Applied Cognitive Psychology, *Appl. Cognit. Psychol.* **27**: 718–724 (2013)
Published online 11 November 2013 in Wiley Online Library (wileyonlinelibrary.com) **DOI**: 10.1002/acp.2967

Effects of Image Preprocessing on Face Matching and Recognition in Human Observers

CHANG HONG LIU[1]*, WENFENG CHEN[2], HU HAN[3] and SHIGUANG SHAN[3]

[1]*School of Design, Engineering & Computing, Bournemouth University, Poole, UK*
[2]*State Key Lab of Brain and Cognitive Science, Institute of Psychology, Chinese Academy of Sciences, Beijing, China*
[3]*Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China*

Summary: In person identification, recognition failure due to variations of illumination is common. In this study, we employed image-processing techniques to tackle this problem. Participants performed recognition and matching tasks where the face stimuli were either original images or computer-processed images in which shading was weakened via a number of image-processing techniques. The results show that whereas recognition accuracy in a memory task was unaffected, some of the techniques significantly improved the identification performance in a face-matching task. We conclude that relative to long-term face memory, face matching is more susceptible to discrepancy of shading in different images of a face. Reducing the discrepancy by certain preprocessing techniques can facilitate person identification when original face images contain large illumination differences. Copyright © 2013 John Wiley & Sons, Ltd.

EFFECTS OF IMAGE PREPROCESSING ON FACE MATCHING AND RECOGNITION

Recognition of unfamiliar faces depends heavily on illumination conditions. A change of illumination from one image to another has been found to drastically undermine recognition performance (Adini, Moses, & Ullman, 1997; Braje, 2003; Braje, Kersten, Tarr, & Troje, 1998; Hill & Bruce, 1996; Liu, Bhuiyan, Ward, & Sui, 2009; Longmore, Liu, & Young, 2008; Tarr, Georghiades, & Jackson, 2008; Tarr, Kersten, & Bülthoff, 1998). This illumination effect is likely to have detrimental implications for forensic investigations in which face recognition is used to identify a suspect. Images of the same person are not always taken under similar illumination conditions, and therefore, matching images or mug shots against each other may be very challenging. This is likely to be particularly true for images taken by security cameras under dim street lightings, which may result in strong shadows. One way to tackle this problem may be to take photos of each offender under various illuminations. However, because the number of necessary illumination conditions is difficult to specify, this multiple photo approach is likely to be impractical.

The problem of illumination is equally challenging for machine vision. It has been found that variations caused by illumination are more significant than the inherent differences between individuals (Adini et al., 1997; Chen, Yin, Zhou, Comaniciu, & Huang, 2006). As Adini et al. (1997) point out: 'the variations between the images of the same face due to illumination and viewing direction are almost always larger than the image variations due to a change in face identity.' This means that images of different people can appear more similar than images of the same person. This could lead to false identification. Engineers have invested substantial energy on this problem in the past decades (Zou, Kittler, & Messer, 2007). Recent research has shown that preprocessing images prior to the stage of recognition can be a useful solution. Examples of the original and preprocessed images are given in Figure 1. The first row of faces in the figure shows the original images. Other rows show images processed by different algorithms. Face images created by these processing techniques are more likely to be recognized successfully by automatic systems in comparison with original (unprocessed) images (Chen, Er, & Wu, 2006; Chen, Yin, et al., 2006; Jobson, Rahman, & Woodell, 1997).

However, it is currently unknown whether the same preprocessing methods are also beneficial for human face recognition. Image preprocessing is analogous to the step of normalization in the human visual system. It may function as a step to reveal image-invariant facial features. As Figure 1 shows, all preprocessing methods appear to make the key facial features more visible relative to the unprocessed images. Having a less shadowy appearance, the image variation due to illumination changes is greatly reduced in the processed images. As image similarity is an important predictor for face recognition in humans, it is likely that the same image-processing techniques can also benefit human recognition performance. The main aim of the current research is to examine this possibility.

In order to assess whether image-processing techniques facilitate both long- and short-term memory of faces, old/new recognition and sequential matching tasks were employed. Unlike the matching task, where a face is shown and stored in short-term memory before it is compared with another face, the old/new recognition task also requires storing the learnt faces in long-term memory. Transferring a face to long-term memory involves a consolidation process that may be carried out differentially for various aspects of trained faces. In other words, whereas some features of the learned face may be selectively maintained in memory, others may be weakened or lost. Because of this, the decay of certain physical characteristics of face stimuli could occur at different rates. The illumination effect may vary accordingly. Therefore, preprocessed images may affect the two tasks differently. In addition to our interests in this theoretical issue, we also aimed to assess whether the image-preprocessing techniques could be useful in forensic practice. In situations such as manhunts, television appeals and surveillance inspections, both long-term recognition memory and face-matching accuracy would be required.

