



Anti-spoofing for speaker verification: the OCTAVE Delta

Nicholas Evans



The OCTAVE Project has received funding from
the *European Union's Horizon 2020 Research and Innovation Program*,
under Grant Agreement No. 647850



Overview



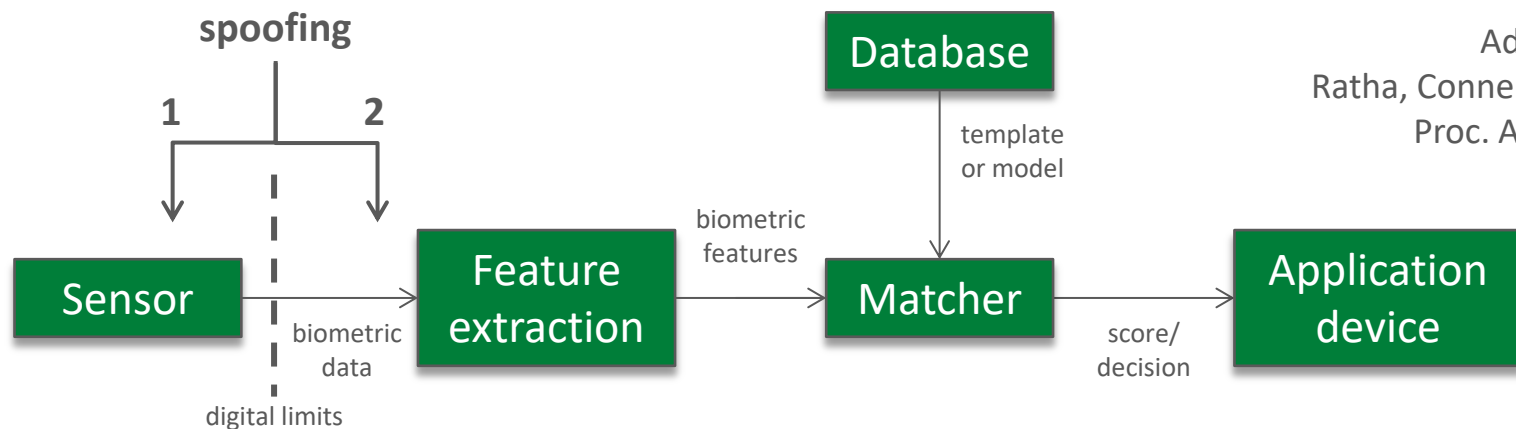
- spoofing and ASV
- the OCTAVE Δ
- OCTAVE
- ASVspoof 2017
- ASVspoof 2015
- wrap-up



Spoofing



- a.k.a. presentation attacks (ISO / IEC)
- “persons masquerading as others in order to gain illegitimate access to sensitive or protected resources” [Hadid et al., IEEE SPM, 2015]



Adapted from
Ratha, Connell and Bolle,
Proc. AVBPA, 2001

- before and after microphone – a contentious issue



Automatic speaker verification



Attack	ASV FAR	
	Baseline	Spoofing
Impersonation [Gonzales-Hautamäki 2013]	9%	12%
Replay [Villalba 2011]	1%	68%
Speech synthesis [De Leon 2012]	0%	86%
Voice conversion [Matrouf 2006]	8%	100%

What is the ASV system?

What is the FRR?

How comparable are these results?

How do these results influence the design of spoofing countermeasures?



