

004.4:65.28

*The article is a study of methods and models that can analyze text. In an analysis of methods defined method that fully satisfies the task, but requires completion – is the theory of character recognition. The mathematical model and algorithm for checking the uniqueness of the text.*

**Keywords:** *method, algorithm, unique, text, pattern recognition.*

[1].

1.

2.

3.

[2].

[3].

( ) .

 $S$  –

$$S = \{k_1 i_1, k_2 i_2, \dots, k_n i_n\},$$

$k_1, k_2, \dots, k_n$        $i_1, i_2, \dots, i_n$

 $S_f$  –

$$S_f = \{k_{f1} i_1, k_{f2} i_2, \dots, k_{fn} i_n\},$$

$k_{f1}, k_{f2}, \dots, k_{fn}$        $i_1, i_2, \dots, i_n$

$$i_1, i_2, \dots, i_n \quad S, \quad S_f$$

$$\sum_{i=1}^n k_i = 1, n = |I|,$$

$$\sum_{i=1}^n k_{fi} = 1, n = |I_f|.$$

1):

$$\sum_{i=1}^n k_i = n,$$

$$\sum_{i=1}^n \frac{k_i}{n} = 1.$$

 $k_{fi}$ . $\tilde{S}_i$ 

$$\tilde{S}_i = \frac{k_{i\min}}{k_{i\max}} \cdot k_i,$$

$$k_{i\min} = k_i, k_{i\max} = k_{fi} \Leftrightarrow k_i < k_{fi},$$

$$k_{i \min} = k_{f_i}, k_{i \max} = k_i \Leftrightarrow k_{f_i} < k_i.$$

( )  
 $|k_i - k_{f_i}|$   
 $k_{f_i} - k_i$   
 $k_{i \min}, k_i, k_i < k_{f_i}$   
 $k_{f_i}, k_{f_i} < k_i$   
 $\tilde{S}_i$  0 1.

$\tilde{S}_i$ :

$$\tilde{S} = \sum_{i=1}^n \tilde{S}_i.$$

S 0 1.

$S, S_{f_i}$

$i,$   
 $i \in S, i \in S_f,$

$S -$

$S_f -$

$F -$

$F_f -$

( ).

1)  $i$   
 $F, F_f).$

$i \in S_f$

$$A = D(F, i, [r^-, r^+]),$$

$$r^- = r;$$

$$r^+ = r.$$

$$A = D(F, i, [r, r]).$$

A

I:

$$A \rightarrow I,$$

$$I = \{i_1, i_2, \dots, i_n\}.$$

$$A = D(F_f, i, [r, r]),$$

$$A \rightarrow I_f,$$

$$I_f = \{i_1, i_2, \dots, i_n\}.$$

$$I \cap I_f$$

2)

y:

$$y = \frac{|I \cap I_f| \cdot 2}{|I| + |I_f|},$$

, y

0 1: 1 -

0 -

$$\check{S} = \sum_{i=1}^n \check{S}_i \cdot y_i.$$

$$\check{S} > 0.$$

$$A = \{a_1, a_2, a_3, \dots, a_n\};$$

$$B = \{b_1, b_2, b_3, \dots, b_n\},$$

A -

 $a_i -$ 

B -

 $b_i -$ 

5

$$l_i = \{a_i, a_{i+1}, a_{i+2}, a_{i+3}, a_{i+4}\};$$

$$p_i = \{b_i, b_{i+1}, b_{i+2}, b_{i+3}, b_{i+4}\},$$

 $l_i -$  $p_i -$  $(v_i).$  $(h_i)$   
 $h_i = v_i$  $l_i \cdot p_i$ 

$$f = 0.$$

$$q = \left(1 - \frac{f}{n}\right) \cdot 100\%$$

- q –  
n –
4.  
1)  $M(I, R)$ .  
2)  $d(i)$ .  
3)  $d(M(I, R))_{max}$   
( ),  $n$  ( $n$   $S_p$ .  
4)  $S_s$ ,  $S_p$ .  
5)  $S_p$ .  
6)  $S$   $S_f$ .  
7)  $S$ .  
8)  $S$ .  
9)  $S$ .  
10)  $S$ .  
11)  $S$ .  
12)  $h_i, v_i$ .  
13)  $S$ .

1. Joofaq [ ] . – : <http://joofaq.ru>  
2. , . / . , . : – ., 1978.  
3. Gotai [ ] . – : <http://gotai.net>

20.01.2013 .

« -  
»  
12-51