

UNIVERSITATEA TEHNICĂ „Gheorghe Asachi” din IAȘI
FACULTATEA DE AUTOMATICĂ ȘI CALCULATOARE
DOMENIUL: Calculatoare și tehnologia informației
SPECIALIZAREA: Tehnologia informației

Recunoașterea textului din imagini folosind metode de clasificare și segmentare

Student
Paul-Cristian Brînză

Iași, 2019

Cuprins

| | |
|---|----|
| Introducere..... | 1 |
| Capitolul 1. Noțiuni teoretice privind recunoașterea optică a caracterelor..... | 2 |
| 1.1. Istoria recunoașterii optice a caracterelor..... | 2 |
| 1.1.1. Istoria..... | 2 |
| 1.1.2. Rezultate privind sistemele OCR..... | 2 |
| 1.2. Recunoașterea optică a caracterelor (Optical Character Recognition)..... | 4 |
| 1.3. Etapele principale ale OCR..... | 6 |
| 1.3.1. Achiziția imaginilor..... | 6 |
| 1.3.2. Preprocesarea..... | 6 |
| Concluzii..... | 8 |
| Bibliografie..... | 9 |
| Anexe..... | 10 |

Recunoașterea textului din imagini folosind metode de clasificare și segmentare

Paul-Cristian Brînză

Rezumat

Procesarea de imagini este considerată a face parte în zilele noastre, din subiectele favorite din industria IT. Este o ramură a procesării semnalelor digitale și una din multele aplicații ale acesteia este reprezentată de *Optical Character Recognition* (OCR).

Atunci când un obiect ce trebuie identificat ne este prezentat, creierul sau sistemul general de recunoaștere începe să extragă caracteristicile cele mai importante ale obiectului printre care culoarea, lungimea, forma și dimensiunea. Acestea sunt stocate într-o parte a memoriei. Atunci când este necesar, creierul încearcă să găsească cea mai bună asemănare pentru aceste caracteristici extrase, dintr-o colecție de obiecte deja cunoscute. Putem numi această colecție ca fiind o bibliotecă standard. În momentul în care a fost găsită cea mai bună potrivire, obiectul din bibliotecă standard este ales ca rezultat final.

Recunoașterea caracterelor de către o persoană, pare a fi un *task* simplu, dar a programa un computer să analizeze și în final să recunoască în mod corect un caracter, se poate dovedi a fi un *task* dificil. OCR reprezintă o astfel de tehnică ce oferă posibilitatea calculatorului de a extrage o din imagini, de a găsi date importante din acestea și a le face să poată fi editate.

Introducere

Procesarea digitală a imaginilor este un domeniu cu evoluție rapidă, datorită creșterii numărului de aplicații în domeniul științei și ingineriei. Procesarea de imagini oferă posibilitatea dezvoltării unui sistem de ultimă generație care să poată îndeplini funcțiile vizuale ale unei persoane. Noțiunea de prelucrare a unei imagini digitale se referă, în general, la prelucrarea unei imagini bidimensionale de către un calculator, adică modificarea unei imagini existente în modul dorit.

Deoarece procesarea se referă la vizualizarea unei imagini, pasul principal este acela de a o obține. O imagine este în esență un model de pixeli (element al unei imagini), astfel că o imagine digitală reprezintă o serie de numere reale și complexe reprezentate pe un număr finit de biți. Imaginile pot fi citite în diferite formate, cum ar fi:

- JPEG (Joint Photographic Experts Group)
- TIFF (Tagged Image File Format)
- BMP (Windows Bitmap)
- PCX (Windows Paintbrush)

Pentru o viziune mai bună asupra acestei ramuri, trebuie amintite și principalele aplicații. Procesarea de imagini oferă un spectru larg de domenii de aplicabilitate printre care se numără următoarele: teledetecția prin satelit, analiza imaginilor medicale, procesarea imaginilor provenite de la radar, recunoașterea vocală, detecția feței, dar nu în cele din urmă recunoașterea optică a caracterelor (OCR).

OCR (*Optical Character Recognition*) este o tehnologie răspândită de recunoaștere a textului din imagini, cum ar fi documente scanate și fotografii. Această tehnologie este folosită pentru a converti aproape orice imagine ce conține text (în format electronic, imprimat sau uneori scris de mână) în text propriu-zis, ce poate fi prelucrat cu orice program specializat de editare.

