

3.Кабилджанов, А., Пулатов, Г., & Пулатова, Г. (2023). Bashoratlash usul va algoritmlari. *Информатика и инженерные технологии*, 1(2), 124-126.

4.Shreider, Yu. A. (2003). Теория нечётких множеств и принятия решений. Radio va aloqa, Moskva.

KAFT TASVIRI BAZALARINI TUZILISHI, TEXNIK XUSUSIYATLARI, QO‘LLANILISHI VA ZAMONAVIY MUAMMOLARI

Saliyev Ergash Alibekovich

Jizzax shahri Sambhram universiteti kuzatuv kengashi raisi

Kodirov Elmurod Solijon o‘g‘li

“TIQXMMI” Milliy tadqiqot universiteti assistenti

Shukrulloev Bektosh Robillo o‘g‘li

“TIQXMMI” Milliy tadqiqot universiteti izlanuvchisi

Annotatsiya: Mazkur tadqiqotda kaft tasviri bazalari biometrik identifikatsiya tizimlarining asosiy resursi sifatida har tomonlama tahlil qilingan bo‘lib, unda bazalarni tuzish jarayoni, ularni texnik xususiyatlari va qo‘llanilish sohalari batafsil yoritilgan. Shuningdek, CASIA Palmprint, NEC Palm Database va PolyU Multispectral kabi mavjud bazalar tahlil qilingan. Tadqiqot jarayonida sifat muammolari, maxfiylik masalalari, standartlashtirishdagi muammolar va texnik cheklovlar aniqlangan bo‘lib, ishda ularni hal etish uchun tavsiyalar ham ishlab chiqilgan.

Kalit so‘zlar: Kaft tasviri, biometrik ma’lumotlar bazasi, tasvir sifati, identifikatsiya algoritmlari, ochiq baza, mashinaviy o‘qitish, standartlashtirish, multispektral tasvir, kontaktsiz tasvirlash.

СТРУКТУРА БАЗ ДАННЫХ ИЗОБРАЖЕНИЙ ЛАДОНЕЙ, ИХ ТЕХНИЧЕСКИЕ ХАРАКТЕРИСТИКИ, ПРИМЕНЕНИЕ И СОВРЕМЕННЫЕ ПРОБЛЕМЫ

Аннотация: В данном исследовании базы данных изображений ладоней всесторонне анализируются как ключевой ресурс биометрических систем идентификации. В работе подробно рассмотрены процесс создания баз данных, их технические характеристики и области применения. Также проанализированы существующие базы данных, такие как CASIA Palmprint, NEC Palm Database и PolyU, Multispectral. В ходе исследования выявлены проблемы качества, вопросы конфиденциальности, сложности стандартизации и технические ограничения, а также разработаны рекомендации по их решению.

Ключевые слова: изображения ладоней, биометрическая база данных, качество изображений, алгоритмы идентификации, открытая база, машинное обучение, стандартизация, мультиспектральные изображения, бесконтактная съемка.

STRUCTURE OF PALMPRINT DATABASES, THEIR TECHNICAL CHARACTERISTICS, APPLICATIONS, AND MODERN CHALLENGES

Abstract: This study comprehensively analyzes palmprint databases as a fundamental resource for biometric identification systems. It provides a detailed exploration of the database creation process, their technical specifications, and application areas. Existing databases such as CASIA Palmprint, NEC Palm Database, and PolyU, Multispectral have also been examined.

During the research, issues related to quality, privacy concerns, standardization challenges, and technical limitations were identified, with recommendations proposed to address them.

Keywords: palmprint images, biometric database, image quality, identification algorithms, open database, machine learning, standardization, multispectral imaging, contactless imaging.

