QUEST Hierarchy for Hyperspectral Face Recognition

David M. Ryer, Trevor J. Bihl, Kenneth W. Bauer, Steven K. Rogers*
Air Force Institute of Technology, Dept. of Operational Sciences, 2950 Hobson Way, Wright Patterson Air Force Base, OH, USA 45433-7765
*Air Force Research Laboratory, Sensors and Information Directorates, 2241 Avionics Circle
Wright Patterson AFB, Oh, 45433

ABSTRACT

A face recognition methodology employing an efficient fusion hierarchy for hyperspectral imagery (HSI) is presented. A Matlab-based graphical user interface (GUI) is developed to aid processing, track performance and to display results. The incorporation of adaptive feedback loops enhance performance through the reduction of candidate subjects in the gallery as well as the injection of additional probe images during the matching process. Algorithmic results and performance improvements are presented as spatial, spectral, and temporal effects are utilized in this Qualia Exploitation of Sensor Technology (QUEST) motivated methodology.

Keywords: Hyperspectral, facial recognition, fusion, qualia

1. INTRODUCTION

Personal identification in social interaction depends heavily on the naturally developed face recognition ability that most people possess. Due to an increasing focus on security and identity verification during common activities (e.g. airports, banks, buildings, etc.) it is important to extend this ability to our security and surveillance systems. Face recognition is a crucial tool being used in current conflicts in Iraq and Afghanistan by the U.S. Army and Marines to identify and track the enemy [1]. The human recognition process utilizes not only spatial information but also spectral and temporal aspects as well. The incorporation and automatic handling of these types of features using hyperspectral imagery has not been fully investigated or subsequently extended to commercial applications.

The design of a biometric identification system should possess certain attributes to make it an effective operational system. These attributes include universality, distinctiveness, permanence, collectability, performance, acceptability, and circumvention [2]. Face recognition modality suffers from weaknesses in the areas of uniqueness, performance, and circumvention [3]. The ability to mitigate these weaknesses and ultimately match or exceed the recognition capability of a human, is the performance benchmark for computer based face recognition applications.

The Carnegie Mellon University (CMU) hyperspectral imagery (HSI) face database, graciously provided by Dr. Takeo Kanade, was used for this research [4]. Figure 1 depicts an example of this data over several sampled wavelengths. The utilization of HSI and the contextual information contained within these image cubes provide the tools to create a hierarchal methodology to mitigate the challenges face recognition systems must overcome.

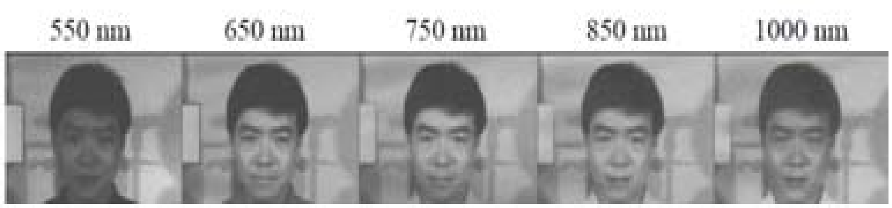

Figure 1. Spectral Layer Example from a CMU HSI Face [4]

In this paper, we describe the various algorithms used to exploit the inherent material reflectance properties in HSI to detect, segment, and identify subjects. A fusion hierarchy is applied to a suite of facial recognition algorithms to produce a cumulative performance improvement over traditional methods. A GUI tool is introduced which facilitates responsive operation as pictorial, numerical, and graphical results from the various algorithms are displayed. Experimental results are presented and recommendations for further research are suggested.

2. FACE RECOGNITION ARCHITECTURE

There are three main focus areas for this research, the application of facial recognition algorithms to HSI, the use of feature and decision fusion for improved results, and adaptive feedback to re-examine and confirm the most difficult matches. This discussion starts with a review of the dataset to understand the dimensionality of the data and exploitation potential.

2.1 Database Description

The CMU database contains visible and near infrared (NIR) images from 450nm to 1100nm containing 65 spectral bands and a spatial resolution of 640x480 pixels [4]. The portion of the CMU database that was available contained images for 54 different subjects with 36 subjects sitting for two sessions on different days. This database subset comprises our gallery, and probe sets (subjects to identify and a gallery to search). Additionally, a subset of subjects from the gallery and probe sets were available for multiple sessions; 3 sessions (28 subjects), 4 sessions (22 subjects), or 5 sessions (16 subjects). These additional images are used to analyze the ability to inject additional images for confirmation of a subject match.

