

УДК 681.518

**О. В. Збруцький, Д. С. Мішкін**

## **ОГЛЯД СУЧАСНИХ СИСТЕМ ТА АЛГОРИТМІВ ТЕХНІЧНОГО ЗОРУ, ЩО ВИКОРИСТОВУЮТЬСЯ НА БЕЗПІЛОТНИХ ЛІТАЛЬНИХ АПАРАТАХ ТА НАЗЕМНИХ МОБІЛЬНИХ РОБОТАХ**

### **Вступ**

На сьогоднішній день все актуальнішим стають питання заміни людини автономними пристроями – роботами. Проте їх насправді широкому використанню заважає недостатня автономність. Дійсно, найкращі моделі вже можуть самостійно пересуватись, визначаючи та оминаючи перешкоди [1]. Проте автономне виконання операцій наразі обмежується картографуванням та патрулюванням або транспортуванням певних стандартних вантажів по заданій програмі. Майже всі інші операції потребують участі оператора не тільки у вигляді контролера, але й як особи, що приймає рішення (у кращому випадку) або безпосередньо керує виконанням операції. Однією з основних причин цього є недосконалість систем технічного зору та розпізнавання образів, що використовуються у робототехніці.

**Метою** даної роботи є систематизація інформації про сучасний стан систем технічного зору у мобільних роботах та безпілотних літальних апаратах, а також алгоритмах, на яких побудовані ці системи та виявлення невирішених проблем у даній галузі для постановки задач майбутньої дисертаційної роботи.

**Об'єктом дослідження** є системи та алгоритми розпізнавання образів.

**Задачами**, що витікають із мети роботи є:

1. Складення переліку систем та алгоритмів технічного зору, що використовуються у сучасній робототехніці.
2. Оцінка ефективності, надійності та швидкодії перелічених систем.
3. Аналіз сильних та слабких сторін, визначення проблем, що потребують вирішення.

Розділимо огляд на декілька частин. Спочатку буде наведено короткий опис найбільш досконалих систем технічного зору, що встановлені на великих (розмір автомобіля) та малих наземних роботах, далі – систем, що використовуються на безпілотних літальних апаратах. У другій частині наведено огляд сучасних алгоритмів розпізнавання та обробки зображення, що використовуються, або можуть

використовуватись у системах технічного зору. У завершенні роботи наведено перелік переваг та недоліків існуючих систем та алгоритмів і зазначені нерозв'язані проблеми.

## **1. Системи технічного зору в наземних мобільних роботах**

### **1.1. Великі мобільні роботи. Змагання DARPA**

Поштовхом для нового етапу розвитку систем технічного зору стали змагання Агенції з перспективних досліджень США (DARPA) – DARPA Grand Challenge [2]. Згідно умов конкурсу, учасники повинні були представити цілком автономний наземний транспортний засіб, здатний самостійно подолати заданий маршрут швидше, ніж за 10 годин. Маршрут довжиною 212 км являв собою декілька ділянок дороги та бездоріжжя, подекуди із непередбаченими перешкодами, які робот повинен сам виявити та оминати.

Умови конкурсу дозволяли використання системи GPS (причому район змагань – пустеля – майже виключав можливість втрати сигналу), а детальний маршрут надавався організаторами змагань [2]. Через це вимоги до системи технічного зору, в основному, зосереджувались на задачі своєчасного та надійного виявлення перешкод та можливості чи неможливості руху через задану ділянку шляху. Задачі з розпізнавання місцевості та визначення місцезнаходження не ставились.

Система технічного зору переможця змагань Stanley складається з відеокамери, п'яти лазерних дальномірів SICK та двох радарів Smart Microwave Sensors 24 GHz [3].

Відеопотік з камери оброблюється за допомогою комп'ютера Pentium M. Із зображення відкидається зображення неба. З того, що залишилось виділяється ділянка, яка відповідає дорожньому полотну та комбінується з інформацією радару (що будує карту перешкод) і лазерних дальномірів. Результатом роботи системи є дво- та тривимірні карти місцевості, що лежить перед автомобілем.

