Research Gaps:

Robust and Novel Acquisition
This topic involves developing innovative methods of acquiring biometric information and novel or emerging biometric modalities. Methods may include contactless sensing from a distance, location specific acquisition with handheld devices, hierarchical collection of data to improve system throughput, and completely automatic and unsupervised collection. Modalities include ear, EKG/EEG, lip movement + speaker verification, 3D face, and potentially others. New methods and modalities need to have at least a hypothetical improvement over existing approaches and a sound scientific basis for that belief. They also need to move, as rapidly as possible, towards substantiating that improvement.

Closing the Gap Between Biometrics and Forensics
Significant research continues to need to be accomplished to close the gap between biometrics and forensics. This includes studies into requirements for forensic certified/verified biometric algorithms for finger/face/iris; data collection/analysis/medical studies to substantiate uniqueness and/or permanence of identifiable biometric features; development of tools for forensic evaluation of faces and irises; and analysis of the use of these tools to learn forensic processes and features used by analysts to uniquely identify individuals.

Biometric/Cryptography or Security Projects
Biometrics need to be integrated with other cryptographic, credential and/or other security approaches and yet there isn't a strongly understood method for doing so and evaluating the resulting performance or increase in security. Many in the security field even state that biometrics aren't secret therefore they are not secure. Projects need to be conducted to evaluate various biometric attributes for their cryptographic "strength;" to fuse biometrics and other security techniques (e.g. PINS, passwords, credentials) and evaluate the resulting performance; and to evaluate software vs. hardware-based techniques to spoof or obfuscate biometric technologies and develop standards and test/evaluation protocols in this area.

Observing the Biometric Menagerie
Traditionally, subjects of a biometric system have been described according to Doddington's zoo or the biometric menagerie. The original scheme included goats, sheep, lamb, and wolves. Here, sheep are subjects who typically match well against themselves (intra-class) and poorly against others (inter-class), goats match poorly against themselves (high FRR), lambs match well against others (high FAR), and wolves prey upon lambs and are successful at impersonating others (high FAR). While the menagerie may provide an interesting analogy, its utility far exceeds explanatory purposes. Careful examination of match score distributions as well as the observance and analysis of problematic subjects shed light onto various aspects of the biometric system in question. It can hint at problems in the biometric capture systems, identify algorithmic flaws, highlight vulnerabilities, and increase the accuracy of system performance assessments. Recently, new additions have been made to the biometric menagerie including worms, chameleons, phantoms, and doves. Conducting such studies may be very important for large-scale multi-biometric systems.

Biometric Capacity Analysis
Despite continued parameter refinement of existing algorithms and development of new algorithms, there is an upper bound on the performance that can be obtained when relying on a single

biometric. This bound is mainly a function of the template used to represent biometric data. The amount of information contained in the template representing the biometric is referred to as the template capacity. Based on the template capacity, one will reach a theoretical system capacity which will define the highest amount of individuals the system could feasibly distinguish between. This theoretical bound will likely be higher than the empirically observed one, which is based both on the template representation and the biometric variations observed during inter-class (templates from different subjects) and intra-class (templates from the same subject) comparisons. Since multi-biometric systems often include information from different and often uncorrelated sources, it is possible to raise both theoretical and empirically observed upper bounds on matching accuracy.

Model Estimation / Update Schemes

Often the data observed in an operational system will vary over time. In some cases, the variations will be significant and may cause the system to deviate greatly from normal performance if left unchecked. In scenarios such as these, it may be necessary to update parameter selection, density estimation, etc. to prevent performance degradation. Often a system may reach a "tipping point" where the sheer number of enrollments or observed probe samples affects matching performance due to the observation of extremely small inter-class differences in biometric traits. Biometric aging, injury, or changes in capture environment can all lead to intra-class variations which could result in an increase the false reject rate (FRR). It is important to consider these possibilities and develop model update schemes which can trigger the re-evaluation of system thresholds, reestimation of relevant densities, or selection of different algorithms for classification. Ideally this process can be automated, but manual model update schemes are also possible. It is also important to decide when and how to do so. Some systems require computationally intensive density estimation procedures which may or may not be performed online. Others have the ability to incrementally update which decreases the time required to re-establish key system parameter values.

Multi-Biometric Indexing Systems

An often integral part of large scale biometric systems is the notion of indexing. Biometric indexing systems can greatly increase the efficiency of identification operations by limiting the percentage of the database that needs to be searched. The classic example of biometric indexing involves the classification of fingerprint types (whorl, loop, arch, etc.). By classifying prints by type, one only needs to match against prints of similar types. There are indexing systems developed for most biometrics with varying degrees of success. One fairly new topic investigates the ability to index based on multiple biometric sources. In any indexing scheme the goal is to minimize the penetration rate (what percentage of the database needs to be searched) without adding additional errors (errors due to incorrect indexing).

Addressing Multi-Biometric Vulnerabilities

Whether considering the notion of biometric spoofing, liveness detection, biometric reconstruction from templates, or generic computer security threats that can be applied to biometric systems, there are many areas that fall within the scope of addressing biometric vulnerabilities. As one of the major drivers of multi-biometric fusion is to prevent circumvention, one must carefully consider how or if multi-biometric systems truly increase the difficulty to circumvent identification, verification, and watchlist processes. Liveness detection methods for most hard biometrics exist but fusion of such liveness detection has not been thoroughly studied. There are a large number of topics that require work in this area.

