

Viability of Hybrid Approaches in Speaker Verification

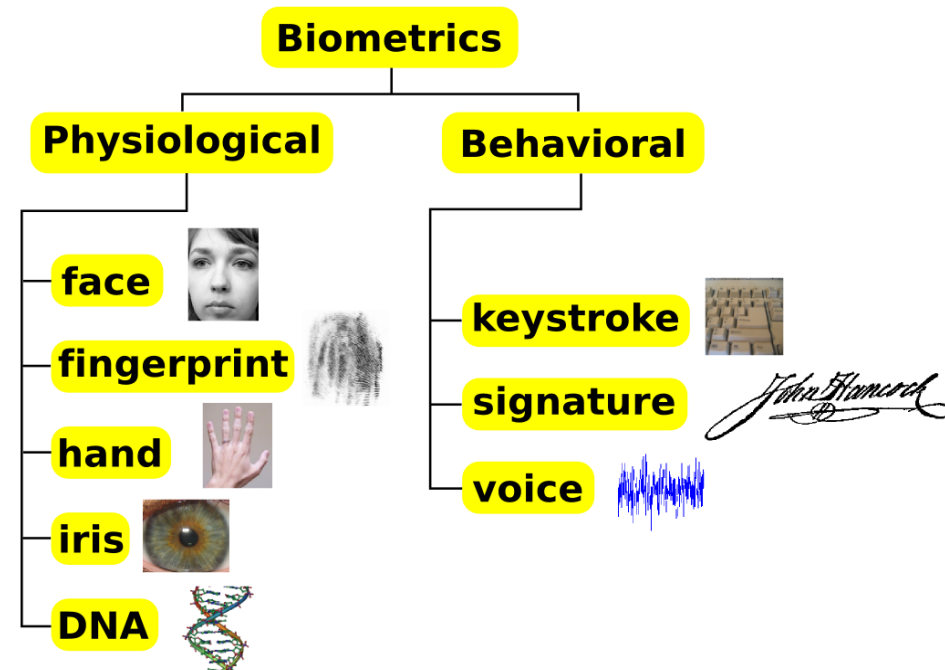
Saeid Safavi

Information Engineering and Processing Architectures Group
Electronics Communications and Electric Division
School of Engineering and Technology
University of Hertfordshire
College Lane Campus, Hatfield AL10 9AB, UK



- Introduction
- Motivation
- Hybrid Architectures
- User convenience vs security
- Conclusion

- Biometrics Applications
 - security access control to physical places
 - secure login to computer systems and mobile devices
 - online banking
 - personalized human-machine interfaces
- Voice biometrics
 - offer convenience to the users
 - rely on microphones not on special sensors



- text-dependent speaker verification
 - users pronounce a pre-determined pass-phrase
 - The pass-phrases are either unique or prompted by the system, e.g. in a screen.
 - High SV performance
 - Vulnerable to spoofing (replay) attacks
- text-independent speaker verification
 - the speech content is not apriori known
 - Lower SV performance
 - Robust to spoofing attacks

<i>Mode of operation</i>		<i>Recognition accuracy</i>	<i>Convenience for clients</i>	<i>Spoofing</i>
Passphrase based voice biometrics (TD)		High	High	Easy
Text prompted voice biometrics	Text-dependent (TD)	High	Medium	Medium
	Text-independent (TI)	Medium	Low	Hard
Intrinsic TI voice biometrics, i.e. verify the client with speech uttered for interaction in automated speech driven systems		Medium	Training: Low	Medium
			Testing: High	