Усі інші фіналісти змагань мали подібну систему технічного зору. Типова схема системи керування [3] із включенням технічного зору показана на рис. 2.

Умови наступного конкурсу передбачали вміння орієнтуватися вже у динамічному середовищі, а саме – на міських вулицях [4]. Роботи-учасники повинні були пройти маршрут із дотриманням усіх правил дорожнього руху в присутності інших автомобілей.

Система технічного зору більшості фіналістів побудована на декількох лазерних сканерах, радарах та стереокамери, причому камера

використовується здебільшого для розпізнавання дорожньої розмітки та дорожніх знаків [5].

Проте деякі учасники використовували стереозображення для побудови тривимірної картини місцевості та визначення відстані до об'єктів [6].

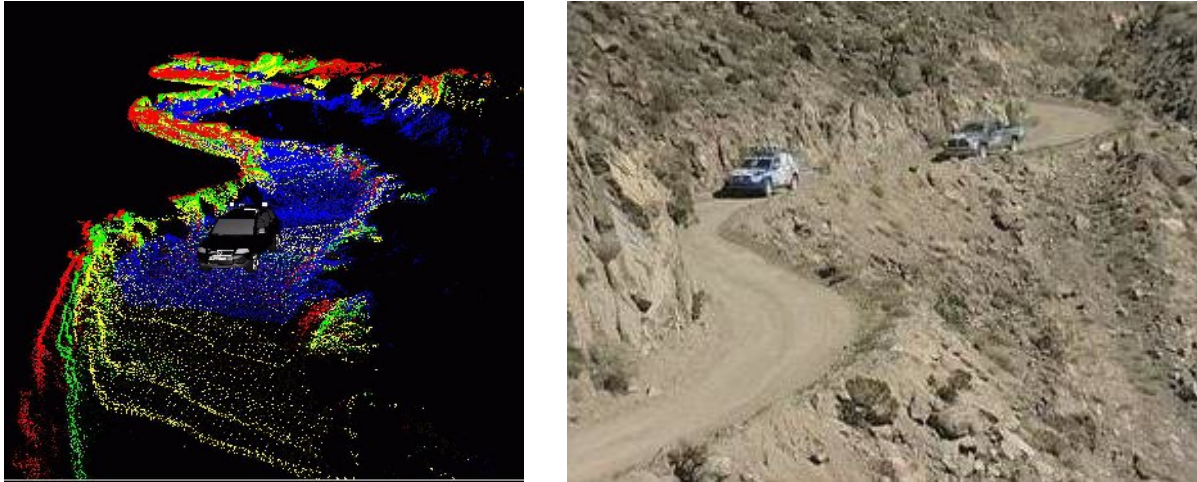


Рис. 1. Результат роботи системи технічного зору Stanley та фотознімок відповідної ділянки шляху [3]

