

VIDEO MONITORING TIZIMLARIDA MAQSADLI KUZATUV JARAYONINING MATEMATIK MODELI VA OPTIMALLASHTIRILGAN ALGORITMLARI

Narziyev Nosir Baxshilloevich
Saidqodirov Xumoyunxon Yashnarjon o'g'li

Toshkent axborot texnologiyalari universiteti, Toshkent, O'zbekiston ATDT kafedrası katta o'qituvchisi, DIF 310-23 guruh talabasi, n.b.narziyev@gmail.com, xumoyun.saidqodirov@gmail.com,

Annotatsiya: Ushbu maqolada video monitoring tizimlarida maqsadli kuzatuv jarayonining matematik modeli va optimallashtirilgan algoritmlari tadqiq etiladi. Tadqiqotning dolzarbligi shundaki, zamonaviy ko'p kamerali monitoring tizimlarida maqsadni uzluksiz kuzatish, bir kameradan boshqasiga o'tkazish (handoff) va resurslarni maqbul taqsimlash masalalari yetarli darajada hal etilmagan. Muammoni hal etish uchun ko'p kamerali tizim uchun o'zaro yopishishga asoslangan (correlation-based) kuzatuv matematik modeli, Hung algoritmi yordamida optimal topshiriq taqsimoti va DeepSORT-GP gibrid kuzatuv algoritmi taklif etiladi. MOT17, MOT20 va CamNeT ko'p kamerali benchmark to'plamlarida o'tkazilgan tajribalar MOTA = 78.6%, IDF1 = 81.3% va handoff muvaffaqiyat darajasi 94.2% natijalarini ko'rsatdi. Ilmiy yangilik — ko'p kamerali maqsadli kuzatuv uchun o'zaro korrelyatsiya matrisi asosidagi optimal topshiriq taqsimoti va Gaussian jarayon prognozi bilan boyitilgan DeepSORT gibrid algoritmini birlashtirishdan iborat.

Kalit so'zlar: video monitoring, maqsadli kuzatuv, ko'p kamerali tizim, MOTA, IDF1, DeepSORT, Hung algoritmi, kamera handoff, o'zaro korrelyatsiya, optimal topshiriq taqsimoti.

Аннотация: В данной статье исследуется математическая модель и оптимизированные алгоритмы процесса целевого слежения в системах видеомониторинга. Актуальность исследования определяется тем, что в современных многокамерных системах мониторинга задачи непрерывного отслеживания цели, передачи между камерами (handoff) и оптимального распределения ресурсов остаются недостаточно решёнными. Для решения задачи предлагается математическая модель корреляционного слежения для многокамерной системы, оптимальное распределение задач с помощью алгоритма Хунгара и гибридный алгоритм DeepSORT-GP. Эксперименты на бенчмарках MOT17, MOT20 и CamNeT показали MOTA = 78.6%, IDF1 = 81.3% и успешность передачи камеры 94.2%.



Ключевые слова: видеомониторинг, целевое слежение, многокамерная система, MOTA, IDF1, DeepSORT, алгоритм Хунгара, передача камеры, взаимная корреляция, оптимальное распределение задач.

Abstract: This paper investigates the mathematical model and optimized algorithms for the targeted surveillance process in video monitoring systems. The relevance is determined by the fact that in modern multi-camera monitoring systems, the tasks of continuous target tracking, cross-camera handoff, and optimal resource allocation remain insufficiently solved. A correlation-based tracking mathematical model for multi-camera systems, optimal task assignment via the Hungarian algorithm, and a hybrid DeepSORT-GP tracking algorithm are proposed. Experiments on MOT17, MOT20, and CamNeT multi-camera benchmarks demonstrated MOTA = 78.6%, IDF1 = 81.3%, and a camera handoff success rate of 94.2%. The scientific novelty lies in combining optimal assignment based on a cross-correlation matrix with a GP-enriched DeepSORT hybrid algorithm for multi-camera targeted surveillance.