Tehnologia de recunoaștere a textului din imagini a devenit cunoscută în prima parte a anilor 1990, în timp ce se încerca digitalizarea ziarelor istorice. Este o ramură a informaticii, care presupune recunoașterea textului scris din diverse tipuri de documente fizice, cum ar fi documente scanate, documente PDF (*Portable Document Format*) sau imagini surprinse cu o cameră digitală și convertirea acestora într-un format care să permită unui sistem de calcul manipularea acestuia [1].

De-a lungul anilor, tehnologia a suferit mai multe îmbunătățiri, fiind un domeniu intens cercetat și având o aplicabilitate mare pentru mai multe domenii (de exemplu domeniul poștal, domeniul economic și cel de securitate). De asemenea, mai poate fi folosită cu succes la tehnologiile utilizate pentru asistarea și sprijinul persoanelor cu dizabilități, mai ales pentru persoanele cu probleme de vedere, la aplicații pentru gestiunea electronică a cărților dintr-o bibliotecă sau pentru conservarea pe termen lung a documentelor tipărite.

Marile companii de IT (*Information Technology*) și dezvoltatori de aplicații au contribuit la progresul și la rezultatele actuale în acest domeniu și au creat aplicații ce oferă facilități de recunoaștere a caracterelor. Companii precum Microsoft sau Google au contribuit în acest domeniu și pun la dispoziție API-uri (*Application Programming Interface*) pentru dezvoltarea aplicațiilor de recunoaștere a textului. Exemple de instrumente ce folosesc aceste API-uri pentru detecția textului din imagini sunt *Google Keep* și *Microsoft Office Lens*.

Un alt instrument folosit în acest domeniu este biblioteca *Tesseract*, un motor *open-source* de recunoaștere a textului, disponibil sub licența Apache. Poate fi folosită în mod direct sau (pentru programatori) utilizând un API pentru extragerea textului din imagini și oferă suport pentru o gamă lingvistică variată [2].

Capitolul 1. Noțiuni teoretice privind recunoașterea optică a caracterelor

1.1. Istoria recunoașterii optice a caracterelor

Pentru a înțelege tehnologia descrisă în această secțiune, este necesar să aruncăm o privire către istoria OCR, dezvoltarea sa, metodele de recunoaștere, tehnologiile informatice și diferența dintre oameni și sistemele de calcul.

1.1.1. Istoria

Este întotdeauna fascinant a fi capabili de a găsi modalități de a permite unui computer să imite funcțiile umane, cum ar fi abilitatea de a scrie, de a vedea lucruri, etc. Cercetarea și dezvoltarea sistemelor OCR pot fi urmărite încă de la începutul anilor 1950, când cercetătorii au încercat să surprindă imagini ale caracterelor și textului, la început prin mijloacele mecanice și optice ale discurilor rotative și fotomultiplicatorului, urmate de fotocelule și matrici de astfel de celule. La început operația de scanare era lentă și o singură linie de caractere putea fi digitalizată prin mișcarea scannerului sau a mediului de hârtie. Ulterior invențiile scannerelor cu tambur și a celor plate au apărut și au extins scanarea pe întreaga pagină. Mai apoi progresul circuitelor digitale integrate au adus matrici cu o densitate mai mare, transporturi mai rapide a documentelor și viteze mai mari în procesul de scanare și de conversie digitală. Aceste îmbunătățiri importante au accelerat rapid viteza de recunoaștere a caracterelor, reducerea costului și au deschis posibilități de prelucrare a unei varietăți mari de formulare și documente. De-a lungul anilor 1960 și 1970, noi aplicații OCR au apărut în domenii precum comerțul cu amănuntul, bănci, spitale, oficii poștale; companii de asigurări, feroviare și aeronautice; editori de ziare și în multe alte industrii [3].