By taking advantage of fundamental properties of HSI (different materials reflect different wavelengths of light differently), skin, hair, and background materials are relatively easy to detect. The advantages of using higher dimensional data compared to grayscale or 3-band 'true' color image includes the ability to detect skin segments since the spectral reflectance properties are well-understood [5]. The segmented portions of the image can be used to provide context that aids traditional face recognition algorithms.

Leveraging the signatures available through HSI, features such as skin and hair can be detected using a straightforward method similar to the Normalized Difference Vegetation Index (NDVI) used in remote sensing to detect live vegetation [5]. A Normalized Differential Skin Index (NDSI) can be computed easily through the sum and difference of key spectral bands [5]. Applying this technique and a variety of edge detection methods, several contextual layers of an individual's face can be extracted automatically from an HSI as seen in figure 2 [6]. For individuals attempting to conceal or alter their appearance, it is now possible to detect inconsistencies such as make-up and prosthetic devices due to the differing reflectance properties [7].

Figure 2. Contextual Layers of a Hyperspectral Image: Skin, Hair, Edges, and Combined Representation

2.2 Previous Hyperspectral face Recognition Research

Robila [8] encompassed both the visible and NIR wavelengths, as he explored the utility of using spectral angles for comparison. Other research involving NIR and visible wavelength faces include Klare [9], who examined matching NIR faces to visible light faces. Bourlai et al. [10] presented an initial study of combining NIR and shortwave IR (SWIR) faces with visible for more complete face representation, in addition to comparing cross-spectral matching (visible to SWIR). Kong et al. [11] delivered an overview of advantages and disadvantages of facial recognition methods with respect to the image wavelengths. Chou [12] used hyperspectral images and experimented with segmenting different tissue types in the human hand. Elbakary [13] used the K-means clustering algorithm to segment the skin surface in hyperspectral images and then measured the Mahalanobis distance between signatures to match subjects. Pan has accomplished the most extensive research utilizing spectral signatures of skin to identify individuals [14], [15], [16] and in a subsequent effort [17] explored the benefit of incorporating spatial measurements at various wavelengths. These efforts produced valuable insight used individually, these techniques did not provide the desired performance for this challenging data set. Denes [4] noted that the prototype camera used for the CMU data was subject to stray light leaks and optical imperfections as he noted that, "better face recognition clearly requires higher definition through a more sensitive, low noise camera or through higher levels of illumination." Viewed from another perspective, this noisy data provided an ideal environment for the development of a fusion strategy. The findings from these previous efforts provide a foundation to construct an intelligent hierarchy to address challenges for face recognition systems.

2.3 Recognition Algorithms

The face recognition suite employed a variety of techniques and methods. Through the segmentation process, two separate algorithms were used to locate face and hair surfaces through NDSI [5]. The skin and hair segments were subsequently fed to supporting algorithms of increasing detail for matching. The first and most straightforward method was to use face measurements such as size, length, width, and eccentricity to create a string of parameters for comparison. Many of the typical manual preprocessing techniques employed in popular face recognition systems [18], such as selecting eyes coordinates, geometric normalization, and masking became unnecessary, as face images are automatically processed and fed to proven algorithms such as eigenface [19] that match holistic features of the grayscale representation. These images either can be the spatial face segments, hair segments, or combined representations. The next step uses spectral signatures of the hair and face segments and compares them using the fast and effective spectral angle comparisons [8], [20]. Finally, the face, hair, and combined representation are feed to the scale and orientation robust Scale Invariant Feature Transform (SIFT) method to compare matching interest points or SIFT keys [21], [22]. To establish baseline performance, these methods are initially used in isolation, and then used in combination to evaluate a range of fusion strategies.

2.4 Fusion Hierarchy

From the field of automatic target recognition, Ando [23] provides a useful hierarchy for processing the hyperspectral face images. At the lowest level, processing includes smoothing and segmenting the image. During mid-level processing, cues such as shading, texture, reflectance, and illumination are integrated. Lastly, high level processing integrates information that is invariant across different viewpoints for final identification. Using this guide, the initial face recognition hierarchy could be achieved though incremental segmentation, processing, and identification steps. These steps would utilize not only information from the spatial dimension of the image, but would use spectral elements to help assist in the tasks of segmenting, processing and identification. Figure 3 illustrates the combined and incremental approach.