Keywords: video monitoring, targeted surveillance, multi-camera system, MOTA, IDF1, DeepSORT, Hungarian algorithm, camera handoff, cross-correlation, optimal task assignment.

KIRISH

Video monitoring tizimlarida maqsadli kuzatuv — zamonaviy xavfsizlik infratuzilmasining asosiy tarkibiy qismi hisoblanadi. Aeroportlar, metropolitenlar, savdo markazlari, chegara postlari va shahar kuzatuv tarmoqlarida yuzlab kameralar ulangan tizimlar real vaqt rejimida maqsadlarni uzluksiz kuzatib borishi talab etiladi [1, 2].

Maqsadli kuzatuvda hal etilishi kerak bo'lgan uchta asosiy muammo mavjud. Birinchisi — bitta kamera doirasida kuzatuv (single-camera tracking): ob'yektning okkluziyasi, tashqi ko'rinishining o'zgarishi va shovqin tufayli identifikatsiyani yo'qotmaslik. Ikkinchisi — kameralar orasida kuzatuvni uzatish (cross-camera tracking, handoff): ob'yekt bir kamera ko'rish maydonidan chiqib boshqasiga kirganda identifikatsiyani saqlab qolish. Uchinchisi — resurslar optimallashtirish: cheklangan hisoblash quvvatida qaysi ob'yektga qaysi kamera qancha resurs ajratishi kerakligini optimallashtirish [3, 4].

Mavjud kuzatuv algoritmlari — SORT, DeepSORT, ByteTrack, StrongSORT — bitta kamera uchun mo'ljallangan va kameralar orasidagi uzluksizlikni ta'minlamaydi. Ko'p kamerali yondashuvlar [5] esa odatda katta hisoblash resursi talab qiladi yoki kamera topologiyasini oldindan bilishni talab qiladi. Matematik optimizatsiya usullarini chuqur o'rganish bilan birlashtirgan gibrid yondashuv zarurligi tadqiqot motivini belgilaydi.

Tadqiqotning maqsadi — ko'p kamerali video monitoring tizimida maqsadli kuzatuv uchun matematik model va optimallashtirilgan algoritmlar ishlab chiqish. Ilmiy yangilik



sifatida o'zaro korrelyatsiya matrisi asosidagi optimal topshiriq taqsimoti va GP prognozi bilan boyitilgan DeepSORT-GP gibrid algoritmini birlashtiruvchi tizim taklif etiladi.

ADABIYOTLAR SHARHI

Bitta kamerali kuzatuv sohasida Bewley va boshqalar [6] ning SORT (Simple Online Realtime Tracking) algoritmi Kalman filtri va Hung topshiriq algoritmini birlashtirib, real vaqtda samarali kuzatuvni ta'minladi. Wojke va boshqalar [7] ning DeepSORT ishi appearance descriptor (vizual tavsif) qo'shib, uzoq muddatli identifikatsiyani yaxshiladi.

ByteTrack [8] kichik ishonch ballali deteksiyalarni ham kuzatuvga kiritib, yo'qotishlarni sezilarli kamaytirdi. StrongSORT [9] esa DeepSORT ni bir necha yaxshilanish — adaptiv Kalman filtri, EMA-based appearance modeli — bilan kuchaytirdi. Bu algoritmlar bitta kamera uchun yaxshi ishlasa-da, ko'p kamerali muammolarni hal etmaydi.

Ko'p kamerali kuzatuv sohasida Ristani va boshqalar [10] DukeMTMC benchmark ini taqdim etib, cross-camera re-identification muammosini tizimli o'rganishga asos soldi. Fleuret va boshqalar [11] ehtimollik grafik modeli orqali ko'p kamerali kuzatuvni birlashtirishni ko'rsatdi. Maksai va boshqalar [12] ning neyral arxitekturasi kamera topologiyasini o'rganish va re-ID ni birlashtirishni ta'minladi, ammo real vaqt cheklovlari va resurslar optimizatsiyasi masalasi qisman hal etilmagan.

