

## НОВ МЕТОД ЗА ИЗБОР НА ЕТАЛОННИ ТОЧКИ

Димитър С. Тянев, Юлка П. Петкова

**Резюме:** Проблемът за търсене на шаблон в изображение е един от съществените в съвременната индустрия. В много случаи този проблем се решава при изисквания за минимално време на обработка. Един от основните подходи за минимизиране на времето за обработка е намаляване броя на точките, представящи изображението на шаблона.

Тук се предлага един нов метод за избор на точките, представящи шаблона. Методът е разработен въз основа на пространствената интерпретация на изображението и идеята за приложение върху него на критерия за D-оптималност на избраното подмножество точки от шаблона. При така формулираните условия става възможно прилагането на облекчен алгоритъм за избор на точки. За целта е дефинирана двупрагова функция за избор, наречена еквипотенциална. При подходящ избор на праговете нива на тази функция може да се регулира броя на точките в търсеното подмножество.

За да се облекчи прилагането на функцията за избор, изображението се подлага на предварителна обработка. Същността на тази обработка е определена според горните условия, като при това отчита промишлената специфика на обработваните изображения. За целта е синтезирана интервална преобразуваща функция, целта на която е да засили пространствеността на изображението, като усилва нелинейно неговите локални екстремни области.

**Ключови думи:** Сравнение с еталон (търсене на шаблон в изображение), Еквипотенциални повърхнини, D-оптималност, Нормирана взаимна корелация.

### Увод

В съвременната индустрия автоматизацията на технологичните процеси има водеща роля. Те са напълно или в много висока степен автоматизирани. Това става възможно с развитието на визуалните системи за управление и контрол. Един проблем, поставен и разискван още през 60-те години на миналия век, лежи в основата на развитието на автоматизацията, а именно "търсене на шаблон в изображение" или сравнение с еталон [3,4,6,8,9,11]. Голямо разнообразие от приложения се нуждаят от някакво решение на споменатия проблем. Някои от тях са:

- регистрация на маркери (знаци) върху печатни материали, ценни книжа, банкноти и др.;
- промишлени работи в електрониката, както и промишлени манипулатори;
- бързо претърсване на сателитни снимки с голям обем за отличителни белези;
- бързо претърсване по признак във все по-големи бази от данни.

Проблемът на търсенето се свежда до избор или синтез на критерий за подобие и изчисляване на неговите оценки и в тази връзка бързо се появяват различни алгоритми за търсене. Търсенето често е поставено в утежняващи го условия, когато изображението е подложено на смущаващи го въздействия като линейни или нелинейни изкривявания, измествания, завъртания, мащабиране и др. Същинският проблем обаче идва от това, че в общия случай трябва да се търси не малък шаблон в доста голямо изображение, което означава, че цифровите данни, които трябва да се обработят, стават твърде големи по обем и много алгоритми се оказват неефективни на практика, поради големите времена, необходими за извършване на изчисленията.

Така възниква нов проблем – бързодействието на алгоритмите. Един от основните подходи към повишаване на скоростта на изчисленията е подходящото намаляване на обема на съществената информация [5]. Целта на настоящата разработка е да предложи нов метод за такава редукция на данните от шаблона, която, запазвайки степента на надеждност, да довежда до съществено намаляване на времето за търсене.

## I. Постановка на проблема

Разглеждаме задачата за откриване на еталонен фрагмент /шаблон/ в черно-бяло изображение с помощта на коефициента на нормираната взаимна корелация, използван в качеството му на мярка за оценка степента на сходство [1]:

$$(1) \quad NC = \frac{\sum_x \sum_y (Img - \overline{Img}) \cdot (Tmp - \overline{Tmp})}{\sqrt{\sum_x \sum_y (Img - \overline{Img})^2 \cdot \sum_x \sum_y (Tmp - \overline{Tmp})^2}}$$

където:

- с  $Img$  е означен цветът на текущия  $(x,y)$  пиксел от текущата рамка. Графичната координатна система  $(x,y)$  на изображението е ортогонална, с начало в горния ляв ъгъл и положителни посоки съответно: на абсцисата отляво надясно и на ординатата отгоре надолу;
- $\overline{Img} = \frac{1}{N} \sum_x \sum_y Img$  – представлява средната стойност на цветовете в текущата рамка;
- с  $Tmp$  е означен цветът на текущия пиксел от шаблона;
- $\overline{Tmp} = \frac{1}{N} \sum_x \sum_y Tmp$  – представлява средната стойност на цветовете в шаблона.

