

Prepoznavanje brojeva

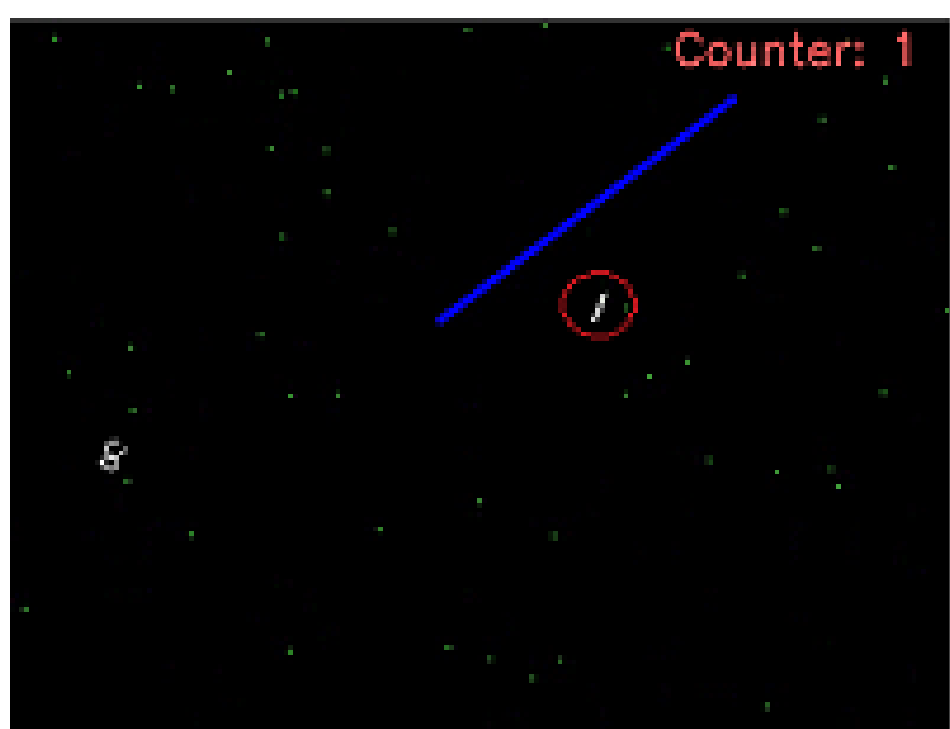
Autor : Acimović Danilo

Broj indeksa : ra37-2013

Asistent: Miroslav Kondić

UVOD U TEMU

Video zapis poseduje jednu pokretnu liniju koja je uvek iste boje. Cifre se kreću za slučajan broj koraka i prolaze iza pokretne linije. Algoritam vrši sabiranje svih cifara koje prođu ispod linije. U tekstu je objašnjen postupak kojim se dolazi do rešenja, kao i tačnost postignutog rešenja. Cilj je bio postići tačnost prepoznavanja od bar 90%



slika 1.1

PODACI

Cifre koje su korišćene u video snimcima preuzete su iz MNIST baze podataka. MNIST baza podataka predstavlja ogromnu bazu ručno pisanih cifara. Ova baza se, u velikoj meri, koristi za obučavanje i testiranje u mašinskom učenju. MNIST baza podataka sadrži 70000 elemenata. Sve slike u MNIST skupu su dimenzija 28x28 piksela. Primeri nekih cifara iz MNIST baze podataka date su na slici 2.1

