Comment on "100% Accuracy in Automatic Face Recognition"

Weihong Deng,* Jun Guo, Jiani Hu, Honggang Zhang

Jenkins and Burton (Brevia, 25 January 2008, p. 435) reported that image averaging increased the accuracy of the automatic face recognition to 100% and thus could be applied to photo-identification documents. We argue that the feasibility of image averaging on identification documents is not fully supported by the presented evidence.

In automatic face recognition, a gallery of facial images is first enrolled and coded for subsequent searching. A probe image is then obtained and compared with each encoded face in the gallery, and a recognition is noted when a suitable match occurs. In a recent study, Jenkins and Burton (*1*) used the photographs of celebrities as probe images to measure the hit rate of the FaceVACS (Cognitec Systems GmbH, Dresden, Germany) face-recognition system used by the genealogy Web site MyHeritage (*2*). Merging the probe images to create an average image for each celebrity raised the overall hit rate for the probe database from 54 to 100%. The authors therefore concluded that the process of image averaging could dramatically boost automatic face recognition and inferred that incorporating average face images into identification documents would greatly reduce the incidence of face-recognition errors.

As Jenkins and Burton suggest in (*1*), it is possible that 100% accuracy was achieved simply because the image averages incorporated some recognizable photos. To allay that concern, the authors reported that a new set of averages using only those photographs that were unrecognized in their first study raised the hit rate from 0 to 80%. They thus reasoned that the improved accuracy could solely be attributed to the averaging process. However, the improvement on the hit rate could be partly attributed to the manual facial registration [see supporting online material for (*1*)] before averaging, which accurately rectified the facial appearance so that all the probe faces were aligned in a standard frontal and upright posture and enclosed by a uniform background. By largely reducing the image variability, the image registration procedure might transform some unrecognized photos into recognizable faces (*3*). Moreover, the standard registered faces might facilitate the automatic face finding and normalization process of the tested algorithm, which may have also boosted the hit rate (*4*). It is thus possible that the registration technique assisted the image averaging to boost the hit rate to a higher level.

Jenkins and Burton correctly suggested that image averaging enhanced the performance by stabilizing the face image. However, the interpretation of this fact was overextended. The conclusion that including average images on identification documents would reduce recognition errors lacks sufficient evidence, especially because it is not an equivalent task to the experiments that were carried out. Specifically, the experiments in (*1*) used the online database as the gallery and the average images as the probe, and the online recognition system only returned the closest matching photo from its database. If the identity of the returned photo matched that of the average image, it was recorded as a hit. Using this methodology, even the 100% hit rate could only ensure that, for each test identity, the system successfully matched the average with "one" gallery photo from that person. However, there were multiple (from 7 to 28) gallery photos for each test identity in the database (*1*). The experiments did not show the number of single (gallery) photos to which the averages could be matched. In contrast, if the average image is incorporated in identification documents, the identity-verification system must be required to suitably match it to every photo from the same person; otherwise any miss on a photo would translate to a recognition error. Therefore, although the recognition algorithm is commutable, the task of identity verification is more demanding than that of Jenkins and Burton's experiments, and the feasibility of using average images for verifying identity requires further testing. Proof of identity is achieved by comparing an individual's appearance to a photo-identification document, where the appearance is captured by any single facial image in diverse locales and different times. The reliability of the proof depends on how stably the single images can be matched to the corresponding photo-identification document. Hence, in order to evaluate the feasibility of average images on identification documents, a refined experiment should be designed to measure the hit rate for single (gallery) photos, showing what proportion of the single images can be matched to the corresponding average. Moreover, the reliability of the proof also depends on the ability to reject the photos of the impostors according to the averages, which also need to be considered. For the scientific methodology, one can refer to United States government–sponsored evaluations, such as the Face Recognition Vendor Test (*5*), which are the standard test beds for face-recognition technologies.

We acknowledge that image averaging contributes to the face-recognition procedures. However, automatic face recognition is a complex pattern-recognition problem involved with early processing, perceptual coding, and cue-fusion mechanisms (*6*). Although countless solid contributions have been made (*7*), 100% accuracy in automatic face recognition in real-world settings remains an ambitious goal.

School of Information Engineering, Beijing University of Posts and Telecommunications, Beijing, 100876, China.

*To whom correspondence should be addressed. E-mail: whdeng@bupt.edu.cn

References and Notes

1. R. Jenkins, A. M. Burton, *Science* **319**, 435 (2008).
2. MyHeritage, www.myheritage.com/face-recognition.
3. I. Craw, N. Costen, T. Kato, S. Akamatsu, *IEEE Trans. Pattern Anal.* **21**, 725 (1999).
4. S. Shan, Y. Chang, W. Gao, B. Cao, *Proc. 6th IEEE Int. Conf. Automatic Face and Gesture Recognition*, 314 (Seoul, Korea, 17 to 19 May 2004).
5. Face Recognition Vendor Test, www.frvt.org.
6. P. Sinha, *Nat. Neurosci.* **5**, 1093 (2002).
7. W. Zhao, R. Chellappa, P. J. Phillips, A. Rosenfield, *Assoc. Comput. Mach. Comput. Surv.* **35**, 399 (2003).
8. The authors are funded by China Scholarship Council, Natural Science Foundation of China (60675001), and National High-Tech Development Plan of China (2007AA01Z417).

10 March 2008; accepted 15 July 2008
10.1126/science.1157523

Comment on "100% Accuracy in Automatic Face Recognition"

Weihong Deng, Jun Guo, Jiani Hu and Honggang Zhang

Science **321** (5891), 912.
DOI: 10.1126/science.1157523