

TASVIRLARDAGI ASOSLANGAN INSON HARAKATINI ANIQLASH OPENPOSE YORDAMIDA ALGORITMI

Fayzullayeva Z.I.

Muhammad al-Xorazmiy nomidagi TATu dotsenti.
zarnigor18z02@gmail.com

<https://doi.org/10.5281/zenodo.16760042>

Mazkur tadqiqot OpenPose vositasidan insonning kalit nuqtalarini generatsiya qilish uchun foydalangan holda, suratlar asosida inson harakatlarini aniqlovchi tizimni taqdim etadi. RGB-ga asoslangan usul tasvirni qayta ishlash bo'yicha oldingi bilimlardan foydalanib, yuqori aniqlikdagi harakatni aniqlash imkonini beradi, lekin hisoblash quvvati va saqlash resurslariga yuqori talablar qo'yadi, shuningdek, fon shovqinlari va yorug'lik o'zgarishlariga sezgir. Aksincha, skeletga asoslangan usul kamroq hisoblash resurslari talab qiladi va yorug'lik yoki fon ta'siridan kamroq ta'sirlanadi, ammo kontekst ma'lumotlarining yetishmasligi tufayli cheklovlarga ega.

Harakatni aniqlash jarayonida bir qator muhim muammolar, jumladan, harakatga aloqasiz fon ma'lumotlari, yorug'lik o'zgarishlari, maxfiylik masalalari va yuqori hisoblash resurslari talabi mavjud. Skelet ma'lumotlar to'plamlari tashqi muhit omillaridan mustaqilligi, masalan, fon va belgilarga e'tibor qaratish qobiliyati va kamroq hisoblash resurslariga ehtiyoj sezishi kabi afzalliklarni taqdim etadi [21][22][12]. Shu bilan birga, skeletga asoslangan harakatni tanib olish usullari kontekstual ma'lumotlarni yo'qotish muammosiga duch keladi va xom video ma'lumotlarga bevosita qo'llanmaydi, chunki dastlab skelet ma'lumotlari olinishi talab qilinadi. Presti Liliana Lo va boshqalar [23] 3D skeletga asoslangan harakatni aniqlash texnologiyalarini tahlil qilib, yangi muammo va motivatsiyalarni muhokama qilgan bo'lib, ushbu muammolar hali ham dolzarb sanaladi. Ren Bin va boshqalar [24] chuqur o'rganishga asoslangan usullarni o'rganib, RNN (takrorlanadigan neyron tarmoqlar), CNN (konvolutsion neyron tarmoqlar) va GCN (graf konvolutsion tarmoqlar) kabi usullarning rivojlanish traektoriyasi va asosiy texnologik yutuqlarini ko'rib chiqqan.

Biz taklif qilayotgan usul yuqorida qilingan ishlardan farqi shundaki ko'p odamli tasvirlarda ishlashi, Part Affinity Fields (PAFs) usulining joriy etilishi va real vaqt rejimida samaradorligidadir. Ilgari mavjud algoritmlar, DeeperCut [1] yoki CPM (Convolutional Pose Machines) [2], odatda bitta odamning pozasini aniqlashga yo'naltirilgan bo'lib, ko'p odamli tasvirlarda kalit nuqtalar noto'g'ri guruhlanardi, ya'ni bir shaxsning bo'g'imlari boshqa shaxsga noto'g'ri bog'lanish holatlari ko'p bo'lardi. OpenPose PAFs usuli orqali bo'g'imlar orasidagi bog'lanishlarni vektor maydonlari sifatida ifodalaydi, bu esa ko'p odamli tasvirlarda kalit nuqtalar guruhlashni aniq amalga oshiradi. Bu usul bir tasvirda 10 tagacha odamning skelet tuzilishini to'g'ri aniqlay oladi.

$$S_j(x, y) = \exp\left(-\frac{\|(x, y) - p_j\|^2}{\sigma^2}\right) \quad (1)$$

Ishonchlilik xaritasi esa quridagi formulasi orqali hisoblanadi:

bu yerda $S_j(x, y)$: j -chi kalit nuqta uchun (x, y) koordinatadagi ishonch darajasi. p_j

-Haqiqiy kalit nuqta koordinatasi. σ -Gauss taqsimotidagi dispersiya parametri.

OpenPose arxitekturasida RGB tasvirlar yoki video kadrlarni ma'lumotlarni kiruvchi ma'lumot sifatida qiladi hamda ushbu kiruvchi ma'lumotlarni, chuqur konvolyutsion neyron tarmoq (CNN) yordamida 2D anatomik nuqtalarni (masalan, 18 yoki 135 ta nuqta, shu jumladan tana, qo'l, yuz va oyoq bo'g'imlari) aniqlaydi. Confidence maps) orqali nuqtalarni joylashuvi aniqlanadi. Tana asosiy nuqtalarini aniqlash hamda pozitsiyani baholash uchun zarur bo'lgan ishonch xaritalari va bo'g'imlararo bog'liqlikni ifodalovchi yaqinlik maydonlari to'plami, har bir shaxs uchun har bir nuqtaning 2D koordinatalari (masalan, yelka, tirsak va boshqalarning x, y pozitsiyalari) hosil qilinadi va (1) formula orqali hisoblanadi.

