

УДК 681.3

А.К. Гігоян, А.С. Чуприна

Харківський національний університет радіоелектроніки, Харків

ДОСЛІДЖЕННЯ МЕТОДІВ ПОБУДОВИ 3D МОДЕЛЕЙ ДЛЯ РОЗПІЗНАВАННЯ СТЕРЕОЗОБРАЖЕНЬ

В статті розглядаються різні методи та алгоритми автоматичної побудови тривимірних моделей об'єктів за стереозображеннями. Проводиться аналіз та порівняльна характеристика щодо їх ефективності, точності та часу отримання результатів. Приведене обґрунтування актуальності дослідження даної теми та приклади практичного застосування описаних методів у різних сферах.

Ключові слова: модель, метод, побудова, стереозображення, алгоритм, ефективність, точність.

Вступ

У сучасному світі комп'ютерні технології розвиваються з кожним днем, та сфери їх застосування поширюються і займають значне місце у житті людини. У зв'язку з цим постійно з'являються проблеми реалізації нових ідей, які потребують вирішення.

Актуальність дослідження тематики 3D технологій та розпізнавання об'єктів зумовлена швидким розвитком сучасної обчислювальної і мультимедійної техніки, їх широким застосуванням в системах штучного інтелекту та технологіях комп'ютерного розпізнавання зорових образів. Методи відновлення просторової конфігурації тривимірних об'єктів за їх стереозображеннями використовуються в багатьох галузях науки і техніки. Побудова тривимірних моделей об'єктів може використовуватися як для автоматизованого отримання цих самих об'єктів у 3D форматах для подальшого їх використання у будь-яких сферах, так і для аналізу та розпізнавання об'єктів присутніх на стереозображеннях, що при якісній реалізації та задовольняючих результатах, може бути дуже корисним в багатьох існуючих системах розпізнавання. Маючи тривимірну поверхню, можна більш якісно та точно виконати процес розпізнавання об'єктів, що знаходяться на них, ніж виконуючи розпізнавання за одним двомірним зображенням. У системах розпізнавання на основі двомірних зображень є ряд істотних недоліків. Наприклад, системи 2D-розпізнавання вельми чутливі до умов освітленості. При нерівномірному освітленні об'єкта достовірність 2D-розпізнавання помітно падає. У той час як для систем 3D-розпізнавання зміни в освітленості впливають лише на текстуру об'єкта, а реконструйована поверхня зберігається незмінною. Для тривимірного розпізнавання об'єктів необхідно саме отримати представлення цих об'єктів у тривимірному просторі. Існує велика кількість методів досягнення цієї мети, але в основному вони потребують багато зусиль та часу з боку людини та нерідко використання дорогого програмного забезпечення.

Останнім часом з'являється все більше автоматизованих систем побудови тривимірних об'єктів, які майже не потребують участі людини у цьому

процесі. Але для отримання якісного результату застосовується складна та дорога техніка. Системи, які використовують тривимірні моделі для розпізнавання є більш точними та надійними. Отримати 3D моделі, що будуть аналізуватися системою, можна різними способами, але найбільш простим та недорогим є побудова тривимірних моделей на основі стереозображень (стереопар) об'єктів. Маючи зображення такого типу можна в подальшому побудувати тривимірну модель об'єкту з них, яка потім може бути розпізнана та класифікована.

Метод побудови 3D моделей за опорними точками

Даний метод являє собою відновлення тривимірної моделі сцени з використанням двох цифрових зображень (стереопари). Процес відновлення тривимірної моделі включає три кроки: **1)** на одному із зображень необхідно вибрати деякі опорні точки, що є елементом зображуваної сцени; **2)** на другому зображенні необхідно відшукати ту ж саму точку; **3)** обчислити тривимірні координати опорних точок.

Метод заснований на ототожненні країв. Яскравість змінюється швидше на краях, ніж на інших ділянках. Різкі зміни (розриви) яскравості, координат кольору або параметрів, характеризують текстуру, є важливими найпростішими ознаками, оскільки вони часто визначають обриси зображених об'єктів. Локальні розриви значень яскравості називаються яскравісними перепадами, або яскравісними контурами. Перепад існує, якщо його кут нахилу і висота більше деякого заданого порогу. Для знаходження перепадів застосовується наступний підхід: кожен піксель і порівнюється з $(i-1)$ -м, знаходиться різниця значення яскравості між ними. Потім різниця значень яскравості порівнюється з порогом і визначається положення елементів з яскраво вираженими перепадами:

$$|E(x_i) - E(x_{i-1})| / E(x_i) \geq A, \quad (1)$$

де $E(x_i)$ – значення яскравості в точці x_i ; $E(x_{i-1})$ – значення яскравості в точці x_{i-1} ; A - значення порога.