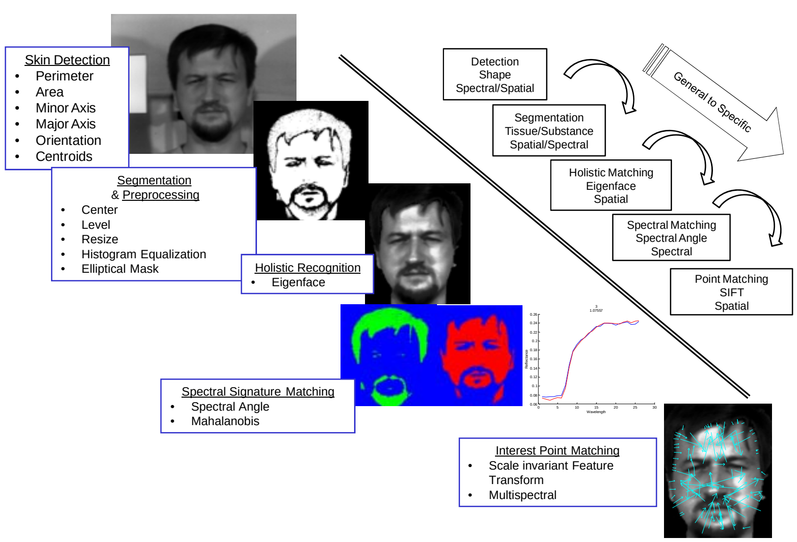

Figure 3. Hyperspectral Face Recognition Fusion Hierarchy

2.5 Qualia Exploitation of Sensor Technology (QUEST) motivated methodology.

Ultimately, the performance and computational demands of working with high dimensional data required a strategy that utilized only the relevant information in a more effective method. Turning to the Qualia Exploitation of Sensor Technology (QUEST) methodology, we attempt to develop a general-purpose computational intelligence system that captures the advantages of qualia-like representations [24]. Qualia can be defined as a representation of the physical environment or a facet included in ones intrinsically available internal representation of the world around them [25]. It is our goal to combine different qualia into a meta-representation, so sensory inputs can be integrated into a model that is adaptable and efficiently functional and can be deliberated repeatedly. A guiding principle of QUEST highlights the use of qualia that map sensory input to more useful and efficient states that complement the reflexive intuition level of processing. The functional requirement for a QUEST system is to possess the ability to detect, distinguish, and characterize entities in the environment [26].

In order to build a QUEST system for our task, it is important to develop and understand the concept of an agent [25]. An agent takes a subset of stimuli from the environment and processes this into relevant information. Information is defined as the reduction of uncertainty in that agent's internal representation. An agent has knowledge of other agents and of their enviromental representation, akin to a Theory of Mind with insight into their needs. The agent transmits selected aspects of its information representation to these neighboring agents. Agents transmit stimuli upward in higher levels of abstraction and can also transmit information downward providing details and context that can influence lower level agents (Figure 4). An entity uses various sets of these agents and their collective knowledge to create an internal representation of its environment.

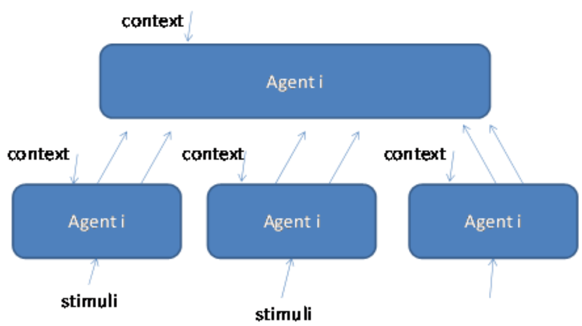

Figure 4. Agents, Information, and Levels of Abstraction

The relevant information or context is compromised of biometric characteristics and cues across the electromagnetic spectrum. Rogers [26] states that the concept of context is the parts of one agent's subjective representation of an entity as it exists in the world. An agent communicates this context to other agents that use this information to improve its internal representation. Context can only be transmitted between agents that are aligned as each agent contains a representation of the other's environment. The combination of fiducial features and higher-level abstracted characteristics creates this context. In the human recognition system, the mind stores data not so much as sensory numbers but as relative comparisons to prior experiences, that can change over time. For a face recognition system, the relative comparisons should serve an equally important role in refining the solution space and guiding the search process. The connections or links in our fusion hierarchy provides the context of the face. There are many links that can connect the internal and external facial features that have proved so important in human recognition research [27]. The links chosen can help incorporate higher levels of abstraction such as important soft biometric [28] cues or can be the connection between spatial and spectral information. Figure 5 illustrates the links employed in our HSI face recogntion system.

