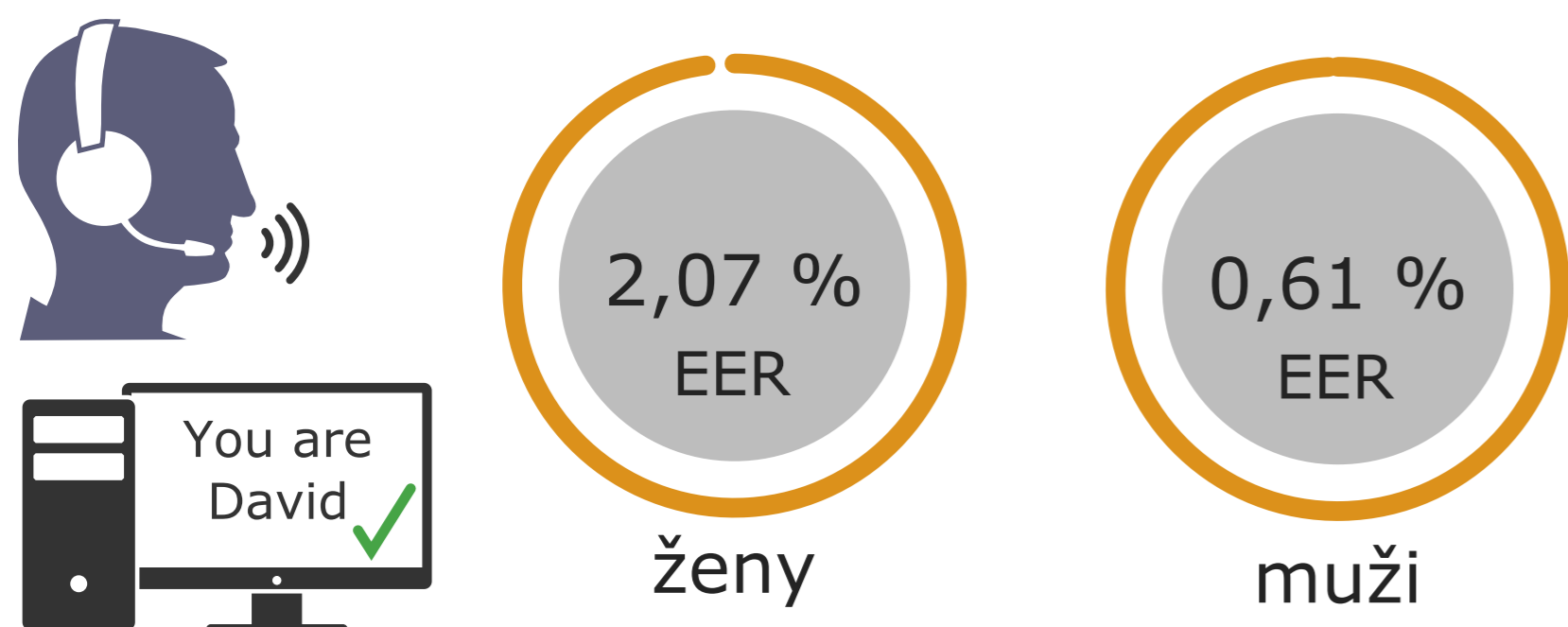


Mikrofonní pole pro rozpoznávání řečníků

Ladislav Mošner, xmosne01@stud.fit.vutbr.cz

