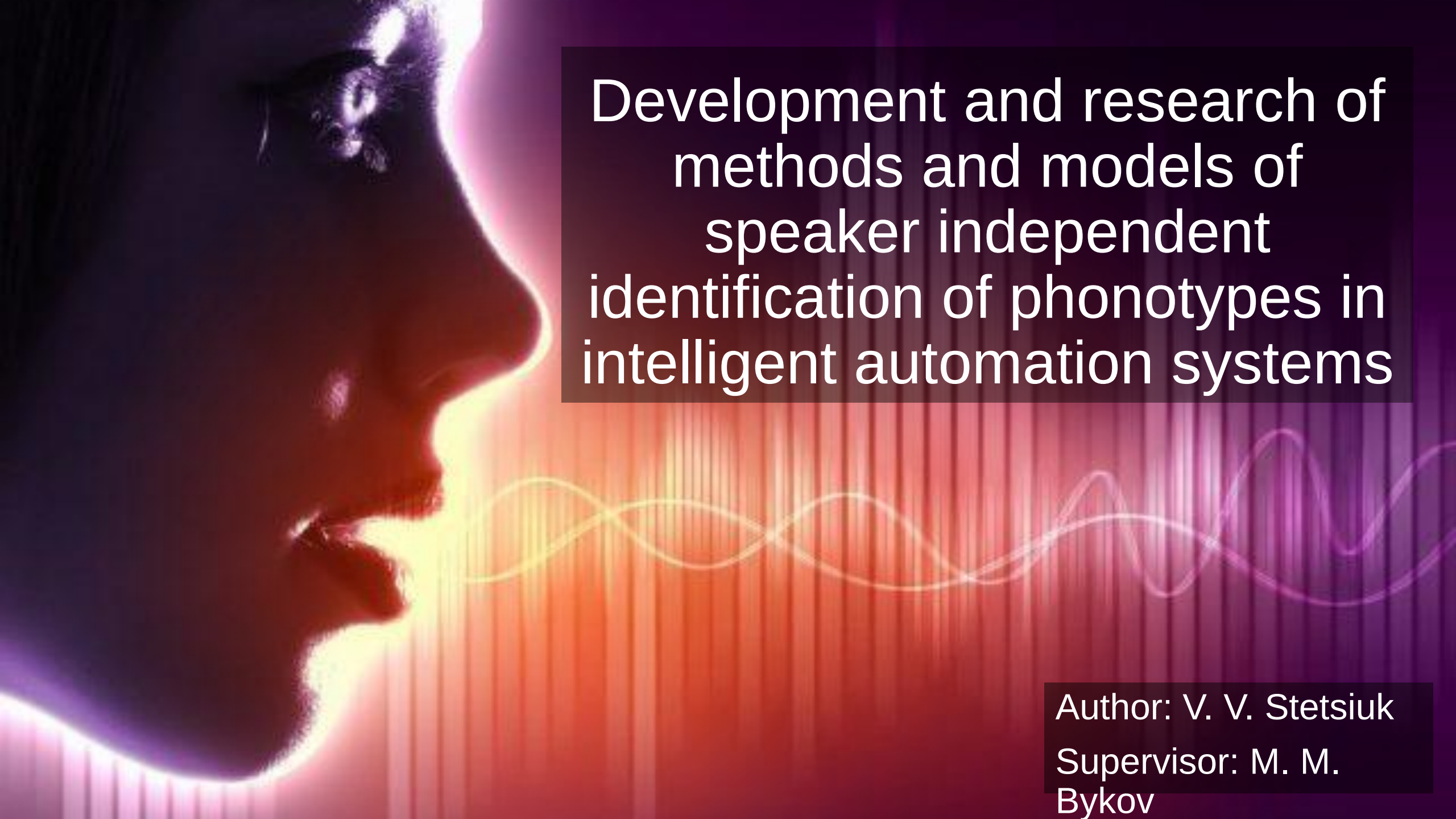


Development and research of methods and models of speaker independent identification of phonotypes in intelligent automation systems