*Correspondence to: Chang Hong Liu, PhD, School of Design, Engineering & Computing, Bournemouth University, Poole BH12 5BB, UK.
E-mail: liuc@bournemouth.ac.uk

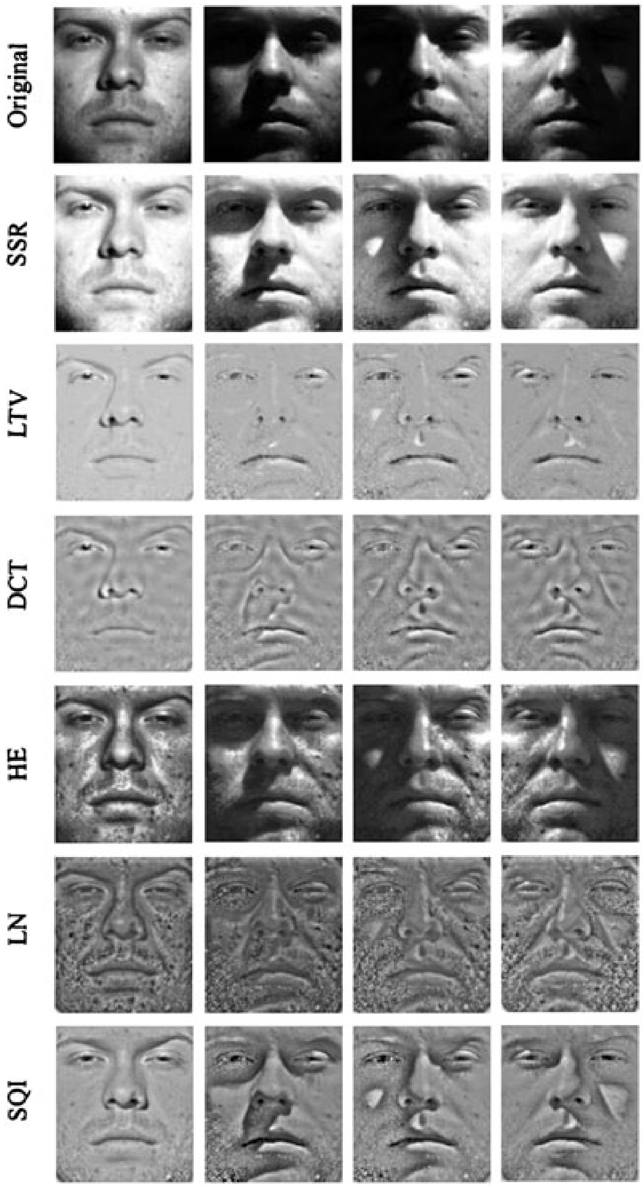

Figure 1. Examples of face stimuli used in the study. Each row represents a different image condition. The first row shows four original, unprocessed images of a face taken under different lighting conditions. The remaining rows show corresponding images transformed from the four originals via a variety of preprocessing methods. Image conditions illustrated in the first four rows were used in Experiment 1, and all seven conditions in this figure were used in Experiment 2. SQI, self-quotient image; LN, local normalization; HE, histogram equalization; DCT, discrete cosine transform; LTV, logarithmic total variation; SSR, single-scale retinex

Therefore, we conducted two experiments to investigate whether preprocessing techniques can facilitate face recognition memory (Experiment 1) and face-matching performance (Experiment 2). In both experiments, we tested the prediction that some image prepossessing techniques could facilitate recognition performance.

EXPERIMENT 1

Our first experiment examined the effect of image-preprocessing techniques on recognition memory. A standard old/new recognition task was employed in which participants were instructed to remember a number of faces presented during a training session. In order to facilitate training, participants were required to pair names with faces. Subsequently, they were asked to identify the trained faces among distractor faces in a test session. The trained faces were presented either in the same or in a new illumination during the test session. The same illumination condition was used as a baseline for the effect of new illumination. In the experimental groups, the face stimuli for both training and test sessions were processed with image-preprocessing techniques. In the control group, the original, unprocessed face images were used for both training and test sessions. Recognition performance of these conditions was then compared. Our main interest in this study was the effect of image manipulation when a face was trained and tested under difficult illumination conditions.

Three preprocessing algorithms were used to manipulate the face stimuli: single-scale retinex (SSR), logarithmic total variation (LTV) and discrete cosine transform (DCT). Examples of these are given in Figure 1. SSR is an image enhancement algorithm that improves the brightness, contrast and sharpness of an image through dynamic range compression. It synthesizes contrast enhancement and colour constancy by performing nonlinear spatial/spectrum transformation that mimics traits seen in the human visual system (Jobson et al., 1997). This method weakens the darkness of shadows and reveals some of the obscured information in the original image regarding face shape and texture. LTV is a model that minimizes the notorious halo (reflectance) artefacts and leaves only the small-scale facial structures (i.e. eyes, nose, mouth and eyebrows; Chen, Yin, et al., 2006). Finally, DCT produces a more thorough transformation of the original image by using a method of image compression that nearly abolishes the contrast created by shading in the original image (Chen, Er, et al., 2006).

