

# Аналіз натовпу людей із застосуванням методів комп'ютерного зору

О.М. Головін

Інститут кібернетики імені В.М. Глушкова НАН України, 03187, м. Київ,  
проспект Академіка Глушкова, 40, [o.m.golovin.1@gmail.com](mailto:o.m.golovin.1@gmail.com)

**O.M. Golovin**

## **ANALYSIS OF CROWD USING COMPUTER VISION TECHNIQUES**

**Abstract.** Three important problems in crowd analysis are considered in the paper: people counting/density estimation, tracking in crowd scenes, and understanding crowd behavior in higher-level analysis. Automated crowd density estimation and crowd counting are actual and important topics in crowd analysis. The article provides an overview of crowd density estimation, methods for calculating crowd density, and methods for tracking crowds and understanding the behavior of groups of people. This review covers two main approaches, direct and indirect. The direct approach aims at the simultaneous determination of the number of people (by identifying individuals) and their localization, and in the indirect approach (also called the map, measurement and feature based) people counting is usually carried out by measuring some of the features. This review categorizes and outlines several density estimation methods used to monitor crowd scenes. It is this characteristic of crowd density that is used in many decision-making processes for events that take place in crowds and thus assess the dynamics of the crowd. And the study of crowd dynamics is of scientific and practical interest because of the diverse phenomena of self-organization that arise from interactions of many participants in large crowds. The immediate detection of dramatic changes in crowd dynamics provide significant opportunities for avoiding crowd-related disasters. Taking into account various aspects, namely, the area of image space, the field of sociology, the level of service according to personal space of the person in a crowd and the area of computer graphics, probable scenarios for a crowd turbulence were considered and the causes of changes in dynamics of the crowd were identified. In the work much attention is paid to methods of estimation of a crowd density, their strengths and weaknesses are identified, which gives the opportunity to choose the appropriate algorithms for further use in certain conditions.

**Key words:** computer vision, crowd density, intelligent video surveillance, image processing.

**Анотація.** Розглянуто три важливі проблеми аналізу натовпу: підрахунок людей/оцінка щільності, відстежування в сценах з натовпом людей та розуміння поведінки натовпу. Автоматизоване оцінювання і підрахунок щільності натовпу – актуальна і важлива тема в аналізі натовпу. В статті представлено огляд оцінки щільності

натовпу, методи її підрахунку, а також методи відстежування натовпу і розуміння поведінки груп людей. Огляд охоплює два основні підходи, а саме: прямий і непрямий. Прямий підхід націлений на одночасне визначення кількості людей (шляхом ідентифікації окремих осіб) та їх місцезнаходження, а у непрямому підході (його також називають підхід на основі карт, вимірювання та ознак) підрахунок людей проводиться зазвичай за допомогою вимірювань деяких ознак. Огляд категоризує та окреслює кілька оцінок щільності та методи підрахунку щільності натовпу, що застосовуються для моніторингу сцен з натовпом людей. Саме така характеристика як щільність натовпу використовується у багатьох процесах прийняття рішень стосовно тих подій, які мають місце у натовпах і, таким чином, дають оцінку динаміки натовпу. А вивчення динаміки натовпу представляють науковий і практичний інтерес через різноманітні явища самоорганізації, що виникають внаслідок взаємодії багатьох учасників великих скупчень людей. Безпосередньо оперативне виявлення і визначення різких змін динаміки натовпу надають значні можливості для уникнення катастроф, пов'язаних з натовпом. Розглянуто ймовірні сценарії розвитку турбулентності натовпу і визначені причини змін динаміки натовпу. Проаналізовано методи оцінки щільності натовпу, визначені їх сильні і слабкі сторони.

**Ключові слова:** комп'ютерний зір, щільність натовпу, інтелектуальне відеоспостереження, обробка зображень.

**Аннотация.** Рассмотрены три важные проблемы анализа толпы: подсчет людей/оценка плотности, отслеживание в сценах с толпой людей и понимания поведения толпы. Автоматизированное оценивание и подсчет плотности толпы – это актуальная и важная тема в анализе толпы. Обзор охватывает два основных подхода оценки плотности толпы, а именно: прямой и косвенный. Прямой подход нацелен на одновременное определение количества людей (путем идентификации отдельных лиц) и их местонахождение, а в косвенном подходе подсчет людей проводится обычно с помощью измерений некоторых признаков. Именно такая характеристика как плотность толпы используется во многих процессах принятия решений в отношении событий, которые имеют место в скоплениях людей и, таким образом, дают оценку динамики толпы. Непосредственно оперативное выявление и определение резких изменений динамики толпы предоставляют значительные возможности для исключения катастроф, связанных с толпой. Рассмотрено много методов оценки плотности толпы, определены их сильные и слабые стороны.

© О.М. ГОЛОВІН, 2019

**Ключевые слова:** компьютерное зрение, плотность толпы, интеллектуальное видеонаблюдение, обработка изображений.

**Вступ.** Людське населення у світі різко збільшується, темпи міграції та урбанізації по всьому світу зростають також. Як результат, ці зміни опосередковано вже привели к зростанню такого явища, як натовп людей. Великі скупчення людей можна спостерігати в закритих місцях (таких як великі зали в будівлях, аеропорти, стадіони, станції метро та інше) і на відкритих майданчиках (таких як пішохідні доріжки, парки, місця проведення спортивних та публічних заходів). При цьому мета скупчень людей має важливий вплив на масштабні властивості та поведінку натовпу. Тому аналіз динаміки та поведінки натовпу викликає великий інтерес у багатьох наукових дослідженнях з психології, соціології, державних служб, безпеки та комп'ютерного зору.

Саме турбулентність натовпу є типовою причиною катастроф при значних скупченнях людей, які виникають внаслідок натискань, паніки, масового руху людей або тисняви, та призводить до загальної втрати контролю над ситуацією. На жаль, печальної статистика про катастрофи, пов'язаної з натовпом, достатньо у відкритій літературі. Для запобігання таких смертельних аварій необхідне якнайшвидше автоматичне виявлення критичних і незвичайних ситуацій у великому натовпі. Як результат, це, безумовно, допоможе прийняти відповідні рішення щодо контролю за надзвичайними ситуаціями та безпекою людей.

З точки зору розпізнавання, моніторингу та аналізу поведінки людей, ще до кінця не вирішеною і важливою темою залишається виявлення та ідентифікація людей в динамічних сценах. Складність розв'язання цієї задачі суттєво ускладнюється наявністю в процесі відстеження однієї і тієї ж людини часткової або повної оклюзії (особливо, коли фіксація об'єкта уваги виконується з бокових камер). У випадку аналізу високого рівня допускається сегментування лише окремих частин тіла людини (наприклад, голова, обличчя, руки та кисті рук) і тоді отримані дані можна використовувати, наприклад, в розпізнаванні жестів або взаємодії «машина-людина». У випадках, коли людина є частиною структур більш високого рівня, наприклад, групи чи натовпу [1] дослідженню підлягає поведінковий аналіз сцен натовпу

людей, який є предметом інтересу в великій кількості таких застосувань, як [2].

- **Керування натовпом:** його можна використовувати для розвитку стратегій керування натовпом людей з метою уникнення катастроф, пов'язаних із скупченням людей та забезпечення громадської безпеки.

- **Дизайн громадського простору:** надання рекомендацій щодо дизайну громадських просторів (приміщень в аеропортах, на залізниці, стадіонах та інше).

- **Віртуальне середовище:** його можна використовувати для перевірки або підвищення швидкодії математичних моделей.

- **Відеоспостереження:** його можна використовувати для автоматичного виявлення аномалій та тривожних ситуацій.