În paralel cu aceste progrese în dezvoltarea hardware, au avut loc cercetări intensive privind recunoașterea caracterelor atât în laboratoarele de cercetare din sectoarele academice cât și cele industriale. Cu toate că atât tehnicile de recunoaștere cât și calculatoarele nu erau atât de puternice în primele zile ale anilor 1960, mașinile OCR au avut tendința de a face multe erori atunci când calitatea imprimării a fost slabă, cauzată fie de fonturi și de asprimea suprafeței hârtiei, fie de panglicile de bumbac ale mașinilor de scris.

1.1.2. Rezultate privind sistemele OCR

O pagină a unui document poate conține caractere tipărite de către mașină (ca această pagină), tipărite de mână sau scris de mână într-un script cursiv. Dintre acestea, cele mai ușor de recunoscut sunt caracterele tipărite de către mașină.

Sistemele de recunoaștere a textului tipărit de mașină, au apărut la sfârșitul anilor 1950 și au fost folosite pe scară largă pe computerele de birou de la începutul anilor 1990. O mare parte din informația lumii este ținută captivă în documentele tipărite. Sistemele OCR eliberează această informație, convertind textul din documente într-un format electronic. O dată adus în această formă, pot fi utilizate sisteme de recuperare a informațiilor pentru a localiza materialele de interes, iar software-ul de procesare a textului poate fi utilizat pentru a îl edita.

Cu toate acestea, sistemele OCR nu realizează această conversie fără probleme. Acestea fac erori, iar versiunea electronică a documentelor poate să nu se potrivească corect cu versiunea de pe hârtie. Tehnologia OCR a avansat până la punctul în care sistemele de astăzi sunt cu adevărat utile pentru prelucrarea unei mari varietăți de documente tipărite. Acuratețea de 99% sau mai mult, este obținută în mod obișnuit pe paginile cu imprimare curată. De reținut este

faptul că până și acuratețea de 99% se traduce în 30 erori pentru o pagină tipică care conține 3000 de caractere [4].

Cu toate acestea, în [4] sunt prezentate 280 de exemple care ilustrează probleme ale unor sisteme OCR¹. Fiecare exemplu prezintă o mică porțiune a unei imagini, prima dată în mărimea ei naturală și apoi mărită de 4 ori pentru a fi mai ușor unui om să discearnă detaliile minuțioase ale caracterelor. Se afișează textul corect care este rezultatul dorit de la un sistem OCR, urmat de rezultatele obținute de la trei sisteme principale, denumite în mod anonim Dispozitivul 1, Dispozitivul 2 și Dispozitivul 3. Un simbol *tilde* (~) obținut la ieșirea unui OCR reprezintă un caracter respins produs de sistem pentru a indica faptul că nu poate recunoaște unul sau mai multe caractere din imagine. Rezultatul „<nil>” este obținut la ieșirea OCR atunci când sistemul nu a produs nici un rezultat.

Astfel, în Figura 1.1 un fragment al imaginii unei pagini este reprodus în dimensiunea sa originală în partea de sus a figurii și este urmat de versiunea mărită de 4 ori. Ieșirea OCR dorită apare în continuare (*BRIDGE PARKWAY*), urmată de ieșirile obținute de la cele trei sisteme OCR.