In addition to the three processing methods, a number of other preprocessing algorithms are also available. However, it is not practical to include all of them in a between-participant design. We chose the three algorithms in this experiment mainly because of their relative merit and success in the computer literature. We aimed to evaluate whether these algorithms are equally effective for human observers. In addition, our choice of the algorithms also took the human face recognition literature into consideration. It is well known that human observers rely on 'mass' or low spatial frequencies that often carry information about shading (e.g. Bruce, Hanna, Dench, Healey, & Burton, 1992; Davies, Ellis, & Shepherd, 1978; Liu, Collin, Rainville, & Chaudhuri, 2000). We decided to use the SSR method because it preserves shading information while revealing some key facial features that are obscured in the original images. The four types of images in this experiment contained different amounts of shading: the original images contained the strongest shading, followed by the SSR, LTV and DCT images, respectively.

Method

Participants

A total of 101 undergraduate students (36 male, 65 female) from the University of Hull participated in this study.

Participants' ages ranged from 18 to 57 years ($Mdn = 20$), and all participants had normal or corrected-to-normal vision. An approximately equal number of participants ($N = 25$ or 26) were randomly assigned to each of the four image conditions (i.e. SSR, LTV, DCT and original images).

Materials

The original images were obtained from the Extended Yale Face Database B (Georghiadis, Belhumeur, & Kriegman, 2001), which contains 28 frontal views of individuals under 64 illumination conditions. We used 24 male faces from the database for this study. Images with extreme lighting angles (110° to 130° from one side or the top of a face) were eliminated from use as these faces are almost completely concealed in dark shadows. In addition, bottom-lit faces were eliminated from use as they are less common in reality and are known to disrupt face processing even without illumination changes (e.g. Hill & Bruce, 1996; Liu, Collin, Burton, & Chaudhuri, 1999). After excluding the extreme lighting angles and bottom-lit conditions, the remaining 21 illumination conditions were used in this study. Each illumination condition was defined by a combination of elevation and azimuth parameters. Elevation was chosen from 0°, 15°, 20°, 40°, 45° and 65°, whereas azimuth was chosen from 0°, ±15°, ±20°, ±25°, ±35°, +50°, ±60° and ±70° lightings. The parameter 0° indicates straight frontal lighting, whereas the negative and positive signs indicate side lightings from the left and right, respectively. With the exception of 50°, all side lightings had both left and right directions. The database does not have a complete combination of these parameters. For example, 0° azimuth was only combined with 0°, 20° and 45° elevation. However, because the purpose of this study was to find out whether image-processing techniques can facilitate recognition performance, the specific combinations of the lighting angles were of little significance. As stated previously, four types of images were used: the original images and the images created by the three preprocessing methods SSR, LTV, and DCT. Details of the these methods are described in Jobson et al. (1997), Chen, Yin, et al. (2006), and Chen, Er, et al. (2006), respectively.

In the face–name training session, a total of 12 common English first names were used (e.g. James, Peter and Frank). These ranged from four to six characters long. All stimuli were shown in greyscale on a neutral grey background. The experiment was run on a Pentium 4 computer, with a 21-in. monitor display (Sony Trinitron, GDM-F520, San Diego, CA, USA). The screen resolution was set at 1280×960. The vertical frequency of the monitor was set at 120 Hz. The experiment was run on a Pentium 4 computer. The software for experimental control was written in MATLAB 6.5 (The MathWorks, Inc., Natick, Massachusetts, USA) for PC, with Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997). The faces were normalized to the same interocular distance and subtended 5×4 degrees of visual angle.

Procedure

Participants were tested individually. Instructions were given on the screen. The task consisted of a training session and a test session. In the first part of the training session, participants were presented with 12 target faces in the centre of the screen, one at a time. A name was presented simultaneously below each face, with a random assignment of names to faces for each participant. Each face–name pair was shown for 5 seconds. Participants were asked to remember the presented pairs. In the second part of the training session that followed immediately afterwards, each presented face was shown again individually for 5 seconds, but this time with a row of four names at the bottom of the screen. One of these names had been paired with the face in the previous session, and the rest were randomly chosen from the 12 names. The order of the names on the screen was random in each trial. The participants had to indicate which name was the correct one for the face by pressing one of four corresponding keys. The names remained on the screen until a response was made. Performance feedback was given, and when an incorrect name was chosen, the correct answer would be shown. The same face–name matching task was repeated after all 12 faces had been shown. Therefore, including the first part of the training session (i.e. the face–name presentation block), each target face was exposed a total of three times during training. Pilot testing suggested that this level of training would prevent both ceiling and floor performances in the subsequent test session.