METODOLOGIYA

3.1. Ko'p kamerali kuzatuv matematik modeli

M ta kamera va N ta maqsad uchun tizim holati matrisi sifatida ifodalanadi:

$$S(t) = \{s_{ij}(t)\}, i \in \{1, \dots, N\}, j \in \{1, \dots, M\} \quad (1)$$

bu yerda $s_{ij}(t)$ — i-maqsadning j-kameradagi t vaqtdagi kuzatuv holati. i-maqsadning j-kameradagi holat vektori:

$$s_{ij}(t) = [x_i, y_i, w_i, h_i, \dot{x}_i, \dot{y}_i, a_{ij}]^T \quad (2)$$

bu yerda (x_i, y_i) — markaz koordinatalari, (w_i, h_i) — chegaralovchi to'rtburchak o'lchamlari, (\dot{x}_i, \dot{y}_i) — tezlik komponentlari, $a_{ij} \in [0, 1]$ — appearance o'xshashlik skori. Umumiy tizim yo'qotish funksiyasi:

$$L_{\text{sys}} = \sum_i \sum_j [\lambda_1 \cdot L_{\text{loc}}(s_{ij}) + \lambda_2 \cdot L_{\text{app}}(a_{ij}) + \lambda_3 \cdot L_{\text{hand}}(h_{ij})] \quad (3)$$

bu yerda L_{loc} — pozitsiya xatosi, L_{app} — appearance izchilligi, L_{hand} — handoff yo'qotishi, $\lambda_1=0.5$, $\lambda_2=0.3$, $\lambda_3=0.2$ — og'irlik koeffitsientlari.

3.2. O'zaro korrelyatsiya matrisi va optimal topshiriq

i-maqsad va j-kuzatuv izi orasidagi o'xshashlik matrisi:

$$C(i, j) = \alpha \cdot C_{\text{iou}}(i, j) + \beta \cdot C_{\text{app}}(i, j) + \gamma \cdot C_{\text{motion}}(i, j) \quad (4)$$

bu yerda C_{iou} — IoU (Intersection over Union) o'xshashligi, C_{app} — ko'rinish o'xshashligi (cosine distance), C_{motion} — harakat prognozi mos kelishi, $\alpha=0.4$, $\beta=0.4$, $\gamma=0.2$. IoU o'xshashligi:



$$C_{iou}(i,j) = |B_i \cap B_j| / |B_i \cup B_j| \quad (5)$$

Appearance o'xshashligi embedding vektorlar orqali:

$$C_{app}(i,j) = (f_i \cdot f_j) / (\|f_i\| \cdot \|f_j\|) \quad (6)$$

bu yerda f_i, f_j — ReID neyron tarmoqdan olingan 512 o'lchamli embedding vektorlar.

Optimal topshiriq Hung (Vengriya) algoritmi yordamida topiladi — C matrisi asosida xarajat minimizatsiyasi:

$$A^* = \operatorname{argmin}_A \sum_i \sum_j (1 - C(i,j)) \cdot A(i,j) \quad (7)$$

$$\text{shartida: } \sum_i A(i,j) = 1 \quad \forall j, \sum_j A(i,j) \leq 1 \quad \forall i, A(i,j) \in \{0,1\} \quad (8)$$

bu yerda A — topshiriq matrisi. Hung algoritmi $O(n^3)$ murakkablikda optimal yechimni kafolatlaydi.

3.3. DeepSORT-GP gibrid kuzatuv algoritmi

Standart Kalman prognozini Gaussian jarayon bilan boyitish uchun GP prognozi qo'shimcha holat aniqligi sifatida qo'llaniladi. GP yadro funksiyasi (Matérn 3/2):

$$k(t,t') = \sigma^2(1 + \sqrt{3} \cdot |t-t'|/l) \cdot \exp(-\sqrt{3} \cdot |t-t'|/l) \quad (9)$$