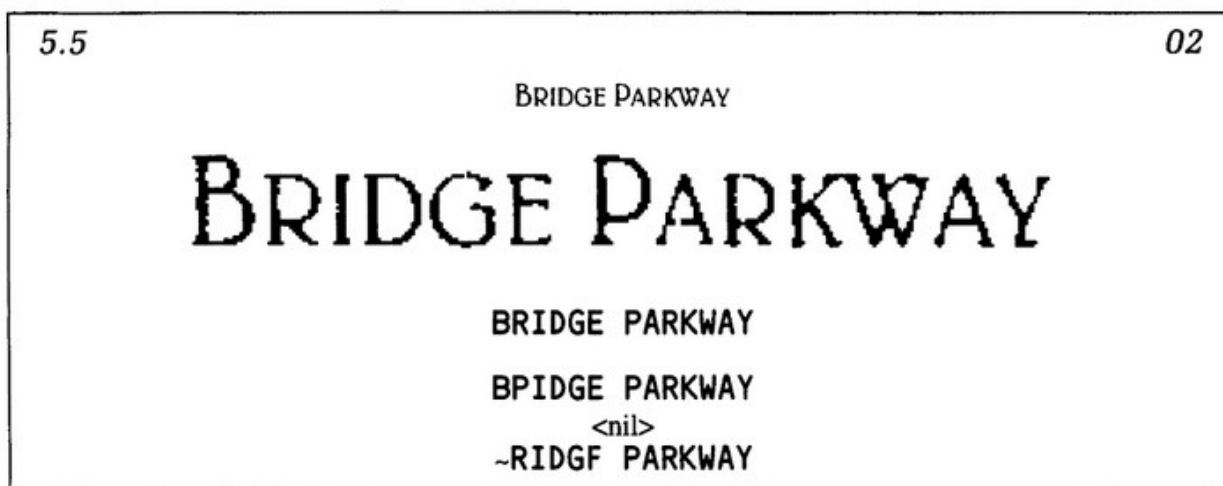


Figura 1.1: Rezultatele obținute de cele trei sisteme OCR (preluat din [4])

Dispozitivul 1 a produs „BPIDGE PARKWAY”, generând în mod eronat un P în loc de R-ul din cuvântul BRIDGE; Dispozitivul 2 a fost confuz și nu a produs nici o ieșire, fapt indicat de prezența simbolului <nil>, iar Dispozitivul 3 a generat un caracter respins pentru B și caracterul F pentru E-ul din BRIDGE.

Cele 280 de exemple analizate fac parte din mai multe categorii de documente printre care documente tehnice, articole, reviste sau scrisori iar unele surse de eroare au apărut mai des la unele tipuri de documente decât la altele. Multe dintre rapoartele tehnice erau fotocopii și prin urmare, au fost mai predispuse la defecte ale imaginii decât alte tipuri de documente, deși urme rătăcite au apărut cel mai des în ziare. Fundaluri întunecate și imagini rotite au apărut cel mai frecvent în articolele din reviste.

Cele trei sisteme OCR fac erori diferite, dar sunt foarte apropiate unele de altele în termeni de precizie generală (Lucru observat nu numai pe cele 280 de exemple ci și pe o colecție mult mai mare). Toate cele trei sisteme nu au produs nici o eroare pentru 13-16% din exemple, o eroare pentru 33-36% și două sau mai multe erori în 51-52% din cazuri. Un test al abilităților unor copii de diferite vârste, de a identifica caracterele din aceste exemple, indică faptul că sistemele informatice sunt mult în urmă:

1 OCR reprezintă abrevierea pentru Optical Character Recognition („Recunoașterea Optică a Caracterelor”)

Tabelul 1.1: Rezultatele obținute de cele trei dispozitive comparativ cu ale copiilor de diferite vârste (adaptat din [4])

| | Nicio eroare | O eroare | Mai multe erori |
|----------------|--------------|----------|-----------------|
| Dispozitivul 1 | 16% | 33% | 51% |
| Dispozitivul 2 | 13% | 36% | 51% |
| Dispozitivul 3 | 15% | 33% | 52% |
| Copil, 7 ani | 58% | 23% | 19% |
| Copil, 10 ani | 86% | 10% | 4% |
| Copil, 14 ani | 92% | 6% | 2% |

1.2. Recunoașterea optică a caracterelor (Optical Character Recognition)

Recunoașterea optică a caracterelor se referă la acea ramură a informaticii care implică citirea textului de pe un document din hârtie și traducerea imaginilor într-o formă pe care calculatorul o poate manipula (de exemplu, în codul ASCII²). Un sistem OCR permite alegerea unei cărți sau a unui articol dintr-o revistă, transformarea acestora într-un fișier electronic pe un computer și apoi editarea fișierului folosind un procesor de cuvinte. Toate sistemele OCR includ un scanner optic pentru citirea textului și un software sofisticat pentru analiza și procesarea imaginilor.

Cele mai multe sisteme de recunoaștere optică a caracterelor utilizează o combinație de hardware (plăci cu circuite electronice specializate) și software pentru a recunoaște caracterele, deși unele sisteme ieftine fac acest lucru în întregime prin software. Sistemele OCR avansate pot citi textul într-o mare varietate de fonturi, dar au încă dificultăți cu textul scris de mână. Reprezintă transformarea mecanică sau electronică a imaginilor scanate a scrisului de mână, scrisului tipărit sau printat, în text editabil codificat pentru calculatoare. Este folosit pe scară largă pentru a converti cărți și documente, în fișiere electronice, pentru a computeriza un sistem de stocare a datelor într-un birou sau pentru a publica textul pe un site web.

