

**РЕПУБЛИКА МАКЕДОНИЈА**



Универзитет "Св. Кирил и Методиј" - Скопје



**ЕЛЕКТРОТЕХНИЧКИ ФАКУЛТЕТ**

Игор Стојановиќ

**ПРЕБАРУВАЊЕ НА ЈРЕГ СЛИКИ И ЛОЦИРАЊЕ НА ОБЈЕКТИ  
ВО СЛИКИТЕ СО МУЛТИРЕЗОЛУЦИСКА АНАЛИЗА**

**- МАГИСТЕРСКИ ТРУД -**

**Скопје, 2002**

Игор Р. Стојановиќ

## **ПРЕБАРУВАЊЕ НА JPEG СЛИКИ И ЛОЦИРАЊЕ НА ОБЈЕКТИ ВО СЛИКИТЕ СО МУЛТИРЕЗОЛУЦИСКА АНАЛИЗА**

*Апстракт:*

Мултирезолуцискиот метод на Harold Stone (1999) за пресметување на прогресивна корелација со Фуриеова анализа во доменот на вејвлет трансформацијата е применет за пребарување и извлекување на слики од колекции на слики во JPEG формат. Процесот на пребарување е базиран врз факторизацијата на DCT матрицата во облик на производ од четири матрици. Нормализирани корелациски коефициенти се користени како мерка за разликите меѓу две слики. Експерименталниот дел, кој опфаќа тестирање на повеќе стотици слики, се однесува на: верификација дали даден објект се наоѓа во некоја слика; одредување на локацијата на објектот во сликата; пронаоѓање и извлекување на дадена слика од библиотека на слики.

*Клучни зборови:* JPEG, локација на објект, пребарување, вејвлети.

## Содржина

1. Вовед .....	3
1.1. Пребарување на слики; стратегии на располагање.....	6
2. Пресметување на корелација со Фуриеовите трансформации на вејвлет трансформациите на сигналите .....	11
2.1. Прелиминарни поими .....	12
2.2. Корелација на сигналите во подопсезите.....	16
2.3. Мултирезолуциска корелација .....	23
3. Алгоритамот на Stone за мултирезолуциска анализа и нормализирани корелациски коефициенти.....	32
3.1. JPEG компресија на слика и факторизација на дискретната косинусна трансформација .....	32
3.2. Алгоритамот на Stone за мултирезолуциска анализа .....	33
3.3. Пресметување на нормализираните корелациски коефициенти .....	35
4. Лоцирање на објекти во JPEG компресирани слики.....	38
4.1. Примери за лоцирање на објекти во JPEG компресирани слики .....	43
5. Лоцирање на JPEG слика во библиотека на слики .....	49
6. MATLAB код .....	57
7. Заклучок.....	64
Литература.....	67

## 1. Вовед

Главниот придонес кон создавањето на Интернетот доаѓа од технолошкиот напредок во областите компјутери и комуникациски мрежи. Јадрото на овој напредок е промената од аналогна во дигитална комуникација заедно со дигитализацијата на звукот и сликите. Денес Интернетот е најголемиот извор на информации, иако WEB страниците и пребарувачите постојат само неколку години.

Сведоци сме на раѓањето на нова индустрија, базирана врз информација, со дигитални библиотеки во нејзиниот центар. Таа индустрија се развива бидејќи информацијата е стока со голема вредност. Интернетот со своите машини за пребарување, пребарувачи и хиперлинкови на WEB страниците обезбедува универзален пристап кон дигиталните библиотеки. Интернетот денес содржи различни значајни колекции, вклучувајќи патенти, возни редови, финансиски податоци, како и комплетни текстови од професионални списанија на научни асоцијации.

Спектакуларното зголемување на WWW<sup>1</sup> може да биде меч со две острици. Постојат огромни количества на информации кои се лесно достапни на WWW, но лоцирањето на специфична информација останува тешка задача. Едни од најпосетуваните WEB страници се машините за пребарување. WEB страниците се составени од текстови и слики. Денес успешно се развиени машините за пребарување на текст, додека машините за пребарување на останатите медиуми на WEB-от како што се сликите, видеото и звукот штотуку почнаа да се појавуваат и се доста примитивни. Lycos, Excite, Alta Vista и Yahoo се примери на корисни машини за пребарување, но сите овие системи главно се дизајнирани за пронаоѓање на текст информации на WEB-от.