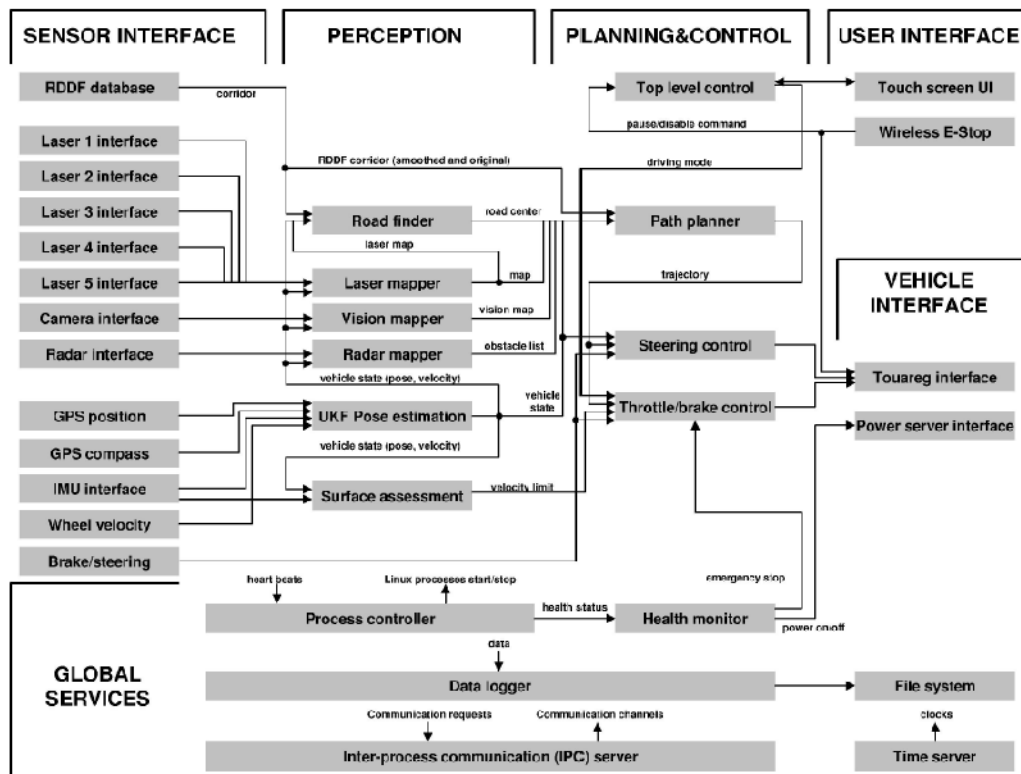


Рис. 2. Схема системи керування роботу Stanley. Система технічного зору представлена блоком PERCEPTION

Система стереозору Alice успішно використовувалась для первинних випробувань для Grand Challenge, проте з невиявлених причин видавала на

систему керування команду до зміщення в один з боків [17]. Крім того, в деяких випадках були хибно ідентифіковані неіснуючі діри в дорожньому полотні глибиною 3 метри, переважно біля середини дороги.

## 1. 2. Малі мобільні роботи

Через невисоку швидкість руху малих мобільних роботів їх використання переважно у приміщеннях та взаємодію з людиною, вимоги до систем технічного зору малих мобільних роботів сконцентровані не на швидкій ідентифікації перешкод, а на класифікації та розпізнаванні конкретних об'єктів, взаємодії з людиною (розпізнавання жестів тощо) [8], [9], [10].

Робот-переможець Semantic Robot Vision Challenge 2008 та 2009 – Curious George – вирішує цю задачу за допомогою алгоритму «областей уваги» [11]. Робот включає в себе стереокамеру для побудови «загального» зображення, лазерного дальноміру для вимірювання відстані до об'єкта, а також цифрову камеру високої розподільної здатності для отримання детального зображення місця, що, ймовірно, містить потрібний об'єкт. Також через те, що вірогідність розпізнавання тривимірного об'єкту суттєво залежить від кута, під яким робот дивиться на об'єкт, система технічного зору може подавати команди системі керування роздивитись об'єкт з різних сторін.

Розпізнавання відбувається за допомогою ознак, отриманих із зображення алгоритмом SIFT [12]. За результатами змагань робот зумів правильно розпізнати 7 із 15-ти об'єктів, зважаючи на те, що еталонні зображення за правилами можна було шукати лише у відкритих інтернет-джерелах протягом декількох годин до початку роботи.

Інший робот-участник змагань Semantic Robot Vision Challenge використовував лише одну відеокамеру та лазерний дальномір [13]. Це зумовило гірші результати, проте у цій системі використаний алгоритм «інформації-з-контексту», що дозволяє робити припущення про вірогідність знайти той чи інший об'єкту певному середовищі [14].

Системи, що представлені іншими розробниками, подібні за можливостями та застосованими алгоритмами.

Попередньо підсумовуючи сучасний стан систем технічного зору для наземних мобільних роботів, варто зазначити наступне:

1. Трьома основними задачами таких систем є визначення перешкод, побудова карти місцевості (приміщення) та розпізнавання заданих об'єктів чи їх класів.
2. Доведена та втілена на практиці можливість визначення перешкод за допомогою лише стереокамери без додаткових датчиків [6]. Проте недостатня точність та надійність таких систем досі не дозволяють використовувати їх окремо від інших.