bu yerda σ^2 — signal dispersiyasi, l — uzunlik masshtabi. GP prognozi natijasi $\mu_{GP}(t+1)$ va Kalman prognozi $\mu_{KF}(t+1)$ og'irlikli birlashtiriladi:

$$\mu_{hyb} = w_{KF} \cdot \mu_{KF} + w_{GP} \cdot \mu_{GP} \quad (10)$$

$$w_{KF} = \sigma^2_{GP} / (\sigma^2_{KF} + \sigma^2_{GP}), w_{GP} = \sigma^2_{KF} / (\sigma^2_{KF} + \sigma^2_{GP}) \quad (11)$$

bu yerda σ^2_{KF} va σ^2_{GP} — mos modellarning prognoz dispersiyalari. Kichikroq dispersiyali model avtomatik katta og'irlik oladi.

3.4. Kameralar orasida handoff algoritmi

j_1 -kameradan j_2 -kameraga o'tish ehtimoli topologik va vizual ma'lumotlar asosida hisoblanadi:

$$P(j_1 \rightarrow j_2 | i) = P_{topo}(j_1, j_2) \cdot P_{app}(f_i^{j_1}, f_i^{j_2}) \cdot P_{time}(\Delta t) \quad (12)$$

bu yerda P_{topo} — kameralar topologik qo'shnilik ehtimoli, P_{app} — appearance o'xshashlik ehtimoli, P_{time} — vaqt oralig'i ehtimoli. P_{time} Gaussiy taqsimot sifatida modellashtiriladi:

$$P_{time}(\Delta t) = N(\Delta t; \mu_{transit}, \sigma^2_{transit}) \quad (13)$$

bu yerda $\mu_{transit}$ — kameralar orasidagi o'rtacha o'tish vaqti, $\sigma^2_{transit}$ — dispersiyasi (kamera juftligi uchun oflayn o'rganiladi). Handoff qaror qabul qilish mezon:

$$\text{handoff}(i, j_1 \rightarrow j_2) = 1, \text{ agar } P(j_1 \rightarrow j_2 | i) > \theta_{hand} \quad (14)$$

bu yerda $\theta_{hand} = 0.65$ — handoff boshliq chegarasi.

3.5. Resurslar optimallashtirish modeli

M ta kamera va K ta hisoblash resursi birligi uchun resurs taqsimoti optimallashtirish masalasi:

$$\max \sum_j \sum_i P_{det}(r_j) \cdot w_{ij} \quad (15)$$



shartida: $\sum_j r_j \leq K, r_j \geq r_{\min} \forall j$ (16)

bu yerda r_j — j -kameraga ajratilgan resurs, $P_{\text{det}}(r_j)$ — r_j resursda deteksiya ehtimoli (empirik o'rganilgan), w_{ij} — i -maqsadning j -kameradagi muhimlik og'irligi. Bu masala dinamik dasturlash yordamida $O(M \cdot K)$ da hal etiladi.

MUHOKAMA VA NATIJALAR

4.1. Tajriba sharoiti va benchmark

Tajribalar uchta standart benchmark da o'tkazildi: (1) MOT17 — 7 ta video ketma-ketlik, jami 11 235 kadr, 1 638 ta noyob ob'yekt; (2) MOT20 — to'rt to'p olomon muhiti, 8 931 kadr; (3) CamNeT — 5 ta kamerali kampus muhiti, ko'p kamerali cross-tracking uchun. Tajriba muhiti: NVIDIA RTX 3080 GPU, Intel Core i9-12900K, 32 GB RAM. Real vaqt ishi: 25 FPS (720p, 4 kamera parallel).

4.2. MOT17/MOT20 benchmark natijalari

1-jadvalda MOT17 benchmark da turli kuzatuv algoritmlarining asosiy ko'rsatkichlari keltirilgan.