Author: V. V. Stetsiuk

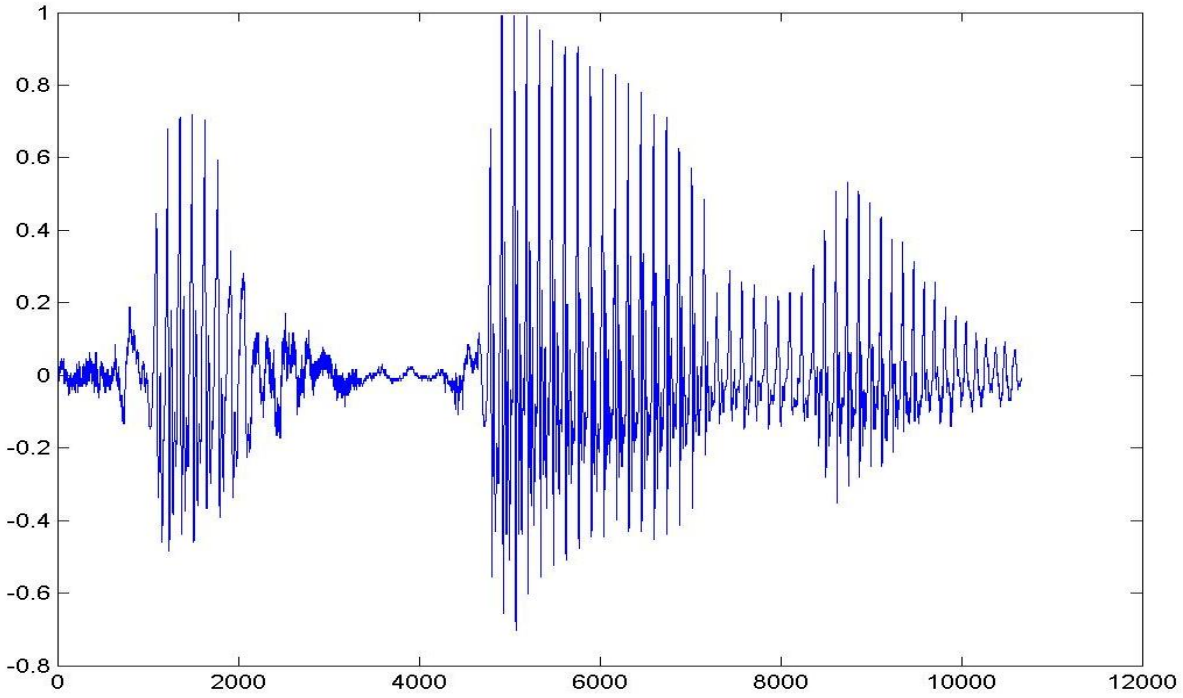
Supervisor: M. M. Bykov

Purpose: increasing the efficiency of the process of phonotypes parametric identification in the module of speech recognition in intelligent automation systems.

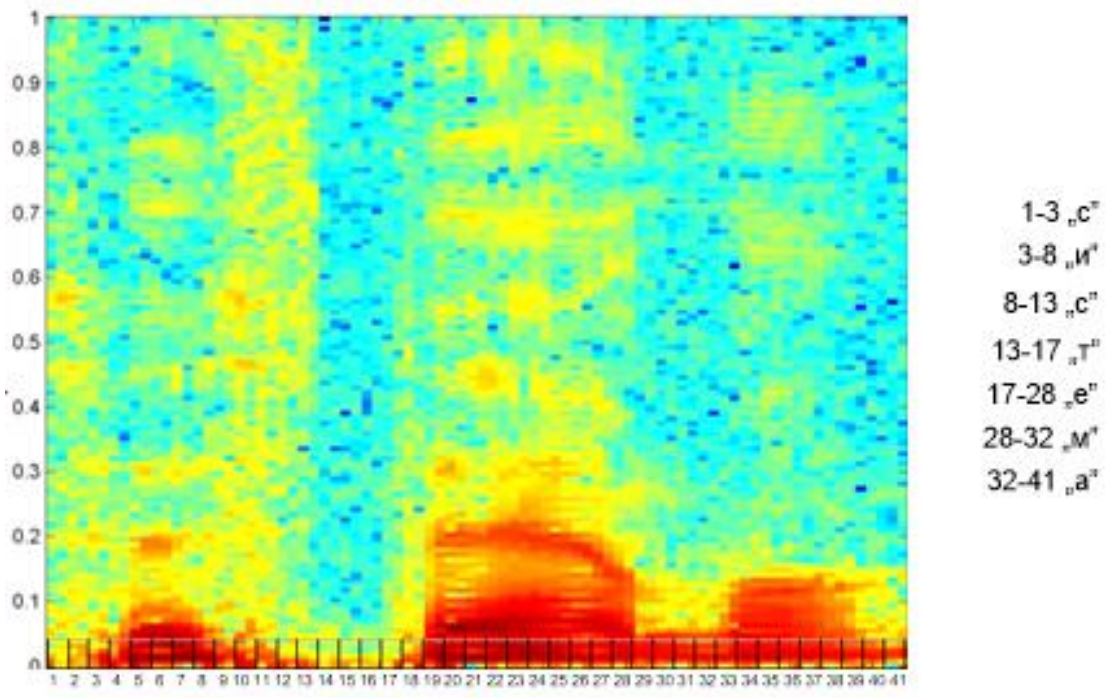
Ways to solve:

- development of a combined correlation-logical method for automation of the process of selection of informative features of speech patterns
- development of algorithms and programs for automation informative features selection
- development of a method of speaker-independent description of speech patterns on the basis of the model of "quasi-frequency modulator" and method of segmentation of continuous signal on the basis of model of "quasi-frequency" segmenter
- development of algorithms and programs for implementation of "quasi-