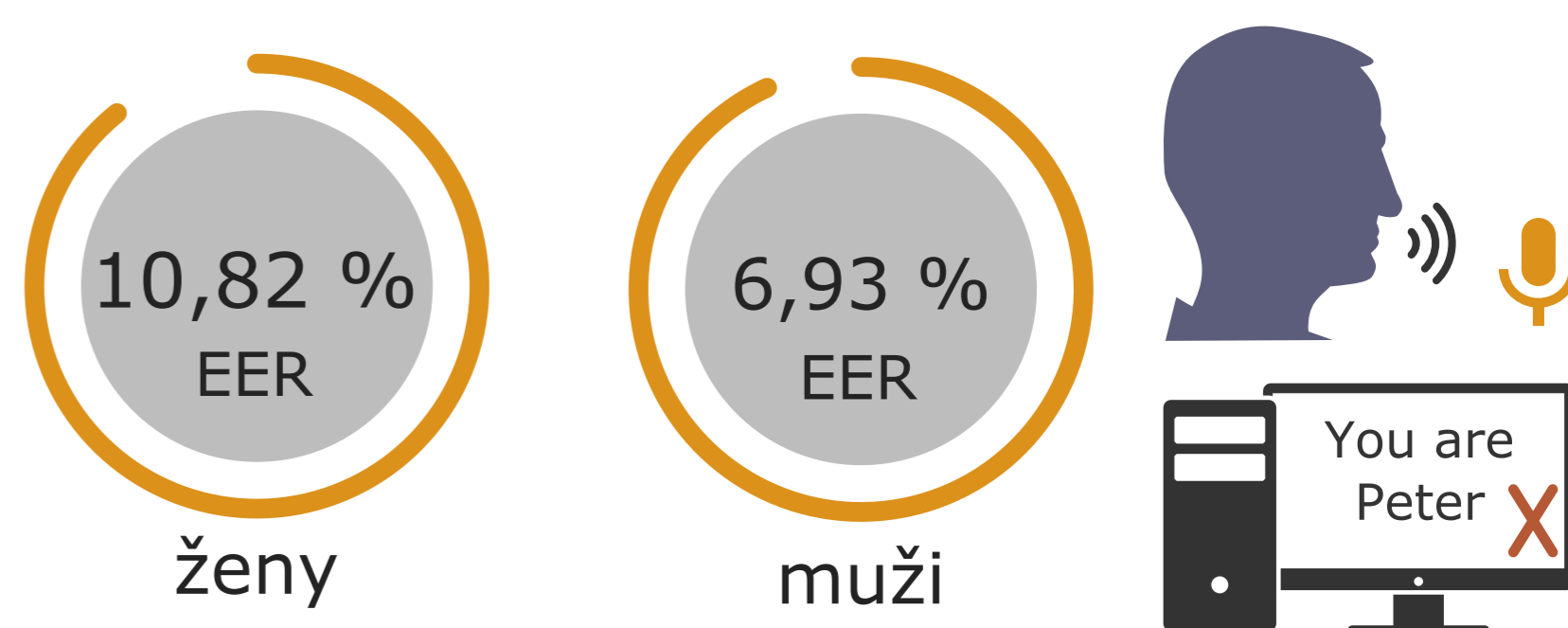
MOTIVACE

Mikrofon blízko úst



Současné systémy pro rozpoznávání řečníků pracují velice přesně.