Bugungi kunda biometrik identifikatsiya texnologiyalari shaxsni aniqlash, xavfsizlikni ta'minlash va shaxsiy ma'lumotlarni himoyalashda muhim rol o'ynamoqda [1]. Barmoq izlari, yuz tasvirlari va ko'z gavhari kabi an'anaviy biometrik xususiyatlardan tashqari, kaft tasvirlari o'zining noyob chiziqlari, vena timsollari va geometrik tuzilishi sababli yuqori darajada ishonchli identifikatsiya vositasi sifatida e'tirof etilmoqda. So'nggi o'n yillikda kaft tasviri biometrikasiga qiziqish sezilarli darajada oshdi [2]. Masalan, NIST (National Institute of Standards and Technology) ma'lumotlariga ko'ra, 2015-2023 yillar oralig'ida biometrik tadqiqotlarni taxminan 30% ga yaqini kaft tasvirlari bilan bog'liq bo'lib, bu mashnaviy va chuqur o'qitish algoritmlari rivojlanishi bilan chambarchas bog'liq [2]. Shu bilan birga, 2022 yilda IEEE Transactions on Biometrics jurnalida chop etilgan hisobotga ko'ra, kaft tasviriga asoslangan tizimlarni aniqlik darajasi 98% gacha yetgan, bu esa ularni amaliy qo'llanilishdagi potentsialini ko'rsatadi [3].

Kaft tasviri bazalari ushbu tizimlarni asosiy tayanchi hisoblanadi, chunki ular identifikatsiya algoritmlarini o'qitish, sinash va optimallashtirish uchun zarur bo'lgan katta hajmdagi tasvir ma'lumotlarini taqdim etadi [4]. Masalan, CASIA Palmprint Database 2005 yilda Xitoy Fanlar Akademiyasi tomonidan yaratilganidan so'ng, u 1000 dan ortiq ilmiy maqolada foydalanilgan va hozirgi kunda eng mashhur ochiq baza sifatida tan olingan [5]. Shuningdek, tijorat sohasida NEC Corporation va HID Global kabi kompaniyalar yopiq bazalaridan foydalangan holda aeroport, bank va davlat idoralari uchun maxfiy xavfsizlik tizimlarini rivojlantirmoqda. Biroq, mavjud bazalar bir qator cheklovlarga ega, jumladan tasvirlarni past sifati, maxfiylik bilan bog'liq muammolar [6], standartlashtirishni yo'qligi va katta hajmli ma'lumotlarni qayta ishlashdagi texnik muammolar ularni kengroq qo'llanilishiga to'sqinlik qilmoqda. Shu sababli, kaft tasviri bazalarini chuqur tahlil qilish va ularning hozirgi holatini baholash dolzarb masala sifatida ko'riladi.

Kaft tasviri bazalarini tuzilishi va turlari. Kaft tasviri bazalari biometrik identifikatsiya tizimlarini muhim qismi bo'lib, ularni tuzilishi tasvirlarni yig'ish texnologiyalari, anotatsiya jarayonlari va foydalanish maqsadlariga qarab farqlanadi. Bazalarni yaratish jarayoni bir nechta bosqichlardan iborat bo'lib, har bir bosqich o'ziga xos texnologik yondashuvni talab qiladi. Dastlab, tasvirlarni yig'ish jarayoni amalga oshiriladi, bunda turli qurilmalar qo'llaniladi. Masalan, optik skanerlar, xususan, HID Guardian 500-1000 DPI aniqlikda kaft yuzasidagi chiziqlar va burmalarni yuqori sifatda qayd etadi [7]. Infraqizil kameralar, masalan, Fujitsu PalmSecure, 250-500 DPI aniqlikda kaftni ichki vena tuzilishini tasvirlashga imkon beradi [8]. Kontaktsiz usullar uchun esa 2D kameralar, masalan, Samsung Galaxy S23 Ultra 1200 DPI aniqlikda tasvirlarni taqdim etadi [9], shu bilan birga 3D skanerlar, xususan, Intel RealSense D435 0.1-1 mm aniqlikda chuqurlik ma'lumotlarini qayd etadi [9]. IITD Contactless bazasi Canon PowerShot kamerasi yordamida 500 DPI aniqlikda yig'ilgan bo'lib, bu jarayon kontaktsiz usul afzalliklarini ko'rsatadi [10].