FEATURE SELECTION METHOD

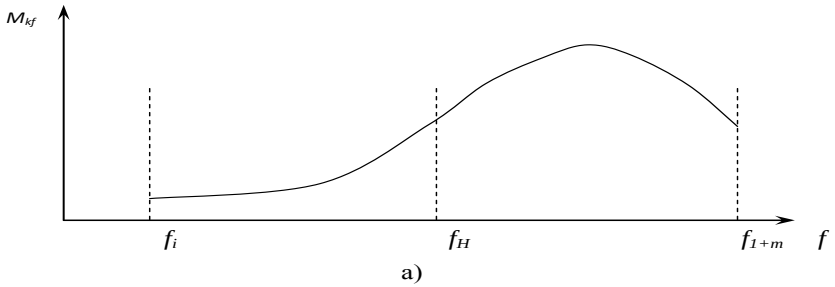


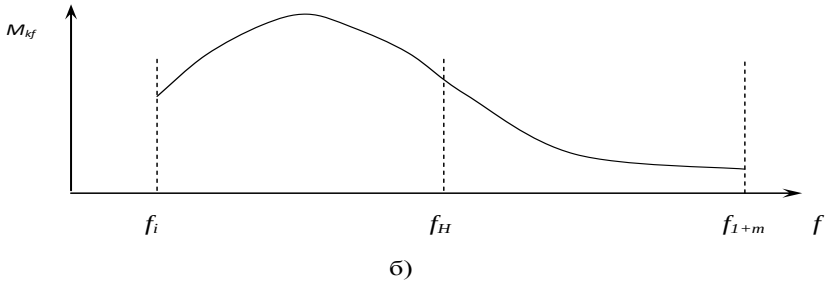
The signal of the word "система" as a function of time

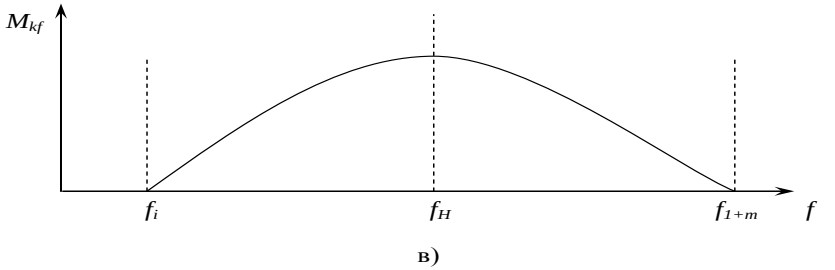


Spectrogram of speech signal

MODEL OF QUASI FREQUENCY MODULATION







a) high frequency torque value
b) low frequency torque value
c) average frequency torque value

State of signal frequency moments:

$$M_{kf} = \frac{\int_{F_{k-1}}^{F_k} A_f \cdot f df}{\int_{F_{k-1}}^{F_k} f df} \qquad M_{kf} = \frac{\sum_{i=1}^{1+m} A_i \cdot f_i}{\sum_{i=1}^{1+m} f_i}$$

Frequency detection function Q_g:

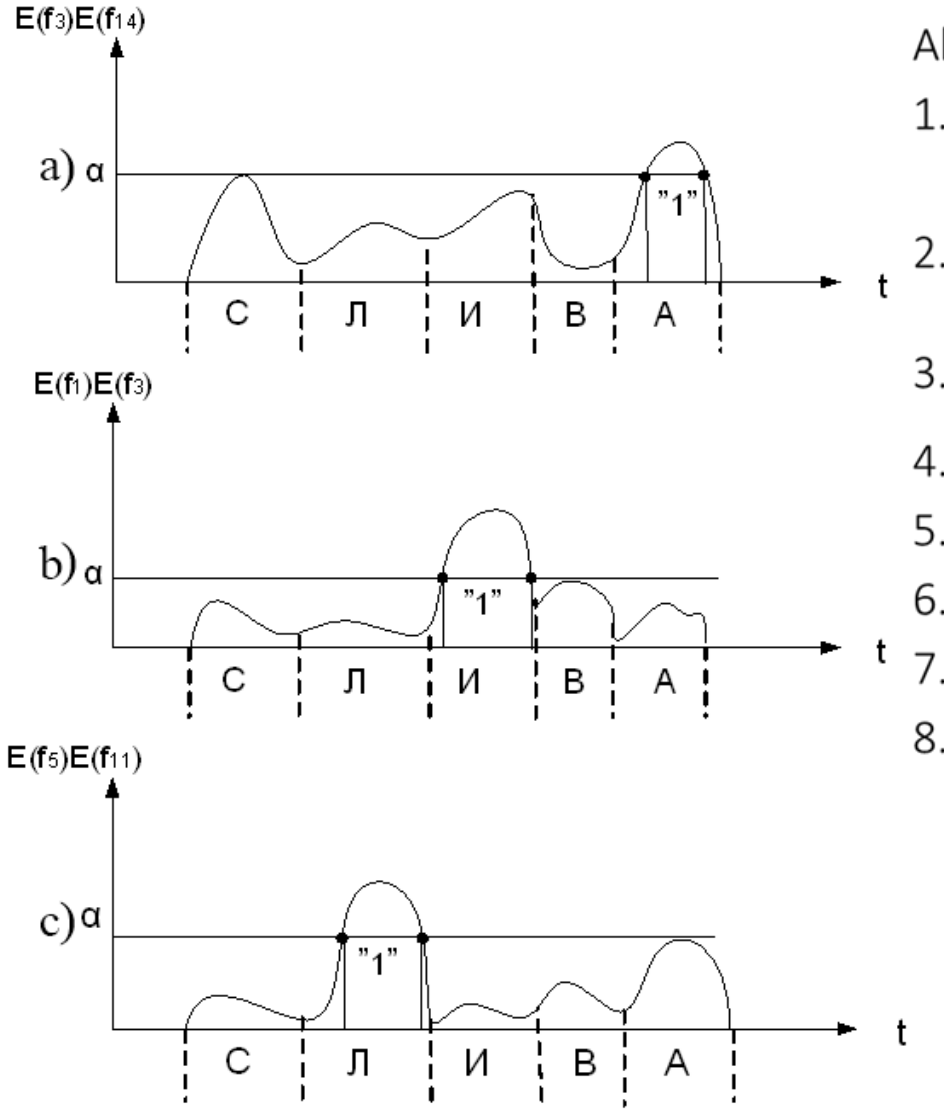
$$Q_g = Y_{i=1}^{2} \sigma\left(M_k^i \alpha M_k^{i+1}\right)$$

$$\sigma\left(M_k^i \alpha M_k^{i+1}\right) = 1, M_k^i > M_k^{i+1};\ 0, M_k^i \le M_k^{i+1}$$

