

ხელნაწერი სიმბოლოების შედარების სიმძიმის ცენტრის მეთოდის ალგორითმი

იოსებ ქართველიშვილი, თეა თოდუა
საქართველოს ტექნიკური უნივერსიტეტი

რეზიუმე

ნაშრომში წარმოდგენილია ხელნაწერი სიმბოლოების შედარების ალგორითმი სიმძიმის ცენტრის მეთოდით. თითოეული ალგორითმული ბლოკი გათვალისწინებულია ავტომატიზებული ამოცნობის გარკვეული ფუნქციის შესასრულებლად. ალგორითმული ბლოკები დეტალიზებულია ისეთი სახით, რომ შემდგომში შესაძლებელი და გაადვილებული იყოს დაპროგრამების პროცესი.

საკვანძო სიტყვები: ხელნაწერი სიმბოლოები, ავტომატიზებული ამოცნობა, სიმძიმის ცენტრი.

1. შესავალი

ბოლო პერიოდში მნიშვნელოვანი პროგრესი შეიმჩნევა ნაბეჭდი ტექსტების ამოცნობის სფეროში. შემუშავებულია ალგორითმები და მათი შესაბამისი პროგრამული პროდუქტები, რომლებიც საშუალებას იძლევა თითქმის 100%-იანი სიზუსტით იქნეს ამოცნობილი ნაბეჭდი ტექსტები. ნაბეჭდი ტექსტების ამომცნობი კომპიუტერული სისტემა შექმნილია მსოფლიოს 186 ენისათვის. სულ სხვაგვარადაა საქმე ხელნაწერი ტექსტების ამომცნობი პროგრამებისა და ალგორითმების შემთხვევაში. ხელნაწერი ტექსტების ამოსაცნობი პროგრამული პროდუქტები ძირითადად განკუთვნილია ხელით შევსებული სპეციალური ფორმების ან ანკეტების ტექსტების ამოსაცნობად, ასევე სხვა სახის ვიწრო, სპეციალიზებული ამოცანების გადასაწყვეტად.

წინამდებარე ნაშრომში, ავტომატიზებული ამომცნობი სისტემების თეორიის საფუძველზე, წარმოდგენილია ხელნაწერი სიმბოლოების ავტომატიზებული ამოცნობის ალგორითმები, განსაზღვრულია და თითოეული ალგორითმისთვის დადგენილია შემაჯავლი და გამომავალი მონაცემები, ჩამოყალიბებულია ორიგინალი და უცნობი სიმბოლოების მონაცემთა მასივები და ბაზები, რომელთა საფუძველზეც შემუშავებული ალგორითმის მიხედვით განხორციელებულია ამოცნობა.

თითოეული ალგორითმული ბლოკი გათვალისწინებულია ავტომატიზებული ამოცნობის გარკვეული ფუნქციის შესასრულებლად. ამომცნობი სისტემის ალგორითმიზაციის ეტაპზე საჭიროა თითოეული ალგორითმული ბლოკის ისეთი სახით დეტალიზაცია, რომ შესაძლებელი და გაადვილებული იყოს დაპროგრამების პროცესი.

2. ძირითადი ნაწილი

ნორმალიზებული სიმბოლოების ბაზების საფუძველზე ხდება ხელნაწერი სიმბოლოების შედარების ალგორითმის შემუშავება სიმძიმის ცენტრის მეთოდის მიხედვით, რომელიც თავისი

