

## PREPOZNAVANJE DVODIMENZIONALNIH OBLIKA I IZGOVORENIH RIJEČI KORIŠTENJEM SLIČNIH ALGORITAMA

---

*U radu se opisuju dva slična algoritma: jedan za prepoznavanje dvodimenzionalnih oblika i drugi za prepoznavanje izgovorenih riječi. Po tim algoritmima izrađeni su programi za verifikaciju. Pokazano je da se slični algoritmi mogu koristiti za prepoznavanje dvodimenzionalnih oblika i izgovorenih riječi. Istaknuti su elementi sličnosti algoritama.*

*Algoritam; kompjutor; oblik; prepoznavanje; sličnost.*

---

### UVOD

Umijeće prepoznavanja jedna je od temeljnih čovjekovih karakteristika. Čovjek prima različite podražaje iz vanjskog svijeta preko svojih osjetila vida, sluha, opipa, mirisa, dodira, zatim preko osjeta bola, osjeta u vezi s ravnotežom itd. Prepoznavanje se sastoji u tome da se podražajima pridruže odgovarajući pojmovi.

Usko grlo u komunikaciji čovjeka i stroja je unos podataka u kompjutor. Da bi se to riješilo, potrebno je na neki način "naučiti" kompjutor da sam u sebe unosi podatke, npr. "čitajući" štampani tekst. Osim odgovarajuće ulazne jedinice za digitalizaciju predmeta prepoznavanja, npr. optičkog skenera, potreban je i program za prepoznavanje.

Problem kod raznih oblika prepoznavanja pomoću kompjutera je u tome što predmeti prepoznavanja skoro nikada nisu identični sa zapisima koji postoje u memoriji kompjutera, oni mogu biti samo više ili manje slični. Zbog toga se traže metode za utvrđivanje stupnja sličnosti, odnosno različitosti predmeta prepoznavanja sa zapisima s kojima se uspoređuje.

Pri izradi programa pošao sam od vlastite ideje koju sam želio realizirati. Na donekle slične postupke prepoznavanja oblika naišao sam u knjizi (L1. str. 125-153). Program za

prepoznavanje dvodimenzionalnih oblika, koji je opisan u ovom radu, predstavlja realizaciju te ideje. Kasnije sam primijetio da bi se ista ideja mogla iskoristiti i za prepoznavanje izgovorenih riječi. Program za prepoznavanje izgovorenih riječi nastao je po analogiji s programom za prepoznavanje dvodimenzionalnih oblika. Tu je kao ulazna jedinica upotrijebljen mikrofon, a način zapisa izgovorenih riječi u memoriji kompjutera ponešto se razlikuje. Međutim, postupci prepoznavanja potpuno su analogni, sve osnovne faze prepoznavanja postoje u oba programa.

## 1. PREPOZNAVANJE DVODIMENZIONALNIH OBLIKA

### 1.1. POSTUPAK PREPOZNAVANJA OBLIKA

Faze postupka prepoznavanja oblika su slijedeće:

1. pretraživanje ekrana sve dok se ne nađe na jednu osvijetljenu točku (točku u boji tinte).<sup>1</sup>
2. poziva se potprogram za bojanje oblika u boju inverznu boji tinte. Na taj način oblik koji želimo prepoznati izdvaja se iz okoline. Isti potprogram određuje odgovarajući okvir (prozor) oblika.
3. ako je veličina okvira manja od minimalne (3\*3 točke), tada se radi o šumu.<sup>2</sup> Šum (tj. oblik koji je manji od minimalnog) treba izbrisati, a postupak nastaviti od koraka (1);
4. preslikavanje oblika koji se nalazi unutar okvira (obojen je u boju inverznu boji tinte) u standardni prozor veličine 32\*32 točaka. Istovremeno s preslikavanjem briše se originalni oblik kako ne bi smetao kod prepoznavanja slijedećeg.
5. uspoređivanje redom dobivenog oblika koji se nalazi u standardnom prozoru sa znakovima koji su već zapisani u memoriji (takoder u formatu 32\*32 točke). To se radi tako da se ustanovi stupanj sličnosti (izražen u postotku) matrice dobivenog oblika pojedinačno s matricom svakog oblika koji se nalazi zapisan u memoriji. Najsličniji je onaj (otprije zapisan) oblik uz koji se dobije najviši stupanj sličnosti. Taj oblik i njemu pridruženi ASCII kod ujedno su i rezultati čitavog postupka.