Допълнително ще уговорим, че под цвят  $Img$  или  $Tmp$  в случая следва да се разбира стойността на интензивността, варираща в интервала  $[0 \div 255]$  на сивата (полутоновата) скала.

Разпознаващото правило, с помощта на което даден фрагмент от едно изображение може да бъде приет за сходен с еталона, при предварително избрана степен на сходство  $p_e$ ,  $p_e \in (0,1)$ , се основава на неравенството:

$$(2) \quad NC > p_e .$$

Изчислителната сложност на алгоритъма за изчисляване на корелационния коефициент (1) при транслирани изображения има следната оценка [1,4]:

$$(3) \quad K = O(n^4) ,$$

където с  $n$  е означен броят на обработваните пиксели.

Може да се каже, че този класически и много надежден критерий за оценка на сходството е доста трудоемък, ето защо в условията на високите скоростни изисквания е актуална задачата за търсене на възможности, позволяващи да бъде намален броят на изчислителните операции, без при това да се отказваме от самия критерий. Разбира се, избраният тук критерий за подобие не е задължителен, а само необходим за провеждане на изследването. Последното може да бъде проведено при използването и на други критерии по преценка.

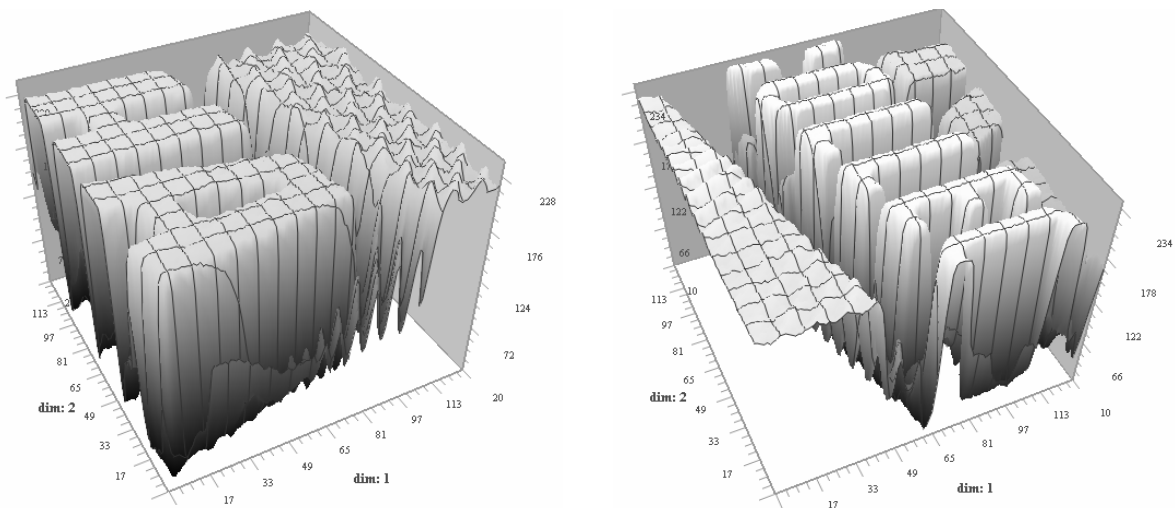
## II. Същност на еквипотенциалния избор

Изхождайки от споменатото по-горе разбиране, естествено достигаме до извода, че броят на изчислителните операции, необходими за изчисляване на оценката (1), може да бъде намален, ако бъде намален броят на пикселите, представящи изображението. Ето защо тук ще изложим и обосновем едно ново за целта решение, както и ще коментираме получените при неговото приложение резултати.