The average frequencies $f_{НК}$ are determined by the value of the formant frequencies in the neutral position of the path:

$$f_{\text{НЖ}} = (2k-1) \cdot \frac{c}{4 \cdot 1_{\text{М}}}$$

COMBINED METHOD OF AUTOMATED FEATURE EXTRACTION



Algorithm:

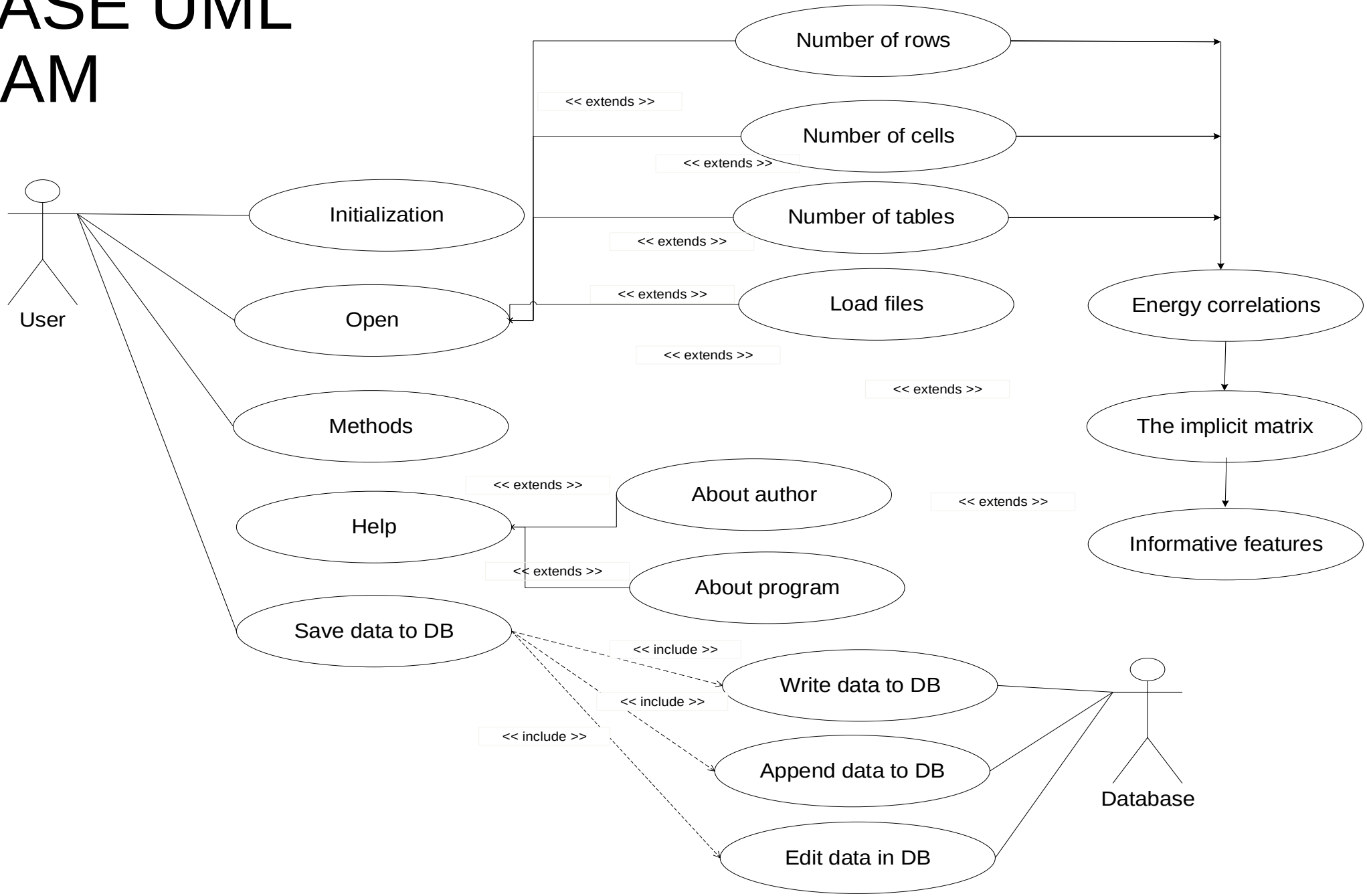
1. Formation of frequency ranges: their amount, frequencies center, bandwidth, filter type are selected.
2. Filtrating
3. Calculation of the energy $E_i(t) = \sqrt{\sum_{k=1}^{N} e_k^2}$
4. Correlation
5. Finding a-levels for each correlation and forming "1" intervals
6. Finding those "1" overlapping intervals
7. Formation of an implicit matrix for them
8. Formation by implicit matrix of ranges, energy which will be informative features (by conorm operation)