The test session began immediately after the training session. Brief instructions were given on the screen. In this session, the 12 trained (target) faces were mixed with 12 new (distractor) faces and shown one at a time in the centre of the screen. However, the names were not presented. To minimize a possible recency effect, the order of presentation of the target faces was the same as in the training session, but distractor faces were randomly inserted into the sequence between targets. Participants were asked to decide whether the faces presented at the test session had been shown during the training session. They did this by pressing a key labelled 'Yes' if the face was seen during the training session or a key labelled 'No' if it had not. The test face remained on the screen until a response key was pressed, which initiated a new test trial.

Because facial distinctiveness and pairing of lighting conditions used at the training and test sessions could have variable effects on recognition memory, we assigned the same set of face images to all experimental and control groups. In the training session, 12 pairs of azimuth-elevation lighting parameters were randomly chosen from the pool of 21 pairs, with the constraints that they consisted of three frontal and nine left/right lighting conditions. Each of these 12 lighting parameters was randomly assigned to one of the 12 target faces. In the test session, six target faces had the same lighting parameters as in the training session (i.e. identical images), and six target faces were shown with new lighting parameters. The six pairs of replaced parameters were assigned randomly to six distractor faces. The new lighting parameters for the six target faces were also assigned to the six remaining distractor faces. The matrix of 21 lighting conditions was shuffled for every group of four participants (one participant for each image condition). This was to guarantee that participants in different image conditions saw exactly the same faces under the same lighting conditions. The purpose of this was to ensure that any differences among the results of the four image conditions were not due to some facial characteristics or some lighting conditions being easier to remember than others.

To maximize the applicability of our findings to different faces and lighting conditions, we randomly assigned a set of 12 target and 12 distractor faces to every four participants, three of these for the experimental groups and one for the control group. This guaranteed that a variety of lighting conditions were included in the study and that the faces used in the four groups were identical. The presentation order of the faces in the four conditions was also identical but was randomized for each newly assigned set of faces.

Design

We employed a mixed design, where image type (original, SSR, LTV and DCT) was a between-participant variable and illumination (same, different) a within-participant variable. The dependent variables in this study were sensitivity (d') and criterion (c), which were computed from hits and false alarms.

Results and discussion

An alpha level of .05 was adopted for all statistical analyses. The mean d' results are shown in Figure 2. The main effect of image condition was marginally significant, $F(3, 97) = 2.58$, $p = .058$, $\eta^2 = .07$. Responses for the same illumination were more accurate than for different illumination, $F(1, 97) = 97.93$, $p < .001$, $\eta^2 = .50$. No interaction was found between the two variables, $F(3, 97) = 0.59$, $p = .62$, $\eta^2 = .02$. Multiple comparison of means with Bonferroni correction showed that the marginal effect was mainly due to the overall better performance for the SSR than for the DCT condition. Other differences were statistically negligible.

The criterion data, along with hit rates and false alarm rates and d', are shown in Table 1. The criterion results were not affected by image type, $F(3, 97) = 0.61$, $p = .61$, $\eta^2 = .02$. However, the mean criterion results for different illumination were more conservative than same illumination, $F(1, 97) = 132.11$, $p < .001$, $\eta^2 = .58$. The interaction between the two variables was not significant, $F(3, 97) = 1.27$, $p = .29$, $\eta^2 = .04$.

The results suggest that although some processed face images (SSR) are better than others (DCT), there was no difference between the results of unprocessed original images and the best results of processed images (SSR). The lack of interaction suggests that sensitivity in all image conditions was equally affected by a change of illumination from learning to test. That is, image-preprocessing