Sistemele de recunoaștere a caracterelor permit operațiile de editare a textului, de căutare a unui cuvânt sau a unei propoziții, de afișare sau imprimare a unui document fără artefactele scanării și de aplicare a unor tehnici precum traducerea automată, transformarea textului în vorbire și redarea acestuia. OCR este un domeniu de cercetare în recunoașterea modelelor, vederea și inteligența artificială. Recunoașterea optică a caracterelor (folosind metode optice ca oglinzi și lentile) și recunoașterea digitală a caracterelor (folosind scanere și algoritmi pe calculator) au fost, inițial, considerate domenii diferite. Deoarece puține aplicații folosesc tehnici optice, termenul OCR include și procesarea digitală a documentelor.

Procesul de recunoaștere optică a caracterelor necesită calibrare pentru a citi un anumit font, primele versiuni fiind programate cu imagini ale fiecărui caracter în parte și funcționând doar pentru un singur font la un moment dat. Sistemele „inteligente” cu un grad ridicat al preciziei recunoașterii mai multor fonturi sunt acum comune. Unele dintre ele sunt capabile de a reproduce rezultate formate care aproximează strâns pagina originală scanată, incluzând imaginile, coloanele sau chiar alte componente non-textuale.

Domeniul este unul de interes, reprezentând una dintre cele mai populare teme de cercetare în recunoașterea modelelor datorită potențialului său imens de aplicare. Cu toate

2 ASCII este acronimul pentru American Standard Code for Information Interchange („Codul Standard American pentru Schimbul de Informații”) și reprezintă un sistem de codificare a caracterelor, bazat pe alfabetul englez. Codurile ASCII reprezintă caractere text pentru computere, echipamente de comunicație și echipamente care lucrează cu text.

acestea, majoritatea metodelor disponibile se ocupă de recunoașterea caracterelor alfabetului roman și a caracterelor unor scrieri orientale precum Kanji, Kana etc.

Scopul acestui mecanism de recunoaștere este de a lua caracterele scrise de mână sau tipărite ca date de intrare și a le procesa, de a antrena algoritmul rețelei neuronale, de a recunoaște modelele și a obține în final caracterele de la intrare într-o variantă îmbunătățită. Unul dintre mijloacele primare prin care computerele imită abilitățile umane este reprezentat de utilizarea unei rețele neuronale. Rețelele neuronale sunt deosebit de folositoare pentru rezolvarea problemelor care nu pot fi exprimate ca o serie de pași, cum ar fi recunoașterea modelelor, clasificarea lor în grupuri, predicția seriilor și extragerea datelor [5].

Există o multitudine de direcții în care cercetarea privind recunoașterea caracterelor a fost efectuată și de aceea mai multe tipuri de sisteme OCR au apărut ca rezultat al acestor cercetări. Ele se clasifică după modul de achiziție a imaginilor, a conexiunii caracterelor, a restricțiilor pentru fonturi etc.

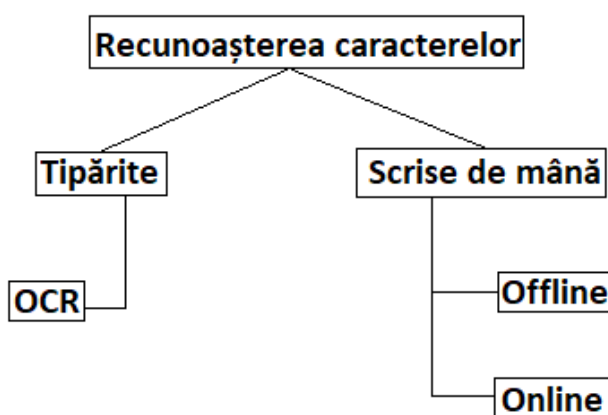


Figura 1.2: Clasificarea sistemelor de recunoaștere a caracterelor

Pe baza tipului de date de intrare, sistemele OCR pot fi împărțite în sisteme de recunoaștere a scrisului de mână și sisteme de recunoaștere a caracterelor tipărite, așa cum se poate observa și în Figura 1.2. Cea din urmă reprezintă o problemă relativ mai simplă, deoarece caracterele sunt de obicei de dimensiuni uniforme, iar pozițiile acestora în pagină pot fi precise. În schimb, recunoașterea caracterelor scrise de mână reprezintă o provocare mai mare, din cauza stilurilor diferite de scriere ale utilizatorilor, precum și a mișcării diferite realizate de către un utilizator, pentru același caracter.