a) selection of logical values "1" for correlation of $E(f_3)E(f_{14})$

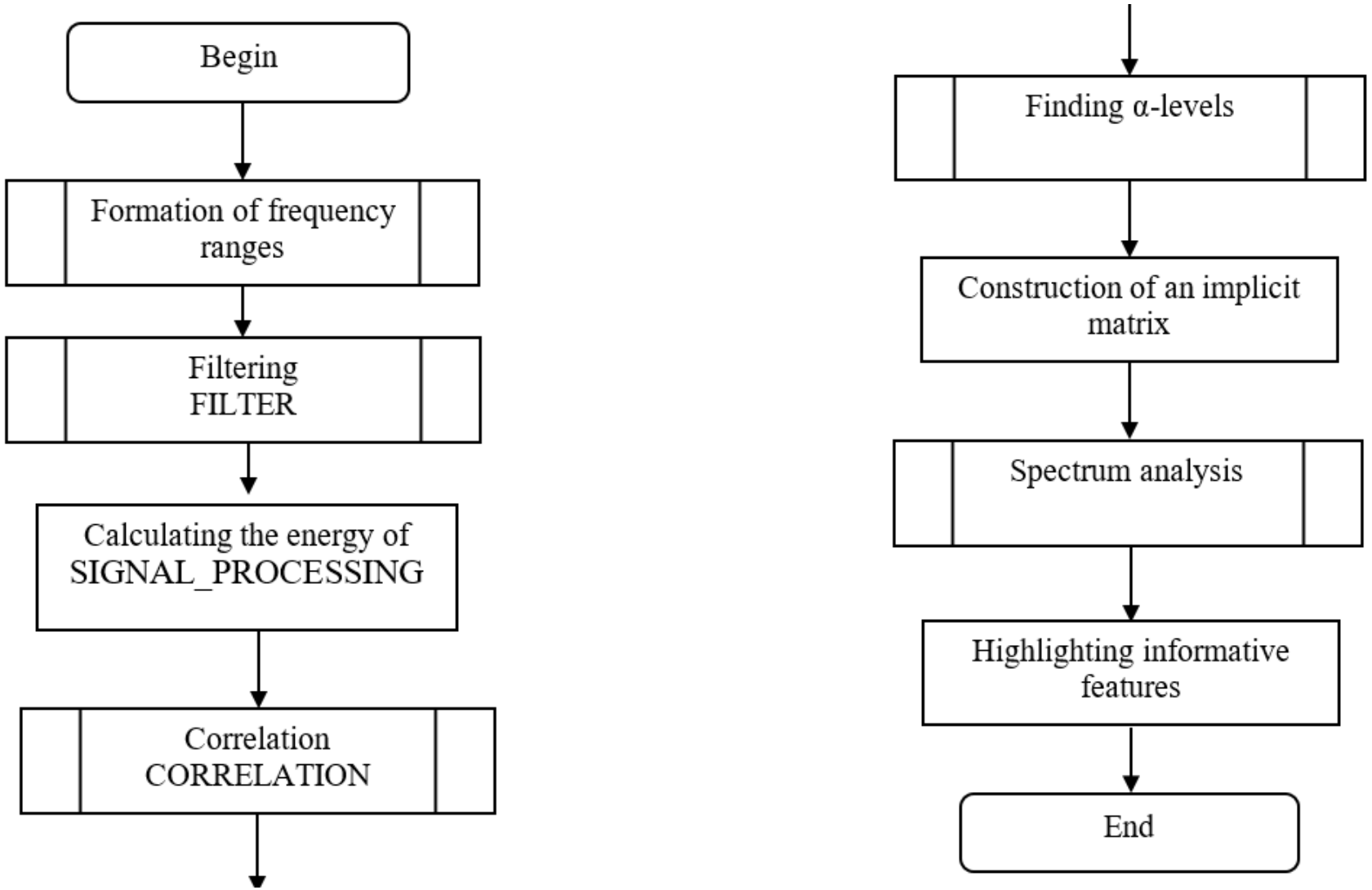
b) selection of logical values "1" for correlation of $E(f_1)E(f_3)$