- **Інтелектуальні середовища:** за їх допомогою можна приймати рішення про те, як розділити натовп людей в музеї, на виставках, масових заходах та інше, виходячи з поведінки окремих людей, груп та натовпу.

Незважаючи на значний потенціал застосувань для аналізу натовпу з алгоритмами комп'ютерного зору, більшість існуючих робіт для виявлення та ідентифікації людей, груп людей або навіть для оцінки окремих частин тіла (в аналізі високого рівня) були зосереджені на ситуаціях, не пов'язаних з натовпом. На жаль, у випадках, коли аналізуються відеопослідовності з дуже щільним натовпом людей, звичайні методи комп'ютерного зору не спрацьовують відповідним чином і через це не є доцільними до застосування в подібних умовах [2]. Некоректно вважати, що просте розширення методів, розроблених для звичайних (без натовпу) сцен, може бути придатним і для вирішення ситуацій з натовпом через те, що натовп – це щось більше і складніше, ніж просто сума людей.

В контексті використання методів комп'ютерного зору найбільш вдалим визначенням поняття натовпу є визначення Дж. Макіоніса [3], який розглядає натовп як **«тимчасове зібрання людей, які розділяють спільну увагу і які впливають один на одного»**. Натовп може поєднувати різні та складні форми поведінки, які притаманні окремим його індивідуумам. Поведінка натовпу є зрозумілою, якщо їй притаманні колективні характеристики, які можуть бути описані в загальних рисах. Наприклад, такі описи, як «розлючений натовп» або «мирний

натовп» добре зрозумілі і можуть бути описані визначеним набором характеристик [4].

Розв'язання задач, пов'язаних з поведінкою натовпу людей, вводить додаткові проблеми в методи комп'ютерного зору, які застосовуються для розуміння поведінки людини. В загальному вигляді першим етапом у системах автоматичного аналізу поведінки людини є відстежування людини у часі і просторі. В ситуації з натовпом неможливо точно сегментувати та відстежувати кожного індивідуума, перш за все, через сильні оклюзії. В реальній ситуації точність традиційних методів для відстеження об'єктів, як правило, із збільшенням щільності людей у натовпі значно зменшується. Це призводить до потреби подолання зазначених обмежень і вимагає розробку нових методів аналізу.

З цих причин альтернативний підхід полягає у представленні натовпу як єдиного цілого замість розгляду кожного індивідуума натовпу як його окремої частки. Такий тип аналізу в більшості випадків використовується у двох основних сферах застосування, а саме для вимірювання рівня комфорту натовпу та виявлення деяких конкретних подій (наприклад, незвичайна поведінка групи людей). Головна мета методів аналізу натовпу на основі комп'ютерного зору – отримання інформації з відеопослідовності про щільність скупченості людей, яку в подальшому можна використати на користь великій кількості застосувань (наприклад, спостереження, проектування просторів загального користування з можливим скупченням великої кількості людей, аналіз небезпеки натовпу людей).

Інша важлива і складна проблема, пов'язана з явищем натовпу, – це моделювання самого натовпу, яке потребує відтворення реалістичних картин цього явища на основі методів комп'ютерної графіки. Задачі анімації натовпу знаходять застосування у багатьох сферах, включаючи розваги (наприклад, анімація великої кількості людей у фільмах та іграх), створення ілюзії занурювання у віртуальному середовищі та оцінювання методів керування натовпом (наприклад, моделювання потоку людей, які залишають футбольний стадіон після матчу).

Слід зауважити, що в минулому синтез та аналіз натовпу людей виконувався у незалежній один від одного манері. Але погляд на ці два процеси змінився і комплексне використання комп'ютерного зору та комп'ютерної графіки

стрімко зближує ці дві проблеми. Наприклад, дуже важливо при аналізі натовпу за допомогою комп'ютерного зору отримати достовірне підтвердження результатів. З одного боку, не просто знайти повністю еталонні дані про реальні сцени з натовпом (особливо в надзвичайних ситуаціях), і дані, згенеровані за допомогою методів комп'ютерної графіки, що можуть бути використані для перевірки або навчання алгоритмів комп'ютерного зору [5]. З іншого боку, моделі натовпу можуть бути більш реалістичними, якщо вони покладаються на інформацію, отриману з реального життя. Наприклад, у праці [6] траєкторії окремих індивідуумів у нелюдних ситуаціях отримуються автоматично із застосуванням методів комп'ютерного зору та використовуються для генерації екстрапольованих карт швидкості, які у подальшому поступають в імітаційну модель з основними напрямками та швидкостями з зафіксованого відеокамери середовища. При тому, що велика кількість методів моделювання натовпу запропонована багато років назад, багато проблем все ще залишаються невирішеними через складність людської поведінки та велику кількість ступенів свободи, наявних у реальних сценаріях. У тих випадках, коли для алгоритмів комп'ютерного зору як еталонна картина реальної ситуації приймаються результати методів моделювання натовпу, виникають такі питання, як «Якого рівня реалізму можна досягти в моделюванні натовпу?» та «Які поведінки можуть бути реалістично змодельовано за допомогою симуляторів натовпу?».

## I. Динаміка натовпу

Натовп можна охарактеризувати, приймаючи до уваги чотири різні аспекти, а саме:

- область простору зображення;
- область соціології;
- рівень сервісу (як описано в роботі [7]);
- область комп'ютерної графіки.

У просторі зображення очевидним є те, що натовп визначається тоді, коли щільність людей настільки велика, що окремий індивідуум втрачає свої видимі особливості і вилучається з подальшого розгляду, а видиме у просторі зображення кількість людей сприймається і ідентифікується як окрема група. Відповідно до цієї концепції [8] така характеристика, як щільність людей, використовується у багатьох процесах прийняття рішень стосовно тих подій,

які мають місце у натовпах, наприклад, при моделюванні уникнення зіткнень.

У соціологічній галузі поведінку груп людей науковці вивчають протягом кількох років і їх в основному цікавить результат, який з'являється, коли люди з однаковою метою стають однією єдиною сутністю і отримують назву «натовп» чи «маса». У цьому випадку люди можуть втрачати свою індивідуальність і приймають поведінку натовпу, ведучи себе по-іншому, ніж якби вони були одні [9]. Головним чином, це означає, що колективна сутність може з'являтися залежно від багатьох аспектів, наприклад, цілі людей, спостережуване оточення, подія, що сталася, а також інші вхідні дані.

Рівень послуг [7] представляє деякі умови для натовпу з точки зору щільності людей та її часової еволюції, яка є дуже корисною при проектуванні місць громадських зборів. J. Fruin представляв різні рівні сервісу, описуючи характеристики людей на кожному рівні, з точки зору щільності, потоків та руху популяції.

У галузі комп'ютерної графіки в останні роки запропоновано багато різних варіантів моделювання натовпу [10]. Деякі з них можуть досягати рівня реалістичної поведінки натовпу, але модель, завдяки якій можна було б імітувати всі можливі варіанти поведінки натовпу, досі ще не запропонована. В дійсності велика проблема в моделюванні натовпу полягає у знаннях про справжні натовпи і тому багато варіантів поведінки реальних натовпів ще досі не спостерігалися, не доводилися або не роз'яснювалися.

Неможливо оминати ще один такий важливий аспект у динаміці людей, що рухаються, як особистий простір. В роботі [11] для опису соціального використання простору було запропоновано термін "проксеміка". Проксеміка, або особистий простір, визначається як область з невидимими межами, що оточують тіло людини. Ця область являє собою зону комфорту під час міжособистісного спілкування і може бути зменшена у конкретних середовищах (переповнених середовищах).