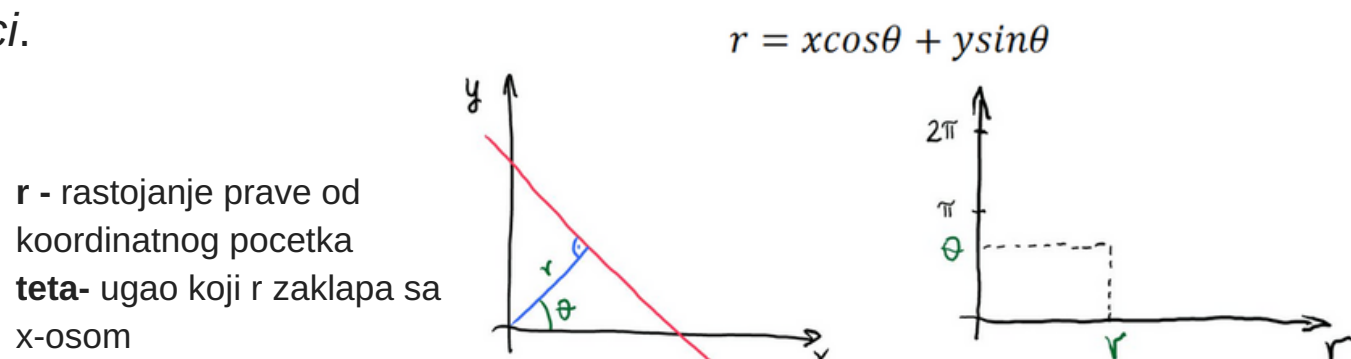


slika 2.1

PREPOZNAVANJE OBJEKATA

Na datom video snimku potrebno je prepoznati dve vrste objekata. Prvo je potrebno detektovati liniju plave boje i izračunati njene koordinate, a zatim je potrebno detektovati objekte, koji predstavljaju cifre i koji su prošli ispod linije. Za detekciju linije je korišćena **HOUGH TRANSFORMACIJA**.

Hough transformacija služi za detektovanje bilo koje figure, koja može da se predstavi u matematičkoj formi pa tako i za *detekciju pravih linija u rasterskoj slici*.



Prostor digitalne rasterske slike

Parametarski prostor

slika 3.1 Hough transformacija: **detekcija linije**

Nakon primene hough transformacije dobijaju se koordinate početka i kraja linije na video snimku, koje nam omogućavaju da odredimo koja cifra je prošla liniju, a koja nije. Za detekciju cifara koristi se MNIST skup. Prilikom detekcije potrebno je eliminisati sve objekte sa slike koji nisu cifre, da bismo to postigli potrebno je izvršiti thresholding operaciju. (Ako je vrednost piksela veća od threshold vrednosti dodeljena je vrednost jedan, u suprotnom je dodeljena vrednost 0.) Zatim je potrebno izvršiti eroziju i dilataciju. Nakon primene ovih operacija, pronalazimo regione na slici. Svaki region predstavlja jednu cifru. Regioni mogu biti različite veličine. Formiramo nove slike za svaki region, dimenzija 28x28 piksela, kako bismo

mogli izvršiti prepoznavanje u MNIST bazi podataka.

Na novoj slici cifra se nalazi u centru. Prepoznavanje sa MNIST skupom je vršeno, tako što se slika transformiše da se cifra nalazi u gornjem levom uglu. I MNIST skup se transformisao na isti način.

Nakon toga se vrši poređenje slike sa slikama iz MNIST skupa. Slika koja poseduje najmanju razliku sa prosledjenom slikom predstavljaće pronađeni broj, odnosno, njen target iz MNIST skupa ce dati cifru koja odgovara prosleđenoj slici.

REZULTATI

Testiranje ispravnosti rešenja je vršeno na primeru od 10 video snimaka. Tačnost se proveravala tako što se vrši suma dobijenih zbrova cifara koje su prešle ispod linije, za svaki video, a zatim se ta suma podeli sa tačnom sumom zbrova iz svakog video.. Primenom prethodno navedenog postupka dobijena je tačnost od 94.77% .U pojedinim video snimcima došlo je do odstupanja od tačnosti rešenja, dok je u većini slučajeva tačnost bila visoka, pa je zbog toga postignut visok procenat tačnosti. Za još veću tačnost, potrebno je obezbediti prepoznavanje preklopljenih objekata, prilikom prelaska ispod linije. Preklopljeni objekti nisu u potpunosti rešeni ovim algoritmom.