c) selection of logical values "1" for correlation of $E(f_5)E(f_{11})$

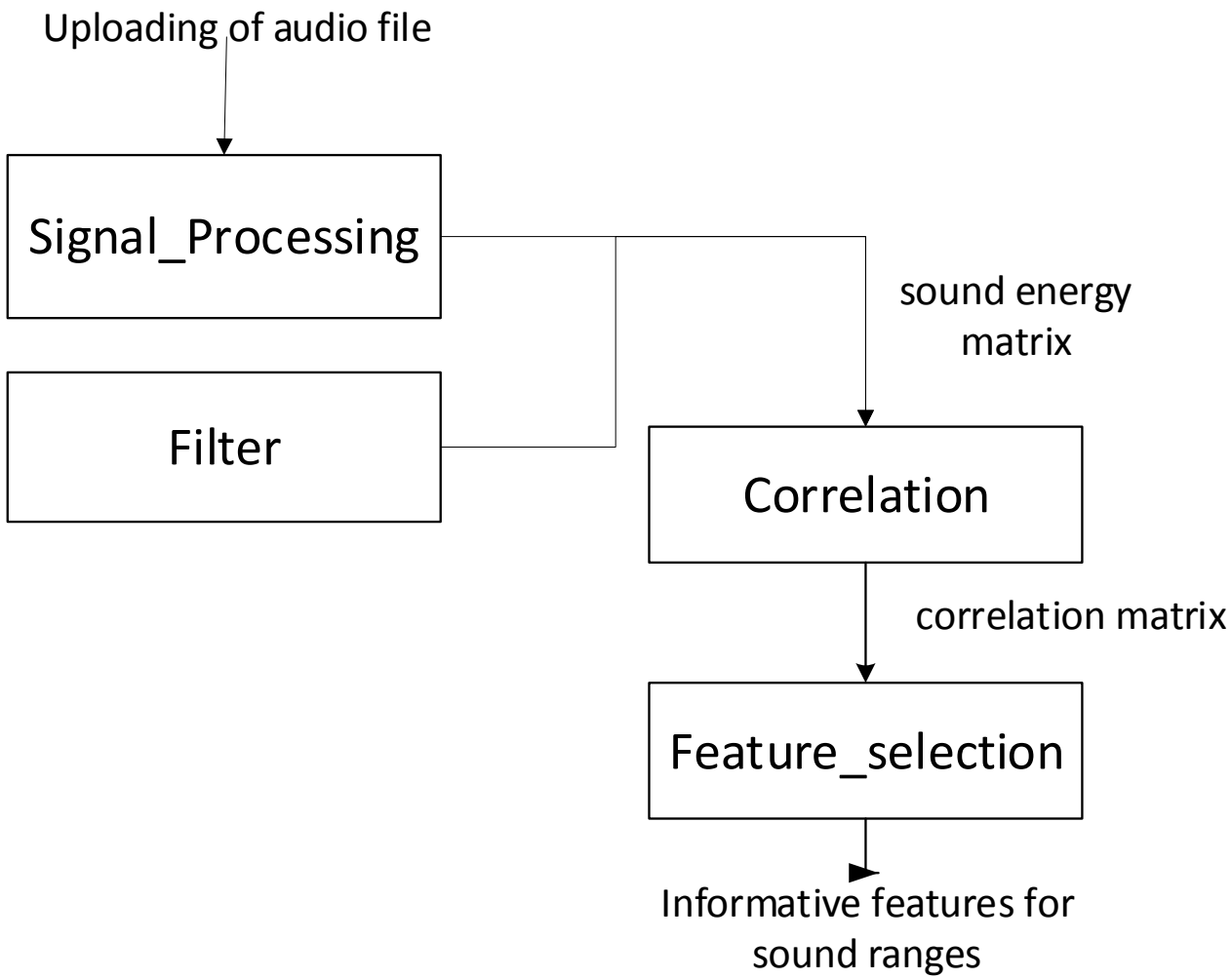
USE-CASE UML DIAGRAM



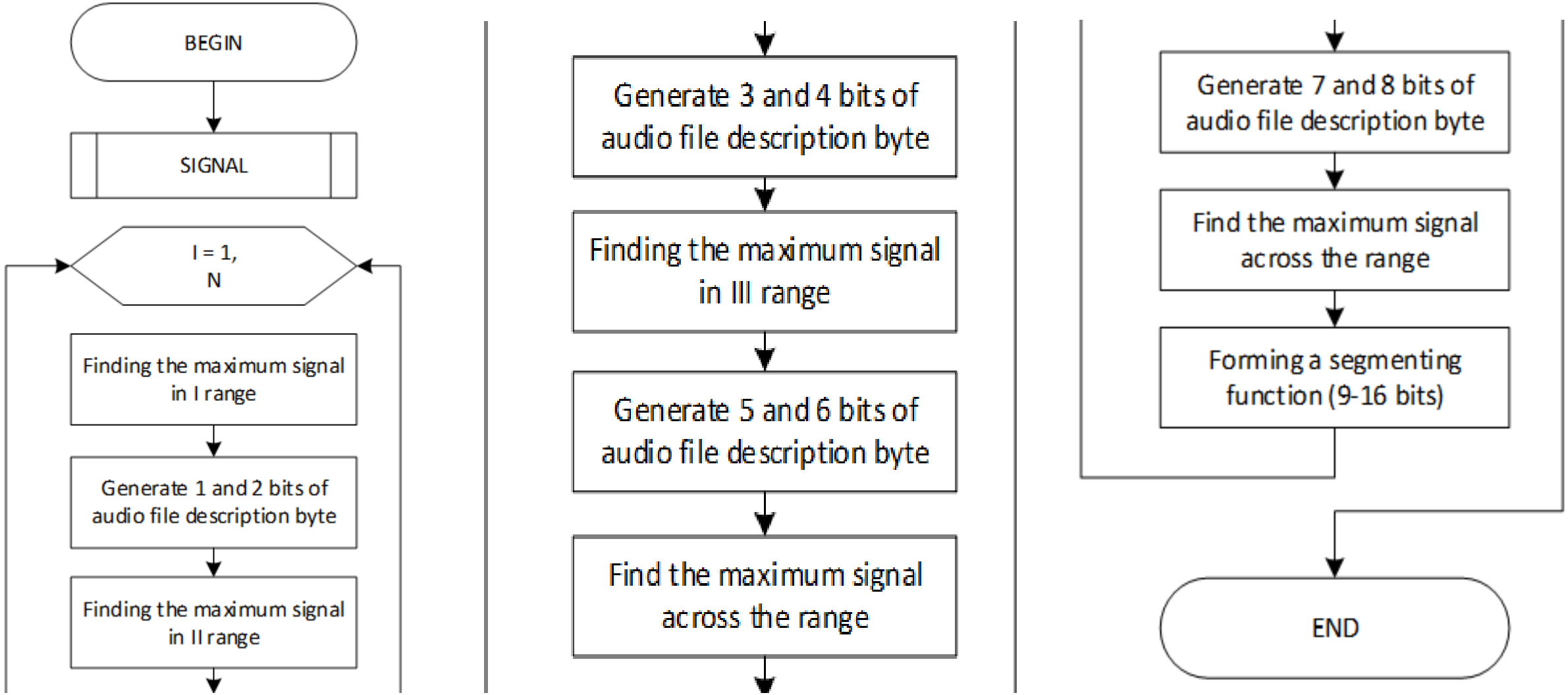
ALGORITHM OF INFORMATIVE FEATURES SELECTION BY THE FUZZY-LOGICAL METHOD



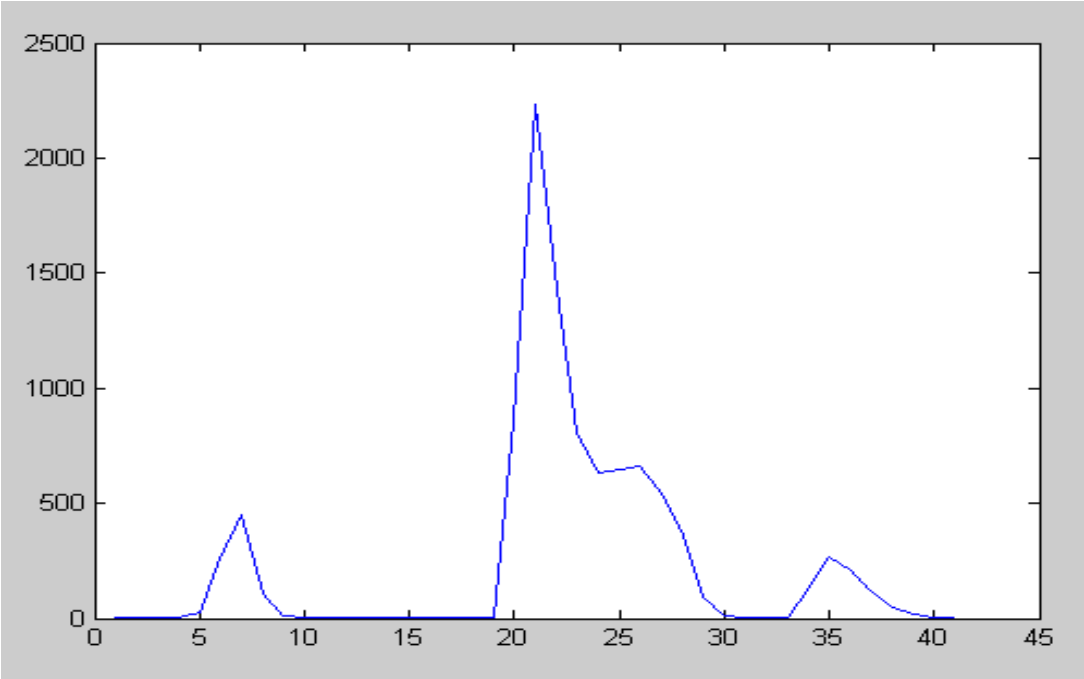
INTERACTION OF SOFTWARE MODULES



ALGORITHM OF FEATURES SELECTION BY QUASI-FREQUENCY METHOD



FUZZY-LOGICAL METHOD FOR HIGHLIGHTING FEATURES



Selection of logical values of "1" for correlation $E(f_1)$ $E(f_2)$

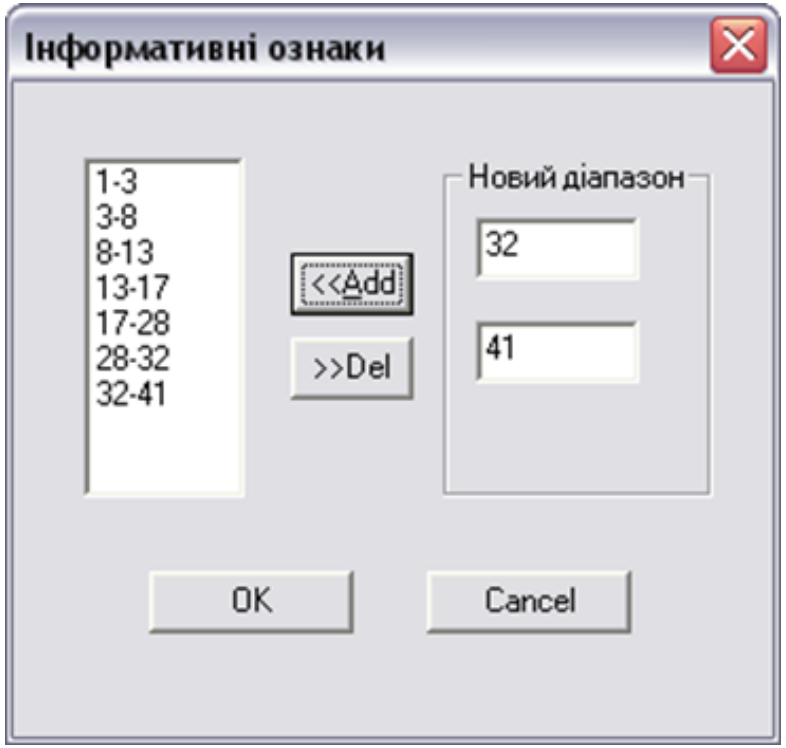
Table

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
E1E1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0
E2E1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
E3E1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
E4E1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
E5E1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
E6E1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
E7E1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
E8E1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
E9E1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
E10E1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
E11E1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
E12E1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
E13E1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
E14E1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
E15E1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
E1E2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0
E2E2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0
E3E2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
E4E2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
E5E2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
E6E2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
E7E2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