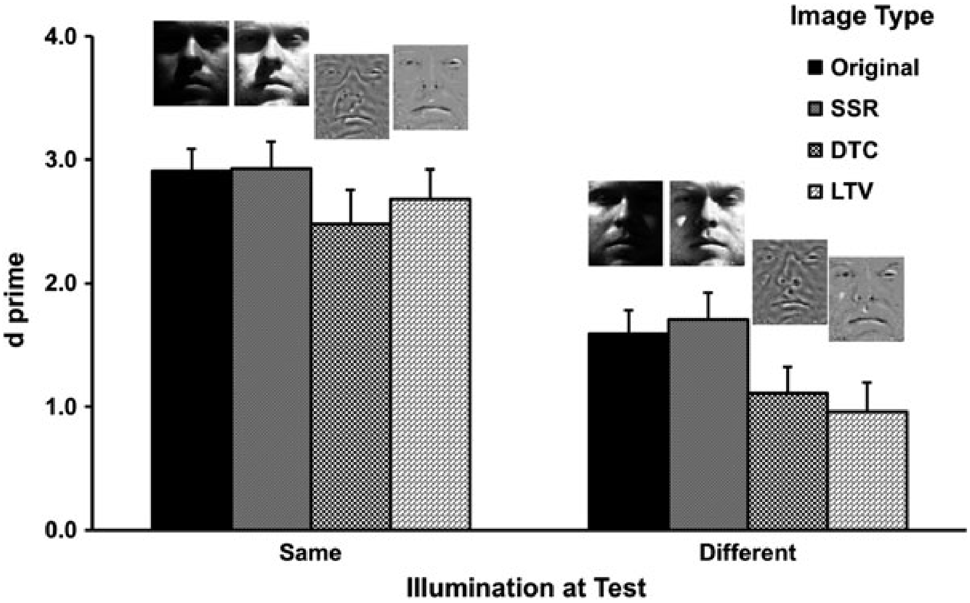

Figure 2. Mean d' results as a function of image type and illumination change in Experiment 1. Error bars represent one standard error above the means. SSR, single-scale retinex; DCT, discrete cosine transform; LTV, logarithmic total variation

Table 1. Means and standard deviations (in parenthesis) for proportions of hits and false alarms, d' and criterion (c) as a function of image type and illumination in Experiment 1

	Illumination at test							
	Same				Different			
	Original	SSR	DCT	LTV	Original	SSR	DCT	LTV
Hits	0.91 (0.12)	0.93 (0.14)	0.88 (0.13)	0.90 (0.13)	0.84 (0.16)	0.80 (0.15)	0.76 (0.20)	0.79 (0.19)
FAs	0.42 (0.21)	0.37 (0.18)	0.43 (0.17)	0.53 (0.20)	0.14 (0.13)	0.16 (0.20)	0.27 (0.27)	0.23 (0.19)
d'	2.91 (0.90)	2.93 (1.09)	2.48 (1.38)	2.68 (1.22)	1.59 (0.96)	1.71 (1.07)	1.11 (1.06)	0.96 (1.19)
c	−0.16 (0.65)	−0.40 (0.54)	−0.25 (0.53)	−0.31(0.55)	0.56 (0.50)	0.50 (0.54)	0.32 (0.75)	0.55 (0.56)

Note: SSR, single-scale retinex; LTV, logarithmic total variation; DCT, discrete cosine transform; FAs, false alarms.

techniques did not improve recognition performance. The overall more impaired accuracy and more conservative response bias for faces trained and tested in the different illumination conditions are consistent with the literature (e.g. Liu et al., 2009). The key finding from this experiment, however, is that these illumination effects on human long-term face memory are not easily eradicated by image-preprocessing techniques despite their proven effectiveness for machine face recognition.

EXPERIMENT 2

The purpose of this experiment was to investigate whether reducing or removing dark shadows with image-processing techniques alleviates recognition impairments due to change of illumination in a sequential matching task. We only tested the conditions where there was a change of illumination in this experiment, because matching identical images without involvement of long-term memory was expected to create ceiling performance.

In addition to the face stimuli used in Experiment 1, three new types of preprocessed images were included in this experiment: histogram equalization (HE), local normalization (LN) and self-quotient (SQI). This was to ensure that as many as currently available preprocessing techniques were evaluated. Unlike the old/new recognition task, the sequential matching task enables this because participants can be easily tested in a within-participant design. Examples of the three additional image types are provided in the bottom three rows of Figure 1. HE allows for areas of lower local contrast to gain a higher contrast (Gonzalez & Woods, 1992). This is accomplished by effectively spreading out the most frequent intensity values. LN reduces or removes shading or uneven illuminations and highlights the markings on the face (Xie & Lam, 2006). But unlike HE, it removes shading at the same time. Finally, SQI weakens or removes shading (Wang, Li, & Wang, 2004). However, compared with other methods, it preserves shading in face images with extreme illumination conditions, producing images that are visually similar to those processed by SSR.

Method

Participants
A total of 52 undergraduate students (16 male, 36 female) from the University of Hull took part in this study. Participants' ages ranged from 18 to 41 years ($Mdn = 19$), and all participants had normal or corrected-to-normal vision.

Materials and procedure
The apparatus and face stimuli were identical to those in Experiment 1 except that three additional types of images (i.e. HE, LN and SQI) were also included.

The seven types of image stimuli were tested in seven blocks of trials. The images within each block were of the same type, and each block consisted of 16 matching trials. In addition, there were four practice trials at the beginning of the experiment. The order of the seven blocks was counterbalanced by a Latin squares design. The faces in each block were presented in a random order for each participant.

Each matching trial consisted of a learn face and a test face of the same image type presented one after the other in the centre of the screen. Trials began with a 500-millisecond central fixation cross, followed by a 500-millisecond blank screen. Following this, a learn face was presented for 3 seconds, and then, a test face appeared after a 500-millisecond blank screen. Half of the test faces were the same person as the learn face (i.e. targets), and the remaining half were different from the learn face (i.e. distractors). The learn face and the test face were always presented in different illumination conditions. Participants were instructed to judge whether the two sequentially presented face images were of the same person and to indicate their decision by pressing one of two keys labelled 'Yes' or 'No'. The test face remained on screen until the participant responded.