## II. Аналіз натовпу

Аналіз натовпу з використанням комп'ютерного зору зосереджується на трьох основних аспектах стосовно автоматизованого аналізу сцен із скупченням людей:

- а) моделі підрахунку/оцінювання щільності людей;
- б) стеження в сценах із скупченням людей та оцінка основних потоків натовпу;
- с) розуміння колективної поведінки.

### *A. Моделі підрахунку/оцінювання щільності людей*

Підрахунок або оцінка щільності людей в натовпі з відеопослідовності є важливим процесом в аналізі натовпу оскільки щільність є одним з основних описів статусу натовпу. Автоматизованому оцінюванню і підрахунку щільності натовпу приділяють багато уваги для контролю безпеки і, особливо, у моніторингу та керуванні потоками людей. Його можна використовувати не тільки для розвитку постачальників послуг у громадських місцях, а і для вимірювання рівня комфортності натовпу та виявлення потенційних ризиків з метою запобігання катастрофам, пов'язаним із скупченням людей. В системах візуального моніторингу розмір натовпу є одним з важливих первинних показників для виявлення таких загроз як заворушення, насильницький протест, бійка, масова паніка та хвилювання [12]. Інформація про динаміку щільності натовпу дозволяє визначити області, де формуються натовпи. Місця, де можливе нарощування натовпу, слід визначити до самої події або на підступах до місця проведення громадських заходів. Це важливо, оскільки натовпи зазвичай існують у певних місцях або в конкретний час доби. Звісно, що існують «стабільні» місця з щільним потоком людей, щільність якого у часі змінюється незначно. Подібні ситуації мають місце в аеропортах, на залізничних станціях і метрополітенах. На жаль при виникненні навіть незначної аномалії в поведінці окремого індивідуума натовпу або виникненні аномального явища ця «стабільність» може швидко перерости в катастрофу, а щільність натовпу при цьому перевищить певний прогнозований рівень [13]. Відеопотоки спостереження подібних «стабільних» сцен, отримані великою кількістю камер, які відстежують ситуацію в зоні своєї відповідальності, повинні постійно продивлятися обмеженою кількістю операторів-спостерігачів, а алгоритми комп'ютерного зору можуть суттєво посилити їх можливості в оперативному виявленні аномальних подій та попередженні спостерігачів. Такі ймовірні ситуації потребують

безперервного проведення автоматичної оцінки щільності натовпу в реальному часі.

Крім того, щільність натовпу та рух натовпу можуть бути пов'язані. Проведеними дослідженнями [14] було доведено:

- що щільність оточуючого натовпу впливає на швидкість руху людини;
- рівняння руху людей, яке визначало реакцію натовпу на натиски з різних напрямків (спереду, ззаду та з боків). Це свідчить про те, що натиски ззовні суттєво впливають на швидкість людей у натовпі;
- що вплив швидкості руху натовпу під впливом тиску між людьми в бічному напрямку значно нижче, ніж у напрямку зпереду назад.

Взагалі проблему оцінки щільності людей та підрахунку натовпу можна розділити на прямі та непрямі підходи [15].

Прямі підходи націлені на одночасне визначення кількості людей (шляхом ідентифікації окремих осіб) та їх місцезнаходження. Після цього підрахунок кількості людей є тривіально досяжним поки сегментація людей виконується коректно і впливи перспективи та щільності скупчення людей відсутні. З появою умов, коли в сцені наявні натовпи та оклюзії, процес виявлення людей стає більш складним завданням і результати стають недостовірними. Як варіант, ці проблеми вирішувалися шляхом адаптації детекторів, наприклад, детектор на основі визначення окремих частин людини (лише голови), детектор форми Омега ( $\Omega$  – літера, яка схожа на голову і плечі людини), або детектор пішохода [16].

У непрямому підході (його також називають методом на основі карт, вимірювання та ознак) підрахунок людей проводиться зазвичай за допомогою вимірювань деяких ознак за допомогою алгоритмів навчання або статистичного аналізу всього натовпу [17]. Методи непрямого підходу зазвичай отримують кілька локальних та цілісних рис із груп людей на зображенні переднього плану. Ці методи є більш ефективними, оскільки виявити ознаки простіше, ніж виявити людей. З цієї причини багато таких ознак з пікселів переднього плану, як область переднього плану, ознаки текстури, гістограми орієнтації контурів або підрахунок ребер, використовуються для оцінки щільності натовпу за допомогою функції регресії, наприклад, лінійної, гауссового процесу або нейронної мережі. Усі ці методи в

основному показали, що взаємозв'язок між площею переднього плану та кількістю людей у сценах майже лінійний. Однак це співвідношення зазвичай не виконується через наявність оклюзій та проблеми перспективи.

Для оцінки кількості осіб у сценаріях із скупченням людей розроблені моделі з застосуванням методів комп'ютерного зору можна поділити на три категорії, а саме: на основі пікселів, текстури і на рівні об'єктів.

**Аналіз на рівні пікселів зображення.** Методи на основі пікселів покладаються на дуже локальні особливості (наприклад, як індивідуальний аналіз пікселів, отриманий за допомогою моделей віднімання заднього фону або виявлення краю) і застосовуються для оцінки кількості людей у сцені. Оскільки в даному випадку використовуються ознаки дуже низького рівня, то цей клас здебільшого зосереджений на оцінці щільності, а не на точному підрахунку людей.

В роботі [4] вперше використано методи комп'ютерного зору для автоматичного отримання певної інформації про натовп. В ній автори запропонували підхід до оцінки щільності натовпу, використовуючи інформацію на рівні пікселів зображення. Для оцінки щільності натовпу поєднано метод віднімання заднього фону та виявлення країв на зображенні.

В роботі [18] представлено метод, заснований на навчанні оцінювати кількість людей у натовпі. Орієнтація краю та гістограма ділянок об'єкта (видобуті з об'єктів переднього плану за допомогою алгоритму віднімання заднього фону) використовуються як ознаки зображення. Для врахування перспективи камери проводиться процедура нормалізації, а модель навчання, яка використовується для встановлення співвідношення виявлених ознак з кількістю людей, ґрунтується на нейронній мережі прямого розповсюдження. Цікавою характеристикою такого підходу є тренування нормалізованих ознак і, таким чином, зміни в налаштуваннях відеокamera не потребують нового етапу тренувань.

Останнім часом все більшої популярності набувають алгоритми комп'ютерного зору, в яких підрахунок натовпу базується на оцінці щільності, а не на принципі «виявлення, а потім підрахунок». До них належать методи, засновані на глибокому навчанні, і які за своєю природою покладаються на дані про значення пікселів. А оскільки значення пікселів, отриманих з відео-

сенсорів, використовуються як вхідні дані, то згортки, як правило, застосовуються на початкових шарах, використовуючи локальну інформацію. Методи на основі оцінки щільності натовпу все більше використовують згорточні нейронні мережі (Convolution Neural Network, CNN) [19] для отримання карти значень щільності (по одній на кожен піксель вхідного зображення). Остаточна оцінка підрахунку може бути отримана шляхом підсумовування по прогнозованій карті щільності. На відміну від підходу «виявлення, а потім підрахунок» [20], вихід підходу оцінки щільності для кожного пікселя не обов'язково є двійковим. Доведено, що методи оцінки щільності є більш надійними, ніж ті, що застосовують підхід «виявлення, а потім підрахунок», тому що перший не повинен брати на себе зобов'язання приймати рішення з двох можливих варіантів на ранній стадії.

