates are interactions, controls, algorithms, dataset, author, and year.

TABLE VII

ALGORITHMS ASSOCIATED WITH FINDINGS SUMMARIZED IN THIS PAPER.

Key	Algorithm	Type	Institution	Citation
Fus(3)	FRVT2006 Fusion	Modern	NIST	[8]
SVM(4)	SVM+*	Summary	”Siemens,UCLA,UMD”	[26]
NJIT	Gabor+KFA	Modern	New Jersey Inst. Of Tech.	[27]
Anon1	FRVT 2006 A	Modern	anonymous	[9]
Anon2	FRVT 2006 B	Modern	anonymous	[9]
Anon3	FRVT 2006 C	Modern	anonymous	[9]
Imagis	Imagis	Commercial		
Viisage	Viisage	Commercial	Viisage	
VisSph	VisionSphere	Commercial	Vision Sphere Technologies Inc	
DrmMIRH	DreamMIRH	Commercial	”DREAM MIRH CO., LTD”	
FaceIt	FaceIt	Commercial	Visionics	
Gabor	2D log polar Gabor transform	Modern	WVU	[43]
Cog1	FaceVACS	Commercial	Cognitec	
Cog2	Cognitec	Commercial	Cognitec	
Idtx	Identix	Commercial	Identix	
Eyem	Eyematic	Commercial	Eyematic	
Unkn1	FRVT2000 Best	Modern	Unknown	
EBGM	EBGM	Modern	Colorado State Univ.	[11]
Neven	NevenVision	Commercial	NevenVision	
Lau	Lau Technologies	Commercial	Lau Technologies	
C-VIS	C-VIS	Commercial	Computer Vision and Automation GmbH	
Ts2	Tsinghua2	Modern	Tsinghua University	
SAIT	SAIT	Commercial	SAIT	
Neven1	NevenVision1	Commercial	NevenVision	
Idtx1	Identix1	Commercial	Identix	
Cog1	Cognitec1	Commercial	Cognitec	
Sagem1	Sagem1	Commercial	Sagem	
Sagem2	Sagem2	Commercial	Sagem	
PCA1	PCA	Subspace	Colorado State Univ.	[11]
PCA2	PCA	Subspace	Univ. North Carolina	[41]
PCA3	PCA	Subspace	NIST 1 (Angle between Vectors)	[31]
PCA4	PCA	Subspace	NIST 2 (L1 norm)	[31]
PCA5	PCA	Subspace	NIST 3 (L2 norms)	[31]
PCA6	PCA	Subspace	NIST 4 (L1 + Mahalanobis)	[31]
PCA7	PCA	Subspace	NIST 5 (Angle + Mahalanobis)	[31]
PCA8	PCA	Subspace	NIST 6 (L2 + Mahalanobis)	[31]
PCA9	PCA	Subspace	NIST 7 (Mahalanobis)	[31]
PCA10	PCA	Subspace	MIT (L2)	[30]
PCA11	PCA	Subspace	NIST/SAIC	
CMU	CFA	Subspace	Carnegie Mellon Univ.	[49]
ICA	ICA	Subspace	UC San Deigo	[6]
Bayes1	Bayesian Intra/Inter subspace	Subspace	MIT	[40]
Bayes2	Bayesian Intra/Inter subspace	Subspace	Colorado State Univ.	[11]
Bayes3	Bayesian Intra/Inter subspace	Subspace	MIT	[30]
LDA1	PCA+LDA	Subspace	University of Twente	[45]
PCA12	PCA	Subspace	Tsinghua University	[46]
LDA4	LDA	Classifier	Tsinghua University	[46]
LDA2	PCA+LDA	Classifier	Euclidean Distance : Univ. catholoque de Louvain	[14]
LDA3	PCA+LDA	Classifier	Cos : Univ. catholoque de Louvain	[14]

TABLE VIII

DATASETS ASSOCIATED WITH FINDINGS SUMMARIZED IN THIS PAPER.

Key	Collection Institution	Citation
FRGC	Notre Dame	[36]
ND2002	Notre Dame	[16]
FRVT2006	Notre Dame	[39]
FERET	NIST	[37]
HumanID	USF	[38]
HCINT	Visa Serv. - Dept. of State	[35]
MCINT	”NIST, NSWC, USF”	[35]
FGNet Aging Db.	FG-NET	[15]
MORPH	Univ. Norch Carolina	[41]
XM2VTS	University of Surrey	[29]
Dahlgren	”NIST, 2000”	[10]
GE	GE	[48]
WVU	West Virginia University	[44]

TABLE IX

CONTROLS USED BY SOME STUDIES TO ADDRESS THE INFLUENCE OF OTHER COVARIATES.