Multibiometric Fusion Approaches

We believe that there is significant work yet to be conducted in multi-biometric fusion and its applications. Especially if we need to do one-to-many multi-biometric identifications for non-contact applications. Continued research needs to be conducted in various fusion approaches for Multibiometrics. This may not even be fully inclusive, nor is the level of detail here intended to imply that it is a top priority to DHS.

Dynamic Decisional Fusion: likely to incorporate ancillary, meta, and biometric data as seen in the other topics. However, the focus here is to adapt the nature of biometric process flow, matching thresholds, etc. based on any number of parameters. These parameters can include passenger throughput, incoming flight risk, biometric quality, and various environmental factors. By dynamically adjusting biometric system processing based on these parameters it may be possible to significantly decrease error rates, adaptively handle the flow of individuals, and account for variations in the collection of biometric samples based on known environmental profiles.

Hierarchical Fusion: entails the fusion of multiple sources of biometric information as necessary or as they become available. For instance in a hierarchical multi-modal system one may initially perform biometric recognition on fingerprint data. If the returned matching result is unclear (evidence does not strongly support either class), a hierarchical multi-modal system may subsequently apply facial recognition to increase the certainty of the recognition result. In an alternative notion of hierarchical fusion, a system may fuse multiple sources of biometric information as they become available. For instance in an unconstrained biometric environment, an acquisition system might capture facial biometric information at 30 feet and then as the subject approaches within 10 feet, the iris modality may become available.

Quality Enhanced Fusion Schemes: both uni-variate and bi-variate measurements of biometric quality are of interest. Primary investigations of biometric quality have focused at the image level of the hard biometrics including face, fingerprint, and iris. A number of efforts have provided strong evidence that fusing information provided from biometric quality metrics can be used to significantly improve the accuracy and certainty of biometric decisions. Work in this topic area would include further development of such fusion schemes. Additionally, further development of individual or collective biometric quality metrics is included within the scope of this topic.

Fusion Incorporating Meta / Ancillary Data: Instead of data associated with the biometric itself, meta-data can be described as information associated with "identity" of a subject. Information of this nature may include biographical data gathered to acquire various documents such as passports, visas, driver's licenses, passenger manifests, temporal-based geographical locations, etc. The fusion of such meta-data with biometric data can lead to increased certainty of identification decisions. Alternatively, ancillary data such as soft biometric features about the subject including height, weight, sex, etc. may be used in the fusion process to produce similar results.

Hybrid Fusion: involves any combination of two or more types of fusion (multimodal, multi-sensor, multi-algorithm, multi-instance, multi-sample). While any combination of types is feasible, most combinations include multiple modalities in addition with another type of fusion, i.e. multi-sensor or multi-algorithm. Another example of a hybrid approach would be one that performs fusion of hard (face, fingerprint, iris, etc.) and soft (height, weight, race,

etc.) biometric information. This approach could incorporate both multiple "modalities," multiple sensors, and potentially multiple algorithms to arrive at final identification decisions. Systems such as those described likely represent the highest level of complexity considered but also afford the greatest opportunity to discriminate between individuals.

Sensor-Level Fusion: is the first opportunity where multi-biometric fusion can take place. In this case, raw data from multiple sources are fused. This fusion involves combining multiple overlapping signals (usually images) through the process of mosaicing / registration. For instance, Kong et. al registered information from the visual spectrum with thermal signatures to arrive at a fused image which was then passed through feature extraction and matching blocks. In a different type of sensor level fusion, a number of "passive sensing" approaches have been proposed which combine multiple 2-dimensional face images to develop a 3-dimensional face model for matching. Sensor level fusion offers the highest level of information availability as it deals with the original sensed signals. While sensor level fusion may have the highest amount of information available, it is also represents the most challenging level in which fusion can take place. Many reasons account for this fact including: the noise associated with the constituent signals being fused and incompatibilities in the sensed signals.

Rank-Level Fusion: can only be applied to systems operating in identification mode. In these systems, ranking reflects the order of identities sorted in ascending order by the comparisons of probe and gallery samples most likely to match. Since ranking corresponds only to potential identities enrolled in a system, there is no reason to normalize or transform the constituent rank inputs from the multiple sources. This makes rank level fusion simple to implement. Furthermore, the availability of rank scores is usually high in commercial systems. While fusion at the rank level is relatively simple to implement, it is subject to information loss as rank is not as informative as match score.

Multi-Sensor Fusion: combination of multiple sensors to extract different biometric information from a single biometric characteristic/trait is another example of multi-biometric fusion. Multi-sensor fusion has similar thrusts as multi-modal fusion. It is similar in that it can allow for greater discriminatory power over single biometric systems as multiple sensors can provide orthogonal (and potentially complementary) information about the biometric trait. Complementary sources of information can be used to increase matching performance and thwart spoof attacks. The drawbacks are also similar to multi-modal fusion schemes as multiple sensors could result in increases in costs, both financial and temporal. Research opportunities in multi-sensor fusion exist in all hard biometrics as well as less common biometrics. Furthermore, multi-sensor fusion could feasibly integrate nicely into a hybrid fusion system.