Kirish tasvirini aniqlash jarayonida konvolyutsion va birlashtiruvchi qatlamlar ketma-ketligidan foydalangan holda xususiyat xaritalari to'plamini olish uchun oldindan qayta ishlanadi. Bu kirish tasviridan past va yuqori darajali xususiyatlarni ajratib olishga yordam beradi. Har bir ishonch xaritasi tasvirning ma'lum bir joyida ma'lum bir asosiy nuqtaning mavjud bo'lish ehtimolini ifodalaydi. Xususiyat xaritalari, tananing turli qismlari o'rtasidagi fazoviy munosabatlarni ifodalovchi qismga yaqinlik maydonlarini yaratish uchun ishlatiladi. Har bir qismning yaqinlik maydoni tasvirning ma'lum bir joyida ikkita tana qismining ulanishi ehtimolini ifodalab beradi. Aniqlash jarayonida ketma-ket kadrlar orasidagi nuqta pozitsiyalarining o'zgarishi harakat dinamikasini aniqlash uchun hisoblanadi. Bo'g'imlar orasidagi masofa va burchak o'zgarishlari (masalan, tirsakning yelka nisbatan masofasi va yo'nalishi). Ishonchlilik xaritasi VGG-19 yoki MobileNet kabi CNN (2) orqali hisoblanadi:

$$F = CNN(I), \quad S = f(F)_{(2)}$$

I - kiritish tasviri (RGB $H \times W \times 3$)

F - xususiyatlar xaritasi (feature map)

f - Konvolyutsion qatlamlar.

Bu xususiyatlar ko'rish burchagi va shaxsiy tana farqlariga chidamli bo'lib, shaxsdan mustaqil va ko'rish burchagiga bog'liq bo'lmagan faoliyat aniqlash uchun mos keladi.

$$L_c(x, y) = \begin{cases} \text{agar}(x, y) \text{ bo'g'im segmentida bo'lsa} \\ 0, & \text{aks holda} \end{cases}$$

V - Bo'g'imlar orasidagi birlik vektor (masalan, tirsakdan bilakka yo'nalish).

$L_c : C$ - chi bo'g'im juftligi uchun PAFs

OpenPose "bottom-up" yondashuvidan foydalanadi, ya'ni avval barcha kalit nuqtalar aniqlanadi, keyin PAFs yordamida ular shaxslarga guruhlanadi. Bu hisoblash samaradorligini oshiradi, chunki shaxslar sonidan qat'i nazar, bir martalik tahlil amalga oshiriladi. Part Affinity Fields (PAFs) ko'pgina algoritmlar bo'g'imlar orasidagi bog'lanishlarni aniqlashda qo'shimcha qatlamlarga yoki post-processing bosqichlariga tayanadi, bu esa aniqlikni pasaytirishi yoki hisoblashni sekinlashtiradi.

OpenPose'ning asosiy farqi uning bottom-up yondashuvi, PAFs usuli va real vaqt samaradorligidir. U ko'p odamli stsenariylarda yuqori aniqlik, MobileNet kabi yengil modellar orqali tezlik va TensorFlow'ning moslashuvchanligini ta'minlaydi. Algoritm ishonchlilik xaritalari va PAFs orqali kalit nuqtalar guruhlashni aniq amalga oshiradi, bu esa boshqa algoritmlardan (DeeperCut, Mask R-CNN) ustunlik qiladi. OpenPose sport, tibbiyot,

xavfsizlik va o'yin industriyasida keng qo'llaniladi va kelajakda 3D poza aniqlash va harakat tahlili sohasida yanada rivojlanib bormoqda.

References:

Используемая литература:

Foydalanilgan adabiyotlar:

1. A Robust Human Activity Recognition Approach Using OpenPose, Motion Features, and Deep Recurrent Neural Network Farzan Majeed Noori^{1(B)}, Benedikte Wallace^{1,2}, Md. Zia Uddin¹, and Jim Torresen^{1,2}.
2. Graph transformer network with temporal kernel attention for skeleton-based action recognition Yanan Liu, Hao Zhang, Dan Xu *, Kangjian He
3. E. Insafutdinov et al., "DeeperCut: A Deeper, Stronger, and Faster Multi-Person Pose Estimation Model," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2016.
4. S.-E. Wei et al., "Convolutional Pose Machines," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2016.
5. K. He et al., "Mask R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, 2017.
6. Z. Cao et al., "Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2017.
7. E. Insafutdinov et al., "DeeperCut: A Deeper, Stronger, and Faster Multi-Person Pose Estimation Model," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2016.
8. G. Moon et al., "V2V-PoseNet: Voxel-to-Voxel Prediction Network for Accurate 3D Hand and Human Pose Estimation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2018.
9. OpenPose GitHub Repository, <https://github.com/CMU-Perceptual-Computing-Lab/openpose>.
10. B. Xiao et al., "Simple Baselines for Human Pose Estimation and Tracking," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018.
11. H. Joo et al., "Panoptic Studio: A Massively Multiview System for Social Motion Capture," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, 2015.
12. J. Liu et al., "NTU RGB+D: A Large-Scale Dataset for 3D Human Activity Analysis," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2016.
13. A. F. Bobick and J. W. Davis, "The Recognition of Human Movement Using Temporal Templates," *IEEE Trans. Pattern Anal. Mach. Intell.*, 2001.
14. S. Patel et al., "A Review of Wearable Sensors and Systems with Application in Rehabilitation," *J. Neuroeng. Rehabil.*, 2012.
15. R. Poppe, "A Survey on Vision-Based Human Action Recognition," *Image Vis. Comput.*, 2010.
16. M. Zyda, "From Visual Simulation to Virtual Reality to Games," *Computer*, 2005.