---

1 U tu svrhu koriste se različiti algoritmi, ovisno o opciji.

2 Pretpostavka je da bilo koji smisljeni znak na ekranu mora biti veći od navedenog minimuma.

## 1.2. IZDVAJANJE OBLIKA IZ OKOLINE

Jedan od glavnih problema pri realizaciji prepoznavanja oblika je izdvajanje oblika iz okoline. U tu svrhu treba sve točke koje čine jednu cjelinu (tj. pripadaju obliku) na neki način označiti. Osim toga, potrebno je odrediti koordinate krajnjih točaka najmanjeg pravokutnog okvira (prozora) koji opisuje oblik koji želimo izdvojiti iz okoline.

Oboje je riješeno pomoću istog potprograma, a uvjet je da se oblik sastoji od samo jednog dijela, te da se nigdje ne dodiruje sa susjednim oblikom. Potprogram oboji sve međusobno povezane točke u boji tinte u boju inverznu boji tinte (također vidljivu na ekranu), te ujedno pronalazi i krajnje točke oblika, tj. određuje okvir<sup>3</sup>.

## 1.3. TRANSFORMACIJA OBLIKA

Transformacija oblika je postupak kojim se oblik pronaden na ekranu preslikava u okvir standardne veličine (32\*32 točaka). To se radi zato da bi se ovaj oblik mogao uspoređivati s oblicima koji su već zapisani u memoriji.

Da bi se postupak mogao provesti, potrebno je najprije originalni oblik smjestiti u odgovarajući okvir. Postupak transformacije ovisi o odnosu veličina okvira originalnog oblika i standardnog okvira.

## 1.4. USPOREĐIVANJE OBLIKA

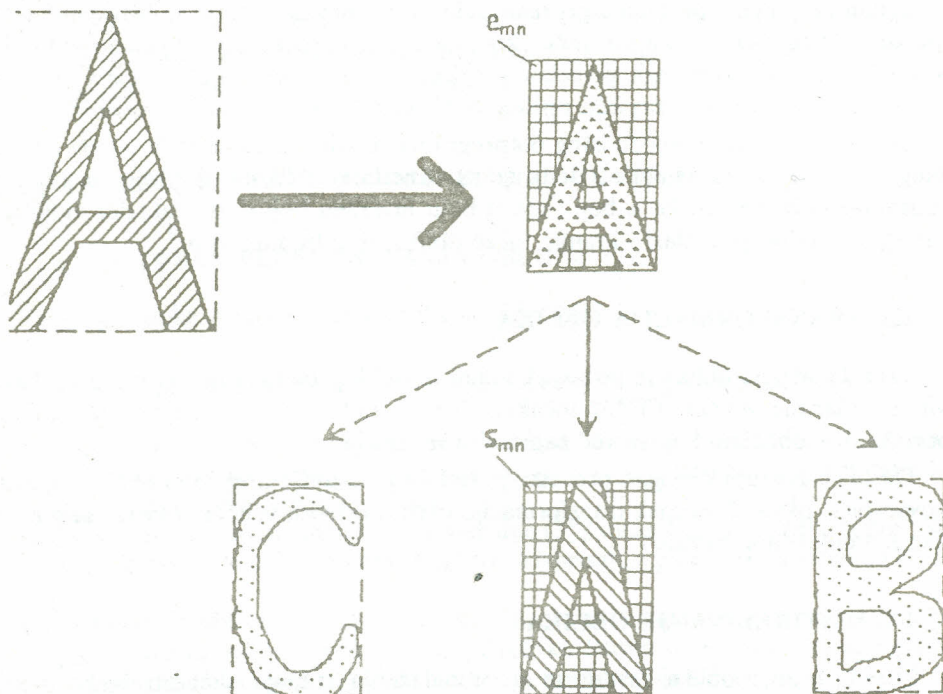
Uspoređivanje oblika sastoji se u pronalaženju stupnja sličnosti svake matrice (32\*32) koja se nalazi zapisana u memoriji (sadrži bitnu mapu jednog znaka) s matricom (također 32\*32) koja je dobivena transformacijom oblika koji se prepoznaje. Stupanj sličnosti izračunava se na bazi preklapanja bitnih mapa uspoređivanih znakova.