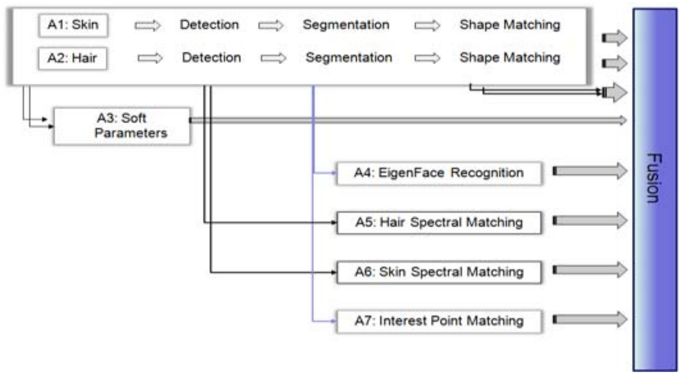

Figure 5. Recognition Agent Interface

2.6 Adaptive Feedback

To incorporate the ability to make relative comparisons over time, adaptive feedback loops were added to the established fusion hierarchy and are depicted in figure 6. The first feedback loop is included to examine the improvement potential of changing the dimensionality of the candidate gallery. This procedure involves reducing the gallery size by removing the lowest scoring subjects. This process is applied only for subject matching scores that fall below a user specified threshold. The multi-look functionality adds the capability to test additional probe images if and when they become available. This facet represents a temporal dimension that comes with multiple probe images or with hyperspectral video that obtains a series of face images over time. Both feedback loops can be active or applied individually. Finally, there are several control variables for the selection and weighting of the agents used in the fusion process. Research by Chawla [29] and Kuncheva [30] have highlighted the importance of randomness and diversity in the creation of classifier ensembles, so the controlled and random selection of these active agents is a current area of research.

At any stage of the hierarchy presented earlier in figure 3, a Libet level answer similar to intuition is created and is integrated at the higher or meta-representation levels of the hierarchy. The incorporation of qualia occurs as deliberation is made over the combined evidence from prior agents. The qualia based Cartesian theater that is created through the fusion representation provides an engineering advantage in the confidence assessment.

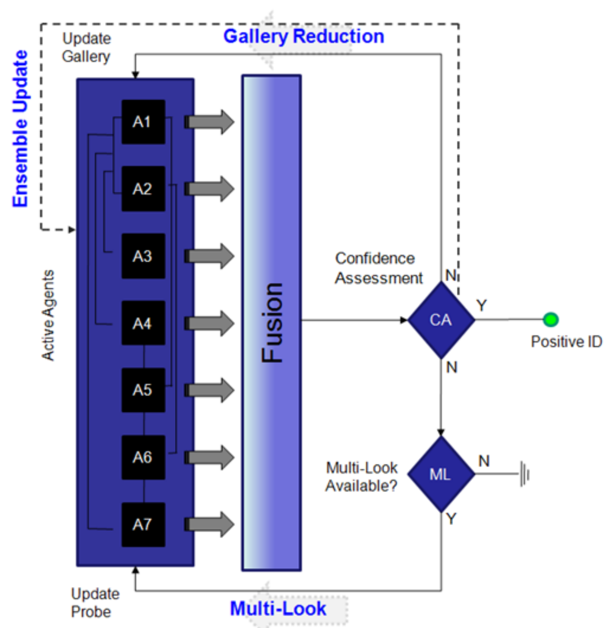

Figure 6. Architecture of Facial Recognition Adaptive Logic

3. GRAPHICAL USER INTERFACE TOOL

To facilitate interpretation of data analysis and assist with the visualization of results a Matlab-based GUI tool was designed to operate and test the facial recognition software. The GUI tool, pictured in figure 7, is a direct parallel to the architecture presented in figure 6. A user can select the active agents, enable feedback loops, and define the fusion weighting scheme while simultaneously analyzing results.

The GUI displays the probe to be matched and the best current match directly opposite. Below these displays, the top ten matches are displayed in thumbnail depiction along with their relative rankings and scores. Viewing the results of each algorithm is permitted by selecting the algorithm of interest in the “Results to Display” drop down menu. If feedback loops are employed, a user can select which result set to view, accompanied by the dimensionality of the gallery, in the “Gallery Set Results to View” menu. The pictorial results can be viewed in either grayscale or color images.

