

Animacija virtualnih likova pomoću neuronske mreže OpenPose

Projekt iz predmeta Neuronske mreže

Nikola Špehar
*Fakultet elektrotehnike i
računarstva*
Zagreb, Hrvatska
nikola.spehar@fer.hr

Vedran Lončar
*Fakultet elektrotehnike i
računarstva*
Zagreb, Hrvatska
vedran.loncar@fer.hr

Ljubomir Jerončić
*Fakultet elektrotehnike i
računarstva*
Zagreb, Hrvatska
ljubomir.jeroncic@fer.hr

Lucija Krezo
*Fakultet elektrotehnike i
računarstva*
Zagreb, Hrvatska
lucija.krezo@fer.hr

I. UVOD

Sve je veća potreba za analizom i virtualizacijom ljudskih pokreta u području računalne grafike. Tradicionalne tehnike animacije često su vremenski zahtjevne i zahtijevaju skupu opremu te stručnost animatora. Zbog toga se sve veća pažnja posvećuje metodama koje se temelje na analizi standardnih 2D slika i videozapisa. Jedan od najpoznatijih alata za prepoznavanje ljudskih pokreta iz slike ili videozapisa je OpenPose, koji koristimo u sklopu našeg projekta. OpenPose pomoću neuronske mreže prepoznaje i prati poze osoba te kao rezultat generira mrežu točaka i linija koje predstavljaju položaj tijela. Dobiveni podaci mogu se koristiti za analizu pokreta ili za njihovu vizualizaciju u obliku pojednostavljenih skeletnih prikaza. Cilj projekta je razviti sustav koji koristi OpenPose za procjenu poze iz videozapisa te omogućuje 2D skeletnu vizualizaciju pokreta i osnovnih izraza lica u Blenderu. Ključni izazovi uključuju:

- preciznu detekciju i praćenje poze u realnom vremenu,
- stabilnu vizualizaciju pokreta kroz vrijeme,
- utjecaj zaklanjanja tijela tijekom izvođenja pokreta.

II. PREGLED POSTOJEĆIH RJEŠENJA

Analiza i vizualizacija ljudskih pokreta temelje se na nekoliko ključnih područja: 2D procjeni poze, 3D rekonstrukciji poze te na prijenosu pokreta. U kontekstu ovog projekta fokus je stavljen na metode 2D procjene poze i njihovu primjenu u vizualizaciji pokreta, bez animacije punog 3D modela.

2D procjena poze

2D procjena poze predstavlja ključni temelj suvremenih sustava koji animiraju videozapise. Cilj ovih metoda je odrediti položaje ključnih točaka ljudskog tijela i lica. OpenPose je jedan od najpoznatijih sustava za 2D procjenu te koristi koncept Part Affinity Fields, tj. mehanizam koji otkriva veze među zglobovima i omogućuje detekciju više osoba u stvarnom vremenu. Najnovije verzije dodatno proširuju taj kapacitet na cijelo tijelo (lice, šake, stopala).

Vizualizacija 2D poze

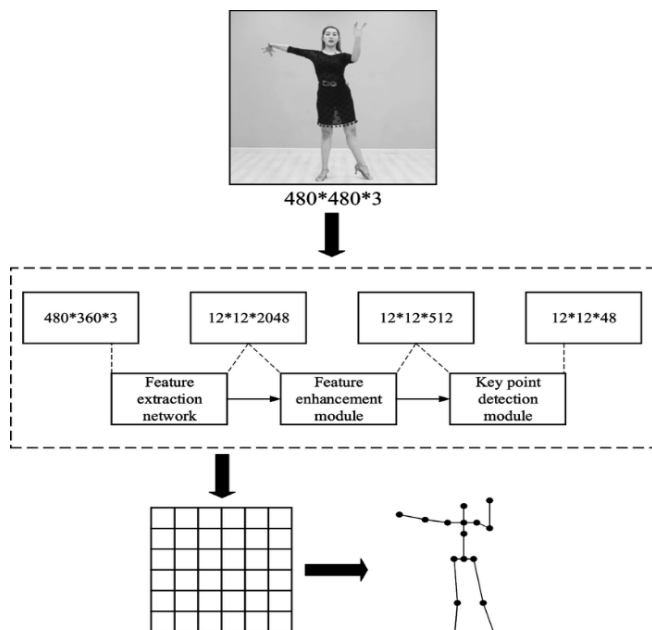
Za potpunu animaciju 2D likova potrebna je informacija o Dobiveni rezultati 2D procjene poze najčešće se koriste za vizualizaciju pokreta u obliku skeletnih prikaza. Takvi prikazi sastoje se od točaka koje predstavljaju zglobove i linija koje ih povezuju. Prednost ovakvog pristupa je jednostavnost i preglednost. U praksi se najčešće koriste za analizu pokreta, sportske aplikacije, edukacije, itd.