Key	Author	Year	Source	Subject	Image	Application
Ctr1	Beveridge	2009	[9]	Age, Gender, Race, Glasses	Focus, Tilt, Resolution	Algorithm, Elapsed Months, Indoor/Outdoor
Ctr2	Beveridge	2009	[7]	Age, Gender, Race, Glasses, Expression	Focus, Size Ratio, Tilt, Resolution, Focus, Fragmentation, Novelty	Elapsed Months, Indoor/Outdoor
Ctr3	Beveridge	2008	[8]	Age, Gender, Race, Glasses	Focus, Tilt, Resolution	Elapsed Months, Indoor/Outdoor
Ctr4	Ho	2007	[23]			Age Balanced Gallery
Ctr5	Givens	2005	[20]	Age, Gender, Bangs, Facial Hair, Eyes		
Ctr6	Givens	2004, 2003	[19],[21]	Age, Gender, Race, Skin, Glasses, Facial Hair, Makeup, Bangs, Expression, Mouth, Eyes		

TABLE X

RELATING AGE OF A PERSON TO FACE RECOGNITION PERFORMANCE.

Older Easier								Younger Easier							
+17.5%	15%	10%	5%	0%	5%	10%	15%	+17.5%	Interactions with other covariates	Controls for other covariates (Table IX)	Algorithm (Table VII)	Dataset (Table VIII)	Author	Year	Source
		*	*						Outdoor/Indoor	Ctr1	Anon1	FRVT2006	Beveridge	2009	[9]
	*	*							Outdoor/Indoor	Ctr1	Anon2	FRVT2006	Beveridge	2009	[9]
			*							Ctr1	Anon3	FRVT2006	Beveridge	2009	[9]
		*								Ctr2	NJIT	FRGC	Beveridge	2009	[7]
			*							Ctr2	PCA1	FRGC	Beveridge	2009	[7]
*										Ctr2	CMU	FRGC	Beveridge	2009	[7]
*											Bayes2	FG-NET	Park	2008	[34]
*											Gabor	FG-NET + WVU	Singh	2007	[44]
				*						Ctr4	PCA1	FERET	Ho	2007	[23]
	*									Ctr5	PCA1	FERET	Givens	2005	[20]
			*							Ctr6	EBGM	FERET	Givens	2004	[19]
		*								Ctr6	PCA1	FERET	Givens	2004	[19]
		*								Ctr6	Bayes2	FERET	Givens	2004	[19]
		*								Ctr6	PCA1	FERET	Givens	2003	[21]
	*								Gender		Cog	HCINT	Phillips	2002	[35]
*									Gender		Idtx	HCINT	Phillips	2002	[35]
		*							Gender		Eyem	HCINT	Phillips	2002	[35]
			*						Gender		Imagis	HCINT	Phillips	2002	[35]
		*							Gender		Viisage	HCINT	Phillips	2002	[35]
		*							Gender		VisSph	HCINT	Phillips	2002	[35]
		*							Gender		C-Vis	HCINT	Phillips	2002	[35]
				*					Gender		DrmMIRH	HCINT	Phillips	2002	[35]

TABLE XI

RELATING ELAPSED TIME BETWEEN IMAGES TO FACE RECOGNITION PERFORMANCE.

Less Time								More Time									
+17.5%	15%	10%	5%	0%	5%	10%	15%	+17.5%	Time Span	Same Day	Interactions with other covariates	Controls (Table IX)	Algorithm (Table VII)	Dataset (Table VIII)	Author	Year	Source
			*						120	No			SVM(4)	FG-NET	Ling	2007	[26]
*									120	No			PCA2	MORPH	Ricanek	2005	[41]
*									108	No			Bayes1	Passports	Ramanathan	2006	[40]
*									37	Yes			PCA2	FERET	Ricanek	2005	[41]
*									37	Yes			Imagis	HCINT	Phillips	2002	[35]
*									37	Yes			Viisage	HCINT	Phillips	2002	[35]
*									37	Yes			VisSph	HCINT	Phillips	2002	[35]
		*							37	Yes			DrmMIRH	HCINT	Phillips	2002	[35]
		*							37	Yes			Cog2	HCINT	Phillips	2002	[35]
		*							37	Yes			Idtx	HCINT	Phillips	2002	[35]
	*								37	Yes			Eyem	HCINT	Phillips	2002	[35]
*									37	Yes			C-VIS	HCINT	Phillips	2002	[35]
*									24	Yes	Expression		PCA13	FERET	Bartlett	1998	[6]
*									24	Yes			ICA	FERET	Bartlett	1998	[6]
		*							8	Yes		Ctr2	NJIT	FRGC	Beveridge	2009	[7]
			*						8	Yes		Ctr2	PCA1	FRGC	Beveridge	2009	[7]
				*					8	Yes		Ctr2	CMU	FRGC	Beveridge	2009	[7]
*									3	Yes			EBGM	HumanID	Liu	2007	[28]
			*						2.5	No			PCA1	ND2002	Flynn	2003	[16]

TABLE XII

RELATING A PERSON'S GENDER TO FACE RECOGNITION PERFORMANCE.