Підрахунок натовпу можна формулювати як проблему оцінювання щільності на основі згорточної нейронної мережі, але ця проблема може бути складною для однієї мережі через величезну різницю значень щільності в пікселях різних зображень. Оцінка щільності натовпу на піксель є складним завданням через велику різницю значень щільності натовпу. Як показано на рисунку, деякі зображення містять сотні людей, а інші лише декілька. Одній згорточній нейронній мережі CNN важко обробити весь спектр щільності натовпу. Ранні роботи [21, 22] вирішили цю проблему, використовуючи багатоклонну архітектуру чи комутаційну архітектуру CNN.

Ці архітектури CNN складаються з трьох паралельних гілок CNN з різними розмірами сприйнятливої області. В таких архітектурах гілка з меншими сприйнятливими полями може добре обробляти зображення натовпу з високою щільністю, тоді як гілка з більшими сприйнятливими полями може обробляти зображення натовпу з низькою щільністю. Не так давно запропонована архітектура CNN [23] з п'ятьма гілками, де три гілки нагадували попередню багатоклонну CNN [21], а решта гілок виступала як глобальні та локальні контекстні оцінювачі. Ці гілки контекстного оцінювача заздалегідь були навчені відповідній задачі класифікації зображення на різні категорії щільності. Декілька ключових висновків із зазначених попередніх підходів є:



РИСУНОК. Показано два зображення із набору даних Shang- haitech, які мають дуже різну щільність натовпу

- використання багатоклонної моделі CNN із змінними розмірами ядра покращує ефективність оцінки щільності натовпу;
- розширення набору ознак за допомогою тих, які навчені з завдання, пов'язаного з оцінкою щільності, як то класифікація діапазону ідрахунків, покращує ефективність задачі оцінки щільності.

**Аналіз на основі текстури.** Алгоритми, що спираються на аналіз текстури, вивчають більш грубі зерна порівняно з методами на основі пікселів, оскільки моделювання текстур вимагає аналізу невеличких ділянок (плям) зображень. Хоча цей клас досліджує ознаки більш високого рівня в порівнянні з підходами на основі пікселів, він здебільшого використовується також для оцінки кількості людей на сцені, а не для визначення людей.

В роботі [24] проаналізовано чотири методи, що використовують в аналізі текстури, та три класифікатори для вирішення проблеми оцінки щільності натовпу. Щодо аналізу текстури, було проведено порівняння чотирьох методів, у основі яких: матриця залежності рівня сірого,

сегменти прямих ліній, аналіз Фур'є та фрактальна розмірність. Для порівняння класифікаторів взято три методи: нейронна мережа, статистика (байєсівська) та підхід на основі емпіричної функції. Найкращі результати отримані при використанні методу на основі матриці залежності рівня сірого у поєднанні з класифікатором Байєса. При цьому забезпечувалися кращі контрастність і однорідність, як ознаки текстури. Однак слід зазначити, що автори генерували еталонну інформацію емпірично, а це, безумовно, могло впливати на результати порівняння. Проведена оцінка щільності натовпу в одному з п'яти наступних класів: дуже низька щільність, низька щільність, середня щільність, висока щільність і дуже висока щільність. При цьому було встановлено, що метод не може дуже добре дискримінувати різницю між високою та дуже високою щільністю натовпу.

В роботі [25] представлено автоматичний метод виявлення аномальної щільності натовпу за допомогою текстурного аналізу та навчання. Для цього використовується модель перспективної проекції і створюється серія клітинок зображення з різною роздільною здатністю для кращої оцінки щільності натовпу в сцені уваги. Розмір клітинок нормалізується для отримання рівномірного представлення особливостей текстури, а для зменшення нестабільності вимірювань текстурних особливостей застосовується техніка пошуку екстремуму в просторі Харіса – Лапласа. З метою встановлення співвідношення текстурних ознак з фактичною щільністю натовпу сцени для кожної клітинки вхідного зображення формуються вектори ознак текстури, а метод опорних векторів застосовується для вирішення проблеми регресії обчислення щільності натовпу. Нарешті, на основі оцінених векторів щільності знову використовується метод опорних векторів для вирішення класифікаційної проблеми виявлення аномального розподілу щільності. Експерименти на реальних даних показали ефективність запропонованої системи, але недоліком цього підходу є необхідність перенавчання класифікатора на основі методу опорних векторів сценаріїв з різними налаштуваннями камер, оскільки клітинки щільності сильно залежать від параметрів камери.

A. Chan та ін. [26] розробили алгоритм підрахунку натовпу, заснований на техніці сегментації руху на основі текстури та регресії

гауссівського процесу. Автори спочатку сегментують натовп на різні напрямки руху, використовуючи суміш динамічних текстур, і для кожного кластеру руху вони визначають ознаки сегментів (площа, периметр, орієнтація країв периметра, співвідношення периметр/площа), ознаки внутрішніх країв (пікселі загального краю, орієнтація країв, розмірність Міньковського) та ознаки текстури (однорідність, енергія та ентропія). Всі ці ознаки нормалізуються з огляду на перспективу камери, а регресію гауссівського процесу використовують для визначення кількості людей на сегмент.

В наступній роботі автори [27] досліджували подібну ідею, але натомість використовували байєсівську регресію Пуассона (яка є більш адекватною для таких дискретних процесів, як підрахунок людей). Регресія застосовувалась в обох роботах і дала гарні результати, але на жаль вони залежать від процесу сегментації, який може не спрацювати у випадках з неструктурованими натовпами і з помилковим рухом, тобто який не відповідають заданому опису.

**Аналіз об'єктного рівня.** Методи, що спираються на аналіз об'єктного рівня, намагаються визначити індивідуальні об'єкти в сцені. Вони, як правило, дають більш точні результати порівняно з підходами аналізу на рівні пікселів або на основі текстури, але ідентифікація осіб в окремому зображенні або в відеопослідовності, в основному, можлива в натовпах з більш низькою щільністю. В щільніших натовпах, при наявному захараченні та оклюзії вирішити проблему підрахунку людей майже неможливо, незважаючи на останні досягнення комп'ютерного зору та методів розпізнавання шаблонів.

В роботі [28] запропоновано алгоритм для оцінки щільності у три етапи. На першому етапі виконується пошук об'єктів з контуром у формі зображення голови людини в просторі зображення, використовуючи вейвлет перетворення Хаара. На другому етапі аналізуються характерні риси об'єкта, використовуючи метод опорних векторів і маючи на меті класифікувати його як голову чи ні. На третьому етапі виконується перспективна трансформація з метою кращої оцінки щільності всього натовпу.

T. Zhao та R. Nevatia [29] запропонували байєсівський підхід до виконання сегментації людей у натовпі. Рішення отримується за допомогою ефективного метода Монте-Карло марковських ланцюгів, який використовує знан-

ня області як імовірності пропозиції. Знання різних аспектів, включаючи форму людини, зріст людини, модель камери та підказки з зображення, розділення переднього плану та фону (задній план) інтегровано в єдину теоретично обґрунтовану схему. Однак у натовпах високої щільності представлення повного тіла зазвичай не дуже корисне, оскільки суворі оклюзії, як правило, приховують більшу частину тіла (і частину голови). Дану роботу можна було вдосконалити або розширити кількома способами. Наразі вірогідність базується лише на двійковій масці переднього плану, інші підказки з зображення, такі як край або колір, можуть використовуватися для зменшення деяких неоднозначностей, але за рахунок збільшення витрат на обчислення.

