

**Large Data Analysis: automatic visual personal identification  
in demography of 1.2 billion persons**

John Daugman  
University of Cambridge (UK)

**Abstract**

The largest biometric deployment in history is now underway in India, where the Government is enrolling the iris patterns (among other data) of all 1.2 billion citizens. The purpose of the Unique Identification Authority of India (UIDAI) is to ensure fair access to welfare benefits and entitlements, to reduce fraud, and to enhance social inclusion. Only a minority of citizens in India have bank accounts; only 4 percent have passports; and less than half of all aid money reaches its intended recipients. A person who lacks any means of establishing their identity is excluded from entitlements and does not officially exist; thus the slogan of UIDAI is: "To give the poor an identity." This ambitious program enrolls a million people every day, across 36,000 stations run by 83 agencies, with a 3 year completion target for the entire national population. To date almost 500 million persons have been enrolled. In order to detect and prevent duplicate identities, every iris pattern that is enrolled is first compared against all others enrolled so far; thus the daily workflow now requires 500 trillion (or 500 million-million) iris cross-comparisons. Avoiding identity collisions (False Matches) requires high biometric entropy, and achieving the tremendous match speed requires phase bit coding. Both of these requirements are being successfully delivered by wavelet methods developed by the author for encoding and comparing iris patterns, which will be the focus of this presentation.