Vzdálený mikrofon

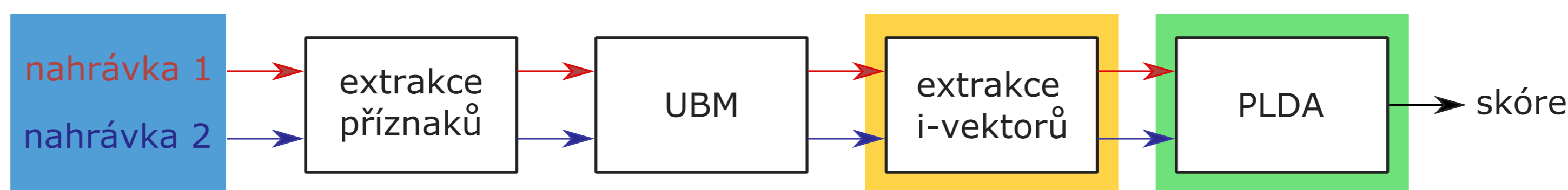
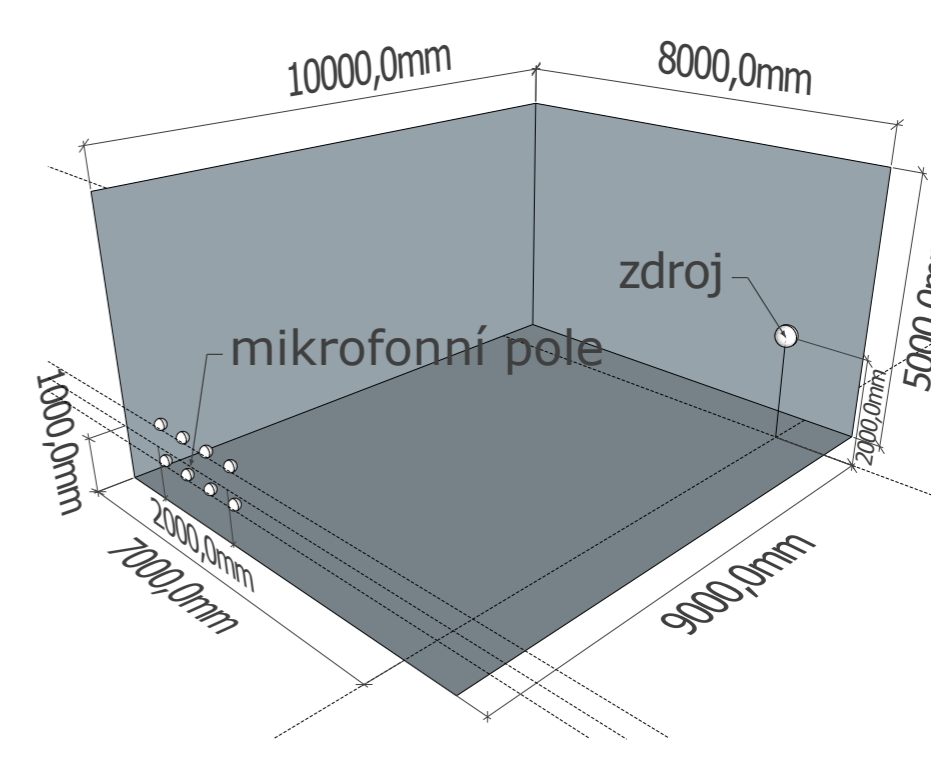
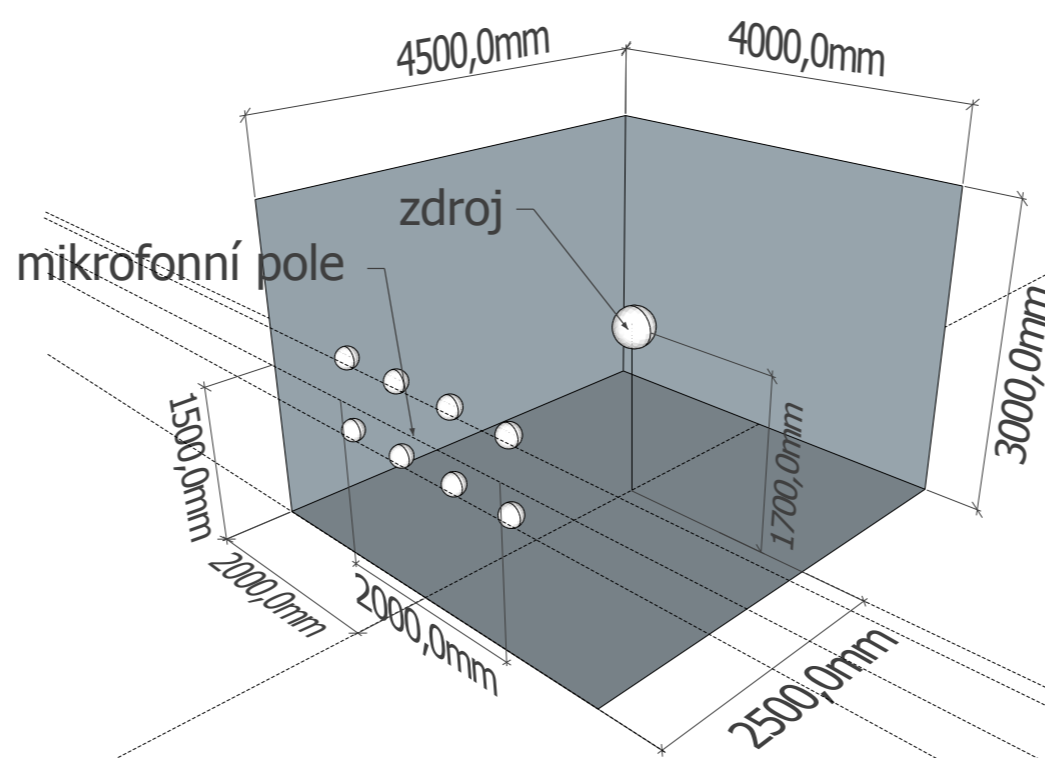