Е. Leibe та інші [30] запропонували схему виявлення пішоходів, використовуючи підхід сегментації зверху вниз. Насправді вони досліджують поєднання локальних особливостей (інваріантна за масштабною версією модель неявної форми) та глобальних особливостей (відстань Шамфера) для отримання ймовірності присутності людини, яка вимірюється порівнянням невеликих навчених ділянок зображень зовнішності людини та частоту їх розповсюдження. Цей алгоритм може надійно виявляти та локалізувати пішоходів у відносно переповнених інших людськими сценах з ймовірними варіантами оклюзії. Однак при установці відеокамери збоку створюється занадто багато оклюзій (часткових або повних) в дуже переповнених сценах і в таких сценаріях їх неможливо адекватно обробити.

V. Rabaud та інші [31] представили метод сегментації людей у натовпі. Вони використовують алгоритм відстеження ознак, а саме трекер Kanade – Лукас – Томасі для виявлення рухомих об'єктів у сцені. Потім трекер поєднується з тимчасовим і просторовим фільтром, а алгоритм кластеризації використовується для групування подібних ознак у траєкторію, яка пов'язана з одним об'єктом. Автори підтвердили свої результати, використовуючи три набори даних, що містять еталонну інформацію, сформовану фахівцем. Вони також використовують відеопослідовність із натовпом клітинок, які вміщують у себе еталонну інформацію, і націлені на демонстрацію надійності запропонованого підходу до сегментації індивідів у натовпах різних сутностей, але однорідних між собою. Одне чітке обмеження цього підходу

стосується нерухомих натовпів, де інформацію про рух неможливо дослідити.

В роботі [32] представили байєсівський метод кластеризації без учителя для виявлення незалежних рухів у натовпі. Їх гіпотеза полягає у тому, що пара точок, які рухаються разом, повинна бути частиною однієї сутності. Алгоритм оптичного потоку поєднаний з алгоритмом повного перебору на основі нормованої крос кореляції використовується для відстеження деяких ознак зображення. Тоді алгоритм кластеризації Бейеса без учителя застосовується до групи таких ознак, спрямовуючись на ідентифікацію кожного індивіда, що рухається у натовпі. Цікавою характеристикою є те, що алгоритм не вимагає жодного етапу тренувань чи моделі зовнішності для відстеження осіб. Однак, оскільки передбачається рух індивідуума, як твердого тіла, то алгоритм може давати збій при наявності сильних рухів руки.

М. J. Jones та D. Snow [33] розробили класифікатор для виявлення пішоходів з використанням просторово-часової інформації. Цей класифікатор включає три типи особливостей: Хаар-подібні функції застосовуються безпосередньо до кожного кадру, абсолютна різниця Хаар-подібних ознак у сусідніх кадрах та фільтр зміщеної різниці, який спрямований на зйомку руху пішоходів (було застосовано вісім зміщених напрямків). Потім використовується Adaboost для побудови м'якого класичного каскадного класифікатора на основі набору вручну помічених навчальних зображень (фактично вісім класифікаторів були навчені працювати з вісьмома різними напрямками руху). Цей підхід успішно відрізняє пішоходів від транспортних засобів, але, як правило, не справляється в сценаріях з порівняно щільним натовпом. Також слід зазначити, що зміни в налаштуваннях камери, що стежить за сценою, вимагають нової процедури тренувань, і, таким чином, поява пішоходів, які пройшли навчання з Adaboost, залежить від розташування камери.

#### *Б. Відстежування в сценах з натовпом*

Наступна важлива проблема з аналізу натовпу – відстеження людей, що полягає у визначенні позиції однієї і тієї ж людини в послідовності кадрів. Знання траєкторії окремих індивідуумів у натовпі можуть бути використані для виявлення основних потоків натовпу або аномальної поведінки. При тому, що існує



декілька підходів до загальної проблеми відстеження об'єктів, завади та суворі часткові або навіть повні оклюзії роблять відстеження окремої людини в більш щільному натовпі складною проблемою.

Важливо зазначити, що проблеми підрахунку та відстеження людей тісно пов'язані, оскільки вони мають мету визначити учасників натовпу. Однак проблема підрахунку зазвичай вимагає лише оцінки кількості людей, незалежно від їх місцезнаходження (та тимчасового переміщення). З іншого боку, вирішення проблеми відстежування включає визначення положення кожної людини на сцені як функція часу. Тим не менш, деякі об'єктно-орієнтовані підходи для підрахунку людей, які описані в розділі "Аналіз натовпу", можуть бути використані для ініціалізації алгоритмів відстежування або навіть розширені для виконання як підрахунку, так і відстежування людей.

Описаний в роботі [34] підхід підрахунку людей, також представляє і розширення для відстежування. У своєму підході кожен трек моделюється кольоровим підписом, шаблоном зовнішнього вигляду та маскою імовірнісною цілі, яка є авторегресійною оцінкою інформації переднього плану. Люди, що йдуть близько один до одного, об'єднуються в «групові доріжки (треки)», а окремі доріжки в межах однієї групи згладжуються за допомогою фільтра Калмана з постійною швидкістю. Цей підхід може вирішувати проблему короткочасної оклюзії між ізольованими треками, але, як правило, не справляється коли щільність натовпу стає дуже високою.

В статті [35] запропоновано підходи для відстеження людей у сценаріях із структурованими натовпами з високою щільністю. В даному підході кожен кадр відеопослідовності ділиться на клітинки, кожна клітина представляє лише одну частинку. Людина на зображенні складається з набору частинок, і на кожну людину впливає компонування сцени (перешкоди та бар'єри), а також рух інших людей. Цікавим аспектом цієї роботи є використання концепцій, пов'язаних з моделюванням натовпу (перешкоди та стосунки із сусідами). З іншого боку, оскільки для ініціалізації кожної доріжки потрібна ручна ідентифікація осіб, такий підхід не є досить зручним для автоматичного підрахунку людей. Крім того, оскільки клітини, які використовуються для відстежування, мають фіксований

розмір, проблеми можуть виникати в дальньому полі бічних камер.

У випадках для боротьби з неструктурованими середовищами, тобто коли рух натовпу виявляється випадковим, або коли різні учасники рухаються в різних напрямках у часі (наприклад, на перехресті) можна використовувати трекінговий підхід [36]. В цьому випадку застосовують корельовану тематичну модель (correlated topic model), яка дозволяє кожному розташуванню сцени мати різні поведінки натовпу. В такому підході послідовність відео поділяється на кліпи, що не перетинаються. В кожному з цих кліпів обчислюється оптичний потік, а положення і вектор швидкості квантуються для генерування слова в кодовій книзі, необхідній для корельованої тематичної моделі. Передбачається, що слова руху виникають із генеративного процесу, параметри якого оцінюються за допомогою колекції навчальних відеопослідовностей. Даний підхід справді здатний боротися з дуже щільними натовпами, але, як описано раніше в [35], існує проблема ініціалізації треку.

#### *В. Моделі розуміння поведінки натовпу*

Поведінковий аналіз натовпу – важлива тема дослідження комп'ютерного зору. Загалом для оцінки поведінки натовпу в заданому середовищі використовується така часова інформація, як основні напрямки [36], швидкості [4] та незвичайні рухи [32]. І хоча вже запропоновано велику різноманітність підходів для вирішення аналізу та розуміння натовпу, складним завданням залишається проблема підтвердження достименності отриманих результатів при роботі з сценами натовпу, оскільки наземні відео матеріали, що містять специфічні аномальні варіанти поведінки у більш щільних натовпах, в основному недоступні. Для подолання цієї проблеми деякі дослідники [37] використовують алгоритми моделювання натовпу для генерування керуваних і заздалегідь прогнозованих або еталонних ситуацій для перевірки своїх алгоритмів. Насправді, головною метою досліджень концепцій, пов'язаних з моделюванням натовпу, є виявлення нормальної та аномальної поведінки [38].