---

<sup>1</sup> World Wide Web

---

Интернетот ги задоволува нашите желби во многу области, но во неколку области, особено во оние поврзани со сликите, нашите можности денеска се мали и ни даваат мал дел од она што го посакуваме. Сликите го сочинуваат главниот дел од пренесените бајтови денес. Се предвидува тој дел да се зголемува во иднина. Комерцијалните алатки за пристап кон сликите денес се базирани врз дескриптори, кои претставуваат голем чекор напред, но сепак остануваат примитивни во однос на она што е возможно. Алатките базирани врз дескриптори се развиваат во насока на инкорпорирање и други софистицирани техники како што се релевантната повратна врска и некои аспекти од “разбирање на слика” (image understanding). Пребарувањето базирано врз пиксели претставува алтернативна насока за апликации кои имаат потреба од висока резолуција. Пребарувањето на слики може да користи и комбинации од дескриптори, разбирање на слика, мултирезолуциска техника базирана врз пиксели и други техники. Целта е да се креира корисна алатка за општа практична употреба.

Денешните апликации за пребарување на слики се развиваат заедно со усовршувањето на алатките за пребарување. Се очекува пребарувањето на слики во блиска иднина да биде исто така ефикасно како што е пребарувањето на текстот денес. Машините за пребарување ќе ги интегрираат двата вида функции: пребарување на текст и пребарување на слики. Типични фајлови со слики за обичните корисници ќе бидат видео e-mail и дигиталното видео.

Поради ова очекувано распространето користење на дигиталните слики во иднина, алатките за сместување на слики и алатките за пронаоѓање на сочуваните слики кога тоа е потребно, ќе бидат основни и општи. Владите, агробизнисмените, луѓето кои работат за заштита на природната средина од загадување и последиците од загадувањето се потпираат врз сателитски слики за да ги разберат географските услови и да ги откријат промените и можните причини за тие промени. Медицинските библиотеки со слики се исто така многу важни, бидејќи на корисникот му дозволуваат да постави прашање чиј што одговор зависи од стотици или илјадници слики од библиотеката на слики.

Технолошката поддршка за WEB пребарување - мемории, комуникации и процесори - се развива многу брзо. Последниве години цените и димензиите на мемориите се намалуваат побрзо од цените и димензиите на процесорите и тој тренд продолжува. Се очекува капацитетот на денешните ефтини мемории од десетина гигабајтови да достигне терабајтови за неколку години.

---

Фреквенцискиот опсег и времето на пристап на уредите за комуникации исто така се менува брзо, но тие можат да бидат главното тесно грло за пристап кон слики во годините кои доаѓаат. Инфраструктурата на Интернетот мора да се надградува да може да одговори на потенцијалните барања за што е потребно време и капитални инвестиции. Техниките за компресија на слики вградени во стандарди како што е, на пример, MPEG-2 за видео помогнаа ефикасен пренос на слики со постоечката инфраструктура. Новите стандарди, такви како MPEG-7, којшто беше предложен да поддржува пребарување на мултимедијални податоци, можат да ги направат поефикасни и мрежите за комуникации и пребарувањата. Процесорите се веќе адекватни за едноставни пребарувања и нормалниот развој на индустријата за производство на чипови ќе овозможи далеку покомплексни пребарувања во наредните неколку години.

Интернетот денес може да се разгледува како колекција од над 2 469 940 000 страници (извор [www.google.com](http://www.google.com)) и овој број се зголемува рапидно. Брзината на зголемување на бројот на страниците може да се согледа од податокот дека кон крајот на 1997 година на располагање имало околу 320 милиони страници. На екранот од посетените страници на Интернет доминантен дел од просторот зазема текстуалниот дел, иако значаен дел е опфатен со слики и анимирани објекти. Интернетот исто така го поддржува пристапот кон аудио и видео клипови, но тие не се често користени денес. Друга многу важна мерка е односот на бајтовите посветени на сликите и бајтовите посветени на текстот. Тој однос, испитуван за најпосетуваните страници, е меѓу 60% и 90% [1].