OCTAVE



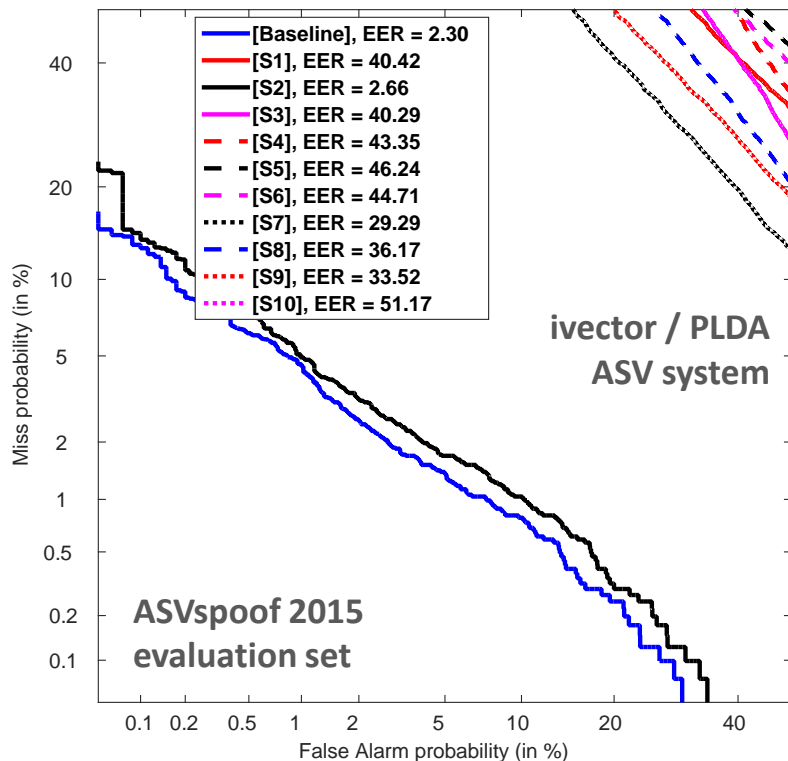
- improve the research methodology
- advance the state of the art in spoofing countermeasures for ASV
- OCTAVE: reliable, secure, robust



ASVspooof 2015



7 voice conversion, 3 speech synthesis spoofing attacks, 100+ speakers



Train set
ground-truth



Development set
ground-truth

Evaluation set
NO ground-truth
unknown attacks



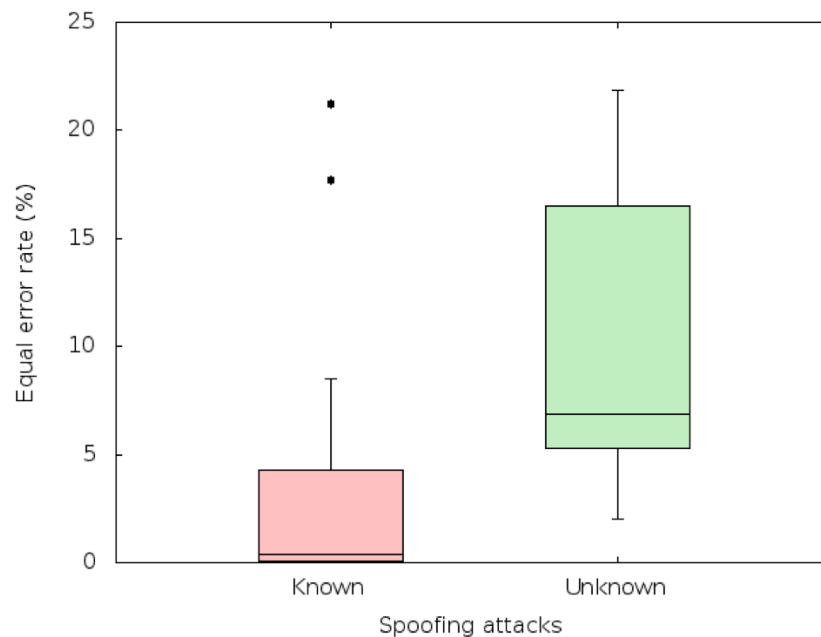
ASVspoof 2015



Top-ranked submissions (of 16)

Team	Known attacks (S1 - S5)	Unknown attacks (S6 - S10)	Average (all)
DA-IICT	0.408	2.013	1.211
STC	0.008	3.922	1.965
SJTU	0.058	4.998	2.528
NTU	0.003	5.231	2.617
CRIM	0.041	5.347	2.694
F	0.358	6.078	3.218
G	0.405	6.247	3.326
H	0.670	6.041	3.355
I	0.005	7.447	3.726
J	0.025	8.168	4.097