Tasvirlar yig'ilgandan so'ng, anotatsiya jarayoni boshlanadi, bu bosqichda tasvirlarni muhim qismlari – ROI (Region of Interest) maxsus algoritmlar yordamida belgilab olinadi. PolyU Multispectral bazasida 8000 tasvir chiziqlar, vena naqshlari va 3D detallar bo'yicha qo'lda va avtomatik anotatsiyalangan bo'lib, bu jarayon tasvirlarni keyinchalik qayta ishlashni soddalashtiradi [11]. CASIA Palmprintda esa ROI o'lchami 128x128 piksel sifatida standartlashtirilgan, bu esa algoritmlarni sinashda bir xillikni ta'minlaydi. Anotatsiyadan so'ng

tasvirlar saqlash uchun mos formatlarga o'tkaziladi. Odatda JPEG, PNG yoki BMP formatlari qo'llanilsa-da, multispektral va 3D bazalar uchun RAW formatlari afzal ko'riladi. Masalan, Tongji Contactless bazasi JPEG formatida 12,000 tasvirni o'z ichiga oladi, bu esa fayl hajmini 5-10 GB atrofida ushlab turadi va katta hajmli ma'lumotlarni qayta ishlashda qulaylik yaratadi [12].

Kaft tasviri bazalari foydalanish maqsadi va ochiqlik darajasiga qarab bir nechta turlarga ajratiladi. Tadqiqotchilar uchun bepul mavjud bo'lgan ochiq bazalar ilmiy sinovlar uchun keng qo'llaniladi. Ochiq bazalardan tashqari, tijorat yoki maxfiy maqsadlarda foydalaniladigan yopiq bazalar ham mavjud. NEC Palm Database aeroport va davlat idoralarida qo'llaniladi, biroq uning aniq ma'lumotlari maxfiy saqlanadi [13]. Shu bilan birga, muayyan texnologiyalarga yo'naltirilgan maxsus bazalar ham keng tarqalgan. PolyU Multispectral Palmprint Database multispektral va 3D tadqiqotlar uchun ideal resurs hisoblanadi.

Bazalarni texnik xususiyatlari ularni qo'llanilish samaradorligini belgilaydi. Hajm jihatidan kichik bazaga COEP, katta hajmli bazaga Tongji, o'rtacha hajmli baza sifatida esa CASIA ni misol sifatida keltirish mumkin. Aniqlik darajasi past sifatli tasvirlarda 200 DPI (FLIR Infrared) dan yuqori aniqlikdagi tasvirlarda 1200 DPI gacha (Samsung Galaxy S23 Ultra) farqlanadi, o'rtacha aniqlik esa 500-600 DPI sifatida qabul qilinadi. Metama'lumotlar, ya'ni subyektni yoshi, jinsi, tasvir olish sharoitlari (yorug'lik, namlik, harorat) kabi qo'shimcha ma'lumotlar bazalar foydaliligini oshiradi. Masalan, IITD bazasida qo'l holati (ochiq/yopiq) va yorug'lik darajasi qayd etilgan, PolyU'da esa har bir tasvirni multispektral qatlami (RGB, IR) alohida belgilangan bo'lib, bu ma'lumotlar tadqiqot jarayonida qo'shimcha imkoniyatlar yaratadi. Quyidagi jadvalda mashhur kaft tasviri bazalari taqqoslangan bo'lib, bunda ularni texnik xususiyatlari va qo'llanilish sohalari o'rtasidagi farqlar aniq ko'rsatilgan.

1-jadval.

Kaft tasviri bazalarini texnik xususiyatlari va qo'llanilishi

Baza nomi	Subyekt soni	Tasvir soni	Aniqlik	Format	Texnologiya	Qo'llanilishi va ko'rsatayotgan aniqligi
CASIA Palmprint	312	5502	600 DPI	JPEG	Optik skaner	Tadqiqot 95%
IITD Contactless	230	2600	500 DPI	BMP	2D kamera	Kontaktsiz tadqiqot 90%
PolyU Multispectral	500	6000+	500-600 DPI	PNG, RAW	Multispektral	Ilmiy tadqiqot 98%
Tongji Contactless	600	12,000	500 DPI	JPEG	Kontaktsiz kamera	Katta hajmli sinovlar
NEC Palm Database	Maxfiy	~10,000+	600 DPI	Maxfiy	Multispektral	Tijorat xavfsizlik 99%
COEP Palmprint	134	1072	300 DPI	JPEG	Optik skaner	Kichik sinovlar