Otprije zapisana matrica, uz koju se dobije najviši stupanj sličnosti, smatra se najbližijom. ASCII kod slova spremljen u za to predviđenoj tablici s istim rednim brojem kao i najbližija matrica je rezultat uspoređivanja oblika.

---

3 Za izradu potprograma nisam koristio literaturu. Međutim, sličan postupak koristi većina programa za crtanje u okviru operacije FILL (jedan je opisan u članku J. Skuljan: Operacija fill, RAČUNARI br. 12 - 2/1986).

Uspješnost prepoznavanja se povećava ako je jedan te isti znak definiran više puta (Slika 1.) .



Slika 1. Uspoređivanje oblika

### 1.5. DEFINIRANJE NOVIH OBLIKA

Kompjutor prepoznaje oblike tako što nove uspoređuje s onima koji već postoje zapisani u memoriji. Zbog toga korisnik programa mora definirati nove oblike i spremiti ih u memoriju u istom formatu (matrica 32\*32). U tu svrhu koristi se potprogram za crtanje. Postupak unošenja novog oblika u memoriju je slijedeći:

#### 1.-4. (vidi:POSTUPAK PREPOZNAVANJA OBLIKA, koraci 1-4).

5. Ako u memoriji nema ni jednog prijašnjeg zapisa, tada se izvođenje nastavlja od koraka (6). U suprotnom se vrši uspoređivanje upravo definiranog oblika s onima koji su već upisani u memoriji. Iscrta se najslićnija matrica, ispiše stupanj sličnosti (brojkom i grafički) i znak s ASCII kodom nadenog znaka.

6. Očekuje se pritisak na odgovarajuću tipku na tastaturi kako bi se dobivenoj matrici pridružio ASCII kod.

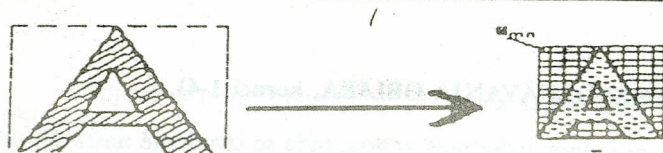
Definirani oblik se u preoblikovanom obliku - kao matrica 32\*32 sprema u memoriju. U posebne tabele spremaju se brojevi bitova spremljene matrice koji su postavljeni na 1 i ASCII kod definiranog oblika.

#### 1.6. PREPOZNAVANJE TEKSTA

Prepoznavanje teksta je glavna opcija za prepoznavanje oblika.

Tekst koji se prepoznaje može biti napisan u više redova, a rezultat se ispisuje na dnu ekrana (Slika 2.). Postupak je slijedeći:

1. Brojač redova kod ispisa postavi se na 19 (prvi red ispod prostora za crtanje oblika).
2. Pretražuje se prostor ekrana predviđen za crtanje od gornjeg ruba prema donjem, red po red, tako dugo dok se ne naiđe na prvu točku u boji tinte (crnu).
3. Oblik kojemu pripada nadena točka izdvoji se iz okoline. Ako je izdvojeni oblik manji od minimalnog, tada se postupak nastavlja od koraka (1).
4. Izvrši se prepoznavanje oblika (vidi: POSTUPAK PREPOZNAVANJA OBLIKA, koraci (4) i (5)).
5. Karakter s ASCII kodom prepoznatog oblika ispiše se na sredinu reda ekrana ispod prostora za crtanje oblika na koji pokazuje brojač redova.
6. Pretraživanje ekrana lijevo od nadenog oblika; traži se oblik (znak) lijevo od upravo prepoznatog. U slučaju da se kod pretraživanja naiđe na rub ekrana, tada se postupak nastavlja od koraka (9).
7. Oblik kojemu pripada nadena točka izdvaja se iz okoline. Ako je okvir oblika manji od minimalnog, tada se ponavlja korak (6).
8. Na temelju prosječne širine već prepoznatih oblika izračunava se broj blanko znakova između zadnjeg prepoznatog i upravo izdvojenog oblika. Blankovi se ispišu na donjem dijelu ekrana.
9. Pretraživanje ekrana desno od znaka koji je prvi pronađen u redu. Postupak je posve analogan koracima (3) do (8), samo što se pretražuje udesno (do desnog ruba dijela ekrana predviđenog za crtanje oblika).
10. Brojač redova kod ispisa povećava se za 1, a postupak se nastavlja od koraka (2) tako dugo dok se ne dođe do donjeg ruba dijela ekrana predviđenog za crtanje oblika.