- Proposed architecture 1

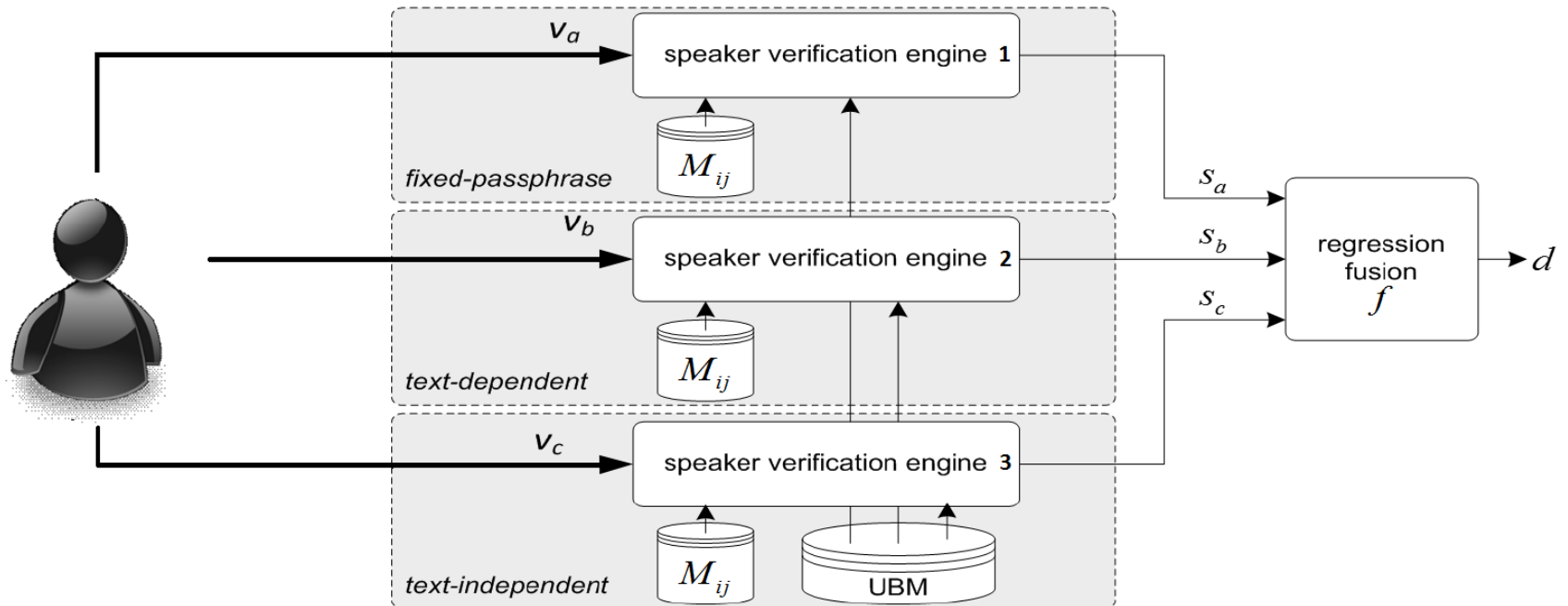


TABLE I . SPEAKER VERIFICATION EQUAL ERROR RATE (EER %) OF THE EVALUATED REGRESSION FUSION ALGORITHMS.

	<i>A</i>	<i>B</i>	<i>C</i>	<i>LR</i>	<i>MLP</i>	<i>SVR</i>
<i>GMM-UBM</i>	20.73	20.86	22.78	17.50	16.96	17.34
<i>HMM-UBM</i>	8.69	11.31	29.70	6.94	6.75	6.89
<i>i-vector</i>	20.91	27.32	29.08	19.23	18.42	18.62

