

Prepoznavanje i sumiranje brojeva u videu

Predmet: Soft Computing

Student: Marko Zorić RA187/2014

Profesor: Vuk Malbaša

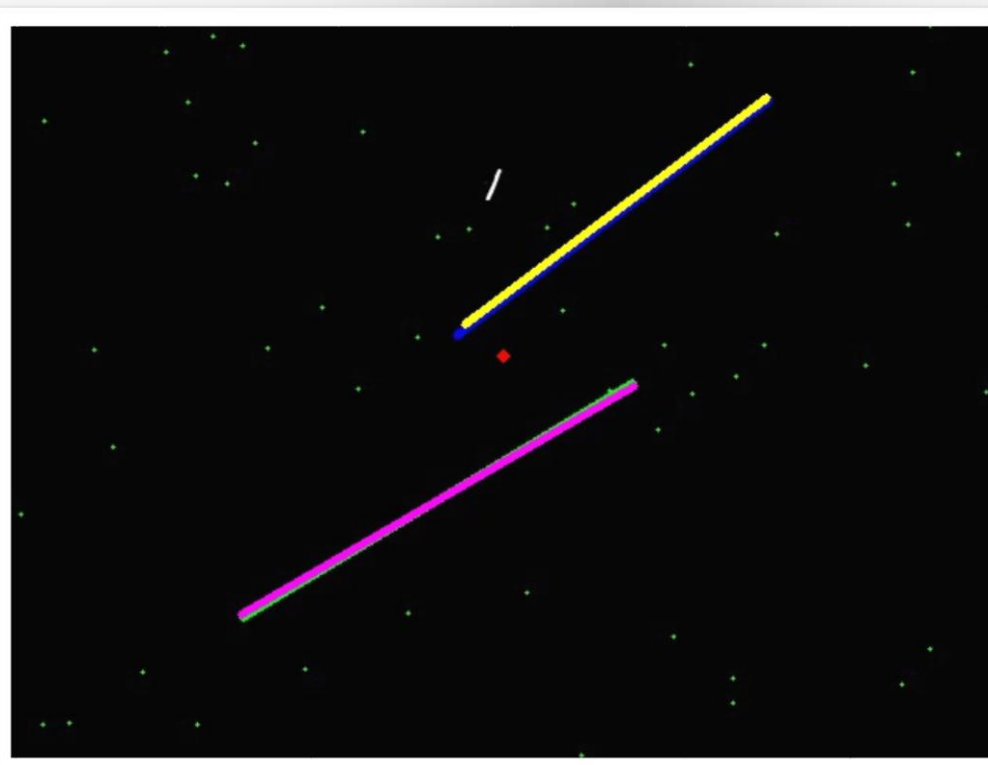
Asistent: Ivan Perić

OPIS PROBLEMA

Imamo deset video klipova trajanja 30 sekundi. Svaki video klip sadrži dve nepokretne linije, jednu zelenu i jednu plavu koje se nalaze u donjoj polovini ekrana. Tokom trajanja video klipa iz gornje ili leve ivice videa se pojavljuju ručno pisane cifre. Cifre se kreću prema dvema linijama proizvoljnim putanjama. Neophodno je detektovati cifre koje pređu preko linija. Cifre koje pređu preko jedne linije se sabiraju sa krajnjom sumom, a cifre koje pređu preko druge linije se oduzimaju od krajnje sume. Problem dodatno otežava sneg (u vidu zelenih tačkica) koji se pojavljuje na proizvoljnim mestima na videu.

NAČIN REŠAVANJA

Linije se detektuju pomoću Hough transformacije. Skup podataka pomoću kojeg ćemo detektovati cifre je MNIST database predstavlja skup od 70000 ručno napisanih cifara. U pitanju su 70000 crno-belih slika (cifre su bele, pozadina slike crna) gde se cifre nalaze u centru slike. Video ćemo obrađivati frejm po frejm. Svaki frejm ćemo filtrirati da bismo uklonili sneg sa njega. Svakoj cifri na videu će biti dodeljen id i vršićemo praćenje pozicije svake cifre na ekranu. Ukoliko se centar cifre poklopi sa jednom ili obema linijama ili ukoliko se nalazi u neposrednoj blizini linije, cifra će biti sabrana ili oduzeta u zavisnosti od linije koju je presekla.



ALGORITAM

Kao što je rečeno, linije ćemo naći na osnovu Hough transformacije. OpenCV ima dve ugrađene funkcije za Hough transformaciju - standardnu i probablističku. Svaka prava se može opisati moću jednačine $y=kx+n$ ili pomoću $r=x\cos(\theta) + y\sin(\theta)$ gde je r normala koordinatnog početka na pravu, a θ je ugao koji normala zaklapa sa x-osom. Standardna Hough transformacija kao svoje rešenje vraća vektor vrednost (θ, r) . Međutim, mi koristimo probabalističku Hough transformaciju. Ona za svoje rešenje vraća vektor početnih i krajnjih tačaka nađenih linija.

Potrebno je odrediti koje od nađenih linija pripadaju prvoj, a koje drugoj liniji sa videa i pomoću njih odrediti dve linije sa svojim početnim i krajnjim tačkama koje najbliže odgovaraju dvema linijama sa videa. Ovo je odrađeno tako što je pronađena srednja tačka maksimalne razdaljine između dve tačke koje je vratila Hough transformacija. Sve tačke koje se nalaze iznad srednje tačke pripadaju jednoj liniji, ostale pripadaju drugoj.

Cifre iz MNIST database-a ćemo smestiti u slike veličine 28x28. Te slike ćemo upoređivati sa svih 70000 podataka iz MNIST database-a. Ona koja se podudara u najvećoj meri biće vraćena kao prepoznata.

REZULTATI

Procenat uspešnosti prepoznavanja varira od videa do videa. Prepoznavanje cifara otežava preklapanje dve ili više cifre. Ukoliko dođe do tog slučaja u videu, verovatno će se desiti pogrešno prepoznavanje u programu. Ukupna tačnost svih videa je oko 80%.