Използвайки стойността на сивия цвят във всяка точка на изображението като негова трета координата, ние можем да представим пространствения му вид. По-долу на фигура 1 и на фигура 3 са представени примерните черно-бели изображения, а на фигура 2 и на фигура 4 – техните примерни интерпретации.



Фиг. 1 Изображение №1 и изображение №2



Фиг. 2 Пространствен вид на изображение №1 и на изображение №2

В светлината на това представяне на изображението, формула (1) може да се интерпретира като своеобразна оценка на средното евклидово разстояние. Оценката на разстоянието на едно изображение до въведения еталон в този смисъл позволява същото да бъде припознато като такова или отхвърлено като различно.

Разделянето на две множества в тяхното пространство с равнина се основава на принципа на минималното разстояние до еталона. Когато класифицирането или разпознаването се основава на този принцип, при който се изхожда от разбирането за възможно най-пълно съвпадение на изображението с еталона, то този подход се обобщава като корелационен [2]. Ето защо ние считаме, че пространствената интерпретация на изображението е подходяща основа за прилагане върху нея на метода за формиране на D-оптимизирани подмножества [13], като подмножества, минимизиращи изчисленията, свързани с оценката на корелацията. С други думи тук се предлага еталонното изображение да бъде представено не чрез пълното си множество пиксели, а чрез подходящо избрани такива, които при това ще запазят и представят в достатъчно висока степен на съответствие пространствената му структура.

Критерият за D-оптималност се изразява формално така:

$$(4) \quad \det(V^t \cdot V)^{-1} \Rightarrow \min ,$$

където с  $V$  е означена матрицата на избраното подмножество, наричано още обучаваща извадка.

От [13] е известно още, че всяка процедура за линейно разделяне, която се основава на критерия за D-оптималност на избраното подмножество, получава най-добрите линейни оценки за коефициентите на разделящата функция. Последното означава такъв избор на елементи в избраното подмножество, при който по най-добър начин се представя пространствената структура на изходното изображение. Анализът на така формираните подмножества показва, че пространственото разположение на елементите им формира изпъкналата периферна обвивка на структурата на обекта.

Като изхождаме от тези окончателни изводи, с цел намаляване обема на изчисленията при получаване на корелационния коефициент (1) чрез минимизиране на елементите в наблюдаваното множество от точки в еталона на изображението, тук, в условията на тримерното пространство, предлагаме подхода на екипотенциалните равнини. При подходящ избор на праговете нива на потенциала (ниско и високо), т.е. на стойността на сивия цвят по третата координатна ос, чрез тези равнини могат да бъдат избирани елементите за търсеното подмножество от точки, които очевидно ще бъдат в пълно съответствие със същността на критерия за D-оптималност, тъй като ще представят пространствената структура на изходното изображение по-добре от всички останали.

За избор на елементи от даденото множество дефинираме следната потенциална функция:

(5) **if**

$$\left( \left[ \begin{array}{l} ( PixelColor(i) \leq Potential_{max} ) \cap \\ ( PixelColor(i+1) > Potential_{max} ) \end{array} \right] \cup \left[ \begin{array}{l} ( PixelColor(i) \geq Potential_{max} ) \cap \\ ( PixelColor(i+1) < Potential_{max} ) \end{array} \right] \right) \cup \\ \cup \left( \left[ \begin{array}{l} ( PixelColor(i) \leq Potential_{min} ) \cap \\ ( PixelColor(i+1) > Potential_{min} ) \end{array} \right] \cup \left[ \begin{array}{l} ( PixelColor(i) \geq Potential_{min} ) \cap \\ ( PixelColor(i+1) < Potential_{min} ) \end{array} \right] \right) = True$$

**then**

$$PixelColor(i) \in \mathfrak{S} ,$$

**end if ;**

където: с  $i$  е означен поредният елемент от изходното множество, чийто цвят  $PixelColor(i)$  се причислява като пореден елемент на новоформиращото се подмножество  $\mathfrak{S}$ ; В същност на практика, фактът, че този елемент е включен в  $\mathfrak{S}$  се състои в запомняне на неговите графични координати в системата  $(x,y)$ .

