

# Формантне розпізнавання мовлення

доповідач –  
Тарануха В. Ю.

2011

# Що таке форманта?

Термін введено німецьким фізіологом Людигмаром Гераном у 1894 році.

Форманта - певна частотна область, в якій внаслідок резонансу посилюється деяке число гармонік тону, виробленого голосовими зв'язками, тобто в спектрі звуку форманта є досить чітко виділяється областю посилених частот, яка визначається за усередненою частотною величиною.

# Фізичний зміст

Кількість формант співставляється з кількістю резонансних порожнин в мовному тракті. Кожна з формант визначається всіма ділянками мовного тракту, хоча ступінь впливу в кожному конкретному випадку неоднакова.

Фактично феномен форманти є прояв роботи активного смугового фільтра в складі мовного тракту.

# Інформаційна насиченість

У більшості випадків для розрізнення голосних звуків достатньо перших двох формант, однак практично завжди кількість формант в спектрі звуку більше двох, що вказує на більш складні зв'язки між артикуляцією та акустичними характеристиками звуку, ніж за умови розгляду тільки двох перших формант.

Вважається, що для характеристики звуків мови досить виділення чотирьох формант –  $F_1$ ,  $F_2$ ,  $F_3$ ,  $F_4$ , які нумеруються в порядку зростання їх частоти.

# Мовлення за звучанням

- Нормальне мовлення.
- Шепіт. Основна проблема в розпізнаванні - зникнення основного тону.
- Спів. Основна проблема в розпізнаванні – спотворення звичних довжин звуків.

# Інформація, яку можна здобути з формант

- Це інформація про голосні звуки.
- Не завжди надійно виражаються у формантах приголосні, особливо такі як: [л], [м], [н], [с] – ті приголосні, де наявні частоти утворені тому числі резонансом Гельмгольца.

# Стійкість формантних ознак

- Незалежно від того, як відбувається мовлення, тоном, пошепки, чи співом людина добре розпізнає голосні звуки, які в свою чергу найсильніше визначаються через співвідношення між  $F1$  і  $F2$ , при можливому ігноруванні  $F0$  (частоти основного фону).

# Використання формант

- Визначення голосних звуків.
- Визначення меж звуків.
- Визначення диктора.
- Розпізнавання зашумленого мовлення.
- Тощо.



# Визначення голосних звуків

- Формантні характеристики голосних звуків найбільше досліджені.
- Наявні усереднені значення для звуків.
- Основний напрямок досліджень – розбивка на класи-зразки, з метою зменшити взаємоперетини відображень звуків.

# Визначення голосних звуків

звук	F1	F2	F3
А	600	1200	2400
Э	450	1700	2500
О	450	800	2500
У	350	750	2200
Ы	300	1800	2450
И	300	2200	3000
Е	400	1850	2600
Я	450	1500	2450

# Визначення голосних звуків

Взаємні перетини пар сумішей нормальних розподілів (%)

звук	А	Э	О	У	Ы	И	Е	Я
А	100	6	40	18	7	8	7	17
Э	6	100	10	6	45	38	42	45
О	40	10	100	46	9	12	9	16
У	18	6	46	100	6	10	7	8
Ы	7	45	9	6	100	64	51	25
И	8	38	12	10	64	100	65	24
Е	7	42	9	7	51	65	100	33
Я	17	45	16	8	25	24	33	100

# Визначення голосних звуків

Взаємні перетини сумішей нормальних розподілів, %  
(«Свої» проти «чужих»).

звук	А	Э	О	У	Ы	И	Е	Я
%	23	39	32	19	40	45	45	37

# Визначення меж звуків

- Попри те, що вимовлені без тону назальні приголосні погано виражаються у формантах, при використанні тону, саме формантні частоти є найбільш дієвими ознаками меж звуків.
- Для порівняння: фрикативні ([с],[ш]) виділяються по характерним частотам спектру, дзвінкі змичні ([г']) – частоті радіального резонансу.

# Визначення меж звуків

- При використанні послідовного розподілу компонент і алгоритму максимізації правдоподібності.
- Середня похибка визначення положень меж склала 4,52 мс, середнє число пропусків меж склала 0,95%.

# Визначення диктора

- Відбувається за умови, що диктор хоче, щоб його впізнали.
- На основі функціоналу близькості між двома наборами формант
- В розглянутій роботі число використаних формант сягає 9.
- Оптимум вказаний при числі формант – 8

# Визначення диктора

Поріг (умовний)	19	28	29	30	42	57	71
Хибні позитивні	0.677	0.3967	0.377	0.3525	0.1148	0.0393	0.0115
Хибні негативні	0	0.0068	0.0096	0.0147	0.1164	0.3700	0.6331



# Розпізнавання зашумленого мовлення

- У широкосмугових спектрограмах зашумленої мови легко виділяється тільки загальне уявлення про розподіл енергії звуків мови, в тому числі траєкторії формант.

# Розпізнавання спотвореного мовлення

- Наявні роботи по розпізнаванню мовлення людей з вадами мовного тракту.

# Додаткові можливості

- Наявні роботи, де крім формант вживаються “антиформанти”, як області де очікуються спади енергії коливань.
- Наприклад, наявне дослідження для звуку [с].

# Основні недоліки методу

- Форманти не завжди визначаються однозначно в спектрі.
- Одних лише формант не досить для розпізнавання.
- Навіть при використанні додаткових звукових ознак, лише на основі звуку не вдається визначити, що ж було вимовлено з достатньою точністю.

**Дякую за увагу**