Slika 2. Prepoznavanje teksta na ekranu

## 2. PREPOZNAVANJE IZGOVORENIH RIJEČI

### 2.1. POSTUPAK PREPOZNAVANJA IZGOVORENIH RIJEČI

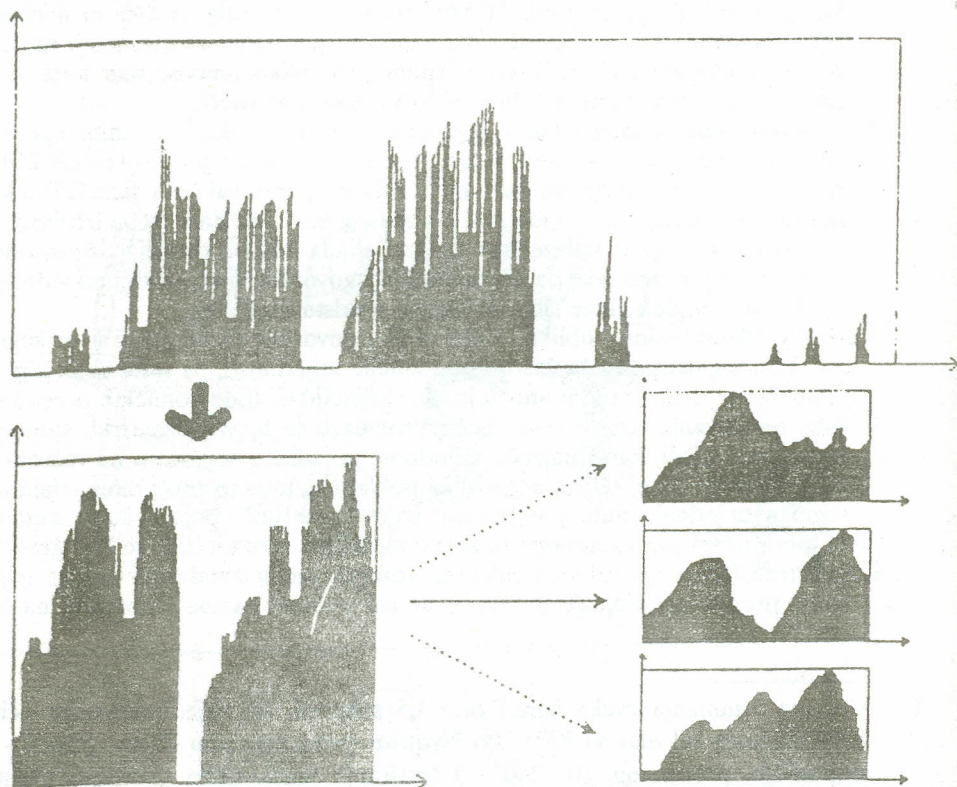
Ovo poglavlje ujedno je opis rada programa koji prepoznaje izgovorene riječi. Postupak prepoznavanja izgovorenih riječi sastoji se od slijedećih faza (Slika 3.):