Results and discussion

Figure 3(A) shows the d' results. Analysis of variance (ANOVA) on these sensitivity results showed a significant main effect of image type, $F(6, 306) = 6.75$, $p < .001$, $\eta^2 = .12$. Bonferroni tests showed that the matching performance for SSR and SQI was significantly better than for

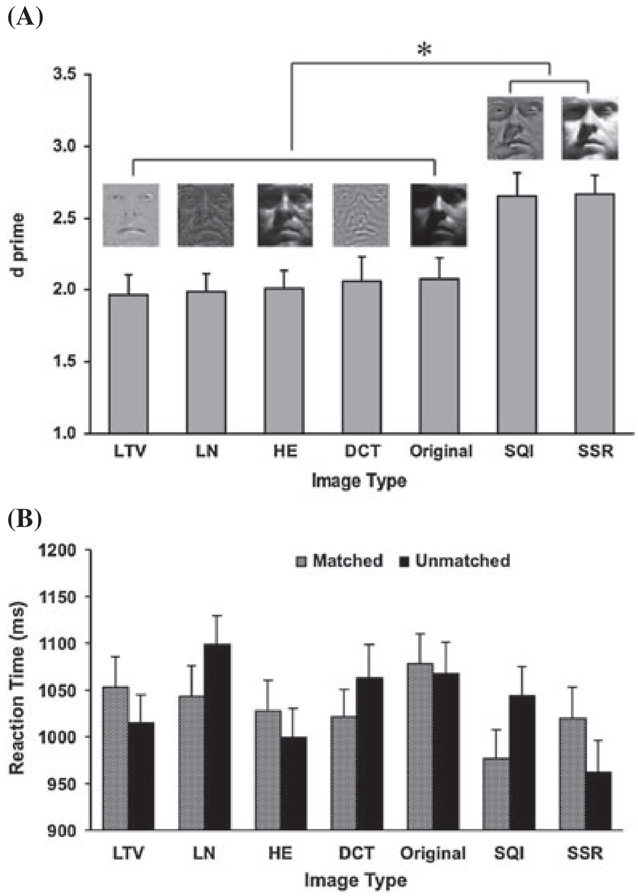

Figure 3. Results as a function of image type in Experiment 2. (A) Mean d' results and (B) mean reaction time results. Error bars represent one standard error above the means. LTV, logarithmic total variation; LN, local normalization; HE, histogram equalization; DCT, discrete cosine transform; SQI, self-quotient image; SSR, single-scale retinex

the originals, DCT, HE, LTV and LN. Performance for SSR was also significantly better than for DCT, but SQI produced similar results to DCT. Results for SSR and SQI were comparable. The performance for the original, DCT and other types of preprocessed images were not different from each other.

Figure 3(B) shows the reaction time results for correct responses in matched and unmatched trials. We conducted a two-way repeated-measure ANOVA, using image type and trial type (matched vs unmatched faces) as within-participant variables. We found a significant image type, $F(6, 294) = 2.93$, $p < .01$, $\eta^2 = .06$, and a significant two-way interaction, $F(6, 294) = 3.65$, $p < .01$, $\eta^2 = .07$. The main effect of trial type was not significant, $F(6, 294) = 0.07$, $p = .79$, $\eta^2 = .001$. Simple main effects analyses revealed that the interaction was due to the lack of image-type effect for the matched trials, $F(6, 300) = 2.02$, $p = .06$, $\eta^2 = .04$, but a strong effect of this for the unmatched trials, $F(6, 300) = 4.64$, $p < .001$, $\eta^2 = .09$. Post-hoc comparison of means with Bonferroni correction showed no difference among the results for the matched trials. For the unmatched trials, response times for both SSR and HE were faster than for LN. Other pairwise comparisons did not reveal any significant difference. These RT results suggest that the improvements of accuracy in the image-type conditions were unlikely because of a speed–accuracy trade-off.

The criterion results, along with proportion hits and false alarms, and d', are shown in Table 2. ANOVA of the criterion data showed a significant main effect of image type, $F(6, 306) = 11.40$, $p < .001$, $\eta^2 = .18$. The responses for the original, HE and SSR conditions were more conservative than for the LN, LTV, SQI, and DCT conditions. Other pairwise comparisons did not yield significant difference. The presence of stronger shadows in the original, HE and SSR conditions appeared to create a stronger conservative bias. That is, participants tended to judge a pair of faces as being different persons when the shading on the faces was more visibly different. This conservative response bias is consistent with the criterion effect in Experiment 1 and with the effects of illumination change reported in the literature (e.g. Liu et al., 2009).