Male Easier								Female Easier							
+17.5%	15%	10%	5%	0%	5%	10%	15%	+17.5%	Interactions with other covariates	Controls (Table IX)	Algorithm (Table VII)	Dataset (Table VIII)	Author	Year	Source
			*							Ctr1	Anon1	FRVT2006	Beveridge	2009	[9]
		*	*	*					Outdoor/Indoor	Ctr1	Anon2	FRVT2006	Beveridge	2009	[9]
		*								Ctr1	Anon3	FRVT2006	Beveridge	2009	[9]
			*	*					Outdoor/Indoor	Ctr3	Fus(3)	FRVT2006	Beveridge	2008	[8]
				*						Ctr2	PCA11	FRGC	Beveridge	2009	[7]
			*							Ctr2	CMU	FRGC	Beveridge	2009	[7]
		*								Ctr2	NJIT	FRGC	Beveridge	2009	[7]
	*									Ctr5	PCA1	FERET	Givens	2005	[20]
			*							Ctr6	EBGM	FERET	Givens	2004	[19]
			*							Ctr6	PCA1	FERET	Givens	2004	[19]
			*							Ctr6	Bayes2	FERET	Givens	2004	[19]
			*							Ctr6	PCA1	FERET	Givens	2003	[21]
*	*	*	*						Age		Cog2	HCINT	Phillips	2002	[35]
	*	*	*	*	*				Age		Idtx	HCINT	Phillips	2002	[35]
*	*	*	*						Age		Eyem	HCINT	Phillips	2002	[35]
				*	*				Expression and Lighting		FaceIt	AR	Gross	2001	[22]

TABLE XIII

RELATING A PERSON'S EXPRESSION TO FACE RECOGNITION PERFORMANCE.

Same								Different							
+17.5%	15%	10%	5%	0%	5%	10%	15%	+17.5%	Interactions with other covariates	Controls (Table IX)	Algorithm (Table VII)	Dataset (Table VIII)	Author	Year	Source
*										Ctr2	PCA1	FRGC	Beveridge	2009	[7]
		*								Ctr2	NJIT	FRGC	Beveridge	2009	[7]
		*								Ctr2	CMU	FRGC	Beveridge	2009	[7]
			*								PCA1	FERET	Givens	2004	[19]
			*								EBGM	FERET	Givens	2004	[19]
				*							Bayes2	FERET	Givens	2004	[19]
		*									PCA1	FERET	Givens	2003	[21]
Neutral vs. Smile								Smile vs. Neutral							
				*						Ctr2	PCA1	FRGC	Beveridge	2009	[7]
				*						Ctr2	CMU	FRGC	Beveridge	2009	[7]
			*							Ctr2	NJIT	FRGC	Beveridge	2009	[7]
Regular vs. Regular								Regular vs. Regular							
					*						Idtx	Dahlgren	Blackburn	2001	[10]
					*						Lau	Dahlgren	Blackburn	2001	[10]
					*						C-Vis	Dahlgren	Blackburn	2001	[10]

TABLE XIV

RELATING IMAGE RESOLUTIONS TO FACE RECOGNITION PERFORMANCE.

