How to assess the iris aging efficiently?

Adam Czajka

Biometrics and Machine Learning Group
Warsaw University of Technology, Poland
aczajka@elka.pw.edu.pl

Iris workshop (panel discussion)
Biometric Consortium Conference – BCC 2013
September 18, 2013
Definition of template aging

1. Popular definition
   - “template aging (...) refers to the increase in error rates caused by time-related changes in the biometric pattern, its presentation, and the sensor”¹
   - scientific literature often assumes monotonic decrease of reliability

2. Aging understood as non-stationarity of matching scores
   - aging may manifest by non-monotonic changes in recognition accuracy
   - while observing stable average of scores, other statistics (e.g. variance) may still vary and impact the recognition accuracy

3. “Scientific” vs. “industrial” understanding of aging
   - science typically categorizes sources of aging (e.g. by possibility of their mitigation) and focuses on iris texture, rarely on its surroundings (e.g. systematic eyelids droop or pupil constriction in elderly people)
   - industry is simply interested how time lapse impacts the recognition reliability in general (independently of the source, and assuming that mitigation of some basic aging factors is possible) and what countermeasures should be applied

Aging databases

1. Adequacy of the existing datasets
   - difficult to assure simultaneously: ample dataset size (i.e. large number of subjects and regular presentations), large time lapse (e.g. a decade), homogeneity of equipment and collection environment, high ethnic heterogeneity and controlled, well documented health of volunteers
   - datasets we currently have are often small and sequestered, or large yet not prepared for aging studies, sometimes arbitrary censored

2. Development of aging datasets
   - collection (of new sets): takes time, equipment and capture protocols must be stable for long time periods, volunteers must be available during the collection, and social resistance may appear (e.g. biometrics and children)
   - connection (of the existing sets): requires international exchange of sensitive data (and sometimes solving intricacies of heterogeneous data protection law systems)
Evaluation procedures

1. Multitude of heterogeneous evaluation procedures
   • allow latitude in interpreting results
   • may provide contradictory conclusions, in spite of correct usage of the same datasets and statistical means (see e.g. NIST IREX VI\textsuperscript{2} vs. Univ. of Notre Dame\textsuperscript{3}, or Univ. of Ulster\textsuperscript{4} and J. Daugman commentary\textsuperscript{5})

2. Evaluation procedures dedicated for aging studies
   • analysis of conclusions from studies to date
   • development of interoperable and thoroughly discussed methodologies
   • respecting expectations of end users (i.e. industry as well as science)
   • involvement of standardization bodies (e.g. NWIP in ISO/IEC JTC1/SC37)

---

\textsuperscript{2} Grother et al. IREX VI: Temporal Stability of Iris Recognition Accuracy, NIST Int. Rep. 7948, 2013
\textsuperscript{3} Graham-Rowe, Ageing eyes hinder biometric scans, Nature/News, May 2012
\textsuperscript{4} Rankin et al., Iris recognition failure over time: The effects of texture, Pat. Rec. 45 (145–150), 2012
\textsuperscript{5} Daugman et al. No change over time is shown in Rankin et al. “Iris recognition failure over time: The effects of texture”, Pat. Rec. 46 (609–610), 2013
1. Aging phenomenon seems to be too complicated to be fully understood in research solitude

2. Joint development of evaluation procedures, cooperation with medical community, adaptation and merging of the existing datasets, analyzing of results and providing the conclusions guarantee faster discovery of the truth

3. Lest we forget about the aim: understand the aging phenomenon and propose adequate countermeasures (not force the results harmonizing with particular interests)