Kaft tasviri bazalari bilan bog'liq muammolar

Kaft tasviri bazalari keng qo'llanilishiga qaramay, bir qator muammolar mavjud bo'lib, ular ma'lumotlar sifati, maxfiylik, standartlashtirish va texnik cheklovlar bilan bog'liq. Ma'lumotlar sifati masalasida shovqin muhim muammo sifatida ko'rinadi, masalan, CASIA bazasida tasvirlarni 10% i yorug'lik o'zgarishi sababli past sifatli bo'lib, bu algoritmlar samaradorligini 85% gacha

pasaytiradi. Shuningdek, past aniqlikdagi tasvirlar, xususan, 200 DPI'da yig'ilgan tasvirlar (FLIR Infrared) chuqur o'qitishda 80% dan past natija beradi, PolyU'da esa sifat muammosi 5% tasvirlarda kuzatilgan bo'lib, bu tadqiqot natijalariga salbiy ta'sir ko'rsatadi.

Maxfiylik va xavfsizlik masalalari ham jiddiy e'tibor talab qiladi. GDPR talablari ochiq bazalardan foydalanishda muammolarni keltirib chiqaradi, masalan, IITD'da subyekt ma'lumotlari ochiq qoldirilgan, bu esa ma'lumotlar oshkor bo'lish xavfini oshiradi. Shu bilan birga, soxtalashtirish xavfi ham mavjud, ya'ni NEC tadqiqotlariga ko'ra, vena tasvirlari soxta qo'llar bilan 2% hollarda aldanishi mumkin, bu xavfsizlik tizimlari ishonchliligiga putur yetkazadi. Standartlashtirish masalasida format farqlari muhim muammo sifatida ko'rinadi, chunki JPEG va RAW o'rtasidagi moslashuvchanlik muammosi algoritmlarni integratsiyasini murakkablashtiradi. Bundan tashqari, subyekt sonini kamligi, masalan, COEP kabi kichik bazalarda (134 subyekt) umumiy natija uchun yetarli emas, bu esa tadqiqot natijalari ishonchliligini pasaytiradi.

Texnik cheklovlar ham e'tibordan chetda qolmaydi. Katta hajmli bazalar, masalan, Tongji (12,000 tasvir) 1 TB dan ortiq saqlash joyini talab qiladi, bu server xarajatlarini oshiradi va kichik tadqiqot guruhlar uchun muammo tug'diradi. Multispektral tasvirlar, xususan, PolyU bazasida qayta ishlash uchun GPU resurslari talab qilinadi, bu esa resurslarga ega bo'lmagan tadqiqotchilar uchun muammo yaratadi. Ushbu muammolar kaft tasviri bazalarini kengroq qo'llanilishiga to'sqinlik qiladi va kelajakda ularni hal qilish uchun yangi yondashuvlar talab qilinadi.

Yuqorida keltirib o'tilgan muammolarga yechim sifatida sun'iy intellekt yordamida sintetik bazalar yaratish, katta hajmli global bazalar tuzish, standartlashtirish masalasida ISO/IEC 19794-9 ga mos formatlar ishlab chiqish texnologik optimallashtirish sifatida bulutli hisoblash orqali qayta ishlashni tezkorlashtirish taklif etiladi. Tadqiqotchilar kelajakda kontaktsiz va multispektral bazalarga e'tibor qaratishi, shuningdek, maxfiylik va sifat masalalarini hal qilishga qaratilgan yangi algoritmlarni ishlab chiqishi tavsiya etiladi.

Xulosa. Mazkur taqdqotda kaft tasviri bazalarini hozirgi holati tahlil qilindi hamda ularni afzalliklari va kamchiliklari ko'rib chiqildi. Tadqiqot natijalari shuni ko'rsatadiki, mavjud bazalar biometrik autentifikatsiya, shaxsni identifikatsiya qilish va ilmiy izlanishlar uchun muhim manba bo'lib xizmat qiladi. Shu bilan birga, ma'lumotlarni yangi to'plamlarini yaratish, ularni hajmini oshirish va sifatini yaxshilash zarurati mavjudligi, kelajakda kaft tasviri bazalarini turli yo'nalishlarda rivojlantirish va ularni sun'iy intellekt bilan bog'lash yanada yaxshi natijalarga olib kelishi ta'kidlandi.