A box plot is displayed for each probe under consideration to provide continuous score distribution feedback when viewing results for each face and method. Additionally, gallery matches for each probe are scrollable to enable the visual evaluation of results for the entire score distribution. To review the quantitative results, the user can choose from cumulative match score plots, box plots, or histogram depiction of the relative scores and statistics.

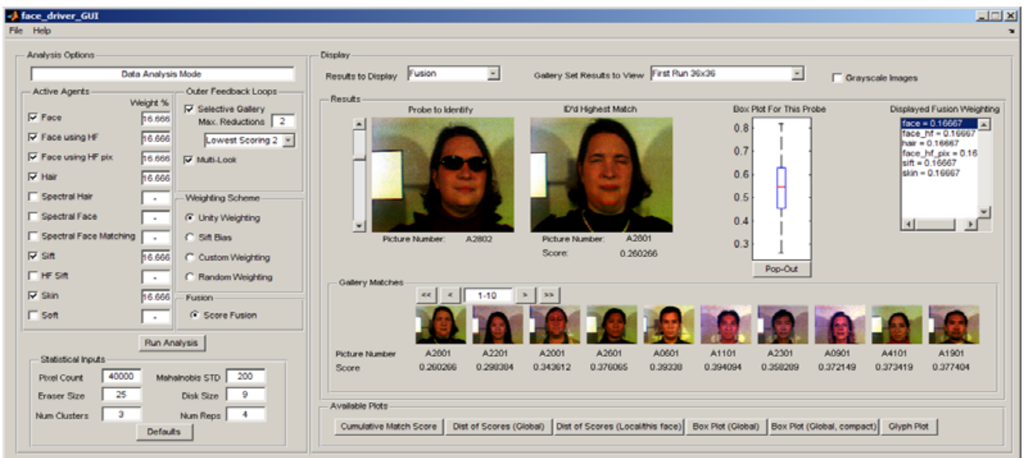

Figure 7. HSI Facial Recognition Graphical User Interface (with unity weighting)

For processing purposes, Matlab’s multiple processor pooling was employed on a dual quad core computer with 16GB of RAM. The processing requirements of the hyperspectral data along with the chosen methods necessitate such the use of parallel processing at times. However, an additional utility tool, allows the user to view saved results of any prior run by simply loading a results file. This file will display the algorithms used, type of feedback loops used, weighting schemes, and permit a user to view all results and face matches. The user is notified if the selected computer can support running the complete suite of software tools by viewing the status bar.

4. PERFORMANCE ASSESSMENT

4.1 Algorithm and Fusion Performance

During the initial testing of the CMU data, many of the same algorithms were utilized from previous HSI research [8], [13], [14], [17]. The results confirmed some of the challenges present in the CMU data. The difference being the quality between the CMU data and the grayscale AT&T data [30] or the more recent CAL HSI data [14] obtained with more modern equipment. Although the performance level of these algorithms were not replicated, the value of the various techniques is not reduced. A comparison of the previously published performance versus that obtained through our initial testing is shown in figure 8, to establish a preliminary performance threshold.