Přesnost se však razantně zhorší v případě vzdáleného nahrávání jedním mikrofonom.

ŘEŠENÍ

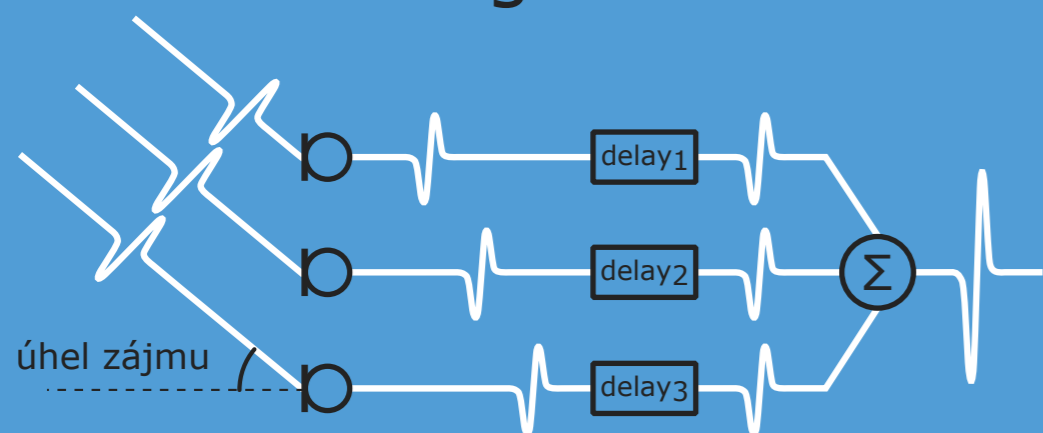
Mikrofonní pole

- prostorově uspořádaná sada mikrofonů
- prostorový filtr
- experimenty probíhaly v simulovaných místnostech: pokoj, hala



Delay-and-sum

- zpracování signálů z mikrofonního pole
- směrování pohledu na konkrétní místo
- zarovnání signálů