Ograničenja 2D pristupa

Iako 2D procjena poze omogućuje jednostavnu i brzu analizu, ona ima svoja ograničenja. Budući da ne sadrži informaciju o dubini, sustav ne može precizno odrediti pokrete u trećoj dimenziji. Uz to, zaklanjanje dijelova tijela ili brzi pokreti mogu negativno utjecati na točnost.

III. OPIS RJEŠENJA PROJEKTOG TIMA

U svrhu ovog projekta razvijen je sustav za animaciju virtualnog lika koji se temelji na analizi ljudskih pokreta iz videozapisa pomoću neuronske mreže OpenPose. Sustav omogućuje prijenos pokreta i izraza lica stvarne osobe na model, pri čemu se koriste standardni videozapisi, bez potrebe za posebnom opremom za snimanje pokreta. Poseban naglasak projekta stavljen je na animaciju lica, sukladno projektnom zadatku.



Slika 1. Shematski prikaz sustava za izvlačenje poze i vizualizaciju pokreta.

Radno okruženje i korišteni alati

Za izvođenje OpenPose sustava korišten je Docker, čime je osigurana jednostavna instalacija potrebnog softvera. Korištena je CPU-only verzija OpenPose-a budući da GPU verzija nije kompatibilna s dostupnim programima. Iako je izvođenje na procesoru bilo nešto sporije, pokazalo se dovoljnim za obradu videozapisa korištenih u ovom projektu.

OpenPose je pokretan unutar izoliranog Docker okruženja, pri čemu su ulazni i izlazni direktoriji mapirani na lokalni datotečni sustav. Na taj je način omogućeno jednostavno upravljanje videozapisima i dohvat rezultata obrade.

Snimanje i priprema ulaznih videozapisa

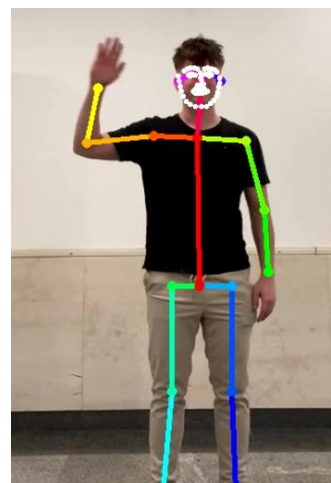
Ulazni podaci dobiveni su snimanjem jednostavnih ljudskih pokreta. Snimani pokreti uključivali su mahanje rukom, dodirivanje glave ramena, koljena i stopala, te nekih učestalih "sportskih" radnji kao što su proslava, imitiranje vođenja košarkaške lopte itd. Videozapisi su snimani u kontroliranim uvjetima kako bi se osigurala što veća točnost detekcije. Tijekom snimanja pazilo se da je cijelo tijelo osobe vidljivo unutar kadra, te da je prostorija dobro osvijetljena. Takvi uvjeti omogućuju OpenPose-u preciznije određivanje ključnih točaka tijela i lica.



Slika 2. Kadar ulaznog videozapisa.