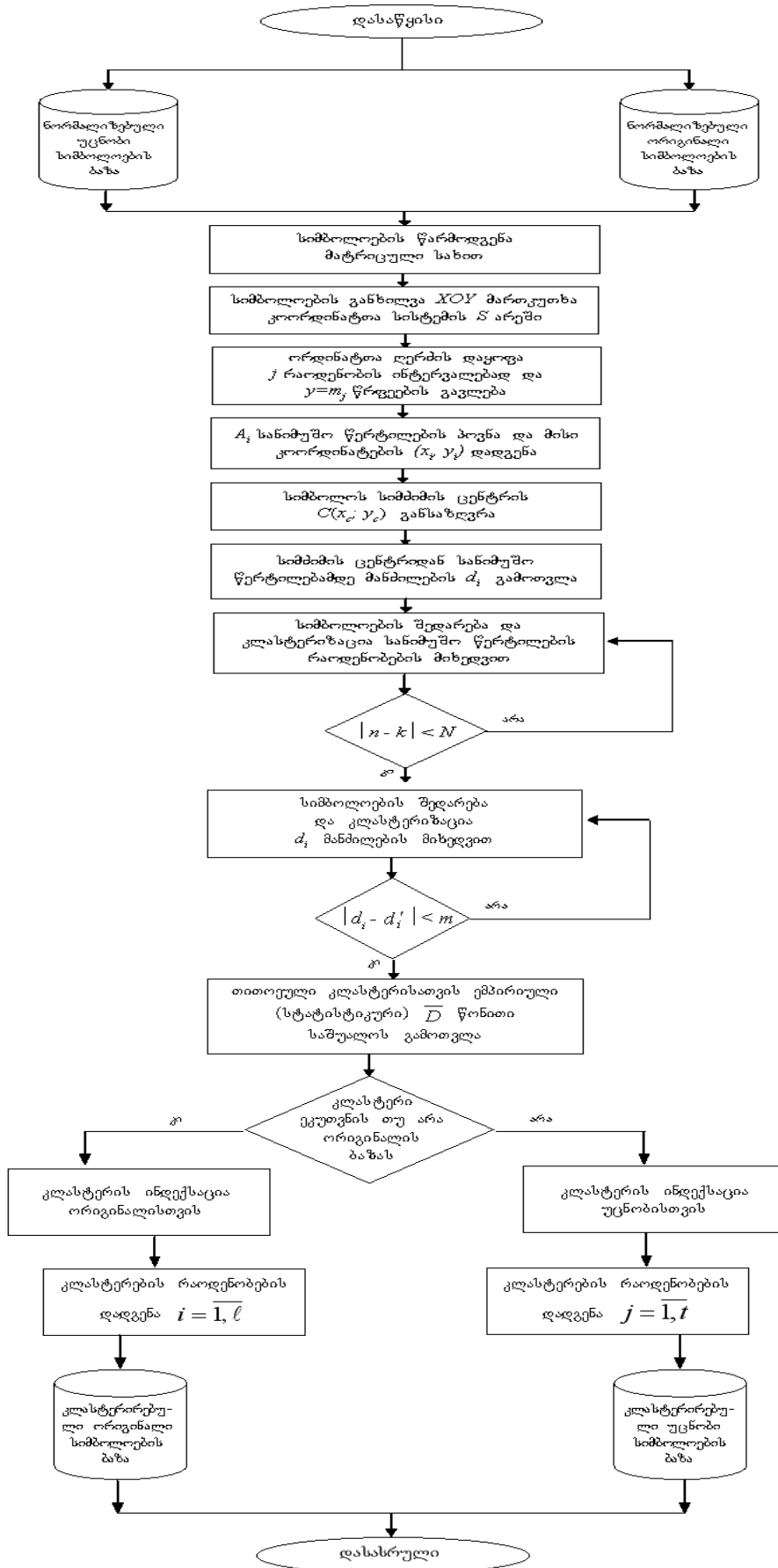
ფუნქციური დანიშნულებით, ავტომატიზებული ამომცნობი სისტემის ერთ-ერთ მთავარ შემაღვენელ ნაწილს წარმოადგენს. მისი საშუალებით, ალგორითმით გათვალისწინებული ხელნაწერი სიმბოლოების შედარების პროცესის გამოყენებით, ხდება ორიგინალი და უცნობი ნორმალიზებული სიმბოლოების კლასტერიზაცია დამახასიათებელი ნიშნების მიხედვით და ხორციელდება კლასტერირებული ორიგინალი და უცნობი სიმბოლოების ბაზების ფორმირება, რომლებიც მთავარ როლს თამაშობენ სიმბოლოების ამოცნობის პროცედურებში.

განვიხილოთ სიმბოლოების შედარების ალგორითმი სიმძიმის ცენტრის მეთოდით (ნახ.1).