Усі відомі підходи до аналізу поведінки натовпу можна розбити на дві великі групи [38], а саме «об'єктно-орієнтовані», в яких натовп трактується як сукупність людей, і "цілісні" підходи, які трактують натовп як єдине ціле, не

потребуючи сегментації кожної окремої людини. Зрозуміло, що в дуже щільних сценах дуже важко відстежувати окремі компоненти в натовпі, і тому другий підхід, як правило, є більш доцільним.

**Об'єктно-орієнтовані методи розуміння поведінки натовпу.** В об'єктно-орієнтованих методах розуміння поведінки натовпу здійснюється через сегментацію та виявлення людей для подальшого аналізу групової поведінки людей. Наприклад, ідентифікація окремої людини, яка рухається проти домінуючого потоку (наприклад, одна людина, яка намагається вийти на спортивну арену після закінчення матчу), може бути пов'язана з потенційно небезпечною ситуацією. Під час скоєння викрадення злодій зазвичай підходить до жертви ззаду, а автоматичне виявлення цієї події вимагає індивідуальної ідентифікації обох треків (злодія та жертви). При тому, що об'єктно-орієнтовані підходи до розуміння поведінки натовпу дозволяють виявляти події високого рівня, вони стикаються із значною складністю при виділенні окремих людей у більш щільних натовпах, через що ці підходи більше підходять до застосування для сцен з низькою або помірною щільністю скупчення людей.

У роботі [39] запропоновано алгоритм виявлення та класифікації груп з використанням камери, направленої зверху вниз для відстежування людей, і застосуванням діаграм Вороного для кількісної оцінки соціологічної концепції особистого простору. Тимчасова еволюція діаграм Вороного використовується для ідентифікації груп у сцені, а частка особистого простору в полі зору кожного індивідуума використовується для класифікації груп. Цікавим аспектом [39] є використання соціологічних аспектів для аналізу натовпу, але одним недоліком є необхідність відстеження кожної людини на сцені, що може бути складним у більш щільних сценах.

А.М. Cheriyadat і R. Radke [36] запропонували підхід для кластеризації набір функцій руху низького рівня в траскторії, аналогічно до [34] та [35], але з використанням додаткових правил у процесі кластеризації, таких як домінуючі рухи, які обчислюються на основі найдовших загальних підрядів.

Однак, хоча в [31] та [32] метою було визначити кожного учасника сцени на основі рухів, головна мета в роботі [36] полягала у виведенні домінуючих моделей руху в скупченій сцені.

Тоді окремі рухи, не узгоджені з домінуючими потоками, можуть бути виділені та позначені як потенційна незвична поведінка. Наприклад, завдяки цьому підходу можна виявити людину, яка робить розворот у сценарії, де більшість людей рухаються в одному напрямку.

**Цілісні методи розуміння поведінки натовпу.** Замість відстежування окремих об'єктів цей погляд зверху вниз розглядає натовп як єдине ціле, яке безпосередньо вирішує проблему щільних оклюзійних натовпів на відміну від об'єктно-методичних методів.

Цілісні підходи зазвичай намагаються отримати більш грубу (глобальну) інформацію, таку як основні потоки натовпу, і вони, як правило, ігнорують локальну інформацію (наприклад, одна людина, що рухається проти потоку).

В роботі [4] запропоновано підхід на основі дискретного перетворення Фур'є (об'єкти, що рухаються швидше, відносяться до високо-частотного рівня, а нерухливі об'єкти – до рівня з постійною) в поєднанні з алгоритмом лінійного перетворення площі для того, щоб відрізнити статичний і рухомий натовпи (корисний показник тисняви або потенційної небезпеки). Його суть полягає у вимірюванні особливостей руху на рівні пікселів або сусідніх пікселів, які потім агрегуються для отримання властивостей руху для більших областей зображення. В подальшому узагальнені результати можуть бути використані для встановлення загальних переважних швидкостей натовпу (напрямок та величина).

R. Mehran та інші [38] розробили підхід для виявлення аномальної поведінки у натовпі, використовуючи модель соціальної сили, яку запроваджено в роботі [40] для моделювання натовпу на основі соціально-психологічних досліджень. Згідно з методом набір частинок накладається на зображення, які переміщуються відповідно до обчисленого оптичного потоку. Потім рух частинок використовується для оцінки соціальних сил. Для виявлення нормальних закономірностей сил у часі, величини векторів сили взаємодії відображаються на площину зображення, отримуючи силовий потік. Далі оцінюється силовий потік ймовірності, використовуючи підхід пакету слів (для нормальних відео), а фіксований поріг оцінюваної ймовірності застосовується для того, щоб класифікувати кадри відеопослідовності як нормальні або аномальні. Цікавим аспектом цієї

роботи є використання соціально-психологічних досліджень для виявлення аномальної поведінки людей у натовпі.

**Висновок.** В роботі розглянуто три важливі проблеми в аналізі натовпу: підрахунок людей/оцінка щільності, відстежування в сценах з натовпом людей та розуміння поведінки натовпу в аналізі вищого рівня, наприклад, еволюція в часі, основні напрямки руху, оцінка швидкості та виявлення аномальних ситуацій.

Алгоритми підрахунку можна поділити на три класи, на основі пікселів, текстури та на рівні об'єктів. Алгоритми об'єктного рівня забезпечують більш точний підрахунок та локалізацію людей на сцені, оскільки вони засновані на індивідуальній ідентифікації. Зазвичай такий клас є адекватним при сцені натовпу з низькою або помірною щільністю, оскільки оклюзії набувають значного значення у сценах натовпу з високою щільністю. З іншого боку, піксельні підходи та ті, які покладаються на аналіз текстури, досліджують ознаки нижнього рівня зображення, не намагаючись ідентифікувати людей у сцені. Ці класи, як правило, менш точні для підрахунку людей, але вони, як правило, працюють краще із сценами натовпу з високою щільністю. Також слід зазначити, що методи, які якимось чином досліджують інформацію камери [26] та [27] є більш гнучкими щодо налаштування камери, яка використовується для моніторингу сцени, а також можуть виправити спотворення перспективи.

Стосовно алгоритмів, орієнтованих на відстежування людей у сценах натовпу слід зазначити, що підходи, які якимось чином вивчають очікувану поведінку натовпу (як взаємодія з перешкодами, що використовується в [38]), обіцяють вирішити суворі оклюзії та захарщення. З іншого боку, ці підходи можуть не бути адекватними для виявлення незвичної поведінки у натовпі (оскільки порушення зазвичай пов'язані з несподіваним рухом). Незважаючи на досягнення для індивідуального відстеження у дуже густих натовпах [38, 39], автоматична ініціалізація кожного треку все ще є проблемою.

Алгоритми розуміння поведінки натовпу можна класифікувати на дві великі групи: на об'єктно-орієнтовані і цілісні. Перша група покладається на знання людей, які формують натовп, що є складним для більш щільних сценаріїв. Крім того, аналіз можливих взаємодій між людьми у натовпі збільшує обчислювальні витрати за мірою збільшення кількості людей.

Отже, подібні алгоритми більше підходять для сцен натовпу з низькою та помірною щільністю. Цілісний підхід розглядає натовп як єдине ціле, тому процес відстежування не є проблемою. Цей підхід є адекватним для сцен натовпу з високою щільністю, але водночас виявлену аномалію у натовпі не можна так добре охарактеризувати як це виконується в групі об'єктно-орієнтованих методів.