Procjena poze i izvlačenje ključnih točaka

Snimljeni videozapisi obrađeni su korištenjem OpenPose sustava, koji za svaki kadar detektira ključne točke tijela i lica te generira njihove dvodimenzionalne koordinate. U ovom projektu korištena je i opcija detekcije lica. Kao rezultat obrade, za svaki videozapis generiran je skup JSON datoteka koje sadrže podatke o položaju ključnih točaka za svaki pojedini kadar. Iako OpenPose omogućuje detekciju cijelog tijela, podaci o tijelu korišteni su prvenstveno za osnovnu orijentaciju pokreta, dok su ključne bile i točke lica.



Slika 3. Prikaz ključnih točaka.

```
"frames": [
  {
    "frame": 0,
    "keypoints": {
      "Head": {
        "x": -0.05805799999999999,
        "y": 0.5187,
        "confidence": 0.910541,
        "type": "body"
      },
      "Neck": {
        "x": -0.06694000000000006,
        "y": 0.36439200000000005,
        "confidence": 0.926306,
        "type": "body"
      }
    }
  }
]
```

Slika 4. Primjer zapisa ključnih točaka tijela i lica za jedan kadar u JSON formatu.

Obrada i utvrđivanje podataka

Budući da OpenPose generira veliki broj datoteka i ključnih točaka, razvijena je Python skripta za dodatnu obradu podataka. Skripta učitava JSON datoteke svih kadrova pojedinog videozapisa te izdvaja samo odabrane ključne točke bitne za animaciju.

Nakon obrade, svi podaci objedinjeni su u jednu JSON datoteku po videozapisu, čime je smanjena količina podataka i pojednostavljen njihov prijenos u alat za animaciju. Takav format omogućio je pregledniju i učinkovitiju obradu kasnije.

Prijenos pokreta i animaciju u Blenderu

Za animaciju virtualnog lika korišten je alat Blender. Unutar Blender okruženja primijenjena je skripta koja učitava objedinjene podatke o ključnim točkama te ih preslikava na kostur i kontrolne točke modela.

Dobiveni podaci korišteni su za animaciju osnovnih pokreta tijela te izraza lica, pri čemu je posvećena pažnja povezivanju kadrova kako bi se postigla glatka animacija. Na taj način virtualni lik reproducira pokrete i izraze lica iz izvornog videozapisa.

IV. OPIS EKSPERIMENTALNIH REZULTATA

Eksperimentalna procjena sustava provedena je na više snimljenih videozapisa koji prikazuju osnovne ljudske pokrete, uključujući mahanje rukom te dodirivanje glave, ramena, koljena i stopala, kao i na nekim malo naprednijim pokretima kao što su imitacija vođenja košarkaške lopte, zamahivanje rukama itd. Cilj eksperimenta bio je ispitati točnost i stabilnost praćenja pokreta te kvalitetu detekcije ključnih točaka lika i njihovu vizualizaciju u alatu Blender.

Praćenje pokreta tijela

Analizom dobivenih rezultata uočeno je da sustav u većini slučajeva uspješno prati osnovne pokrete tijela. Pokreti ruku i nogu jasno su vidljivi u 2D skeletnoj animaciji te su vremenski dobro usklađeni s izvornim videozapisima. Posebno su dobri rezultati kod sporijih i jasno definiranih pokreta, poput mahanja rukom ili dodirivanja pojedinih dijelova tijela.

Manje pogreške u praćenju javljaju se kada je neka ključna točka zaklonjena drugim dijelom tijela tijekom izvođenja pokreta, posebice kod složenijih pokreta (zamahivanje rukama). U tim slučajevima dolazi do odstupanja u položaju pojedinih zglobova. Unatoč tome, ukupni prikaz pokreta ostaje prepoznatljiv.



Slika 5. Kadar dobivene animacije.

Analiza detekcije i animacije lica

Rezultati detekcije lica pokazali su ograničenu uspješnost. Jasno su vidljivi osnovni obrisi lica, pri čemu se prepoznaju položaji očiju i nosa. Ostali detalji lica, poput obrisa usana ili nekih još detalja, nisu izraženi u prikazu. Takvi rezultati su i očekivani s obzirom na korištenje 2D procjene poze. Unatoč ograničenjima, pokreti glave i osnovna orijentacija lica se uspješno prenose.