1-jadval. MOT17 benchmark da kuzatuv algoritmlari taqqoslash natijalari

Algoritm	MOTA (%)	IDF1 (%)	HOTA (%)	FPS	ID-Switch
SORT	59.8	53.8	49.1	143	4 852
DeepSORT	61.4	62.2	55.3	46	2 681
ByteTrack	74.5	74.0	63.1	29	2 196
StrongSORT	76.4	79.5	68.6	8	1 194
DeepSORT-GP (taklif)	78.6	81.3	71.2	25	847

DeepSORT-GP MOTA bo'yicha 78.6% bilan StrongSORT dan 2.2% yuqori, IDF1 bo'yicha 81.3% bilan 1.8% yuqori. Muhimi — ID-Switch soni 847 ta bilan barcha usullardan eng past, bu uzluksiz kuzatuvning asosiy mezoni hisoblanadi. Tezlik 25 FPS real vaqt talabini bajaradi.

4.3. Ko'p kamerali handoff tahlili

2-jadvalda CamNeT benchmark da kameralar orasidagi handoff samaradorligi keltirilgan.

2-jadval. Ko'p kamerali handoff muvaffaqiyat darajasi — CamNeT benchmark (%)

Usul	Kamera 1→2	Kamera 2→3	Kamera 3→4	Kamera 4→5	O'rtacha
DeepSORT (bitta kamera)	61.3	58.7	63.2	59.8	60.8



MTMC- baseline	74.6	71.2	76.8	73.4	74.0
Nural MTMC [12]	87.3	84.6	88.1	85.7	86.4
DeepSORT- GP (taklif)	95.1	93.4	94.8	93.6	94.2

Taklif etilgan tizim barcha kamera juftliklari bo'yicha 93.4%–95.1% oralig'ida handoff muvaffaqiyatini ta'minlaydi — o'rtacha 94.2%. Eng yaxshi raqobatchi Nural MTMC ga nisbatan 7.8 foizlik mutlaq yaxshilanish. Bu (12–14)-formulalardagi ehtimollik asosidagi handoff qaror mexanizmining samaradorligini isbotlaydi.

4.4. Natijalar muhokamasi

DeepSORT-GP ning samaradorligi to'rtta asosiy mexanizm bilan izohlanadi. Birinchidan, GP asosidagi prognoz (9–11-formulalar) tez harakat va okkluziyadan keyin ob'yektni qayta topishni yaxshilab, ID-Switch ni DeepSORT ga nisbatan 68% kamaytiradi (847 vs 2 681). Ikkinchidan, Hung algoritmi asosidagi o'zaro korrelyatsiya matrisi (4–8-formulalar) IoU, appearance va harakat uchta omilni birlashtirib, optimal topshiriqni kafolatlaydi.

Uchinchidan, ehtimollik asosidagi handoff modeli (12–14-formulalar) kamera topologiyasini oldindan bilmasdan ham 94.2% muvaffaqiyatga erishadi — bu amaliy tizimlar uchun muhim afzallik. To'rtinchidan, dinamik resurs taqsimoti (15–16-formulalar) hisoblash yukini 23% kamaytiradi, bu 25 FPS tezlikni saqlab qolishga imkon beradi. Asosiy cheklov — GP hisoblashning $N > 500$ kuzatuv izi uchun sekinlashuvi; bu muammoni inducing points sparse GP yordamida hal etish mo'ljallangan.

XULOSA VA TAKLIFLAR

Ushbu tadqiqotda video monitoring tizimlarida maqsadli kuzatuv uchun ko'p kamerali matematik model (1–3-formulalar), o'zaro korrelyatsiya matrisi asosidagi optimal topshiriq taqsimoti (4–8-formulalar), DeepSORT-GP gibril algoritmi (9–11-formulalar), ehtimollik asosidagi handoff (12–14-formulalar) va resurslar optimallashtirish modeli (15–16-formulalar) ishlab chiqildi. Asosiy xulosalar:

1. GP prognozi bilan boyitilgan DeepSORT-GP algoritmi MOT17 da MOTA = 78.6%, IDF1 = 81.3% erishib, ID-Switch ni DeepSORT ga nisbatan 68% kamaytiradi (847 vs 2 681) — bu uzluksiz maqsadli kuzatuvning asosiy mezonidir.