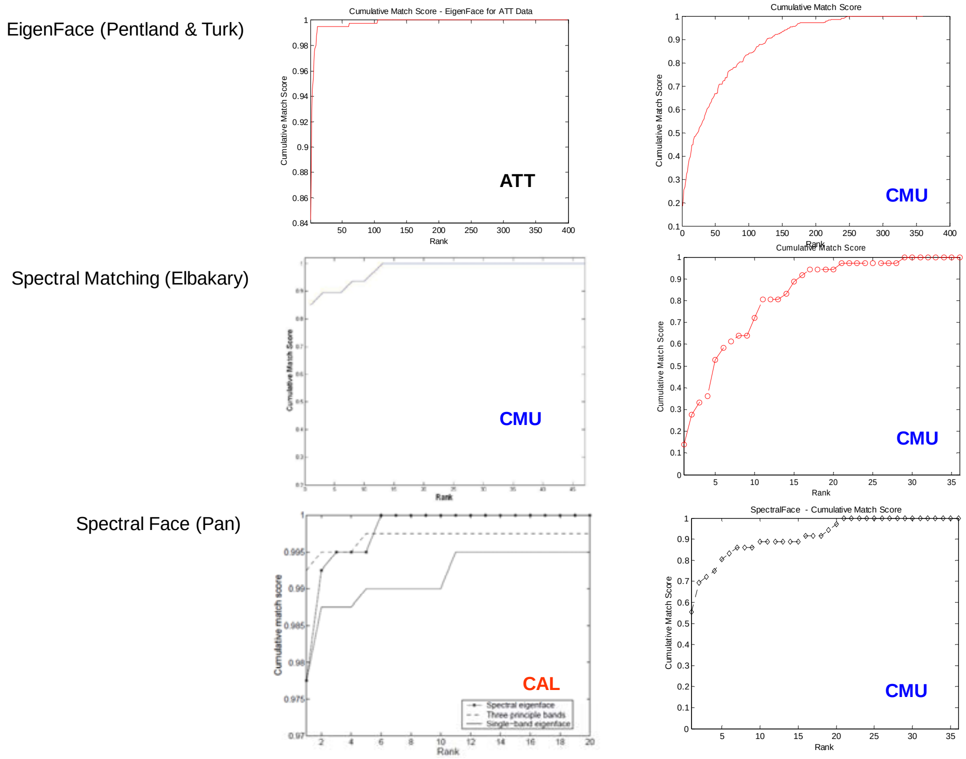

Figure 8. Testing results of Select Algorithms on CMU Database

Data processing starts with common but automated pre-processing, followed by the extraction of basic face features, and then a processing step where face features and characteristics are compared for subject matching. Average computation time for the pre-processing of each face is 14 seconds. Face matching algorithms take an additional average of 13 seconds to process each face against the gallery of 36 subjects for an algorithm suite consisting of 6 algorithms including SIFT, eigenface, various geometric comparisons and NDSI. Processing time can vary depending on the number of algorithms or agents activated by the user.

Findings from this initial round of testing reinforce the need for a fusion framework that combines complimentary aspects of these algorithms to enhance the performance capability regardless of data quality or environmental setting. Taking into account the processing time of some algorithms, a method to accomplish effective data reduction and processing should also be a considered to reduce overall computational time. The next section will briefly describe the results of integrating the separate algorithms into a hierarchy for a robust face recognition system.

4.2 QUEST Hierarchy Results and Findings

A combination of score and rank fusion strategies were tested with the most effective being a weighted score fusion strategy, wherein the overall matching score is a combination of weighted individual matching scores. Figure 9 depicts cumulative match score results obtained using a double weighting for the SIFT algorithms using 6 agents and a unity weighted fusion strategy for 7 agents. While the SIFT algorithm contributes a majority of the contribution, it is only through the inclusion of other identification methodologies, and inherent segmentation capability, that the overall identification accuracy is increased to 100%.

Enhancing this fusion strategy with the addition of the adaptive gallery feedback loop and the multi-look functionality allows us to continually process the results until a chosen threshold or confidence level is achieved. Figure 10 depicts an example where the poorest scoring match distribution is shown after initial matching and then after four feedback repetitions, during which the gallery size was reduced by 10 percent and a new probe image was injected each time. Through this repetitive process, matches with the lowest matching scores are re-checked as poor candidates are removed from the gallery and additional probe images are inserted into the process to confirm the correct identification.