V. DISKUSIJA I USPOREĐIVANJE REZULTATA

Što se tiče konkretnih primjera u kojima se OpenPose koristio za 2D animaciju, ističe se Everybody Dance Now (Chan et al., 2019). U tom radu OpenPose se koristi za detekciju 2D ključnih točaka ljudskog tijela iz videozapisa, koje zatim služe kao ulaz za daljnju animaciju.

U našem projektu, kao i u referentnom primjeru, OpenPose je uspješno detektirao ključne točke koje se mogu koristiti za animaciju pokreta. Razlika između ova dva pristupa očituje se u načinu interpretacije i vizualizacije dobivenih ključnih točaka. Dok Everybody Dance Now koristi detektiranu 2D pozu kao ulaz u generativni neuronski sustav koji proizvodi foto realističan 2D video osobe u pokretu, naš projekt koristi iste podatke za izravno upravljanje 2D animacijom unutar Blender okruženja. Time se izbjegava potreba za dodatnim treniranjem neuronskih mreža, a naglasak se stavlja na kontrolu animacije i jasnoću pokreta.

Analizom rezultata može se primijetiti da je stabilnost animacije u našem projektu usporediva s referentnim primjerima, uz povremenu prisutnost blagog podrhtavanja karakterističnog za 2D sustave temeljene na pozama. Međutim, s obzirom na to da se radi o 2D animaciji, takva odstupanja imaju manji negativni utjecaj na konačni dojam nego u slučaju referentnog rada, gdje i manja pogreška može rezultirati izraženijim vizualnim podrhtavanjem.

Važno je naglasiti da Everybody Dance Now generira vizualno realističniji rezultat, ali uz veću računalnu složenost i manju razinu izravne kontrole nad animacijskim procesom. Suprotno tome, naš pristup omogućuje jednostavniju prilagodbu izgleda lika, lakše uređivanje animacije te jasnu integraciju u standardni animacijski pipeline.

VI. ZAKLJUČAK

U ovom projektu OpenPose je korišten kao alat za detekciju 2D ključnih točaka ljudskog tijela iz videozapisa, koje su se zatim koristile za animaciju unutar Blender okruženja. Korištena je CPU verzija OpenPosea, koja je sporija u odnosu na GPU verziju, ali se pokazala dovoljno učinkovitom za potrebe ovog projekta. Svaki videozapis obrađen je unutar Docker kontejnera, pri čemu su za svaki kadar generirane JSON datoteke s podacima o detektiranim ključnim točkama. Ti su podaci naknadno obrađeni i prilagođeni kako bi se mogli koristiti za izravno upravljanje 2D animacijom u Blenderu.

U ovom projektu OpenPose je korišten kao alat za detekciju 2D ključnih točaka ljudskog tijela iz videozapisa, koje su se zatim koristile za animaciju unutar Blender okruženja. Korištena je CPU verzija OpenPosea, koja je sporija u odnosu na GPU verziju, ali se pokazala dovoljno učinkovitom za potrebe ovog projekta. Svaki videozapis obrađen je unutar Docker kontejnera, pri čemu su za svaki kadar generirane JSON datoteke s podacima o detektiranim ključnim točkama.

Ti su podaci naknadno obrađeni i prilagođeni kako bi se mogli koristiti za izravno upravljanje 2D animacijom u Blenderu.

LITERATURA

- [1] Z. Cao, G. Hidalgo, T. Simon, S.-E. Wei, and Y. Sheikh, "OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields,"
- [2] C. Chan, S. Ginosar, T. Zhou, and A. A. Efros, "Everybody Dance Now," presented at *International Conference on Computer Vision (ICCV)*, 2019.
- [3] S.-E. Wei, V. Ramakrishna, T. Kanade, and Y. Sheikh, "Convolutional Pose Machines," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2016.
- [4] Blender Foundation, "Blender 3D: Noob to Pro," Blender Documentation, Available: <https://docs.blender.org/manual/en/latest/> [Online]. Available: <https://codeocean.com/capsule/4989235/tree>