Аналіз методів розуміння поведінки натовпу доводить, що алгоритми, які використовують досягнення у вивченні психосоціальних аспектів натовпу, наприклад, виявлення та класифікація групування на основі особистого простору [39], або модель соціальної сили, що використовується в [38], є цікавою тенденцією виявлення аномалій у багатолюдних сценаріях.

#### СПИСОК ЛІТЕРАТУРИ

1. Kilambi P., Ribnick E., Joshi A.J., Masoud O., and Papanikolopoulos N. Estimating pedestrian counts in groups, *Comput. Vis. Image Understand.* 2008. Vol. 110, N. 1. P. 43–59.
2. Zhan B., Monekosso D.N., Remagnino P., Velastin S.A., and Xu L.-Q. Crowd analysis: A survey, *Machine Vis. Applicat.* 2008. Vol. 19, N 2. P. 345–357.
3. Macionis John J. *Sociology.* Prentice Hall, 2005.
4. Davies A.C., Yin J.H., and Velastin S.A. Crowd monitoring using image processing, *IEE Electron. Commun. Eng. J.* 1995. Vol. 7, N 1. P. 37–47.
5. Andrade E.L. and Fisher R.B. "Simulation of crowd problems for computer vision," in *Proc. 1st Int. Workshop on Crowd Simulation.* 2005. P. 71–80.
6. Musse S.R., Jung C.R., J. Jacques C.S., Jr., and Braun A. Using computer vision to simulate the motion of virtual agents, *Comput. Animation Virtual Worlds.* 2007. Vol. 18, N 2. P. 83–93.
7. Fruin J. *Pedestrian and Planning Design.* Mobile, AL: Elevator World Inc. 1971.
8. Boghossian B.A. and Velastin S.A. Motion-based machine vision technique for the management of large crowds, in *Proc. 6th IEEE Int. Conf. Electronics, Circuits and Systems.* 1999. Vol. 2. P. 961–964.
9. Benesch H. *Atlas de la Psychologie.* France: Encyclopedies d'Aujourd'hui, 1995.
10. Daniel Thalmann. *Crowd Simulation.* Encyclopedia of Computer Graphics and Games. 2016.
11. Hall E.T. *The Silent Language.* New York: Doubleday, 1959.
12. Dittrich F., Koerich A., Oliveira L. People counting in crowded scenes using multiple cameras. In: *2012 19th International Conference on Systems, Signals and Image Processing (IWSSIP).* IEEE. P. 138–141.
13. Marana A.N., Velastin S.A., Costa L.F., and Lotufo R.A. Estimation crowd density with minkowski fractal dimension. *Proceedings of IEEE,* pages 3521–3524, 1999.
14. Fang Z., Lo S.M., and Lu J.A. On the relationship between crowd density and movement velocity. *Fire Safety Journal.* 2003. 38(3). P. 271–284.

15. Conte D., Foggia P., Percannella G., Tufano F., Vento M. A method for counting people in crowded scenes. In: 2010 Seventh IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS). IEEE. P. 225–232.
16. Hou Y.-L., Pang G.K. People counting and human detection in a challenging situation. IEEE Trans. Syst. Man Cybern. Part A Syst. Hum. 2011. 41 (1). P. 24–33.
17. Zhang Z., Li M. Crowd density estimation based on statistical analysis of local intra-crowd motions for public area surveillance. Opt. Eng. 2012. 51 (4) 047204-1.
18. Kong D., Gray D., and Tao H. “A viewpoint invariant approach for crowd counting,” in Proc. Int. Conf. Pattern Recognition. 2006. Vol. 3. P. 1187–1190.
19. Krizhevsky A., Sutskever I., Hinton G. ImageNet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems. 2012.
20. Hoai M., Zisserman A.: Talking heads. Detecting humans and recognizing their interactions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2014.
21. Zhang Y., Zhou D., Chen S., Gao S., Ma Y. Single-image crowd counting via multi-column convolutional neural network. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016.
22. Sam D.B., Surya S., Babu R.V. Switching convolutional neural network for crowd counting. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017.
23. Sindagi V.A., Patel V.M. Generating high-quality crowd density maps using contextual pyramid cnns. In: Proceedings of the International Conference on Computer Vision. 2017.
24. Marana A., L. da Costa, Lotufo R., and Velastin S. On the efficacy of texture analysis for crowd monitoring,” in Proc. Int. Symp. Computer Graphics, Image Processing, and Vision (SIBGRAP'98), Washington, DC. 1998. P. 354.
25. Wu X., Liang G., Lee K., Xu Y. Crowd density estimation using texture analysis and learning, in Proc. IEEE Int. Conf. Robotics and Biomimetics. 2006. P. 214–219.
26. Chan A., Liang Z., and Vasconcelos N. “Privacy preserving crowd monitoring: Counting people without people models or tracking,” in Proc. IEEE Conf. Computer Vision and Pattern Recognition. 2008. P. 1–7.
27. Chan A.B. and Vasconcelos N. “Bayesian Poisson regression for crowd counting,” in Proc. IEEE Int. Conf. Computer Vision. 2009. P. 1–7.
28. Lin S.-F., Chen J.-Y., and Chao H.-X. “Estimation of number of people in crowded scenes using perspective transformation,” IEEE Trans. Systems, Man, Cybern. A. 2001. Vol. 31, N 6. P. 645–654,
29. Zhao T. and Nevatia R. “Bayesian human segmentation in crowded situations,” in Proc. IEEE Conf. Computer Vision and Pattern Recognition. 2003, Vol. 2. P. 459–466.
30. Leibe E., Seemann B., and Schiele B. “Pedestrian detection in crowded scenes,” in Proc. IEEE Conf. Computer Vision and Pattern Recognition, Washington, DC. 2005. P. 878–885.
31. Rabaud V. and Belongie S. “Counting crowded moving objects,” in Proc. IEEE Conf. Computer Vision and Pattern Recognition. 2006. P. 705–711.
32. Brostow G.J. and Cipolla R. “Unsupervised Bayesian detection of independent motion in crowds,” in Proc. IEEE Conf. Computer Vision and Pattern Recognition, Washington, DC. 2006. P. 594–601.
33. Jones M. J. and Snow D. “Pedestrian detection using boosted features over many frames,” in Proc. Int. Conf. Pattern Recognition. 2008. P. 1–4.
34. Rittsche J., Tu P.H., and Krahnstoeve N. “Simultaneous estimation of segmentation and shape,” in Proc. Computer Vision and Pattern Recognition, Washington, DC. 2005. Vol. 2. P. 486–493.
35. Ali S. and Shah M. “Floor fields for tracking in high density crowd scenes,” in Proc. European Conf. Computer Vision. 2008. P. II:1–14.
36. Rodriguez M., Ali S., Kanade and T. “Tracking in unstructured crowded scenes,” in Proc. IEEE Int. Conf. Computer Vision, Kyoto, Japan. 2009. P. 1389–1396.
37. Andrade E.L., Blunsden S., and Fisher R.B. “Modelling crowd scenes for event detection,” in Proc. Int. Conf. Pattern Recognition, Washington, DC. 2006. P. 175–178.
38. Mehran R., Oyama A., and Shah M. “Abnormal crowd behavior detection using social force model,” in Proc. IEEE Conf. Computer Vision and Pattern Recognition. 2009. P. 935–942.
39. Jacques J.C.S., Jr., Braun A., Soldera J., Musse S.R., and Jung C.R. “Understanding people motion in video sequences using voronoi diagrams,” Pattern Anal. Applicat. 2007. Vol. 10, N. 4. P. 321–332.
40. Helbing D. and Molnár P. “Social force model for pedestrian dynamics,” Phys. Rev. E. May 1995. Vol. 51, N. 5. P. 4282–4286.