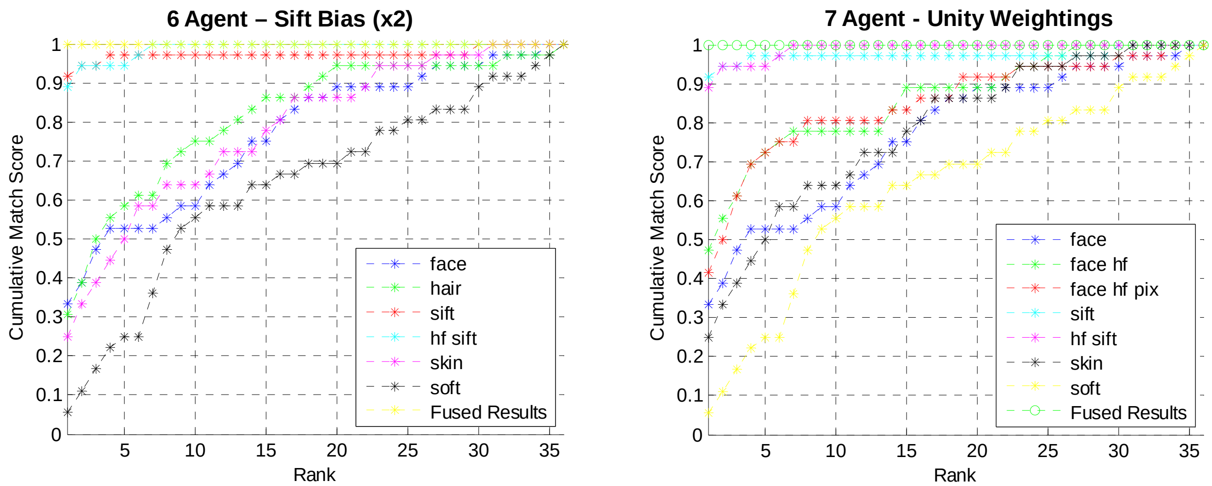

Figure 9. Cumulative Match Score Results Including Score Fusion

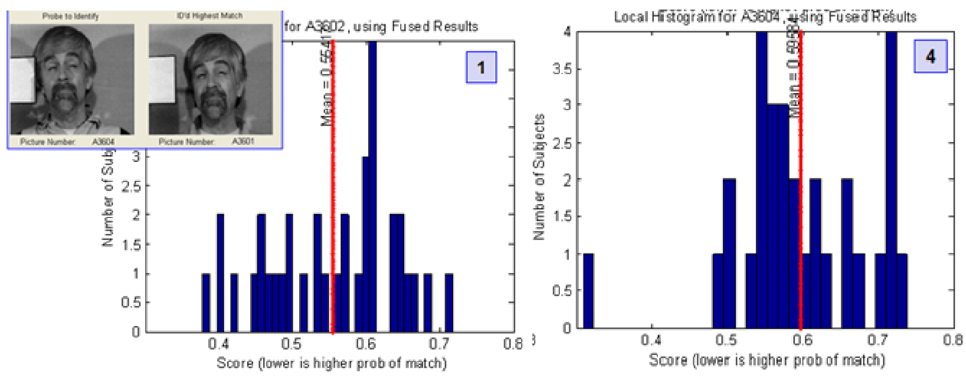

Figure 10. Adaptive Gallery / Multi-Look Results – Distribution of Matching Scores for Poorest Match

5. CONCLUSION

Even with the distinctiveness that comes with every human being, no single metric or feature has demonstrated the ability to identify all individuals in both controlled and uncontrolled environments across large populations using a single modality. This challenge frequently leads to solutions that incorporate multiple modalities that require close proximity and permission that accompany the selected biometrics not to mention the additional equipment and complexity. An alternative to this challenge may be to fuse contextual or complimentary spatial, spectral and temporal information in an efficient architecture that enhances effectiveness and efficiency. The use of hyperspectral imagery and a fusion hierarchy similar to the one presented in this paper offers many opportunities for the improvement of current face recognition systems and can be applied to a wider array of object recognition problems.

REFERENCES

[1] Osborn, Kris, "U.S. Ugrades Biometrics Gear to Spot Terrorists," Defense News, 16 Feb 2010.

[2] Jain, Anil, K., "Biometric Recognition: How Do I Know Who You Are?," *Signal Processing and Communications Applications Conference*, 3-5 (2004).

[3] Jain, Anil, Arun Ross, Salil Prabhakar, "An introduction to biometric recognition", *IEEE Transactions on Circuits and Systems for Video Technology*, 4 - 20 (January 2004).

[4] Denes, Louis J., Peter Metes, and Yanxi. Liu, "Hyperspectral Face Database," *Tech. Report CMU-RI-TR-02-25*,

