**Senate Judiciary Committee**
**Subcommittee on Privacy, Technology and the Law**
**Hearing on "What Facial Recognition Technology Means for Privacy and Civil Liberties"**
**July 18, 2012**

**Questions for the Record from U.S. Senator Al Franken**
**For Dr. Brian Martin**

1. **What are the error rates for verification uses of facial recognition technology, and what are they for identification uses?**

This is a good question, the answer to which is not black and white. There are two error rates that one can make a tradeoff between by varying the match threshold of the system. By turning up the match threshold, you can reduce the **false positive** match rate (the chance of incorrectly saying two faces are the same person when they are not) and in exchange this would increase the **false negative** match rate (the probability that the algorithm fails to match two faces that were from the same person). Therefore, when we talk about error rates, we talk about both error rates simultaneously.

In quoting accuracy or the error rates of a technology, the value will depend strongly on the quality of the images and the strength of the algorithm. You can probably find experiments to support almost any result. Therefore, it is best to look at independently conducted, controlled tests. The National Institute of Standards and Technology (NIST) is arguably the best organization for ensuring independent non-biased faced recognition accuracy tests on sequestered data. In the 2010 NIST Interagency Report 7709, the stated **error rate for Verification is 4% false negative rate when you are operating at a threshold for 0.1% false positives** (Figure 12 in the report). This is based on FBI data; the accuracy can be better or worse depending on the quality of the face images. In real life, this would mean if you used this particular algorithm for face matching to unlock your phone, it would not let you unlock your phone about once in every 25 attempts (or would not work very well for 1 out of 25 people) since the false negative rate is 4%. At the same time it would let one of 1000 other people's faces unlock your phone since the false positive rate is 0.1%.

Identification error rates are more complex to nail down to a single set of values since the results depend on the number of faces in the database or gallery of faces. If the database has 1 person in it, then the error rate is the same as seen with Verification. If the database has 1,000,000 faces, there are 999,999 more chances to falsely match a face and get a false positive. A false positive in this sense means that out of 1,000,000 faces in the gallery, there was at least one false positive. To help separate the fact that this is a different measure, in identification we call it the **false positive identification rate** and the **false negative identification rate**. The false negative rate, however, does not necessarily increase with the larger database size since this rate depends on just a single

image (the person it is supposed to match). In order to have a false positive identification rate of 0.1%, we would need to set our matching threshold to be much higher to discriminate against the 1,000,000 chances of a false positive. In fact, the match threshold needed to ensure a 0.1% false positive rate for identification of 1,000,000 faces will correspond to the Verification threshold of 0.0000001%. Setting the threshold this high, has the tradeoff that the false negative identification rate will increase as compared to the Verification scenario. From the 2010 NIST report, the false negative identification rate for the FBI data is roughly 25% at a corresponding 0.1% false positive identification rate for a database of 10,000 faces. If the database grows to 1,600,000 faces, the false negative identification rate is about 43% at 0.1% false positive identification rate. This is 10 times more false negatives (missed matches) than seen with Verification. Even with good quality data, the expected false negative rate would likely be on the order of 10% for a database of 1,000,000 good quality faces and still an order of magnitude or more worse than with Verification. To compensate for the higher error rates in Identification, a human is put in the loop to validate Identification results. With a human reviewing the top 10 best matches from every identification attempt, and ignoring the match threshold, face recognition can produce a 'hit' over 95% of the time (or inversely a 5% false missed hit rate) on a database of over 1,000,000 records. This investigative measure of accuracy (hit rate) shows that even though the error rate may be relatively high compared to Verification, face recognition is still a useful tool when combined with a human operator.

2. **What happens to error rates—false negative and false positive rates—when you're working with photos of people who don't know they're being photographed?**

For face recognition, as the quality of the face image deviates from the ideal image of a frontal looking face in good focus with good lighting, the accuracy will deteriorate. The extent of the drop in accuracy will depend on the capabilities of the face recognition algorithm (as it can be robust or trained to handle some of the issues). As an example, NIST presented at Biometrics 2010 in London a presentation "The limits of face recognition in law enforcement" where they showed Identification from a database of controlled quality images gave a 'hit rate' of 93% for a vendor, but when the images were less controlled and from a web cam the 'hit rate' dropped to 54%. They did not publish the false positive and false negative identification rates, but at a false positive identification rate of 0.1% I would estimate the false negative identification rate to be on the order of 60%. If the images are more uncontrolled, such as when taken far away when people do not know they are being photographed the error rates could be anywhere from 10% to 100% false negative identification rate – all depending on if they are caught looking at the camera, with good lighting and good focus.