Overfitting – generalisation still missing

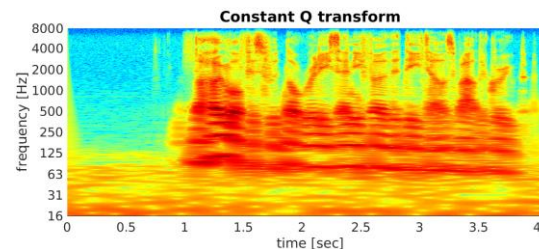
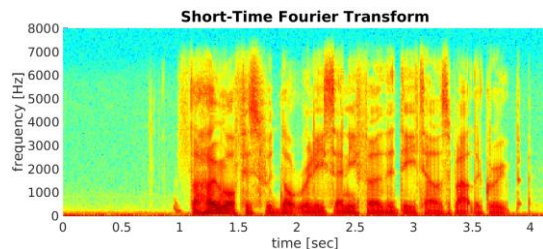




The OCTAVE Δ



“A new feature for automatic speaker verification anti-spoofing: constant Q cepstral coefficients”, M. Todisco, H. Delgado, N. Evans, in Proc. ODYSSEY, 2016.



Winner of best paper award
at the Speaker and Language Recognition Workshop (ODYSSEY) 2016



The OCTAVE Δ



Features	Known	Unknown	Average
LFCC	3.54	4.49	4.01
MFCC	1.50	4.97	3.24
CFCC	1.43	3.11	2.27
CQCC	0.05	0.46	0.26 Δ

ASVspoof 2015 (EER)

Features	Controlled	Variable	Average
LFCC	5.88	4.43	5.11
CQCC	1.80	1.92	1.86 Δ

RedDots Replayed (EER)

AVspoof (HTER)

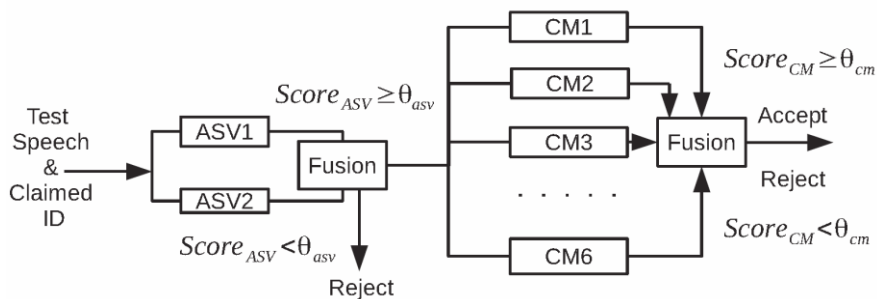
Submission	Known	Unknown	Pooled
IITKGP ABSP	0.98	14.75	1.26
Idiap	1.05	48.36	2.04
SJTUSpeech	2.08	10.46	2.20
CQCC	0.29	12.10	0.67 Δ



The OCTAVE Δ



“Integrated Spoofing Countermeasures and Automatic Speaker Verification: an Evaluation on ASVspooF 2015”, M. Sahidullah, H. Delgado, M. Todisco, H. Yu, T. Kinnunen, N. Evans and Zheng-Hua Tan, in Proc. Interspeech, 2016.



Metric	Male		Female	
	Dev	Eval	Dev	Eval
FRR	5.47	7.40	7.50	7.29
FAR(Z)	5.33	6.24	7.44	6.26
FAR(K)	0.00	0.00	0.00	0.00
FAR(U)	-	0.34	-	1.75

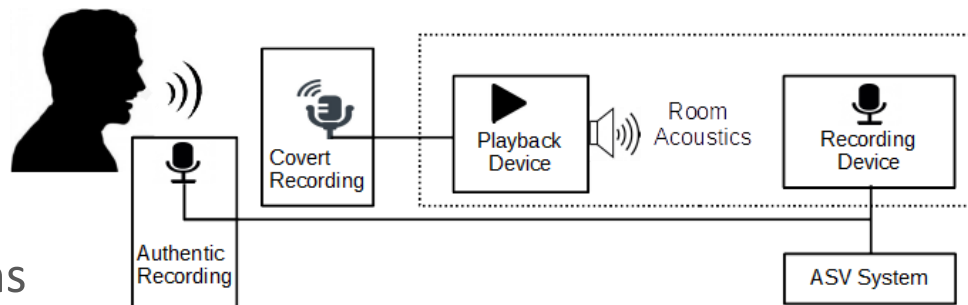
OCTAVE Δ : first integration of ASV with spoofing countermeasures for ASVspooF



