

Obrada video zapisa i sumiranje prepoznatih brojeva

Dragan Dulic RA189/2014

Uvod u problem

Kao sto sam naslov kaze u projektu cemo raditi obradu video zapisa i manipulirati dobijenim podacima. Problem je sledeci- video zapis poseduje jednu pokretnu liniju koja je uvek iste boje. Tokom 30 sekundi video zapisa po ekranu se krecu cifre razlicitih stilova pisanja. Kretanje nije pravolinijsko i svaka cifra se kreće proizvoljnom brzinom. Prepoznavanje dodatno otežavaju tackice u pozadini video zapisa. Nas glavni zadatak je da detektujemo cifre koje prođu ispod linije i da ih saberemo, a ostale cifre ignorisemo. Skup podataka nad kojim cemo vrsiti obradu predstavlja 10 video zapisa i dostupni su zajedno sa projektom kao resenje problema.

Proces

Za pocetak liniju detektova pomocu Hough transformacije. Skup podataka koji ce nam pomoci oko detekcije cifara je MNIST DATABASE sa 60000 slicica dimenzije 28x28. To su crno-bele slike u cijem centru se nalaze cifre. Video cemo izdeliti u frejmove. Zatim cemo filtriranjem slike da uklonimo smetnje(tackice u pozadini) i da ostavimo samo regione sa ciframa. Svakom objektu dodelimo ID i detektujemo cifru cim se pojavi u frame-u. Vrsimo pracenje te cifre sve dok ne napusti okvire frame-a. Kada se centar objekta, na kom se nalazi cifra, poklopi sa centrom linije i njenom blizom okolinom- tu cifru dodajemo u konacni zbir.

Algoritmi

Prvo cemo pomenuti Hough transformaciju za detekciju linije. Svaka linija se moze opisati pomocu jednacine $y=m*x+c$, kao i preko $r=x*\cos(\theta)+y*\sin(\theta)$. r je duzina od koordinatnog pocetka do linije, a θ je ugao koji zaklapa x osa sa r duzi. Uz pomoc funkcije, koju nam je obezbedio open cv, HoughLinesP mi dobijamo vektor linija. Vektor sadrzi dve kordinate iz kojih kroz iteraciju nekoliko iteracija dobijamo one koje odgovaraju krajnjim tackama linije.

Sto se tice algoritma za prepoznavanja cifre prvo je potrebno izdvojiti cifru iz regiona i smestiti je u sliku dimenzija 28x28. Zatim tu sliku uporediti sa svih 60000 podataka u MNIST DATASET-U. Redni broj izlaza koji je najvise pobudjen, odgovara broju sa slike.

Rezultati

Cilj projekta je tacno detektovanje linje i cifara koje su ispod nje prosle. Rezultati detekcije se vide prilikom pokretanja aplikacije. Uspesnost detektovanih cifara je u proseku 92% za svih 10 video klipova. Na nekima od njih tacnost je i do 100%.

Procenat tacnosti se moze povecati ukoliko bi proces dilatacije bio odradjen dodatno za neke video klipove.