Бидејќи технологијата се стреми кон зголемување на ширината на фреквенцискиот опсег по корисник за помала цена, се очекуваат многу повеќе слики на WEB страниците, особено слики со висока резолуција. Зголемената ширина на фреквенцискиот опсег ќе овозможи замена на актуелниот текстуален e-mail со видео e-mail. Станува јасно дека со с# помасовното користење на слики, видео и аудио пребарувањето на текст на WEB-от ќе се намалува и ќе завзема многу мал дел од сообраќајот на Интернетот. Иднината треба да овозможи пронаоѓањето на слики од библиотеките на Интернет и објекти на сликите да се врши со истата леснотија со која денес се пронаоѓа текст.

## 1.1. Пребарување на слики; стратегии на располагање

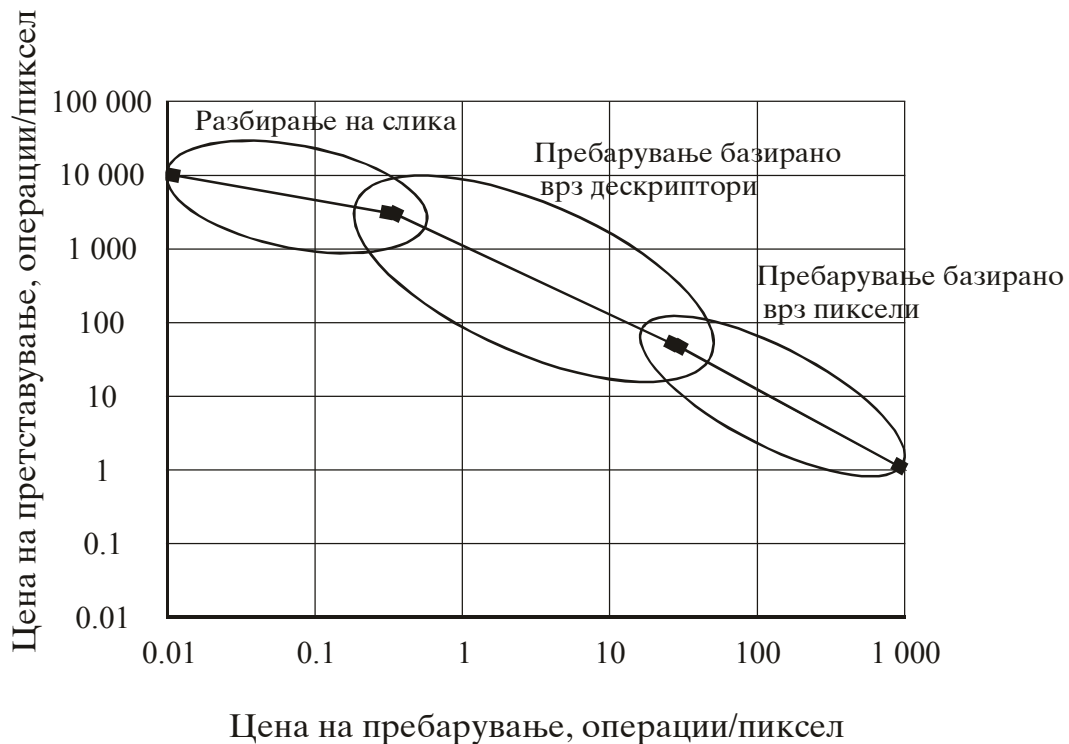
Главната алатка која помогна да го направи Интернетот универзално користен е машината за пребарување на текст. Машините за пребарување на слики кои се на располагање денес се примитивни во однос на машините за пребарување на текст.