Robotics Institute, Carnegie Mellon University (October 2002).
[5] Nunez, Abel, *A Physical Model of Human Skin and Its Application for Search and Rescue.* Air Force Institute of Technology (AU), Wright-Patterson AFB, OH, December 2009.
[6] Ryer, David, *QUEST Hierarchy for Hyperspectral Face Recognition*. Air Force Institute of Technology (AU), Wright-Patterson AFB, OH, December 2010.
[7] Pavlidis, Ioannis, P. Symosek, "The Imaging Issue in an Automatic Face/Disguise Detection System," *Proceedings IEEE Workshop on Computer Vision beyond the Visible Spectrum: Methods and Applications*, 15-24 (2000).
[8] Robila, Stephan, "Toward Hyperspectral Face Recognition", *Proceedings of SPIE – IS&T Electronic Imaging*, 6812: 1-9 (2008).
[9] Klare, Brendan and Anil Jain, "Heterogeneous Face Recognition: Matching NIR to Visible Light Images," Proceedings of the ICPR (2010).
[10] Bourlai, Thirimachos, Nathan Kalka, Arun Ross, Bojan Cukic and L. Hornak, "Cross-spectral Face Verification in the Short Wave Infrared (SWIR) Band," 2010 International Conference on Pattern Recognition (2010).
[11] Kong, Seong G., Jingu Heo, Besma R. Abidi, Joonki Paik and Mongi A. Abidi, "Recent advances in visual and infrared face recognition – a review," Computer Vision and Image Understanding, 97:103-135 (2005).
[12] Chou, Yi-Ting and P. Bajcsy, "Toward Face Detection, Pose Estimation and Human Recognition from Hyperspectral Imagery," Technical Report NCSA-ALG04-0005 (October 2004).
[13] Elbakary, M. I., M. S. Alam, M. S. Aslan, "Face Recognition Algorithm in Hyperspectral Imagery by Employing the K-means Method and the Mahalanobis Distance," Proceedings of SPIE, 6697:1-9 (2007).
[14] Pan, Zhihong, Glenn Healey, Manish Prasad, Bruce Tromberg, "Face Recognition in Hyperspectral Images," IEEE Transactions on Pattern Analysis and Machine Intelligence, 25.12: 1552-1560 (December 2003).
[15] Pan, Zhihong, Glenn Healey, Manish Prasad, Bruce Tromberg, "Hyperspectral Face Recognition in for Homeland Security," Proceedings of SPIE, 5074: 767-776 (2003).
[16] Pan, Zhihong, Glenn Healey, Manish Prasad, Bruce Tromberg, "Hyperspectral Face Recognition under Variable Outdoor Illumination," Proceedings of SPIE, 5425:520-529 (2004).
[17] Pan, Zhihong, Glenn Healey, Bruce Tromberg, "Multiband and Spectral Eigenfaces for Face Recognition in Hyperspectral Images" Proceedings of SPIE, 5779:144-151 (2005).
[18] Beveridge, R., Bolme, D., Teixerira, M., and Draper, B., 2003, The CSU Face Identification Evaluation System User's Guide: Version 5.0, Computer Science Department, Colorado State University.
[19] Turk, Matthew and Alex Pentland, "Eigenfaces for Recognition," *Journal of Cognitive Neuroscience*, 3.1: 71-86 (Winter 1991).
[20] Robila, S. A., "Using Spectral Distances for Speedup in Hyperspectral Image Processing", *International Journal of Remote Sensing*, 26.24: 5629-5650 (2005).
[21] Lowe, David G., "Object Recognition from Local Scale-Invariant Features," *Proceedings of the International Conference on Computer Vision*, Corfu, 1-8 (September 1999).
[22] Lowe, David G., "Distinctive Image Features from Scale-Invariant Keypoints," *International Journal of Computer Vision*, 1-28 (2004).
[23] Ando, Hiroshi, "ATR Human Information Processing Laboratories," International Media Technology Workshop on Abstract Perception, Japan (January 1994).
[24] Rogers, Steven K. "Qualia Exploitation of Sensor Technology." Web log post. Quest Discussion Topics. 28 June 2010. Web. 9 Oct. 2010. <http://qualellc.wordpress.com/2010/06/17/quest-discussion-topics-june-182010/>.
[25] Rogers, Steven K., Types of Qualia, Presentation, November 2009.
[26] Rogers, S., M. Kabrisky, K. Bauer, M. Oxley, A. Rogers, *QUEST: QUalia Exploitation of Sensor Technology* (2008).
[27] Jarudi, Izzat, Pawan Sinha, "Recognizing degraded faces: Contribution of internal and external features," Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Artificial Intelligence Memo 2003-004 and Center for Biological and Computational Learning Memo 225 (2005).
[28] Jain, Anil, Sarat Dass, Karthik Nandakumar, "Can Soft Biometric Traits Assist User Recognition," Proceedings of SPIE, 5404: 561-572 (2004).
[29] Chawla, Nitesh V., Kevin W. Bowyer, "Random Subspaces And Subsampling For 2-D Face Recognition," CVPR, San Diego II: 582-589 (June 2005).
[30] Kuncheva, Ludmila, Combining Pattern Classifiers, Methods and Algorithms, Wiley and Sons, New Jersey, 2004.
[31] AT&T the Database of Faces, AT&T Laboratories Cambridge, http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html