Experimental Results

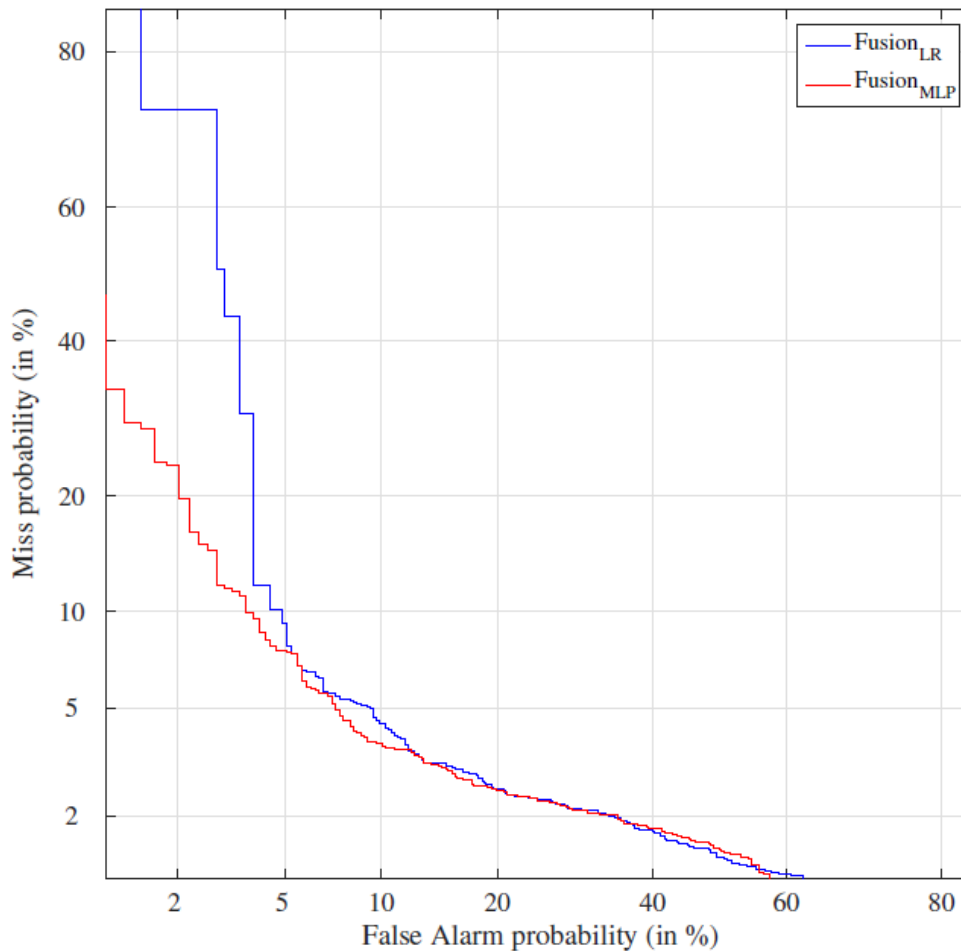
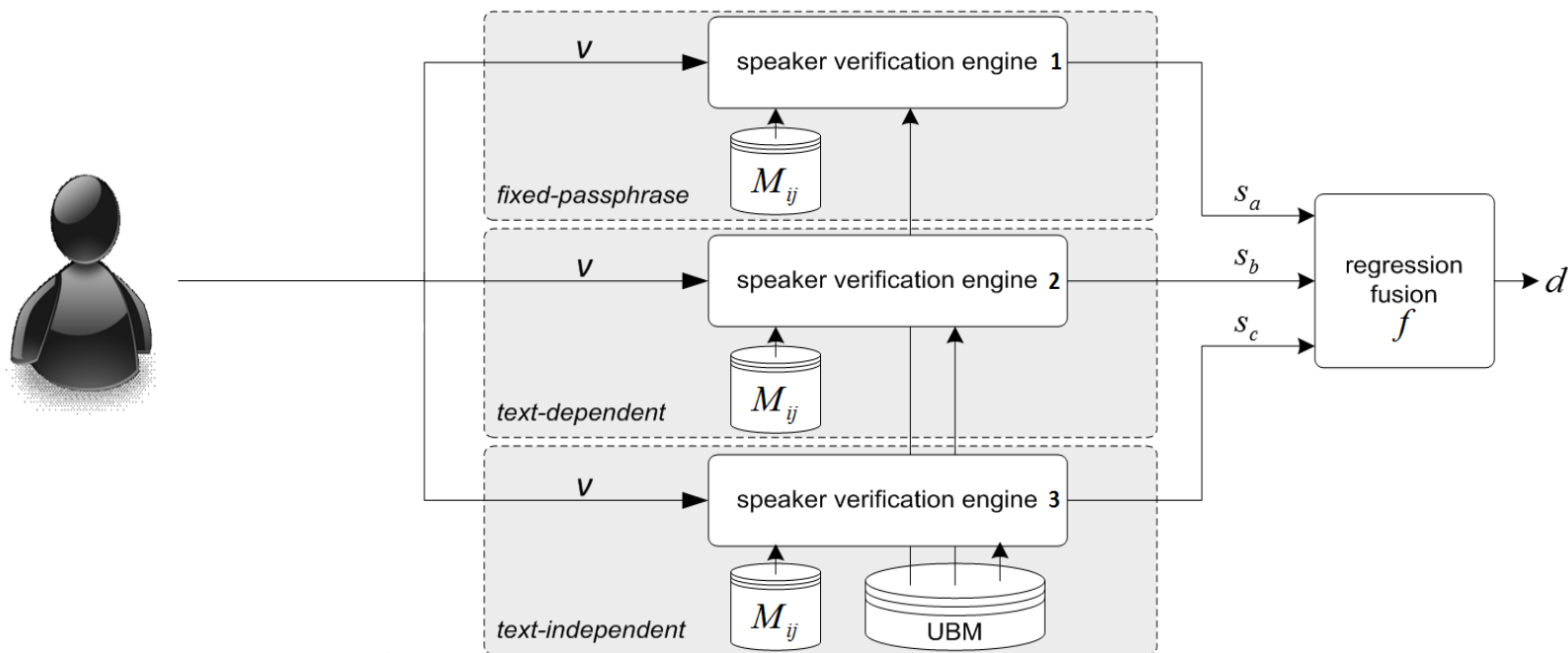


Fig. 1 Speaker verification performance for fusion of modes of operation using **HMM-UBM** single mode engines for mode A and B and using **GMM-UBM** single mode engine for mode C, using **male** speakers of **AvspooF**.

Experimental Setup

- Proposed architecture 2



Experimental Results

Table II. Obtained performance in terms of EER(%) corresponding to different setups for proposed hybrid authentication machine.¹ Fusion applied on scores from best performing single mode engines.

Different Setups	Hybrid Speaker Verification Accuracy		User Convenience
	LR	MLP	Level
Three Inputs (HMM-UBM)	6.94%	6.75%	Low
Three Inputs ¹	6.07%	5.87%	Low
One Input ¹	7.28%	6.96%	High
Baseline	7.78%	7.18%	Low

Viability of hybrid architecture



Convenience



Security



Complexity

- A Hybrid architecture speaker verification technique has been proposed to safeguard the voice-based authentication services
- The architecture tested under different conditions (e.g. Using three inputs from the user and only one input, during test)
- In addition to comparison of the results with single mode performance, the results also compared with more comparable baseline system.
- Hybrid architecture will improve the performance effectively but for the cost of user convenience.
- Which architecture to choose is mainly depends on the type of application.

Thank you for your attention!

University of
Hertfordshire



The research reported in the present paper was partially supported by the H2020 OCTAVE Project (“Objective Control for TAlker VErification” - Grand Agreement number 647850) funded by the European Commission under the Horizon 2020 Programme. Project web-site: <https://www.octave-project.eu//>.