Во типичните библиотеки базирани врз текст, алатките за пронаоѓање на барана информација се потпираат врз индекси на дескриптори. Дескриптори можат да бидат зборови или реченици. Тие можат да бидат и други податоци кои не се појавуваат во рамките на документот, како на пример, имиња на автори, издавачи, датум и слично. Дескрипторите на слики се сосема аналогни. Дескриптор на слики може да биде обичен текст. Дескриптори можат да бидат дигитални описи на сликата и можат да се состојат од податоци кои ги опишуваат боите, објектите, композицијата, текстурата или другите карактеристики на сликата.

Во суштина, има неколку различни нивои на претставување на слики кои можат да се искористат за пронаоѓање на конкретна слика и кои варираат од самите пиксели до визуелната интерпретација на сликата. Сликата 1.1 го илустрира компромисот помеѓу овие нивои. На хоризонталната оска е прикажан бројот на операции по пиксел којшто е потребен да може да се врши пребарување. На вертикалната оска е потребниот број на операции по пиксел за таква репрезентација на сликата којашто овозможува пребарување.

Во областа ограничена со полигонот на долниот десен крај од сликата се застапени техники за пронаоѓање слики кои користат најдетална претстава на сликата. Типични репрезентации од овој вид се нискорезолуциски верзии на оригиналната слика, кои можат да се добијат со мал број на операции по пиксел. Меѓутоа, овие техники на пребарување базирано врз пиксели се најсложени од пресметувачка гледна точка; машината за пребарување мора да врши обемни пресметувања за да утврди дали некоја слика одговара на дадено квери.

---



*сл.1.1. Компромис помеѓу цената на претставување и пребарувањето*

Средната област од сликата 1.1, ограничена со елипса се однесува на пребарувања базирани врз дескриптор. Дескрипторите на дадена слика, како на пример, хистограми на бои, информација за текстура, облик и слично, се полека од нискорезолуциските верзии на сликата, но тие можат да пребаруваат многу побрзо.

На највисокото ниво на репрезентација на сликата којашто овозможува пребарување е информација базирана врз анализа на визуелната содржина на сликата. Таква репрезентација може да биде многу скапа, но таа овозможува пребарување базирано врз содржината на сликата кое би било комплицирано или невозможно со дескриптори на слика или со операции над пикселите.

Како пример за различните нивоа на индексирање да разгледаме слика на човеково лице. Најниското ниво на репрезентација на сликата е претставата на лицето со пиксели, најверојатно со ниска резолуција. На следното ниво со дескриптори сликата е опишана со бојата на кожата на лицето, текстурата на кожата и косата и обликот на лицето. Највисокото ниво, визуелното разбирање на сликата, го содржи името на



личноста чијшто портрет е сликата. Денес пребарувањата на слики се вршат речиси исклучиво со дескриптори, бидејќи техниките за визуелно разбирање на слика се нови и недоволно истражени, а цената на процесирање пребарувања базирани врз пиксели има тенденција да го надмине капацитетот на денешните компјутери.

Најраната и најмногу развиениената машина за пребарување на слики базирана врз дескриптор е IBM QBIC [2] лоциран на [www.qbic.almaden.ibm.com](http://www.qbic.almaden.ibm.com).

Алатките базирани врз содржина постојано се развиваат и стануваат се пософистицирани. Страната Corbis ([www.corbis.com](http://www.corbis.com)) обезбедува пребарување со дескриптор на текст на половина милион достапни слики. Virage ([www.virage.com](http://www.virage.com)) нуди современа машина за пребарување заснована врз дескриптор којшто е многу сличен на QBIC. Пребарувањето на слики кај Alta Vista ([www.altavista.com](http://www.altavista.com)) почнува со дескриптор на текст со цел да ја пронајде првата група погодоци, и потоа користи сличност на сликите за дополнително рафинирање. ImageRover [3] користи ниска резолуција за репрезентација на слика во шест региони со цел да бидат опфатени одредени информации заедно со дескрипторите за тие региони.