The implicit matrix

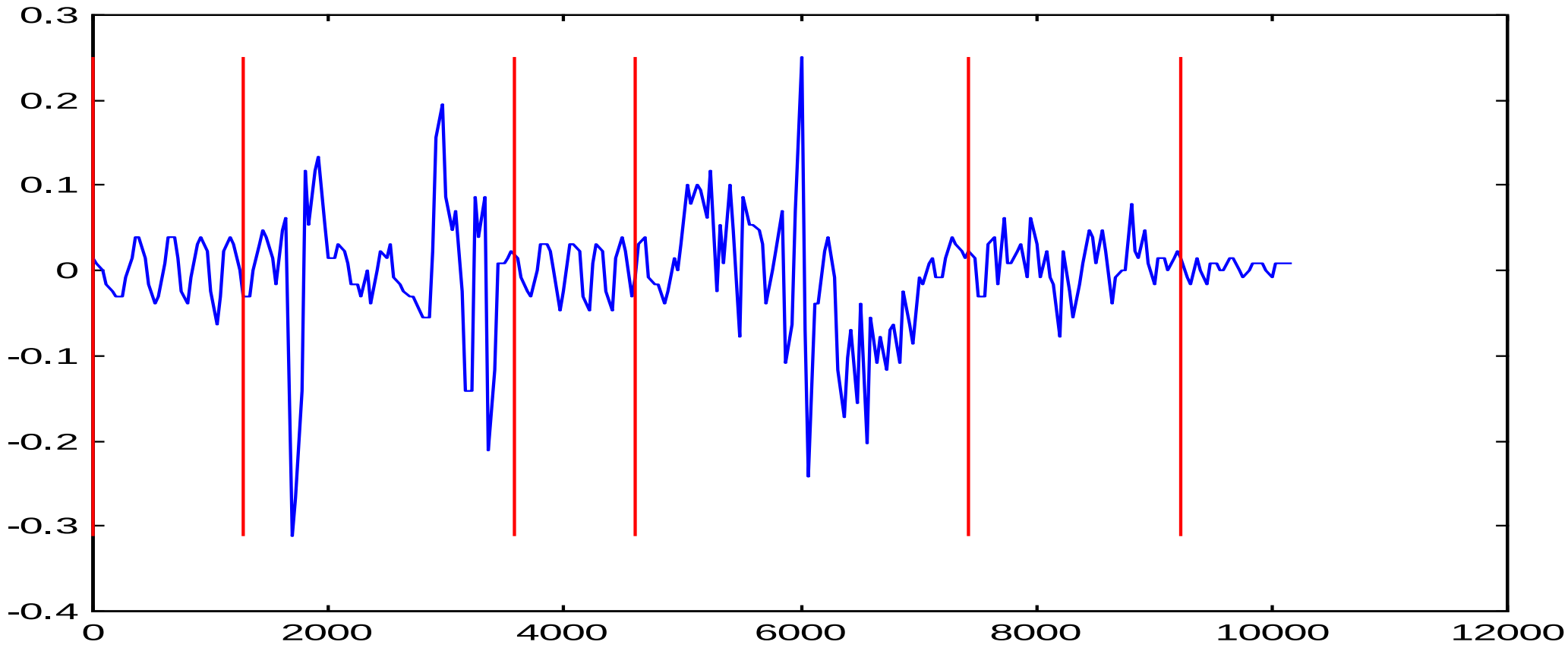
INFORMATIVE FEATURES OF SPEECH SOUNDS

Звукові інтер-вали	Інформативні ознаки мовних звуків
"с" 1-3	
"и" 3-8	
"с" 8-13	E11E15,E11E15,E13E14,E9E15,E10E15,E11E15,E13E15
"т" 13-17	
"е" 17-28	E1E13,E1E11,E1E13,E2E13,E1E11,E2E14,E1E13,E2E14,E3E11,E1E11,E2E12,E3E13,E1E12,E2E13,E3E14,E4E1
"м" 28-32	E1E6,E3E7,E4E6,E4E7
"а" 32-41	



Results of operation of the frequency-segmenting classifier

EXPERIMENTAL RESEARCH
Results of the frequency-segmenting classifier



Segmentation of word “*менше*”

CONCLUSIONS

Section 1

- analyzed the existing problems and ways of their solution
- conducted the feasibility study of the optimal solution to the problem

Section 2

- conducted theoretical researches of parametric description of speech signals
- analyzed fuzzy and logical methods
- developed a structural diagram of the recognition system and mathematical models of the speech signal
  - feature extraction method
  - speaker independent description of speech patterns

Section 3

- developed algorithms and software to extract informative features based on the proposed methods