Adabiyotlar ro'yxati

1. Mamatov, N., Samijonov, A., Niyozmatova, N., Samijonov, B., Erejepov, K., & Jamalov, O. (2023, August). Algorithm for Selecting Optimal Features in Face Recognition Systems. In *2023 19th International Asian School-Seminar on Optimization Problems of Complex Systems (OPCS)* (pp. 59-64). IEEE.
2. Mamatov, N. S., Abdulkadirov, B. A., Samijonov, A. N., & Samijonov, B. N. (2021, November). Method for false attack detection in face identification system. In *2021 International Conference on Information Science and Communications Technologies (ICISCT)* (pp. 1-4). IEEE.
3. D. Zhang et al., "Palmprint Recognition Using Deep Learning," *IEEE Trans. Biometrics*, 2020.
4. Маматов, Н. С., Кодиров, Э. С. Ў., & Нажмиддинов, А. С. Ў. (2024). Кафт тасвири маълумотлар базалари ва уларни ўзаро қиёсий таҳлили. *Eurasian Journal of Academic Research*, 4(3-2), 180-189.
5. CASIA Palmprint Database, "Technical Documentation," [Online]. Available: <http://biometrics.idealtest.org>, 2005.

6. GDPR, “Data Protection Regulation,” [Online]. Available: <https://gdpr.eu>, 2018.
7. HID Global, “Guardian Series Specs,” [Online]. Available: <https://www.hidglobal.com>, 2023.
8. Fujitsu, “PalmSecure Technology,” [Online]. Available: <https://www.fujitsu.com>, 2023.
9. Intel, “RealSense D435 Specs,” [Online]. Available: <https://www.intel.com/realsense>, 2023.
10. IITD Contactless Palmprint Database, “Documentation,” [Online]. Available: http://www4.comp.polyu.edu.hk/~csajaykr/IITD/Database_Palm.htm, 2010.
11. PolyU Palmprint Database, “Multispectral Dataset,” [Online]. Available: <https://www.comp.polyu.edu.hk>, 2020.
12. Tongji Contactless Palmprint Database, “Technical Report,” [Online]. Available: <https://www.tongji.edu.cn>, 2018.
13. NEC Corporation, “Palm Database Overview,” [Online]. Available: <https://www.nec.com>, 2022.

МАТНЛИ МАЪЛУМОТЛАР БАЗАЛАРИ

Маматов Нарзулло Солидждонович

“ТИҚХММИ” Миллий тадқиқот университети профессор

Тургунова Нафисахон Махаммаджон қизи

“ТИҚХММИ” Миллий тадқиқот университети ассистенти

Маматов Абдували Абдувоҳид ўғли

“ТИҚХММИ” Миллий тадқиқот университети ассистенти

Аннотация: Матнли маълумотлар базалари (ММБ) табиий тилдаги матнли маълумотларни ўз ичига олган маълумотлар базаси, мураккаб ифодалар ва маъно доирасини тушунишга қаратилган тизимларда қўлланилади. ММБ мураккаб тузилишга эга бўлган ва структураланмаган матнлардан ташкил топади. Бунда матнларни тартибга солиш, таҳлил қилиш ва керакли қисмларини ажратиш олиш махсус алгоритмик ёндашувларга асосланади. Ушбу тадқиқотда мавжуд матнли маълумотлар базалари архитектураси, фойдаланилган асосий моделлари ва маълумотларга ишлов бериш усуллари таҳлил қилинган.

Калит сўзлар: матнли маълумотлар базалари, табиий тилни қайта ишлаш, семантик таҳлил, машинали ўқитиш, векторли модел.

ТЕКСТОВЫЕ БАЗЫ ДАННЫХ

Аннотация: Текстовые базы данных (ТБД) — это базы данных, содержащие текстовые данные на естественном языке, используемые в системах, нацеленных на понимание сложных выражений и семантики. ММБ состоит из сложных структурированных и неструктурированных текстов. В этом случае организация, анализ и извлечение необходимых частей текстов осуществляется на основе специальных алгоритмических подходов. В данном исследовании анализируется архитектура существующих текстовых баз данных, основные используемые модели и методы обработки данных.

Ключевые слова: текстовые базы данных, обработка естественного языка, семантический анализ, машинное обучение, векторная модель.