3. Більшість систем візуальної навігації використовують лазерні сканери та дальноміри. З одного боку, це суттєво підвищує точність та надійність таких систем, з іншого – на порядки підвищує їхню вартість (ціна простої моделі лазерного сканера складає приблизно \$5000)
4. Системи технічного зору здатні розпізнавати задані об'єкти, проте при погіршенні умов спостереження (інші кути обзору, розмиття, інше освітлення), їх працездатність значно знижується.

## **2. Системи технічного зору на безпілотних літальних апаратах**

В [15] наведено основні відмінності умов функціонування систем технічного зору безпілотних літальних апаратів (БПЛА) від аналогічних систем для наземних роботів, що підвищують складність задачі:

- обмежена максимальна маса датчиків, через це суттєво обмежений діапазон можливих моделей (особливо для легких та надлегких БПЛА);
- неможливість використання одометричних датчиків. Для обчислення шляху потрібно двічі інтегрувати показання акселерометрів, що призводить до швидкого росту похибки;
- суттєво обмежена обчислювальна здатність бортових комп'ютерів;
- швидка динаміка руху;
- необхідність одночасного керування швидкістю та місцеположенням (через залежність підйомної сили від швидкості);
- постійний рух (на відміну від наземних роботів, БПЛА не може зупинись на місці; навіть для гелікоптера та квадрантора це є складною задачею).

Проте є і умови, що полегшують задачу для деяких випадків. Це робота з двовимірним зображенням, що не суттєво змінюється при зміні куту зору (за умови, що БПЛА знаходиться на відкритому повітрі та достатньо високо). Таким чином, якщо мобільні роботи працюють у тривимірному середовищі, то БПЛА (за виключення тих, що призначені для польоту у приміщенні) – переважно у двовимірному.

Це ставить такі основні задачі перед системами технічного зору для БПЛА [10], [16]:

1. Пошук та розпізнавання певного об'єкту.
2. Відслідковування (трекінг) визначеного об'єкту.
3. Визначення свого місцезнаходження за зовнішніми орієнтирами.

У статтях [17] та [12] описана система розпізнавання транспортних засобів та людей, що може використовувати як інформацію з відеокамери, так і з тепловізору. Для детектування людей та автомобілів було використано алгоритм розпізнавання на базі Хаар-подібних ознак.

В роботі [18] розроблено систему, що визначає місцезнаходження БПЛА та об'єкта, що відслідковується, методом розпізнавання місцевості. На відміну від кореляційних методів, цей метод є більш стійким до умов освітлення тощо, та ближче до концепції «розпізнавання» зображення. Похибка для наведених у статті умов тестування складає 2,3 метри при висоті польоту БПЛА в 70 метрів.

Система технічного зору БПЛА WITAS [1] вирішує задачі, аналогічні [17], проте вона побудована на визначенні карт швидкості та напрямку руху окремих пікселів зображення та націлена на визначення та відслідковування рухомих цілей.

Попередньо підсумовуючи сучасний стан систем технічного зору для БПЛА, зазначимо наступне:

1. Задача визначення координат БПЛА та цілі, яку відслідковує БПЛА (за наявності знімків земної поверхні) є практично та теоретично вирішеною з достатньо високою точністю (в межах похибки самої геоприв'язки знімка).
2. Задача трекингу (відслідковування) об'єкта також є вирішеною.
3. Задача класифікації (наприклад, визначення автомобіля як «автомобіля», а не людини) є частково вирішеною – через недостатню надійність існуючих рішень.
4. Задача пошуку та розпізнавання цілі потребує подальшого опрацювання.

### **3. Алгоритми виділення ознак та класифікації, що застосовуються у системах технічного зору**

Розпізнавання об'єкта за його зображенням можна розділити на три етапи [19]:

1. Підготовка зображення (препроцесінг).
2. Робота алгоритму виділення ознак, за якими буде проводиться порівняння.
3. Робота алгоритму порівняння(класифікації).

Зазвичай, у сучасних системах препроцесінг є частиною алгоритму з виділення ознак. Чіткого визначення поняття «ознаки» не існує, тому що в кожному алгоритмі під "ознаками" розуміють різні речі. Проте в загальному випадку "ознака" – це частина зображення чи її похідна, що нас цікавить. Наприклад, ознаками можуть бути: карта яскравостей (тобто саме зображення), контури, окремі пікселі, лінії тощо.