## REFERENCES

1. P. Kilambi, E. Ribnick, A. J. Joshi, O. Masoud, and N. Papanikolopoulos, Estimating pedestrian counts in groups, *Comput. Vis. Image Understand.*, vol. 110, no. 1, pp. 43–59, 2008
2. B. Zhan, D. N. Monekosso, P. Remagnino, S. A. Velastin, and L.-Q. Xu, Crowd analysis: A survey, *Machine Vis. Applicat.*, vol. 19, no. 2, pp. 345–357, 2008.
3. John J Macionis. *Sociology*. Prentice Hall, 2005.
4. A. C. Davies, J. H. Yin, and S. A. Velastin, Crowd monitoring using image processing, *IEE Electron. Commun. Eng. J.*, vol. 7, no. 1, pp. 37–47, 1995.
5. E. L. Andrade and R. B. Fisher, “Simulation of crowd problems for computer vision,” in Proc. 1st Int. Workshop on Crowd Simulation, 2005, pp. 71–80.
6. S. R. Musse, C. R. Jung, J. C. S. Jacques, Jr., and A. Braun, Using computer vision to simulate the motion of virtual agents, *Comput. Animation Virtual Worlds*, vol. 18, no. 2, pp. 83–93, 2007.
7. J. Fruin, *Pedestrian and Planning Design*. Mobile, AL: Elevator World Inc. 1971.
8. B. A. Boghossian and S. A. Velastin, Motion-based machine vision technique for the management of large crowds, in Proc. 6th IEEE Int. Conf. Electronics, Circuits and Systems, 1999, vol. 2, pp. 961–964.
9. H. Benesch, *Atlas de la Psychologie*. France: Encyclopedies d’Aujourd’hui, 1995.
10. Daniel Thalmann. *Crowd Simulation*. Encyclopedia of Computer Graphics and Games. 2016.
11. E. T. Hall, *The Silent Language*. New York: Doubleday, 1959.

12. Dittrich, F., Koerich, A., Oliveira, L., People counting in crowded scenes using multiple cameras. In: 2012 19th International Conference on Systems, Signals and Image Processing (IWSSIP). IEEE, pp.138–141.
13. A. N. Marana, S. A. Velastin, L. F. Costa, and R. A. Lotufo. Estimation crowd density with minkowski fractal dimension. Proceedings of IEEE, pages 3521 {3524, 1999.
14. Z. Fang, S. M. Lo, and J. A. Lu. On the relationship between crowd density and movement velocity. Fire Safety Journal, 38(3):271 {284, 2003.
15. Conte, D., Foggia, P., Percannella, G., Tufano, F., Vento, M.. A method for counting people in crowded scenes. In: 2010 Seventh IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS). IEEE, pp.225–232
16. Hou, Y.-L., Pang, G.K., 2011. People counting and human detection in a challenging situation. IEEE Trans. Syst. Man Cybern. Part A Syst. Hum. 41 (1), 24–33.
17. Zhang, Z., Li, M., 2012. Crowd density estimation based on statistical analysis of local intra-crowd motions for public area surveillance. Opt. Eng. 51 (4) 047204-1.
18. D. Kong, D. Gray, and H. Tao, "A viewpoint invariant approach for crowd counting," in Proc. Int. Conf. Pattern Recognition, 2006, vol. 3, pp. 1187–1190
19. Krizhevsky, A., Sutskever, I., Hinton, G.: ImageNet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems (2012)
20. Hoai, M., Zisserman, A.: Talking heads: Detecting humans and recognizing their interactions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2014)
21. Zhang, Y., Zhou, D., Chen, S., Gao, S., Ma, Y.: Single-image crowd counting via multi-column convolutional neural network. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2016)
22. Sam, D.B., Surya, S., Babu, R.V.: Switching convolutional neural network for crowd counting. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2017)
23. Sindagi, V.A., Patel, V.M.: Generating high-quality crowd density maps using contextual pyramid cnns. In: Proceedings of the International Conference on Computer Vision (2017)
24. A. Marana, L. da Costa, R. Lotufo, and S. Velastin, On the efficacy of texture analysis for crowd monitoring," in Proc. Int. Symp. Computer Graphics, Image Processing, and Vision (SIBGRAPI'98), Washington, DC, 1998, p. 354.
25. X. Wu, G. Liang, K. Lee, Y. Xu, Crowd density estimation using texture analysis and learning, in Proc. IEEE Int. Conf. Robotics and Biomimetics, 2006, pp. 214–219.
26. A. Chan, Z. Liang, and N. Vasconcelos, "Privacy preserving crowd monitoring: Counting people without people models or tracking," in Proc. IEEE Conf. Computer Vision and Pattern Recognition, 2008, pp. 1–7.
27. A. B. Chan and N. Vasconcelos, "Bayesian Poisson regression for crowd counting," in Proc. IEEE Int. Conf. Computer Vision, 2009, pp. 1–7.
28. S.-F. Lin, J.-Y. Chen, and H.-X. Chao, "Estimation of number of people in crowded scenes using perspective transformation," IEEE Trans. Systems, Man, Cybern. A, vol. 31, no. 6, pp. 645–654, 2001.
29. T. Zhao and R. Nevatia, "Bayesian human segmentation in crowded situations," in Proc. IEEE Conf. Computer Vision and Pattern Recognition, 2003, vol. 2, pp. 459–466.
30. E. Leibe, B. Seemann, B. Schiele, Pedestrian detection in crowded scenes, in Proc. IEEE Conf. Computer Vision and Pattern Recognition, Washington, DC, 2005, pp. 878–885.
31. V. Rabaud and S. Belongie, "Counting crowded moving objects," in Proc. IEEE Conf. Computer Vision and Pattern Recognition, 2006, pp. 705–711.
32. G. J. Brostow and R. Cipolla, "Unsupervised Bayesian detection of independent motion in crowds," in Proc. IEEE Conf. Computer Vision and Pattern Recognition, Washington, DC, 2006, pp. 594–601.
33. M. J. Jones and D. Snow, "Pedestrian detection using boosted features over many frames," in Proc. Int. Conf. Pattern Recognition, 2008, pp. 1–4.
34. J. Rittsche, P. H. Tu, and N. Krahnstoeve, "Simultaneous estimation of segmentation and shape," in Proc. Computer Vision and Pattern Recognition, Washington, DC, 2005, vol. 2, pp. 486–493.
35. S. Ali and M. Shah, "Floor fields for tracking in high density crowd scenes," in Proc. European Conf. Computer Vision, 2008, pp. II:1–14.
36. M. Rodriguez, S. Ali, and T. Kanade, "Tracking in unstructured crowded scenes," in Proc. IEEE Int. Conf. Computer Vision, Kyoto, Japan, 2009, pp. 1389–1396.
37. E. L. Andrade, S. Blunsden, and R. B. Fisher, "Modelling crowd scenes for event detection," in Proc. Int. Conf. Pattern Recognition, Washington, DC, 2006, pp. 175–178.
38. R. Mehran, A. Oyama, and M. Shah, "Abnormal crowd behavior detection using social force model," in Proc. IEEE Conf. Computer Vision and Pattern Recognition, 2009, pp. 935–942.
39. J. C. S. Jacques, Jr., A. Braun, J. Soldera, S. R. Musse, and C. R. Jung, "Understanding people motion in video sequences using voronoi diagrams," Pattern Anal. Applicat., vol. 10, no. 4, pp. 321–332, 2007.
40. D. Helbing and P. Molnár, "Social force model for pedestrian dynamics," Phys. Rev. E, vol. 51, no. 5, pp. 4282–4286, May 1995.

Одержано 28.10.2019