Aceste sisteme de recunoaștere pot fi împărțite în două subcategorii, online și offline. Prima este realizată în timp real, pe măsură ce utilizatorii scriu caracterele și este mai puțin complexă, deoarece poate capta informații ce depind de timp (viteza de scriere, numărul de lovituri realizate, direcția de scriere etc.). În plus, nu este nevoie de tehnici de subțiere (*thinning*) a caracterelor, deoarece urmele lăsate de dispozitivul de scriere sunt de câțiva pixeli lățime. Sistemele de recunoaștere offline funcționează cu date statice, adică intrarea este reprezentată de o imagine bitmap. Prin urmare, este foarte dificil să se realizeze recunoașterea.

OCR permite un număr mare de aplicații utile. Pe parcursul primelor zile sistemul a fost utilizat pentru sortarea mail-urilor, verificarea cecului bancar sau verificarea semnăturii. Pe lângă acestea, mai poate fi folosit de organizații pentru procesarea automată a formularelor în locuri unde un număr mare de date este disponibil în forma tipărită.

Alte utilizări includ procesarea facturilor pentru utilități, validarea pașapoartelor, computerizarea scrisului și recunoașterea automată a plăcuțelor de înmatriculare. O altă aplicație a OCR îi ajută pe utilizatorii nevăzători sau cu probleme de vedere să citească textul [6].

1.3. Etapele principale ale OCR

Procesul de recunoaștere optică a caracterelor este o activitate compusă din mai multe etape. Aceste etape sunt următoarele:

- Achiziția imaginilor
- Preprocesarea
- Segmentarea caracterelor
- Extragerea caracteristicilor
- Clasificarea caracterelor
- Postprocesarea

1.3.1. Achiziția imaginilor

Imaginea este reprezentarea unui obiect sau a mai multor obiecte, executată pe o suprafață, prin acțiunea directă a utilizatorului sau cu ajutorul unui echipament. Achiziția de imagini reprezintă etapa inițială în recunoașterea optică a caracterelor ce cuprinde obținerea unei imagini digitale și transformarea acesteia într-o formă adecvată care poate fi ușor procesată de calculator. Asta poate include procedee de digitizare, cuantizare și de compresie a imaginii.

Cuantizarea imaginii reprezintă operația de transformare a acesteia într-un set de numere binare. Un caz special de cuantizare este binarizarea, care implică doar două nivele ale imaginii și care în multe cazuri sunt suficiente pentru a caracteriza imaginea. În cel mai simplu caz, o imagine conține un singur obiect sau mai multe obiecte individuale, separate, de o intensitate relativ mare, vizualizate pe un fundal de intensitate mică. În acest caz, separația dintre obiecte și fundal se poate realiza prin binarizare [7]. Procedeele de compresie pot fi cu pierderi sau fără pierderi.

1.3.2. Preprocesarea

Odată ce imaginea a fost achiziționată, pot fi executați pași pentru a îmbunătăți calitatea imaginii. Printre diferitele tehnici de preprocesare se numără eliminarea zgomotului, *thinning*, *thresholding*, detecția și corecția unghiului de înclinare precum și diferitele operații morfologice.

Una dintre tehnicile de preprocesare este cea numită *thresholding* care vizează binarizarea imaginii pe baza unei anumite valori de prag. Aceasta poate fi stabilită la nivel local sau global.