На сьогоднішній день використовуються наступні методи виділення ознак [19], [20], [21], [22]:

1. Моментні інваріанти (наприклад, моменти другого порядку є аналогом поняття моментів інерції в механіці, тільки замість маси використовується яскравість пікселя). Вони вперше запропоновані в

роботі [23] і досі використовуються для опису форми об'єкта взагалі.

2. Комплексні фільтри [24]. До зображення застосовуються певні перетворення, після яких воно описується за допомогою комплексних чисел. При цьому повороти в площині зображення впливають лише на фазову частину, а амплітуда лишається незмінною – таким чином забезпечується інваріантність до повороту.
3. Диференційні інваріанти [25] є нечутливими до поворотів зображення та певним чином до затінення його частин. Це є результатом застосування до зображення певного гаусівського матричного фільтру.
4. SIFT (scale-invariant feature transform). Багатоетапний запатентований ([12], [26]) метод визначення ознак зображення, що використовується зараз у багатьох системах технічного зору.
5. GLOH (Gradient Location and Orientation Histogram) – завадостійкий, проте більш вимогливий до обчислювальних ресурсів, варіант методу SIFT.
6. SURF (Speeded-Up Robust Features) [27] розкладає зображення за Хаар-вейвлетами та використовує так звані «інтегральні зображення». Вони являють собою додаткову матрицю, кожна комірка  $(x, y)$  якої містить суму інтенсивностей (яскравостей) пікселів від початку координат до відповідного пікселя  $(x, y)$ . Цей підхід дозволяє значно скоротити обчислювальні витрати на застосування числових фільтрів. Це найновіший із розглянутих алгоритмів.
7. Каскад слабких класифікаторів [28]. Це цілий клас алгоритмів, побудованих за принципом послідовного ("каскадного") застосування слабких фільтрів, що у поєднанні мають сильні дискримінаційні властивості.

Алгоритмів порівняння та прийняття рішень існує набагато більше, ніж можна розглянути у статті, наведемо найбільш популярні у системах технічного зору.

1. Порогове значення. Якщо порогове значення перевищене, то об'єкт причислюється до відповідного класу.
2. Метод найближчого сусіда [29]. Клас точки з множини визначається класом більшості його найближчих сусідів (у найпростішому випадку – одним найближчим).
3. Метод порівняння відстані. Міра "подібності" точки до точок визначеного класу визначається однією з відстаней у евклідовому чи неевклідовому просторі. Одним з найпопулярніших є відстань Махаланобіса [30].

4. Набір слів (bag-of-words). Цей метод запозичений із галузі обробки природних мов. Кожна ознака позначається як «слово» у словнику. За цим методом, чим більше однакових «слів» містять зображення, тим більше вони схожі [31].
5. Словникове дерево (vocabulary tree) [32] є розвитком методу «набір слів», проте тепер слова об'єднуються у деревовидну структуру, що дозволяє підвищити швидкість та якість алгоритму.
6. Нейронні мережі [33] в наш час використовуються відносно рідко.

В статті [22] проведені випробовування алгоритмів розпізнавання на точність та відтворюваність. Результати показують, що хоча описані методи і здатні розв'язувати задачу розпізнавання, проте на практиці не завжди спрацьовують навіть в нормальних умовах. За наявності перешкод їх ефективність знижується.

В статті [20] зазначається, що найбільш ефективним і точним з порівнюваних є алгоритм GLOH, від якого не дуже відстає алгоритм SIFT. Інші суттєво програють в точності та відтворюваності результатів.

В статті [21] продемонстровано принципову уразливість усіх існуючих на момент написання статті алгоритмів до зміни точки зору на об'єкт. Жоден із дескрипторів не зміг впізнати об'єкт при його повороті більше, ніж на 25-30 градусів (див. рис. 3).

## **Висновки**

Галузь технічного зору та застосування його до навігації є достатньо розвиненою та поступово втілюваною в технічні рішення.