Сите овие техники ги унапредуваат точноста и полезноста од пребарувањето на слики, и ќе доведат до многу помоќни комерцијални машини во иднина. Прецизноста на барањето има тенденција на подобрување со зголемување на резолуцијата на дескрипторите и со претпроцесирање на сликите за пронаоѓање на посоодветни дескриптори.

Опишаните машини за пребарување се потпираат врз компактна репрезентација на сликата со цел да се добие висока брзина на процесирање. Тие генерално немаат време да вршат анализа на сликите додека трае пребарувањето, и затоа тие настојуваат да инкорпорираат што е можно повеќе корисни информации во претставата на сликите, информации што можат да се стават на располагање кога тоа е потребно.

Техниките за пребарување базирани врз дескриптори се ограничени главно на два начина: тие не можат да дадат одговор за деталите на сликата или не успеваат да ги опфатат севкупните карактеристики во рамките на дескрипторите на слики. Големи комплексни слики можат да содржат многу детали така што нема да може да се најде соодветна компактна репрезентација.

Пребарувањето базирано врз пиксели подразбира пронаоѓање на дадена форма во една библиотека на слики.

---

Одредени параметри вклучени во постапката го специфицираат видот на пребарување, имено дали пребарувањето се однесува на рабови, облици, боја, текстура или на други мерливи релации помеѓу сликите. Популарен критериум е нормализираниот корелациски коефициент, којшто мери колкава е усогласеноста помеѓу дадена форма и слика од библиотека. Со овој критериум се отстранува осетливоста на униформни разлики во осветленоста. Друг често применуван критериум, сумата на квадрати на разликите, може исто така да биде пресметуван со техники на корелација.

Развиени се две независни и комплементарни техники за брзо пресметување на корелацискиот коефициент. Нивната комбинација резултира во забрзувања од 10 000 до 100 000 за типични случаи кога се споредува со пресметување на корелација базарана врз пиксели. Ова е доволно пребарувањето на слики базирано брз пиксели да постане остварливо за истражувачки цели денес и практично во комерцијална употреба во наредните години.

Првата техника го користи поимот нискорезолуциско пребарување. Le Moigne [4] ја применила оваа техника на сателитски слики користејќи го нискофреквенцискиот подопсег на вејвлет трансформацијата. Максималната вредност на корелацијата пресметана на најгруба резолуција се случува на позиција која одговара на иста релативна позиција на најфината резолуција. Со рафинирање на резултатот од грубата резолуција, преку барање на мал број можни точки во високорезолуциската слика, многу брзо може да се утврди усогласеноста помеѓу сликата и дадената форма на најфината резолуција. Ако пребарувањето на груба резолуција не води кон добар кандидат, рафинирањето е непотребно и сликата се отфрла.

Втората техника е мултирезолуциската техника на Harold Stone [5] која вејвлет трансформацијата ја комбинира со Фуриеовата анализа. Вејвлет трансформацијата е погодна за ефикасна компресија, а Фуриеовата анализа на вејвлет трансформацијата овозможува да се добие брза корелација на сликата.

Магистерскиот труд има за цел да го имплементира (софтверски во MATLAB) методот на H. Stone и да ја испита можноста за негова примена во одредување локација на објекти во слики. Извлекувањето на дадена слика од библиотека на слики е исто така опфатено во целите на магистерскиот труд.

---

Содржината на магистерскиот труд е разделена во шест глави. Во главата 2 се дадени основните поими од прогресивна вејвлет корелација со Фуриеова метода. Оваа техника е базирана врз Vaidyanathan-овата теорема [6]. Во подглавата 2.2 е опишано како да се комбинира вејвлет-пакет трансформацијата со Фуриеовата трансформација за да се добие нултиот подопсег од цикличната корелација помеѓу векторите  $x$  и  $y$ . Изложени се и техниката на мултирезолуциска анализа на Н. Stone и начинот како да се пополнат компонентите кои недостасуваат кога оваа техника се користи.

Во главата 3 е изложена факторизацијата на DCT матрицата врз која е базиран процесот на пребарување. Подглавата 3.1 е посветена на пресметувањето на нормализираните корелациски коефициенти како мерка за разликите помеѓу две слики.