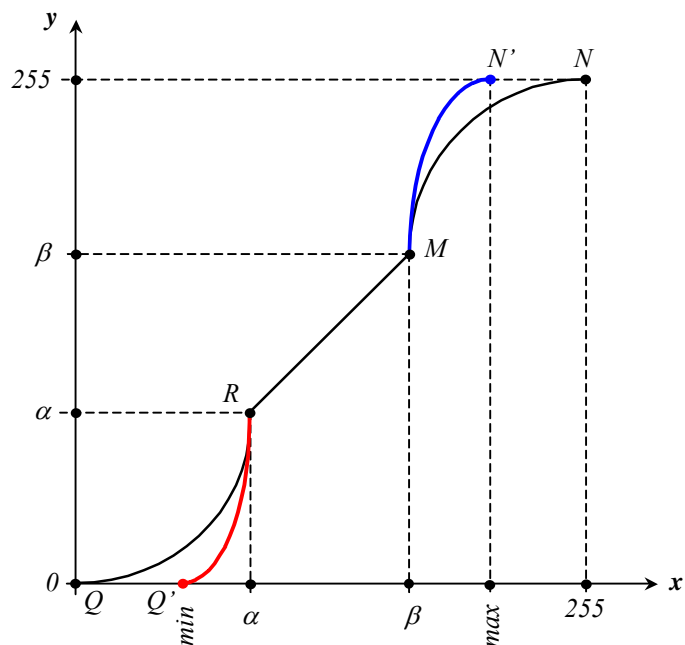
с  $Potential_{min}$  и  $Potential_{max}$  са означени двете прагови потенциални нива (ниско и високо) за избор. Тези две нива се определят адаптивно в програмната реализация на метода така, че да водят до избор на не повече точки от предварително зададения брой. Определянето им става като високото ниво  $Potential_{max}$  се спуска от стойността 255 надолу, а ниското ниво  $Potential_{min}$  се издига нагоре от стойността 0, което продължава до момента, когато броят на избраните точки е достатъчен.

### III. Осигуряване на подходящи условия за екипотенциален избор

Необходимо е да отбележим, че релефното разположение в пространствената координатна система на интерпретираното изображение е произволно, а това затруднява приложението на въведената функция за избор (5). Ето защо преди това е необходимо да се създадат подходящи за целта условия. Като такива определяме условията, когато релефът бъде усилен по вертикалната ос така, че цветовете да достигат едновременно и двата възможни екстремума – стойностите 0 и 255 от една страна и в същото време от друга страна да бъдат подобрени такива характеристики на изображението като яркост и контраст. Определяме тези условия като изхождаме от разбирането, че търсим графичните координатите на пресечните точки на релефа с равнината на съответното ниво  $Potential_{max}$  или  $Potential_{min}$ , а не самата стойност на цвета. Съчетавайки тези условия стигаме до идеята за нелинейно преобразуване в интервалите на ниските и на високите нива. За целта дефинираме преобразуваща функция  $f(x) > 0$ , при  $x > 0$  от следния вид:

$$(6) \quad y = \begin{cases} f_1(x), \text{ за всяко } x \in [\min, \alpha) ; \\ \quad \quad \quad 0 \leq \min < \alpha, \text{ и } y \leq x, \\ \quad \quad \quad \text{а ако } x = \min, \text{ то } y = 0 ; \\ x, \quad \text{ за всяко } x \in [\alpha, \beta] ; \\ f_2(x), \text{ за всяко } x \in (\beta, \max] , \\ \quad \quad \quad \beta < \max \leq 255, \text{ и } y \geq x, \\ \quad \quad \quad \text{а ако } x = \max, \text{ то } y = 255. \end{cases}$$

Целта на това функционално преобразуване е потискане на стойностите на функцията  $y$  за стойностите на аргумента  $x$  под нивото  $\alpha$  и в същото време усилване на стойностите на функцията за стойностите на аргумента над нивото  $\beta$  така, че екстремните стойности на функцията винаги да достигат максимално възможните нива от 255 за високото и от 0 за ниското, при запазване на стойностите ѝ в интервала  $[\alpha, \beta]$ . Геометричната интерпретация на определената функция е представена на фигура 3.