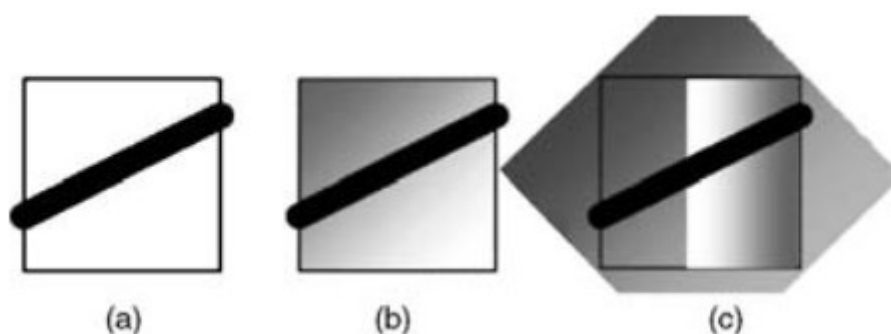


Figura 1.3: Linii ale unor caractere pe diferite tipuri de fundal: (a) Fundal simplu, ușor de binarizat utilizând *global thresholding*; (b) Fundal ce variază ușor, care poate fi îndepărtat folosind *local thresholding*; (c) Fundal ce variază rapid și care poate fi îndepărtat folosind *feature thresholding* (preluat din [8])

În Figura 1.3 sunt prezentate diferite tipuri de *thresholding*:

- *Global Gray Thresholding*, Figura 1.3a: Algoritmii utilizează un singur prag pentru întreaga imagine. Un pixel având un nivel de gri mai mic decât valoarea pragului este etichetat ca aparținând obiectului (negru), iar în caz contrar, fundal (alb).
- *Local Gray Thresholding*, Figura 1.3b: Algoritmii de prag local calculează un prag desemnat pentru fiecare pixel bazat pe o vecinătate a acestuia. Un pixel având un nivel de gri mai mic decât valoarea pragului este etichetat ca aparținând obiectului (negru), iar în caz contrar, fundal (alb). Este folosit pentru un fond care variază lent.
- *Feature Thresholding*, Figura 1.3c: Este la fel ca *Local Gray Thresholding*, dar este aplicat pentru o schimbare bruscă a fundalului.

O altă tehnică este reprezentată de aplicarea algoritmilor de *thinning* (subțiere), care se aplică tuturor caracterelor normalizate. Când caracterele sunt normalizate, unele valori de 1, redundante sunt de asemenea stocate într-o matrice standard. Aceste valori afectează unul dintre parametrii caracterului, adică grosimea. Această „subțiere” reduce grosimea caracterelor până la scheletul inițial (*skeleton*) și elimină acele valori redundante cu ajutorul a opt măști standard de ștergere, câte două pentru fiecare direcție (sus, jos, stânga, dreapta).

Liniile caracterelor sunt reduse la o grosime de un singur pixel. Cu toate acestea, prin ștergerea pixelilor aflați la extremități, pot fi șterși și câțiva pixeli importanți care pot duce la discontinuități ale imaginii. Scheletul obținut trebuie să aibă următoarele proprietăți: să fie cât se poate de subțiri, să fie conectate și de asemenea să fie centrate.

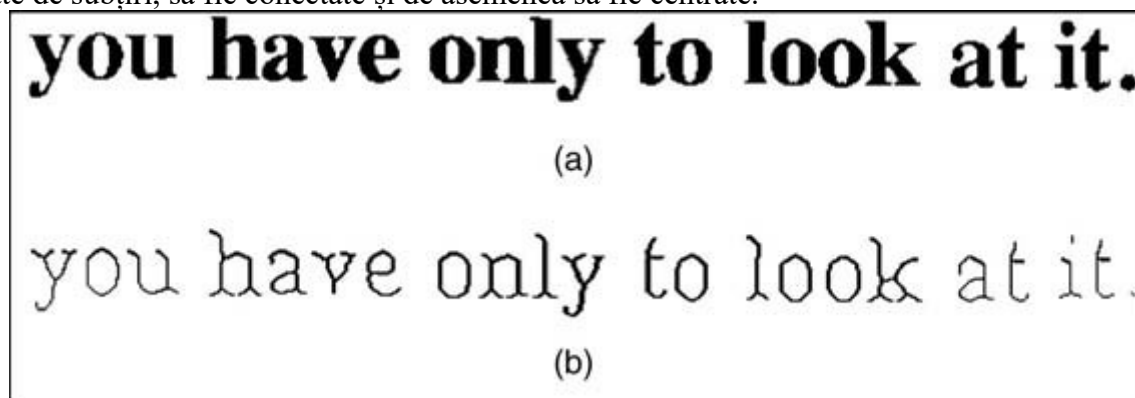


Figura 1.4: Rezultatul aplicării algoritmului de *thinning*;
(a) Textul original; (b) Scheletul liniei de text din (a); (preluat din [8])