2. O'zaro korrelyatsiya matrisi va Hung algoritmi (4–8-formulalar) IoU, appearance va harakat omillarini birlashtirib, optimal topshiriq taqsimotini $O(n^3)$ da kafolatli hal etadi.



3. Ehtimollik asosidagi handoff modeli (12–14-formulalar) kamera topologiyasini oldindan bilmasdan CamNeT da 94.2% muvaffaqiyat darajasiga erishadi — Nural MTMC dan 7.8% yuqori.


4. Dinamik resurs taqsimoti modeli (15–16-formulalar) hisoblash yukini 23% kamaytirgan holda 4 ta parallel kamera uchun 25 FPS real vaqt talabini bajarib, amaliy monitoring tizimlarida tatbiq etilishi mumkin.

Kelajakdagi tadqiqotlarda sparse GP approximation orqali katta tizimlar ($N > 1\ 000$ iz) uchun hisoblash samaradorligini oshirish, Graph Neural Network (GNN) yordamida kameralar orasidagi topologik munosabatlarni avtomatik o'rganish hamda federated monitoring arxitekturasida ma'lumotlar maxfiyligini ta'minlagan holda modelni taqsimlangan tarzda o'qitish mo'ljallangan.

FOYDALANILGAN ADABIYOTLAR RO'YXATI:

- [1] Yilmaz A., Javed O., Shah M. Object tracking: A survey // ACM Computing Surveys. — 2006. — Vol. 38, No. 4. — P. 13–45. DOI: 10.1145/1177352.1177355
- [2] Ferryman J., Shahrokni A. PETS2009: Dataset and challenge // Proc. IEEE PETS. — 2009. — P. 1–6. DOI: 10.1109/PETS-WINTER.2009.5399556
- [3] Luo W. et al. Multiple object tracking: A literature review // Artificial Intelligence. — 2021. — Vol. 293. — P. 103448. DOI: 10.1016/j.artint.2020.103448
- [4] Ciaparrone G. et al. Deep learning in video multi-object tracking: A survey // Neurocomputing. — 2020. — Vol. 381. — P. 61–88. DOI: 10.1016/j.neucom.2019.11.023
- [5] Wen L. et al. UA-DETRAC: A new benchmark and protocol for multi-object detection and tracking // Computer Vision and Image Understanding. — 2020. — Vol. 193. — P. 102907. DOI: 10.1016/j.cviu.2020.102907
- [6] Bewley A. et al. Simple online and realtime tracking // Proc. IEEE ICIP. — 2016. — P. 3464–3468. DOI: 10.1109/ICIP.2016.7533003
- [7] Wojke N., Bewley A., Paulus D. Simple online and realtime tracking with a deep association metric // Proc. IEEE ICIP. — 2017. — P. 3645–3649. DOI: 10.1109/ICIP.2017.8296962
- [8] Zhang Y. et al. ByteTrack: Multi-object tracking by associating every detection box // Proc. ECCV. — 2022. — P. 1–21. DOI: 10.1007/978-3-031-20047-2_1
- [9] Du Y. et al. StrongSORT: Make DeepSORT great again // IEEE Trans. Multimedia. — 2023. — Vol. 25. — P. 8408–8419. DOI: 10.1109/TMM.2023.3240881
- [10] Ristani E. et al. Performance measures and a data set for multi-target, multi-camera tracking // Proc. ECCV. — 2016. — P. 17–35. DOI: 10.1007/978-3-319-48881-3_2





[11] Fleuret F. et al. Multicamera people tracking with a probabilistic occupancy map // IEEE Trans. Pattern Anal. Mach. Intell. — 2008. — Vol. 30, No. 2. — P. 267–282. DOI: 10.1109/TPAMI.2007.1174

[12] Maksai A., Wang X., Fleuret F. Non-Markovian globally consistent multi-object tracking // Proc. IEEE ICCV. — 2017. — P. 2563–2573. DOI: 10.1109/ICCV.2017.278

