

Opportunity Title: A Fusion Approach To Multimedia Biometric Verification

Opportunity Reference Code: IC-16-08

Organization Office of the Director of National Intelligence (ODNI)

Reference Code IC-16-08

How to Apply **Create and release your Profile on Zintellect** – Postdoctoral applicants must create an account and complete a profile in the on-line application system. **Please note: your resume/CV may not exceed 2 pages.**

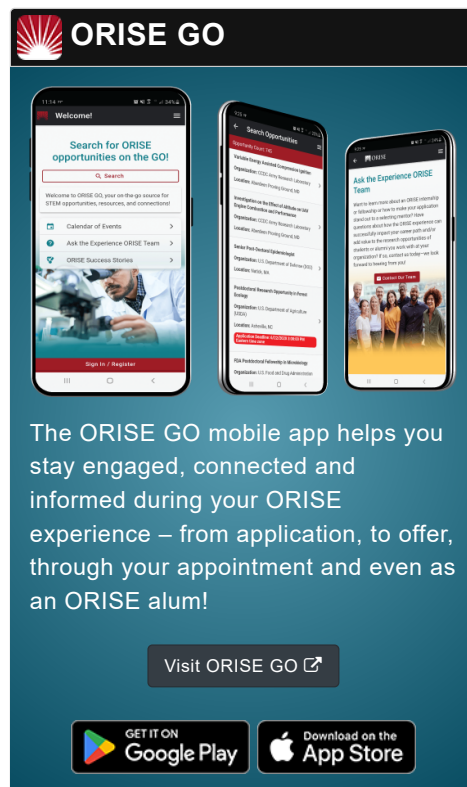
Complete your application – Enter the rest of the information required for the IC Postdoc Program Research Opportunity. The application itself contains detailed instructions for each one of these components: availability, citizenship, transcripts, dissertation abstract, publication and presentation plan, and information about your Research Advisor co-applicant.

Application Deadline 4/15/2016 6:00:00 PM Eastern Time Zone

Description Even those who are not familiar with law enforcement or intelligence practices should, by now, have a basic understanding of biometrics from the popularity of television crime dramas, such as CSI. As the law enforcement and intelligence communities (as well as private industry) move towards using more biometrics to define a user or suspect, the challenge remains that no single biometric method is likely to ever be completely accurate. Not only does this reduce confidence in the accuracy of any given biometric result, but it also limits the ability of the user to adequately triage large volumes of data based on a single biometric modality.

This is where multimodal biometrics comes in. Multimodal biometrics involves the use of multiple characteristics, used in conjunction, to identify an individual. Different biometric technologies utilize a variety of characteristics to identify subjects, including facial appearance (i.e., facial recognition), fingerprints, iris, voice, handwriting, gait, scent, and ear features. All of these, on their own, are subject to limitations. What is of most significance to the intelligence community, however, is that the accuracy of any single given biometric technology may not be adequate to identify a target with high confidence when the input signal is of low quality (e.g., CCTV images or voicemail recording from a call made on a cell phone). What the intelligence community needs is a better way to take advantage of lower quality biometric signatures.

Digital multimedia offers an opportunity to do that. For the law enforcement and intelligence communities, digital multimedia is becoming more common as material to be exploited and it often includes both visual and audio information. Within the law enforcement and intelligence communities there is an increasingly large amount of multimedia data which is often



times examined forensically to identify individuals depicted in this data. This data may be recovered from suspects' computers or from social media sites. Of particular interest, video recordings which depict people speaking (whether directly to the camera or not) are becoming more and more the subject of these forensic examinations. This offers the potential for a multimodal biometric approach utilizing facial recognition and speaker recognition.

Although current algorithms for both face recognition and speaker recognition are relatively robust under highly controlled circumstances, their performance degrades with lower quality data – including the type of data under consideration here. While improvement of algorithms in each modality is likely to be a continuing area of research, a new line of investigation is proposed to actively consider how one can fuse existing face and speaker recognition solutions to generate higher matching confidence.

Therefore, the community at large is seeking a solution to strengthen the ability to create a match using a fusion of face recognition and speaker recognition technologies. Such solutions are not only needed to strengthen the results of one exam, but also to perform automated recognition across large volumes of data (i.e., "Big Data").

Example Approaches

Proposals should consider the fusion of a variety of biometric modalities. Research could consider, but not be limited to, the examples below:

Automated facial recognition and speaker recognition approaches are quite mature at this time. Current commercial implementations of automatic facial recognition are derived from two major approaches: geometric (feature based) and photometric (view based). The latter approach has become quite robust to slight changes in pose, angle, illumination and expression, making practical application of facial recognition a reality in many operational scenarios.

In general, speaker recognition system used in forensic or investigative application is composed of acquisition and preprocessing of voice samples, feature extractor that generates feature vectors, and the classifier (or matching) engine which generates matching scores. Trained human analysts interpret matching scores for match, no match, or inconclusive as their final conclusions. Today's main stream advanced speaker recognition system uses MFCCs_based i-Vector and PLDA classifier to generate matching scores in the form of Likelihood Ratio (LR) or Log Likelihood Ratio (LLR) scores, normalized or non-normalized, and calibrated or non-calibrated.

Opportunity Title: A Fusion Approach To Multimedia Biometric Verification

Opportunity Reference Code: IC-16-08

The fact that face and voice features can be safely assumed to be independent from each other, and the fact that scores from speaker recognition system are generative, and based on the probabilities, voice makes an ideal choice of a biometric modality to be fused with output of the facial recognition system. There are various stages in the recognition process where fusion can take place in voice and face recognition systems, the simplest and the most effective stage appears to be after the matching scores are generated. The ultimate multimodal biometric system we are seeking will be one which ingests scores (LLRs, etc.) from both face and voice recognition systems, fuse the two scores into a single score, normalize it, and calibrate it as a final a score.

**Eligibility
Requirements**

- **Citizenship:** U.S. Citizen Only
- **Degree:** Doctoral Degree.
- **Discipline(s):**
 - **Business** (11 )
 - **Chemistry and Materials Sciences** (12 )
 - **Communications and Graphics Design** (6 )
 - **Computer, Information, and Data Sciences** (16 )
 - **Earth and Geosciences** (21 )
 - **Engineering** (27 )
 - **Environmental and Marine Sciences** (14 )
 - **Life Health and Medical Sciences** (45 )
 - **Mathematics and Statistics** (10 )
 - **Other Non-Science & Engineering** (13 )
 - **Physics** (16 )
 - **Science & Engineering-related** (1 )
 - **Social and Behavioral Sciences** (28 )