Проте лишається цілий ряд невирішених проблем, серед яких однією з найскладніших є створення систем або алгоритмів, які дозволяли б надійно ідентифікувати об'єкт, незважаючи на зміну умов освітлення, перекриття його частини та зміни кута зору. Розробка такої системи чи її окремих елементів є перспективним та актуальним напрямком для наукових досліджень.



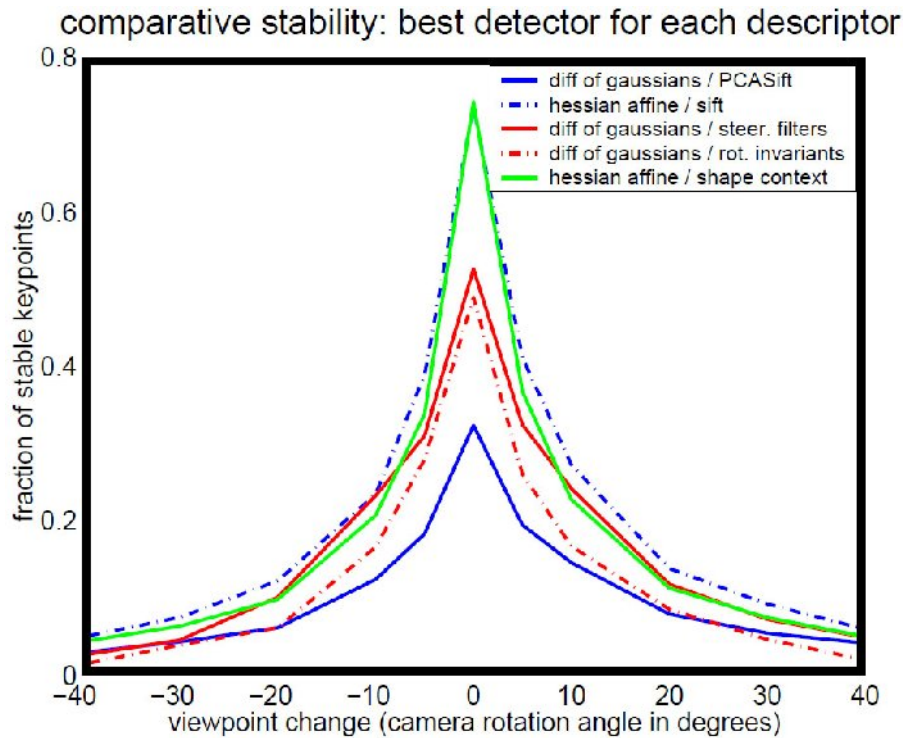


Рис. 3. Залежність частки вірно розпізнаних точок від кута повороту камери для різних методів розпізнавання [21]

### Список використаних джерел

1. *Klas Nordberg Per-Erik Forssen, Johan Wiklund*. A flexible runtime system for image processing in a distributed computational environment for an unmanned aerial vehicle // In the proceedings of the 9th International Workshop on Systems, Signals and Image Processing, IWSSIP . - 2002.
2. *Darpa Grand Challenge 2005 Rules*. - 2004.
3. *Thrun et al*. Stanley: The Robot That Won The DARPA Grand Challenge - Stanford Artificial Intelligence Laboratory, 2006.
4. *Darpa darpa Urban Challenge Rules*. - 2007.
5. *Bohren et al*. Little Ben: The Ben Franklin Racing Team's Entry in the 2007 DARPA Urban Challenge // Journal of Field Robotics. - 2008 - 25p.
6. *Aly et al*. Sensing, Navigation and Reasoning Technologies for the DARPA Urban Challenge. California Institute of Technology, 2007.
7. *Johannesson Erik* Integrated Stereovision for an Autonomous Ground Vehicle. - Lund Institute of Technology, 2005.
8. *SRVC Semantic Robot Vision Challenge Rules*. - 2010.
9. *RoboCup@Home RoboCup@Home Rules & Regulations*. - 2010.
10. *UK Ministry of Defence MOD GRAND CHALLENGE INNOVATION THROUGH CHALLENGE*. - 2008.