1. Faza čekanja. U toj fazi računar očekuje da korisnik programa počne izgovarati riječ. Testira se stanje na EAR utičnici kompjutera (može biti "0" ili "1") tako dugo dok se ne pojavi stanje "1".
2. Faza snimanja zvuka. U određenim intervalima testira se stanje na EAR utičnici kompjutera<sup>4</sup>. Broje se stanja "1". Po završetku intervala ukupan broj stanja "1" u testiranom intervalu sprema se u dio memorije predviđen za spremanje zapisa. Tako se ispita ukupno 256 intervala i ispuni isto toliko memorijskih lokacija (svaka predstavlja jedan segment valnog oblika izgovorene riječi).
3. Faza filtriranja valnog oblika izgovorene riječi. U toku faze snimanja zvuka (2) obično se snime i razni šumovi. Da bi se moglo izvršiti prepoznavanje izgovorene riječi, moraju se najprije odrediti njeni rubovi, tj. interval (vidi: faza (1)) u kojem je počelo izgovaranje riječi i interval u kojem je završilo. Zato treba izbrisati dijelove valnog oblika koji se nalaze ispred i iza dijela koji predstavlja izgovorenu riječ. Pošlo se od pretpostavke da se valni oblik izgovorene riječi sastoji od jednog ili više većih dijelova, dok razne sitne nakupine predstavljaju šum.
4. Faza "gladenja" valnog oblika. Valni oblik izgovorene riječi sastoji se od segmenata. Susjedni segmenti sadrže često vrlo različite vrijednosti, pa valni oblik nije gladak. Gladenje je realizirano pomoću jedne varijacije metode pomičnih prosjeka.
5. Faza omeđivanja izgovorene riječi. Pronalaze se lijevi i desni rub valnog oblika izgovorene riječi, kao i najveća vrijednost pojedinog segmenta na valnom obliku. Ako se lijevi i desni rub valnog oblika poklapaju, tada to znači da je vrijednost svih segmenata jednaka nuli, tj. uopće nije izgovorena riječ - pojavili su se samo šumovi. U tom slučaju prepoznavanje se završava, a kao rezultat daje se kod greške.
6. Faza transformacije valnog oblika izgovorene riječi u standardni format. Riječ koju treba prepoznati moguće je izgovoriti na različite načine s obzirom na brzinu i

4 Vrijeme snimanja zvuka iznosi oko 1,5 sekundi. To daje frekvenciju očitavanja EAR utičnice od oko 44 KHz. Po Nyquist- ovom teoremu (Byte XII/89 - Sound and image processing, str. 240 - ) testiranje ulaza treba provoditi dvostrukom brzinom u odnosu na najvišu moguću frekvenciju ulaza da bi se dobila zadovoljavajuća kvaliteta). Najviša frekvencija koju čuje ljudsko uho iznosi oko 20 KHz.

glasnoću izgovaranja. Da bi valni oblici bili međusobno usporedivi, mora ih se prethodno preoblikovati u standardni format. Korekcijom visine stupaca valnog oblika djelomično se poništava utjecaj različite glasnoće izgovaranja riječi. Korekcijom širine zapisa izgovorene riječi (vidi: faza (5)) na standardnu od 64 bajta poništava se utjecaj različite brzine izgovaranja riječi.

7. Uspoređivanje valnih oblika. Uspoređuju se visine korespondentnih stupaca 64-bajtnog zapisa upravo izgovorene riječi s prvim zapisom u memoriji (također u 64-bajtnom formatu). Na temelju toga izračunava se stupanj različitosti (u %). Otprije zapisani valni oblik, uz kojeg se dobije najniži stupanj različitosti smatra se najslabijim, Kao rezultat prepoznavanja dobije se redni broj najslabijeg valnog oblika.



Slika 3. Postupak prepoznavanja izgovorenih riječi



### 3. SLIČNOSTI POSTUPKA ZA PREPOZNAVANJE OBLIKA S POSTUPKOM ZA PREPOZNAVANJE IZGOVORENIH RIJEČI

U toku izrade programa za prepoznavanje oblika došao sam na ideju da bi se isti postupak prepoznavanja s određenim modifikacijama mogao koristiti i za prepoznavanje izgovorenih riječi. Postupak se u oba slučaja sastoji od slijedećih faza:

1. Digitalizacija predmeta prepoznavanja;
2. Izdvajanje predmeta prepoznavanja iz okoline, uključujući i otklanjanje šumova;
3. Transformacija predmeta prepoznavanja u standardni format i
4. Uspoređivanje predmeta prepoznavanja sa zapisima koji postoje otrpije u memoriji i pronalaženje najbližijeg. Najveća razlika je u izvedbi 1. faze: kod programa za prepoznavanje oblika (u daljnjem tekstu: PPO) to je riješeno pomoću potprograma za crtanje<sup>5</sup>, a kod programa za prepoznavanje izgovorenih riječi (u daljnjem tekstu: PPIR) koristi se potprogram koji očitava stanje Spectrumove EAR utičnice (vidi: POSTUPAK PREPOZNAVANJA IZGOVORENIH RIJEČI, faze (1) i (2)). U ostalim fazama postupak je analogan, a razlika je u tome što se kod PPO oblici prikazuju pomoću binarnih matrica<sup>6</sup>, dok se kod PPIR izgovorene riječi prikazuju pomoću niza segmenata koji zajedno čine valni oblik izgovorene riječi, tj. pomoću jednodimenzionalne matrice čiji elementi mogu poprimati vrijednosti 0 - 255. Analizirajući programe po fazama (počevši od faze (2)) možemo ustanoviti slijedeće: - u fazi (2):

---

5 Tako je riješeno jer program kao takav ima uglavnom samo demonstrativnu svrhu.

6 Binarna matrica je matrica koja ima elemente 0 i 1.

- a.) Kod PPO šumom se smatraju svi oblici čiji okvir je manji od minimalnog, a kod PPIR sve skupine segmenata čije zbrojene vrijednosti čine nakupine koje su manje od minimalne.
- b.) Predmet prepoznavanja kod PPO izdvaja se iz okoline tako da mu se sve pripadajuće točke oboje, uz istovremeno određivanje okvira (prozora), dok je kod PPIR dovoljno pronaći samo krajnji lijevi i krajnji desni segment čija vrijednost je različita od nule<sup>7</sup>, - u fazi (3): Kod PPO izdvojeni predmet prepoznavanja transformira se u binarnu matricu 32\*32, a kod PPIR se transformira u jednodimenzionalnu matricu od 64 elementa, s tim da najmanje jedan element ima maksimalnu vrijednost (255). U oba slučaja kod transformacije može se dio informacije izgubiti, ali to obično nije značajno. - i u fazi (4): Kod PPO ispituje se poklapanje matrica, na temelju čega se utvrđuje stupanj sličnosti, a kod PPIR uspoređuju se apsolutne vrijednosti elemenata jednodimenzionalne matrice, na temelju čega se utvrđuje stupanj različitosti.

## ZAKLJUČAK

U radu su opisani algoritmi za dva različita tipa prepoznavanja uzoraka pomoću kompjutera. Značajno je to da su algoritmi slični. Sličnost se sastoji u tome što su četiri osnovne faze prepoznavanja u oba slučaja iste. Razlike su u načinu realiziranja tih osnovnih faza, premda i tu postoji analogija. Veliki problem je bio nedostatak odgovarajuće literature s područja prepoznavanja uzoraka pomoću kompjutera. Slični postupci spomenuti su u knjizi (L1. str. 125-153). Postupci nisu detaljno opisani, a po tamo navedenim karakteristikama opisani algoritmi spadali bi u tu grupu.

---

7 To je samo zbog toga jer PPIR dozvoljava da se izgovori samo jedna riječ; inače bi postupak izdvajanja bio znatno kompliciraniji.

## LITERATURA

1. L. Gyergyek, N. Pavešić, S. Ribarić: Uvod u raspoznavanje uzoraka, Tehnička knjiga, Zagreb, 1988.
2. M. Kantardžić, A. Delić: Principi raspoznavanja uzoraka, Svjetlost, Sarajevo, 1984.
3. G. Smarte: Sounds (U grupi članaka: Sound and image processing), Byte, 12/89, str. 240.
4. W. Tang: Spectrumov mašinski jezik za apsolutne početnike, Prosveta, Beograd, 1987.
5. J. Skuljan: Operacija fill, RAČUNARI br.12 - 2/1986., str.36-38.
6. P. Robinson: Mastering the 68000 microprocessor, Tab books Inc., 1985.
7. S. Ditrich: Atari ST peeks & pokes, Data Becker, 1985.

*Radošević D. Recognizing Two-dimensional Shapes and Spoken Words Using Similar Algorithms*

## SUMMARY

*There are two algorithms described in this paper: one for the recognition of two-dimensional shapes and the other for the recognition of spoken words. The verification programs were made on the basis of these algorithms. The paper shows that similar algorithms can be used for the recognition of two-dimensional shapes and spoken words. The elements of similarity are stressed.*