Фиг. 3 Графика на преобразуващата функция за интервала  $[0, 255]$ .

#### IV. Функция на преобразуване за интервала $[\min, \alpha)$

Функцията на преобразуване  $f_1(x)$  за интервала  $[\min, \alpha)$  ще търсим като полином от втора степен, който да приема положителни стойности за стойностите на аргумента от интервала  $[\min, \alpha)$ . При тези условия каноничното уравнение на парабола с връх в точка  $Q'(\min, 0)$ , има следния вид:

$$(7) \quad (x - \min)^2 = 2 \cdot p \cdot y, \quad p > 0,$$

където  $y = f_1(x)$ .

Допълнителното условие, което в този случай е известно е, че точка  $R(\alpha, \alpha)$ , която лежи на параболата, винаги има еднакви координати, т.е.  $x_R = y_R = \alpha$ . Тогава координатите на тази точка следва да удовлетворяват уравнението на параболата (7) [10], т.е.

$$(8) \quad (\alpha - \min)^2 = 2.p.\alpha .$$

Оттук определяме параметъра  $p$ :

$$p = \frac{(\alpha - \min)^2}{2.\alpha} .$$

Окончателният вид на функцията  $y = f_1(x)$  след заместването на  $p$  в (7) е следният:

$$(9) \quad f_1(x) = \frac{\alpha}{(\alpha - \min)^2} \cdot (x - \min)^2 .$$

## V. Функция на преобразуване за интервала $(\beta, \max]$

И в този случай функцията на преобразуване  $f_2(x)$  за интервала  $(\beta, \max]$  от стойностите на аргумента ще търсим като полином от втори степен. Десният край на интервала е променлив, но е известно, че стойността му не може да бъде по-голяма от стойността 255. Оттук следва, че графиката на функцията  $f_2(x)$  ще търсим от семейство параболи, всяка от които минава през точка  $M(\beta, \beta)$  и има връх в точка  $N(\max, 255)$ .

Уравнението на параболата с търсеното положение има вида:

$$(10) \quad (x - \max)^2 = -2.q.(y - 255) ,$$

където точката с координати  $(\max, 255)$  представлява върхът на параболата.

Другата точка, за която знаем, че лежи на параболата, е точка  $M(\beta, \beta)$ . Следователно нейните координати трябва да удовлетворяват уравнението (10). Така получаваме:

$$(11) \quad (\beta - \max)^2 = -2.q.(\beta - 255) ,$$

откъдето можем да определим неизвестния параметър  $q$ :

$$(12) \quad -q = \frac{(\beta - \max)^2}{2.(\beta - 255)} .$$

Така уравнението (10), след заместване на параметъра  $q$  с намерената за него стойност, получава окончателния вид:

$$(13) \quad (x - \max)^2 = 2. \frac{(\beta - \max)^2}{2.(\beta - 255)} \cdot (y - 255) .$$

Оттук можем да запишем окончателния вид на преобразуващата функция  $y = f_2(x)$ :

$$(14) \quad f_2(x) = 255 + \frac{(\beta - 255)}{(\beta - \max)^2} \cdot (x - \max)^2 .$$

## VI. Експериментални резултати

Синтезираната интервална нелинейна преобразуваща функция (6) е приложена върху редица тестови изображения. Резултатът от прилагането ѝ се изразява в значителното повишаване на контрастността на изображенията и в по-ясното очертаване на контурите на техните фрагменти. По-долу чрез десните половици на фигури 4 и 5 е представено въздействието на преобразуването върху изходните изображения (левите половици на фигурите).