Přetrénování extraktoru i-vektorů

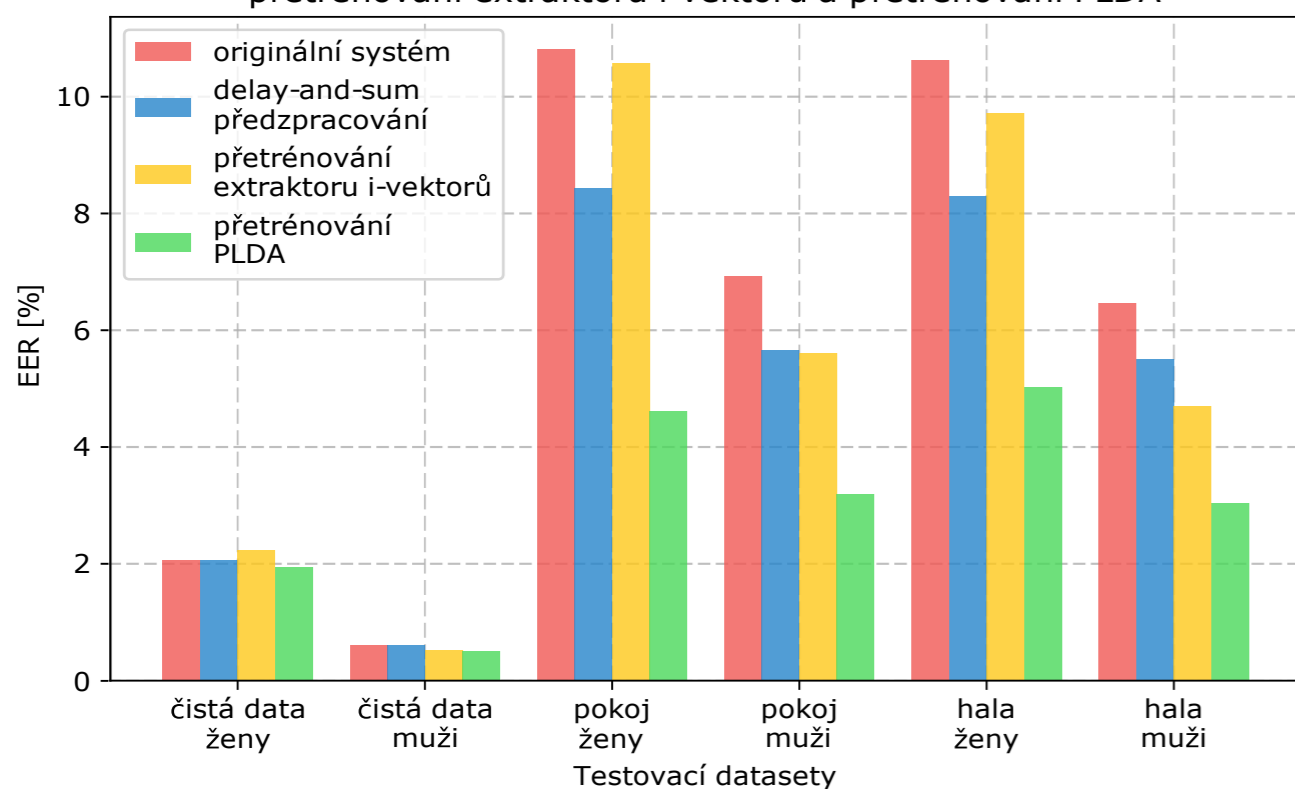
- rozšíření trénovací sady na dvojnásobek
- simulace náhodných místností
- multikanálová data z mikrofonního pole
- provedení delay-and-sum

Přetrénování PLDA

- adaptace klasifikátoru i-vektorů
- rozšíření trénovací sady obdobně jako u přetrénování extraktoru i-vektorů

VÝSLEDKY

Vliv nezávislé aplikace delay-and-sum předzpracování testovacích dat, přetrénování extraktoru i-vektorů a přetrénování PLDA



	Pokoj		Hala	
	Ženy	Muži	Ženy	Muži
DS	27,36	20,09	21,19	16,38
ivec_r	2,83	20,89	10,59	30,17
DS + ivec_r	27,39	34,37	33,21	32,76
PLDA_r	70,99	59,16	65,49	58,62
DS + PLDA_r	75,92	68,03	75,44	60,34
ivec_r + PLDA_r	71,10	54,45	69,97	61,21
DS + ivec_r + PLDA_r	78,86	63,24	77,95	62,12

Čísla v tabulce znamenají zotavení z chyb způsobených akustikou místnosti při vzdáleném snímání jedním mikrofonom v procentech (větší je lepší).

Budoucí práce

- srovnání výsledků s výsledky pro reálné nahrávky
- sledování pozice mluvčího