ალგორითმის საწყის ეტაპზე ხდება ორიგინალი და უცნობი მონაცემების ბაზების ნორმალიზებული სიმბოლოების წარმოდგენა მატრიცული სახით. შემდეგ ეტაპზე ორივე ბაზის სიმბოლოები განიხილება XOY კოორდინატთა სისტემის S მართკუთხა არეში, რომელიც შექმნილია ალგორითმით მოთხოვნილი პირობების გათვალისწინებით. შემდეგ ბლოკში ორდინატთა ღერძი იყოფა j რაოდენობის ინტერვალებად S არის ფარგლებში და ხდება OX ღერძის პარალელური $y=m_j$ წრფეების გაკლება S მართკუთხა არის გვერდის გადაკვეთამდე. შემდეგ ეტაპზე მოიძებნება A_i ეტალონური (სანიმუშო) წერტილები, რომლებიც მიიღება პარალელური წრფეების სიმბოლოს გამოსახულებასთან (მრუდთან) გადაკვეთის შედეგად და დგინდება მათი შესაბამისი კოორდინატები. ეტალონური წერტილების კოორდინატების საშუალებით, მოიძებნება გამოსახულების სიმძიმის ცენტრი და განისაზღვრება მისი კოორდინატები.

სიმძიმის ცენტრის მოძებნისა და მისი კოორდინატების განსაზღვრის შემდეგ ხდება სიმბოლოს დამახასიათებელი სიდიდეების, სიმძიმის ცენტრიდან ეტალონურ წერტილებამდე d_i მანძილების გამოთვლა ეკვიდეს მანძილის ფორმულის საშუალებით და სიმბოლოს ფორმირება d_i მანძილების მიხედვით. შემდგომ ეტაპზე ხდება სიმბოლოების შედარება და კლასტერიზაცია ეტალონური წერტილების რაოდენობების მიხედვით. მოწმდება პირობა, მოცემული და შესადარებელი წერტილების რაოდენობების სხვაობის აბსოლუტური სიდიდე აღემატება თუ არა ემპირიულად დადგენილ ზღვრულ სიდიდეს, თუ ეს სხვაობა ნაკლებია ზღვრულ სიდიდეს, მაშინ ხდება სიმბოლოების შედარება და კლასტერიზაცია d_i მანძილების მიხედვით. წინააღმდეგ შემთხვევაში, სიმბოლოების ეტალონური წერტილების d_i მანძილების საშუალებით თითოეული კლასტერიზაციის ხდება ემპირიული (სტატისტიკური) \bar{D} წონითი საშუალოს გამოთვლა.

ალგორითმის საბოლოო ეტაპზე მიმდინარეობს კლასტერების დაჯგუფება ორიგინალი და უცნობი სიმბოლოების მიხედვით, თითოეულ ჯგუფში ხდება კლასტერების ინდექსაცია და დგინდება მათი რაოდენობები, რის შედეგადაც კლასტერების მიხედვით ხორციელდება ორიგინალი და უცნობი სიმბოლოების ბაზების ფორმირება.



ნახ.1. სიმბოლოების შედარების ალგორითმი სიმბიმის ცენტრის მეთოდით

3. დასკვნა

ჩამოყალიბებულია სიმბოლოების შედარების ალგორითმი სიმბოლის ცენტრის მეთოდით, თითოეული ალგორითმული ბლოკი წარმოდგენილია და დეტალიზებულია ისეთი სახით, რომ შემდგომში გაადვილებული იყოს სისტემის პრაქტიკული რეალიზაცია.

ლიტერატურა:

1. ვერულავა თ. ხუროძე რ. ამომცნობი სისტემების თეორიის საფუძვლები. სტუ. თბილისი, 2001
2. ქართველიშვილი ი. ორიგინალი და უცნობი ხელნაწერი სიმბოლოების შედარების კომპიუტერული ანალიზი. სტუ, შრ.კრ. №4 (432), თბილისი, 2000.

ALGORITHM OF A METHOD OF THE CENTER OF GRAVITY FOR COMPARISON OF HAND-WRITTEN SYMBOLS

Kartvelishvili Ioseb, Todua Tea
Georgian Technical University

Summary

In article the algorithm of comparison of hand-written symbols through the method of the center of gravity is being presented. Each algorithmic block shall perform the definite function of the automated recognition. Algorithmic blocks are detailed to allow and even ease the programming process.

АЛГОРИТМ МЕТОДА ЦЕНТРА ТЯЖЕСТИ ДЛЯ СРАВНЕНИЯ РУКОПИСНЫХ СИМВОЛОВ

Картвелишвили И., Тодуа Т.
Грузинский Технический Университет

Резюме

Представлен алгоритм сравнения рукописных символов с помощью метода центра тяжести. Каждый алгоритмический блок предназначен для выполнения некоторых функций автоматизированного распознавания. Алгоритмические блоки детализованы таким образом, что впоследствии делает возможным и упрощает процесс программирования.