Фиг. 4 Оригинално изображение №1 (отляво) и след преобразуване (отдясно)



Фиг. 5 Оригинално изображение №2 (отляво) и след преобразуване (отдясно)

Засилването на цветовете на изображението по същество представлява изкривяване на изображението, което не винаги е желано в буквалния смисъл. Има обаче редица случаи, когато се търси положението (координатите) на граничните пиксели, а техният цвят е без значение. Успешното им намиране зависи от степента на контраст в изображението. Въпреки че преобразуването внася грешка в координатите на граничните пиксели, то методите за тяхното определяне внасят по-голяма грешка, работейки върху оригиналното изображение. Ето защо, изложената по-горе предварителна обработка на изображението има място в комбинацията от методи за неговата обработка. Степента на грешка, внасяна в координатите на граничните точки, може да се регулира чрез преместване на нивата  $\alpha$  и  $\beta$ . За наша радост това не ни безпокои, тъй като нашата крайна цел е друга.

Резултатът от приложението на потенциалната функция (5) върху показаните предварително подготвени примерни изображения показваме на фигура 6, където в качеството на шаблон е избран централният квадратен фрагмент с размер 64x64 съдържащ общо 4096 точки. На фигурите ясно се виждат формираните затворени контури от избраните точки.



Фиг. 6 Еквипотенциални криви в изображение №1 и в изображение №2

В първото изображение броят на избраните точки е 532 - това е 7,7 пъти по-малко от общия брой точки в изходното множество, който е 4096. Във второто множество избраните точки са 833, което е около 5 пъти по-малко.

## VI. Резултати от търсене на шаблона

За да се оценят възможностите и параметрите на предложения метод са проведени тестове за търсене на предварително избрани шаблони в изображения, подложени на въздействието на равномерно разпределен шум. Сравнителният анализ е направен между класическия алгоритъм за изчисляване на нормираната взаимна корелация с използване на всичките 4096 точки от шаблона и с използване на редуциран брой точки, отделени по предложения метод, а също така и с редуциран брой точки, определени чрез известните оператори на Собел [4,11,12] и Робъртс [4,11,12]. Наблюдаваните параметри са времето за изпълнение (за откриване на шаблона) и прецизност на резултата (определяне на точното местоположение на шаблона).

За провеждане на тестовете е използван е компютър с процесор Pentium II 350 MHz и оперативна памет SDRAM 128 MB.

Резултатите от тестовете са показани в Таблица 1 и Таблица 2.

Таблица 1

| Изображение No1 |              |            |              |            |                   |            |              |            |
|-----------------|--------------|------------|--------------|------------|-------------------|------------|--------------|------------|
| SNR             | Робъртс      |            | Собел        |            | Еквипотенц. точки |            | Всички точки |            |
|                 | NC           | Координати | NC           | Координати | NC                | Координати | NC           | Координати |
| не              | 1,000        | 32 32      | 1,000        | 32 32      | 1,000             | 32 32      | 1,000        | 32 32      |
| 10              | 0,989        | 32 32      | 0,973        | 32 32      | 0,996             | 32 32      | 0,992        | 32 32      |
| 4               | 0,927        | 32 32      | 0,890        | 32 32      | 0,981             | 32 32      | 0,961        | 32 32      |
| 2,5             | 0,852        | 32 32      | 0,757        | 32 32      | 0,948             | 32 32      | 0,911        | 32 32      |
| 2               | 0,791        | 32 32      | 0,689        | 31 32      | 0,937             | 32 32      | 0,904        | 32 32      |
| 1,33            | 0,618        | 31 32      | 0,543        | 31 32      | 0,871             | 32 32      | 0,783        | 32 32      |
| 1               | 0,538        | 31 32      | 0,444        | 31 32      | 0,759             | 30 32      | 0,649        | 32 32      |
| Run-time        | 01:01:85 min |            | 01:02:83 min |            | 00:56:46 min      |            | 07:44:34 min |            |