11. *David Meger et al.* Curious George: An Integrated Visual Search Platform – 2010.
12. *Lowe D.* Distinctive image features from scale-invariant keypoints // International Journal of Computer Vision. - 2003 p. 91-110.
13. *Kristoffer Sjo et al.* Object Search and Localization for an Indoor Mobile Robot . - Sweden Royal Institute of Technology, 2007.
14. *Mark Antonio Torralba and Kevin P. Murphy and William T. Freeman and* Context-Based Vision System for Place and Object Recognition // Ninth IEEE International Conference on Computer Vision. - 2003.
15. *Markus Achtelik Abraham Bachrach, Ruijie He, Samuel Prentice and Nicholas Roy.* Autonomous Navigation and Exploration of a Quadrotor Helicopter in GPS-denied Indoor Environments // First Symposium on Indoor Flight Issues. - 2009.
16. *IARC Rules for the International Aerial Robotics Competition 6th mission -* 2010.
17. *Toby P. Breckon Stuart E. Barnes, Marcin L. Eichner, Ken Wahren* Autonomous Real-time Vehicle Detection from a Medium-Level UAV. – 2009
18. *Conte G. Hempel M, Rudol P., Lundstr?m D., Duranti S., Wzorek M., Doherty P.* High Accuracy Ground Target Geo-Location Using Autonomous Micro Aerial Vehicle Platforms // Navigation and Control Conference and Exhibit. - Honolulu : AIAA Guidance, 2008.
19. *Brown Lisa Gottesfeld* A Survey of Image Registration Techniques // ACM Computing Surveys. - 1992 . – 24 p.
20. *Schmid Krystian Mikolajczyk and Cordelia* Performance evaluation of local descriptors // IEEE Transactions on Pattern Analysis and Machine Intelligence. - 2005 p.. - 27 : v. 10. - p. 1615-1630.
21. *Perona Pierre Moreels and Pietro* Evaluation of features detectors and descriptors based on 3d objects // IJCV. - 2005.
22. *Arnau Ramisa David Aldavert, Shrihari Vasudevan, Ricardo Toledo, Ramon Lopez de Mantaras* Evaluation of Three Vision Based Object Perception Methods for a Mobile Robot. – 2011.
23. *Hu Ming-Kuei* Visual pattern recognition by moment invariants // IEEE Transactions on Information Theory. - 1962- pp. 179-187.
24. *Schaffalitzky Frederik and Zisserman, Andrew* ECCV '02: Multi-view Matching for Unordered Image Sets // Proceedings of the 7th European Conference on Computer Vision-Part I. - 2002.
25. *Mohr Cordelia Schmid and Roger* Local Greyvalue Invariants for Image Retrieval // IEEE Transactions on Pattern Analysis and Machine Intelligence. - 1997 p.. - 19. - p. 530-535.

26. *Lowe David G.* Method and apparatus for identifying scale invariant features in an image and use of same for locating an object in an image [Patent]. - USA, 6 March 2000 p.
27. *Gool Herbert Bay and Tinne Tuytelaars and Luc Van In* // Surf: Speeded up robust features. - 2006.
28. *Paul Viola and Michael Jones.* Rapid object detection using a boosted cascade of simple features // Proceedings of the Conference on Computer Vision and Pattern Recognition . - 2001. - pp. 511-518.
29. *Hart P.* Nearest neighbor pattern classification // IEEE Transactions on Information Theory. - 1967 . -p. 13.
30. *Mahalanobis P. C.* On the generalised distance in statistics // Proceedings National Institute of Science, India. - 1936.
31. *Sivic, Josef and Russell, Bryan C. and Efros, Alexei A. and Zisserman, Andrew and Freeman, William T.* Discovering objects and their location in images // IEEE International Conference on Computer Vision . - 2005.
32. *David Nister Henrik Stewenius* // Scalable Recognition with a Vocabulary Tree. - 2006. - pp. 2161–2168.
33. *Kanade Henry A. Rowley and Student Member and Shumeet Baluja and Takeo* Neural Network-Based Face Detection // IEEE Transactions On Pattern Analysis and Machine intelligence. - 1998 - pp. 23-38.