ASVspoof 2017



- replay spoofing
- aligned to research in ASV
- re-recording of RedDots
- OCTAVE and other collaborations



“ASVspoof 2017: Automatic Speaker Verification Spoofing and Countermeasures Challenge Evaluation Plan”

T. Kinnunen, N. Evans,

J. Yamagishi, K. A. Lee, M. Sahidullah, M. Todisco, H. Delgado

“RedDots Replayed: A New Replay Spoofing Attack Corpus for Text-dependent Speaker Verification Research”,

T. Kinnunen, M. Sahidullah, M. Falcone,

L. Costantini, R. Gonzalez Hautamaki, D. Thomsen, A. Sarkar, Z.-H. Tan,

H. Delgado, M. Todisco, N. Evans, V. Hautamaki, K. A. Lee, in Proc. ICASSP, 2017



ASVspoof 2017



- wide participation
 - ~120 data requests
 - ~30 countries
- evaluation data
 - 1.1 GB in size
 - ~15,000 audio files
 - 112 replay conditions
- 49 evaluations submissions

Top-ranked submissions (preliminary analysis)

Team	C1	C2	C3	C4	C5	C6	Pool. Av.
A	8.77	4.46	6.50	5.71	6.97	25.28	6.73
B	1.70	1.91	8.09	5.18	8.68	25.63	7.39
C	9.75	9.99	10.77	9.78	11.43	20.77	10.54
D	8.65	14.77	10.80	9.89	14.72	23.03	12.39
E	12.59	17.35	12.89	13.04	14.01	29.02	14.31
F	11.92	18.61	12.34	13.21	14.96	27.20	14.93
G	15.81	21.62	15.38	15.74	14.99	22.97	16.35
H	13.81	32.76	14.71	14.17	16.25	27.18	17.62
I	19.10	23.38	19.32	17.47	15.52	18.52	18.07
J	17.89	17.28	15.94	17.33	20.31	22.45	18.33



Progress



small, purpose collected datasets

▲ 2013 Interspeech special session

adapted, standard datasets

common datasets, metrics, protocols

Post-OCTAVE
Among the most successful anti-spoofing initiatives

Pre-OCTAVE
Anti-spoofing research in ASV far behind that in other modalities

common datasets, replay, generalisation, channel variation

▲ ASVspooF 2015

▲ ASVspooF 2017



Summary



- **spoofing is a real threat to ASV reliability**
- **threat diversity: voice conversion, speech synthesis, replay**
- **potential to protect ASV with dedicated spoofing countermeasures**
- **integration and manageable / justified impacts on usability**
- **OCTAVE: reliable, secure, robust**



Acknowledgements



■ OCTAVE members:

- Benoit Fauve, Jasmin Liu, Federico Alegre, Hector Delgado, Massimiliano Todisco, Nicholas Evans, Saeid Safavi, Hock Gan, Iosif Mporas, Reza Sotudeh, Dennis Thomsen, Achintya K. Sarkar, Zheng-Hua Tan, Tomi Kinnunen, Md Sahidullah, Fotis Talantzis, Sotirios Diamantas, Heinz Hertlein, Aladdin Ariyaeenia, Mauro Falcone, Andrea Paoloni, Nicolai Bæk Thomsen, Federica Mangiattordi, Emiliano Pallotti, Martti Pitkanen, Zoe Jeffrey and Hock Gan



■ ASVspoof 2015 and 2017 contributors:

- Zhizheng Wu, Tomi Kinnunen, Junichi Yamagishi, Cemal Hanilci, Md Sahidullah, Aleksandr Sizov, Daisuke Saito, Tomoki Toda, Ali Khodabakhsh, Cenk Demiroglu, Linghui Chen, Zhen-Hua Ling, Hector Delgado, and Massimiliano Todisco