High Resolution +17.5%	15%	10%	5%	0%	5%	10%	15%	Low Resolution +17.5%	Pixels between the eyes	Interactions with other covariates	Controls (Table IX)	Algorithm (Table VII)	Dataset (Table VIII)	Author	Year	Source
*	*	*	*	*	*				202 $\sim$ 134	Outdoor/Indoor	Ctr1	Anon1	FRVT2006	Beveridge	-	[9]
*	*	*	*	*	*				202 $\sim$ 134	Outdoor/Indoor	Ctr1	Anon2	FRVT2006	Beveridge	-	[9]
*	*	*	*	*	*	*			202 $\sim$ 134	Outdoor/Indoor	Ctr1	Anon3	FRVT2006	Beveridge	-	[9]
								*	180 $\sim$ 120		Ctr2	PCA1	FRGC	Beveridge	2009	[7]
	*								180 $\sim$ 120		Ctr2	CMU	FRGC	Beveridge	2009	[7]
*									180 $\sim$ 120		Ctr2	NJIT	FRGC	Beveridge	2009	[7]
*									210 $\sim$ 137		Ctr3	Fus(3)	FRVT2006	Beveridge	2008	[8]
				*					190 $\sim$ 110			Viisage	FRVT2006	O'Toole	2008	[33]
								*	190 $\sim$ 110			Ts2	FRVT2006	O'Toole	2008	[33]
				*					190 $\sim$ 110			SAIT	FRVT2006	O'Toole	2008	[33]
			*						190 $\sim$ 110			Neven1	FRVT2006	O'Toole	2008	[33]
					*				190 $\sim$ 110			Idtx1	FRVT2006	O'Toole	2008	[33]
*	*	*	*	*					400 $\sim$ 75	Controlled/Uncontrolled		Cog1	FRVT2006	Phillips	2007	[39]
					*	*	*	*	400 $\sim$ 75	Controlled/Uncontrolled		Idtx1	FRVT2006	Phillips	2007	[39]
			*						400 $\sim$ 75	Controlled/Uncontrolled		Neven1	FRVT2006	Phillips	2007	[39]
*	*	*	*	*					400 $\sim$ 75	Controlled/Uncontrolled		Sagem2	FRVT2006	Phillips	2007	[39]
		*	*						400 $\sim$ 75	Controlled/Uncontrolled		SAIT	FRVT2006	Phillips	2007	[39]
			*	*					400 $\sim$ 75	Controlled/Uncontrolled		Ts2	FRVT2006	Phillips	2007	[39]
			*	*					400 $\sim$ 75	Controlled/Uncontrolled		Viisage	FRVT2006	Phillips	2007	[39]
*									48 $\sim$ 17			Idtx	GE+FERET	Wheeler	2006	[48]

High Resolution +17.5%	15%	10%	5%	0%	5%	10%	15%	Low Resolution +17.5%	Pixels between the eyes	Interactions with other covariates	Controls (Table IX)	Algorithm (Table VII)	Dataset (Table VIII)	Author	Year	Source
				*					60 $\sim$ 15			Idtx	Dahlgren	Blackburn	2001	[10]
			*						60 $\sim$ 15			Lau	Dahlgren	Blackburn	2001	[10]
	*								60 $\sim$ 15			C-VIS	Dahlgren	Blackburn	2001	[10]

High Resolution +17.5%	15%	10%	5%	0%	5%	10%	15%	Low Resolution +17.5%	Image size	Interactions with other covariates	Controls (Table IX)	Algorithm (Table VII)	Dataset (Table VIII)	Author	Year	Source
				*					128x128 vs 32x32			LDA1	FRGC	Boom	2006	[12]
			*						32x32 vs 16x16			LDA1	FRGC	Boom	2006	[12]
				*					144x108 vs 64x48			PCA12	AR	Wang	2004	[46]
			*						64x48 vs 32x24			PCA12	AR	Wang	2004	[46]
				*					144x108 vs 64x48			LDA4	AR	Wang	2004	[46]
		*							64x48 vs 32x24			LDA4	AR	Wang	2004	[46]
				*					64x64 vs 32x32	Dist. Measure		LDA(2)	XM2VTS	Czyz	2002	[14]
			*	*					64x64 vs 16x16	Dist. Measure		LDA(2)	XM2VTS	Czyz	2002	[14]
*	*	*	*						64x64 vs 8x8	Dist. Measure		LDA(2)	XM2VTS	Czyz	2002	[14]

TABLE XV

RACE TABLE

Caucasian Easier +17.5%	15%	10%	5%	0%	5%	10%	15%	East Asian Easier +17.5%	Interactions with other covariates	Controls (Table IX)	Algorithm (Table VII)	Dataset (Table VIII)	Author	Year	Source
							*			Ctr1	Anon1	FRVT2006	Beveridge	2009	[9]
							*	*	Outdoor/Indoor	Ctr1	Anon2	FRVT2006	Beveridge	2009	[9]
								*		Ctr1	Anon3	FRVT2006	Beveridge	2009	[9]
					*	*			Outdoor/Indoor	Ctr3	Fus(3)	FRVT2006	Beveridge	2008	[8]
					*					Ctr2	PCA1	FRGC	Beveridge	2009	[7]
						*				Ctr2	NJIT	FRGC	Beveridge	2009	[7]
						*				Ctr2	CMU	FRGC	Beveridge	2009	[7]
					*					Ctr6	PCA1	FERET	Givens	2004	[19]
					*					Ctr6	Bayes2	FERET	Givens	2004	[19]
					*					Ctr6	EBGM	FERET	Givens	2004	[19]
					*					Ctr6	PCA1	FERET	Givens	2003	[21]
					*	*	*	*	Dist. Measure		PCA(7)	FERET	Furl	2002	[17]
						*					Bayes3	FERET	Furl	2002	[17]