Главите 4 и 5 се однесуваат на практичната примена на прогресивната вејвлет корелација со Фуриеова метода на JPEG слики и тоа во двете насоки: пронаоѓање на местото на локација на објекти во одредена слика и пронаоѓање на одредена слика во една библиотека на слики.

Последната глава 6 ги содржи компјутерските програми, напишани во програмскиот јазик MATLAB, коишто се користени во експерименталниот дел на магистерската теза.

---

## Литература

- [1] H. S. Stone, "Image Libraries and the Internet," *IEEE Commun. Magazine*, pp. 99-106, Jan. 1999.
  - [2] M. Flickner et al., "Query by image and video content: The QBIC system," *IEEE Comp.*, vol. 28, pp. 23-32, Sept. 1995.
  - [3] S. Sclaroff, L. Taycher, and M. La Cascia, "Imagerover: A content-based image browser for the world wide web," *IEEE Wksp. Content-Based Access of Image and Video Libraries*, pp. 2-9, June 1997.
  - [4] J. Le Moigne, "Parallel registration of multi-sensor remotely sensed imagery using wavelet coefficients," *Proc. SPIE O/E Aerospace Sensing, Wavelet Appl.*, pp. 432-43, Apr. 1994.
  - [5] H. S. Stone, "Progressive Wavelet Correlation Using Fourier Methods," *IEEE Trans. Signal Processing*, vol. 47, pp. 97-107, Jan. 1999.
  - [6] P. P. Vaidyanathan, "Orthonormal and biorthonormal filter banks as convolvers, and convolutional coding gain," *IEEE Trans. Signal Processing*, vol. 41, pp. 2110-2130, June 1993.
  - [7] G. K. Wallace, "The JPEG still-picture compression standard," *Commun. ACM*, vol. 34, no. 4, pp. 30-44, Apr. 1991.
  - [8] H. S. Stone and T. Shamoon, "The use of image content to control image retrieval and image processing," *Int'l. J. Digital Libraries*, vol. 1, no. 4, pp. 329-43, Dec. 1997.
  - [9] I. Stojanovic, M. Bogdanov, "Location of Objects in a JPEG Image with Progressive Wavelet Correlation using Fourier methods", *9th Telecommunications Forum*, pp. 561-564, Belgrade, Yugoslavia, Nov. 20-22, 2001.
  - [10] I. Stojanovic, M. Bogdanov, D. Taskovski, "Pixel-Based Searching for Object Location in a JPEG Image", *MELECON 2002-11<sup>th</sup> Mediterranean Electrotechnical Conference*, on CD-ROM, Cairo, Egypt, May 7-9, 2002.
-

- 
- [11] I. Stojanovic, D. Taskovski, I. Kraljevski, "Normalized Correlation Coefficients for Searching JPEG Images", *7th IT' 02*, pp. 104-107, Zabljak, Yugoslavia, 24 Feb. - 2 Mar., 2002.
- [12] J. S. Lim, *Two-Dimensional Signal and Image Processing*, Englewood Cliffs, N.J., Prentice-Hall, 1990.
- [13] *MATLAB User's Guide*, The Math Works, Inc., Natick, MA, 1996.
- [14] Krauss, T. P., Shure, L., and Little, J. N. *Signal Processing Toolbox User's Guide*, The Math Works, Inc., Natick, MA, 1996.
- [15] *Image Processing Toolbox User's Guide*, The Math Works, Inc., Natick, MA, 1997.
- [16] *Wavelet Processing Toolbox User's Guide*, The Math Works, Inc., Natick, MA, 1996.
- [17] V. Castelli *et al.*, "Progressive classification in the compressed domain for large EOS satellite databases," *Proc. ICASSP '96*, vol. 4, pp. 2199–2202, Apr. 1996.
- [18] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*, Englewood Cliffs, N.J., Prentice-Hall, 1993.
- [19] Gilbert Strang and Truong Nguyen, *Wavelets and Filter Banks*, Wellesley-Cambridge Press, 1996.
-