The key finding in this experiment is that face-matching performance in the SSR and SQI conditions was more accurate than in the original image condition and other processed image conditions. Therefore, the results indicate that some image-processing techniques can facilitate face matching.

GENERAL DISCUSSION

The effectiveness of using image-preprocessing techniques to improve human face recognition was examined in two experiments. Experiment 1 found little advantage of using image-preprocessing techniques for recognition memory. However, in Experiment 2, face-matching performance benefited when certain preprocessing techniques (i.e. SQI or SSR) were applied to the original face images. Overall, the results show that some kinds of preprocessing techniques can be useful for improving face recognition in humans, although the benefit is likely to be limited to a matching task.

It is not clear why SSR showed a positive effect for face-matching performance but not for recognition memory. In the matching task, the positive effect was likely due to improved visibility of the key facial features after the strong shadows were weakened by the preprocessing method. The resulting additional details may have allowed for an increased number of reference points to be utilized during face matching. However, it may be that the same details are not useful for long-term face memory, as some of this information is unlikely to survive the process of memory decay. Because the matching task does not suffer from the same level of decay, more useful information would be available in conditions where more facial features are visible. Strong shadows in the original images can result in fewer useful features for recognition. However, the reduced information may be relatively more robust against continuous memory decay. Greater details of the facial features, on the other hand, may decay at a faster rate following the time of encoding. Whether this account can explain the lack of effect of the same image-preprocessing technique on long-term face memory will require careful future investigations.

Our experiments also show that unlike automated face recognition systems, algorithms that remove shadings from faces do not necessarily improve human recognition performance. In fact, some of the techniques could be even detrimental for humans. For example, results in Experiment 1 demonstrate that recognition was significantly impaired after the original images were processed by the DCT method compared with those processed by the SSR method. A clear difference between these methods is that DCT nearly abolished the contrast created by shading in the original images, whereas SSR preserved the contrast created by shading while improving the visibility of facial features. Results in Experiment 2

Table 2. Means and standard deviations (in parenthesis) for proportions of hits and false alarms, d', and criterion (c) as a function of image type in Experiment 2

	Image type						
	Original	SSR	SQI	DCT	LTV	HE	LN
Hits	0.73 (0.20)	0.80 (0.18)	0.84 (0.16)	0.79 (0.19)	0.79 (0.16)	0.69 (0.20)	0.83 (0.14)
FAs	0.15 (0.15)	0.09 (0.10)	0.13 (0.14)	0.22 (0.18)	0.22 (0.16)	0.14 (0.14)	0.25 (0.16)
d'	2.08 (1.09)	2.66 (0.98)	2.65 (1.18)	2.06 (1.23)	1.97 (1.03)	2.01 (0.89)	1.99 (0.90)
c	0.25 (0.55)	0.26 (0.56)	0.06 (0.54)	−0.03 (0.57)	−0.03 (0.52)	0.39 (0.60)	−0.19 (0.55)

Note: SSR, single-scale retinex; LTV, logarithmic total variation; SQI, self-quotient image; DCT, discrete cosine transform; FAs, false alarms; HE, histogram equalization ; LN, local normalization.

further demonstrate that algorithms drastically removing shading can have detrimental effects on human face recognition. Here, images processed by the SSR and SQI methods were matched significantly better than images processed by all the other methods. Similar to SSR, SQI preserves shading information in the original images, whereas all other methods in our test except for HE focus more on removing shading. As shown in Figure 1, images using the HE method do show preserved shading. However, unlike SSR and SQI but similar to LN, the HE method exaggerates spots or local landmarks on the surface of skin. Although these landmarks may provide useful information about invariant facial features for automatic face recognition systems, they may be unhelpful or even harmful for human recognition. Perhaps this explains why the HE condition did not yield a similar advantage as the SSR and SQI conditions. On the whole, the current results are consistent with the previous finding that human face recognition relies on shading and low spatial frequency information (e.g. Bruce et al., 1992; Davies et al., 1978; Liu et al., 2000). Although removing shading makes two images of the same face look more similar, it also results in a loss of this property required by the human vision.

However, our results also demonstrate that excessive shadows under extreme lighting conditions can also be detrimental to face recognition in humans. Experiment 2 shows that whereas preserved shading in the SSR and SQI conditions facilitated face-matching performance, the strong shadows in the original images were as damaging as the preprocessed conditions where the shadows were nearly absent. This suggests that shading information is only useful when the key facial features are not obscured by strong shadows. This means that the facilitating effect of the SSR and SQI methods is most likely to be restricted to images taken under extreme lighting conditions.