Detecția și corecția unghiului de înclinare (*skew detection and correction*) – În timpul procesului de scanare a documentelor, este posibil ca unele pagini să nu fie introduse perfect drept în scanner, cauzând înclinarea textului în imaginile bitmap ale acestor documente. Deci, înclinarea documentelor apare adesea în timpul scanării sau copierii documentelor. Acest efect apare vizual ca o pantă a liniilor de text în raport cu axa Ox (axa absciselor din sistemul cartezian) și se referă în principal la orientarea liniilor de text.

După detecția unghiului de înclinare al textului din imagine, aceasta trebuie rotită pentru a corecta această înclinare. Un algoritm de rotație trebuie să fie în același timp, destul de rapid și destul de precis. O transformare de rotație a coordonatelor, poate fi folosită pentru a corecta unghiul de înclinare.

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad (1)$$

Având un punct dat, noile coordonate după rotirea întregii imagini în jurul originii sale cu un

unghi (care reprezintă chiar unghiul de înclinare determinat), pot fi calculate folosind (1). Rezultatul acestei rotații aplicate unei imagini pentru corecția înclinării textului, poate fi observat în figura de mai jos:

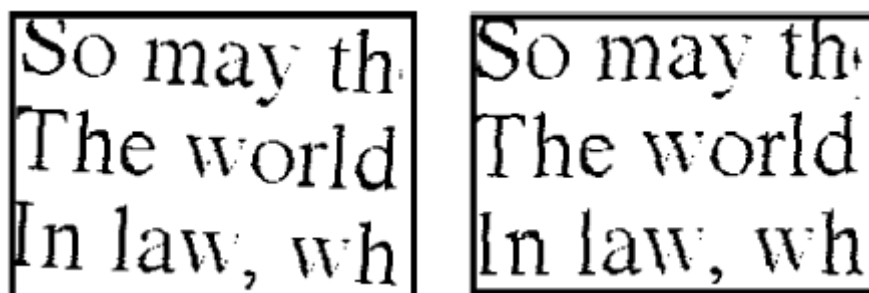


Figura 1.5: Exemplu al corecției înclinării textului (preluat din [8])

Concluzii

Bibliografie

- [1] Abby, What is OCR and OCR technology [Online], Disponibil la adresa: <https://www.abbyy.com/en-ca/finereader/what-is-ocr/>, Accesat: 2019.
- [2] S. Weil, Tesseract open source OCR engine [Online], Disponibil la adresa: <https://github.com/tesseract-ocr/tesseract/wiki>, Accesat: 2019.
- [3] J. Rabinow , „Optical Character Recognition and the Years Ahead”, The Business Press, Elmhurst, Ill, 1969.
- [4] S. Rice, G. Nagy, T. A. Nartker, „Optical Character Recognition: An Illustrated Guide to the Frontier”, Kluwer Academic Publishers, 1999.
- [5] N. A. M. Isheawy, H. Gasan, „Optical Character Recognition (OCR) System”, IOSR Journal of Computer Engineering (IOSR-JCE), vol. 17, pp. 22-26, Mar – Apr. 2015.
- [6] N. Islam, Z. Islam, N. Noor, „A Survey on Optical Character Recognition System”, ITB Journal of Information and Communication Technology, vol. 10, pp. , Dec. 2016.
- [7] G. Danescu, Curs 3: Imagini binare. Proprietati geometrice simple. [Online], Disponibil la adresa: http://users.utcluj.ro/~rdanescu/pi_c03.pdf, Accesat: 2019.
- [8] M. Cheriet, N. Khama, C. Y. Suen, „Character Recognition Systems: A Guide for Students and Practitioners”, John Wiley & Sons, Inc., Hoboken, New Jersey, 2007.

Anexe.

Anexa 1.