Таблица 2

| Изображение No2 |              |            |              |            |                   |            |              |            |
|-----------------|--------------|------------|--------------|------------|-------------------|------------|--------------|------------|
| SNR             | Робъртс      |            | Собел        |            | Еквипотенц. точки |            | Всички точки |            |
|                 | NC           | Координати | NC           | Координати | NC                | Координати | NC           | Координати |
| не              | 1,000        | 32 32      | 1,000        | 32 32      | 1,000             | 32 32      | 1,000        | 32 32      |
| 10              | 0,993        | 32 32      | 0,996        | 32 32      | 0,996             | 32 32      | 0,996        | 32 32      |
| 4               | 0,961        | 32 32      | 0,971        | 32 32      | 0,982             | 32 32      | 0,971        | 32 32      |
| 2,5             | 0,911        | 32 32      | 0,950        | 32 32      | 0,957             | 32 32      | 0,951        | 32 32      |
| 2               | 0,904        | 32 32      | 0,899        | 32 32      | 0,931             | 32 32      | 0,899        | 32 32      |
| 1,33            | 0,783        | 31 32      | 0,820        | 32 32      | 0,850             | 32 32      | 0,821        | 32 32      |
| 1               | 0,649        | 31 32      | 0,718        | 31 31      | 0,799             | 31 32      | 0,718        | 32 32      |
| Run-time        | 01:06:73 min |            | 01:02:40 min |            | 01:16:46 min      |            | 07:44:34 min |            |

Вижда се, че времето за изчисление при използване на редуциран брой точки от шаблона е значително по-малко в сравнение с класическия алгоритъм при запазване надеждността на критерия нормирана взаимна корелация.

### Заклучение

Предложеният метод за отделяне на съществени точки от изображение може да се използва в редица задачи от областта на разпознаването на образи и обработката на изображения. В частност методът е приложим в задачите за търсене на шаблон в изображение с използване на нормираната взаимна корелация като мярка за подобие.

Бъдещите изследвания ще бъдат насочени към оценка на метода при прилагане на други критерии за подобие и сравнителен анализ с други известни методи и алгоритми за отделяне на съществени точки.



**Литература:**

- [1] Brunelli R., Poggio T., *Template Matching. Matched spatial filter and beyond*, A.I.Memo 1549, M.I.T. October 1995
- [2] Gonzalez R.C., R.E. Woods, *Digital image processing*, Addison-Wesley Publishing Company, 1993
- [3] Datacube Inc., "*Intelligent and optimal normalized correlation for high-speed pattern matching*", 300, Rosewood Drive, Danvers, MA 01923, USA 2002
- [4] Julius T. Tou, Rafael C. Gonzalez, *Pattern Recognition Principles*, Addison-Wesley Publishing Company, 1974
- [5] Krattenthaler W., K.J. Mayer, M. Zeiller, *Point correlation: a reduced cost template matching technique*, First IEEE International Conference on Image Processing, 1994
- [6] Lee S. W., Lee D. J., Park H. S., *A New Methodology for Gray-Scale Character Segmentation and Recognition*, IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 18, no. 10, pp. 1045-1050, Oct. 1996
- [7] M. Heath, S. Sarkar, T. Sanocki, K.W. Bowyer, *A Robust Visual Method for Assessing the Relative Performance of Edge-Detection Algorithms*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 19, No. 12, December 1997, pp. 1338-1359
- [8] Wakahara T., Odaka K., *Adaptive Normalization of Handwritten Characters Using Global/Local Affine Transformation*, IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 20, no. 12, pp. 1332-1341, Dec. 1998
- [9] Wang L., Pavlidis T., *Direct Gray-Scale Extraction of Features for Character Recognition*, IEEE Trans. Pattern Analysis and Machine Intelligence, vo. 15, no. 10, pp. 1053-1067, Oct. 1993
- [10] Апостолова М., *Лекции по линейна алгебра и аналитична геометрия*, София, 1992 г.
- [11] Дуда Р., П. Харт, *Распознавание образов и анализ сцен*, Издательство "Мир", Москва 1976
- [12] Павлидис Т., *Алгоритмы машинной графики и обработки изображений*, Издательство "Радио и связь", Москва 1986
- [13] Тянев Д. С., *Алгоритми за разпознаване на образи и използването им в техническата диагностика*, докторска дисертация, 1991, София.