Advanced image-processing techniques have proven to be useful for human face recognition. In addition to the late applications of face composite systems such as EvoFIT (Frowd et al., 2011, 2012), research has shown that computer-synthesized images can reduce the limitation of face training based on a single photograph (Liu, Chai, Shan, Honma, & Osada, 2009). The present study similarly demonstrates that technology for face recognition can play a role in assisting human face identification even though there are fundamental differences between face recognition by certain automatic systems and by human observers. Our findings suggest that image-processing techniques may be useful in manhunts, television appeals and surveillance inspections, when facial features are obscured by dark shadows in closed-circuit television images.

ACKNOWLEDGEMENTS

This study was supported by grants from the British Academy, the National Natural Science Foundation of China (no. 60772071) and the Hi-Tech R&D Program of China (no.2007AA01Z163). We thank Robert Holliday, Tim Hoare and Emma Medford for data collection.

REFERENCES

Adini, Y., Moses, Y., & Ullman, S. (1997). Face recognition: The problem of compensating for changes in illumination direction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *19*, 721–732.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.

Braje, W. L. (2003). Illumination effects in face recognition: Effect of position shift. *Journal of Vision*, *3*, 161–170.

Braje, W. L., Kersten, D., Tarr, M. J., & Troje, N. F. (1998). Illumination effects in face recognition. *Psychobiology*, *26*, 371–380.

Bruce, V., Hanna, E., Dench, N., Healey, P., & Burton, M. (1992). The importance of "mass" in line drawings of faces. *Applied Cognitive Psychology*, *6*, 619–628.

Chen, W., Er, M. J., & Wu, S. (2006) Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain. *IEEE Transactions on Systems, Man, and Cybernetics. Part B, Cybernetics*, *36*, 458–466.

Chen, T., Yin, W., Zhou, X. S., Comaniciu, D., & Huang, T. S. (2006). Total variation models for variable lighting face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *28*, 1519–1524.

Davies, G., Ellis, H. D., & Shepherd, J. (1978). Face recognition accuracy as a function of mode of representation. *Journal of Applied Psychology*, *63*, 180–187.

Frowd, C., Bruce, V., Pitchford, M., Jackson, S., Hepton, G., Greenall, M., … Hancock, P. (2011). The psychology of face construction: Giving evolution a helping hand. *Applied Cognitive Psychology*, *25*, 195–203.

Frowd, C., Skelton, F. C., Atherton, C. J., Pitchford, M., Hepton, G., Holden, L., … Hancock, P. (2012). Recovering faces from memory: The distracting influence of external facial features. *Journal of Experimental Psychology: Applied*, *18*, 224–238.

Georghiadis, A. S., Belhumeur, P. N., & Kriegman, D. J. (2001). From few to many: Illumination cone models for face recognition under variable lighting and pose. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *23*, 643–660.

Gonzalez, R., & Woods, R. (1992). *Digital image processing* (pp. 91–94). NJ, USA: Prentice Hall.

Hill, H., & Bruce, V. (1996). Effects of lighting on matching facial surfaces. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 986–1004.

Jobson, D. J., Rahman, Z., & Woodell, G. A. (1997). Properties and performance of a center/surround retinex. *IEEE Transactions on Image Processing*, *6*, 451–462.

Liu, C. H., Bhuiyan, Md. A.-A., Ward, J., & Sui, J. (2009). Transfer between pose and illumination training in face recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *39*, 939–947.

Liu, C. H., Chai, X., Shan, S., Honma, M., & Osada, Y. (2009). Synthesized views can improve face recognition. *Applied Cognitive Psychology*, *23*, 987–998.

Liu, C. H., Collin, C. A., Burton, A. M., & Chaudhuri, A. (1999). Lighting direction affects recognition of untextured faces in photographic positive and negative. *Vision Research*, *39*, 4003–4009.

Liu, C. H., Collin, C. A., Rainville, S. J. M., & Chaudhuri, A. (2000). The effects of spatial frequency overlap on face recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 956–979.

Longmore, C. A., Liu, C. H., & Young, A. W. (2008). Learning faces from photographs. *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 77–100.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442.

Tarr, M. J., Georghiades, A. S., & Jackson, C. D. (2008). Identifying faces across variations in lighting: Psychophysics and computation. *ACM Transactions on Applied Perception*, *5*(2:10), 1–25.

Tarr, M. J., Kersten, D., & Bülthoff, H. H. (1998). Why the visual system might encode the effects of illumination. *Vision Research*, *38*, 2259–2275.

Wang, H., Li, S., & Wang, Y. (2004). Face recognition under varying lighting conditions using self quotient image. *Proceedings of 6th IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 819–824.

Xie, X., & Lam, K. (2006). An efficient illumination normalization method for face recognition. *Pattern Recognition Letters*, *27*, 609–617.

Zou, X., Kittler, J., & Messer, K. (2007). Illumination invariant face recognition: A survey. *First IEEE International Conference on Biometrics Theory Applications and Systems*, Crystal City, VA.