Вибір порога є одним з ключових питань виділення перепадів. При занадто високому рівні порога

не будуть виявлені структурні елементи з низьким контрастом. Навпаки, занадто низький рівень порога стане причиною того, що шум буде помилково прийнятий за перепад [1]. Пошук сполучених точок ефективніше виконувати кореляційним методом. Кореляційні методи - це знаходження піксельних відповідей шляхом порівняння профілів яскравості в околиці потенційно відповідних точок різних зображень об'єкта. Розглянемо конкретний приклад - стереопару і точку (x_i, y_i) . З вікном розміру $p = (2m + 1) \times (2n + 1)$ з центром на (x_i, y_i) , співвіднесемо вектор $w(x_i, y_i)$, який отримано шляхом порядкового сканування значень вікна (насправді, порядок сканування не має значення, достатньо лише, щоб він був фіксованим). Тепер для даної точки другого зображення $(x_i + d, y_i)$, потенційно відповідній точці (x_i, y_i) , можна побудувати другий вектор $w'(x_i + d, y_i)$ і визначити відповідну нормовану кореляційну функцію як

$$C(d) = \frac{1}{|w - \bar{w}|} \times \frac{1}{|w' - \bar{w}'|} \times (w - \bar{w}) \times (w' - \bar{w}'), \quad (2)$$

де через \bar{w} позначений вектор, всі координати якого рівні середньому від координат вектору w .

Нормована кореляційна функція C , очевидно, приймає значення з діапазону від -1 до +1. Максимум цієї функції відповідає ділянкам зображення, які відрізняються постійним зсувом і позитивним масштабним множником, а відповідні точки стереозображення можна знайти, визначивши максимум функції C в межах деякого зумовленого діапазону диспаратностей [2]. Розрахунок тривимірних координат опорних точок. Система пасивного стереозору, як правило, включає в себе дві камери. Припустимо, у нас є вирівняний і вимірний стереостенд (рис. 1): дві камери, у яких оптичні осі паралельні і стоять на відстані T одна від одної, а фокусні відстані рівні $f_l = f_r$. Пряма, що з'єднує центри об'єктів називається стереобазою. Базова лінія перпендикулярна оптичним осям і паралельна осі x .

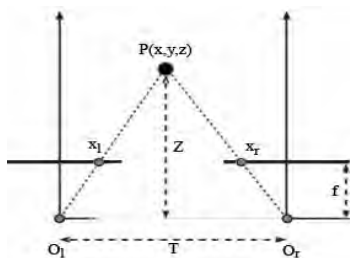


Рис. 1. Схема стереостенду

Координати точки $P(x, y, z)$ визначаються щодо системи координат, початок якої розташований посередині між центрами об'єктів [3]. Є точка $P(x, y, z)$ в реальному світі, яка на зображеннях представляється точками O_l і O_r (сполучена пара). Позначимо координати через (x_l, y_l) і (x_r, y_r) для

лівого і правого зображень відповідно. Тоді враховуючи подібність трикутників отримаємо

$$\frac{z}{f} = \frac{x}{x_l}, \frac{z}{f} = \frac{x - T}{x_r}, \frac{z}{f} = \frac{y}{y_l} = \frac{y}{y_r}. \quad (3)$$

Рівняння моделі описують взаємозв'язок точок у тривимірній системі координат камери з точками на двовимірній дійсній системі координат зображення. Після низки підстановок і перетворень з рівнянь (3) можна отримати невідомі координати x, y, z точки P . Необхідно враховувати перехід від пікселів до лінійних одиниць. Таким чином, підсумкові координати опорних точок в тривимірному просторі запишуться в такому вигляді

$$C = \left[\begin{array}{ccc} \frac{x_l T}{x_l - x_r} & \frac{y_l T}{x_l - x_r} & \frac{-f T}{x_l - x_r} \\ S + \frac{f T}{x_l - x_r} \end{array} \right], \quad (4)$$

де S - коефіцієнт масштабування, зазвичай $S = 1$.

Відстань до об'єкту Z обернено пропорційно диспаратності, причому відстань до близьких об'єктів можна виміряти точніше, ніж до віддалених. Диспаратністю називається відносна відмінність положень об'єктів на стереозображеннях, яка викликається відмінностями в відстані до спостерігача. Зауважимо також, що диспаратність пропорційна відстані між центрами лінз T . Тому при фіксованій похибці вимірювання диспаратності точність визначення глибини буде зростати зі збільшенням бази. На жаль, чим далі рознесені камери, тим менше схожі зображення. Диспаратність пропорційна також ефективній фокусній відстані f , оскільки зображення збільшується зі збільшенням фокусної відстані. На практиці дві камери, за допомогою яких отримують стереопару, не можна точно виставити, як належало при спрощеному розгляді. Важко домогтися строгої паралельності оптичних осей і перпендикулярності до них бази. Якщо дві камери необхідно розташовувати так, щоб в їхньому полі зору потрапляли одні й ті ж об'єкти, можливо, їх слід розвернути так, щоб їх оптичні осі сходилися один до одного в області розташування цих об'єктів. Перш ніж використовувати фотографії, необхідно під час зйомки визначати взаємне положення і орієнтацію камер. Цей процес, званий відносною орієнтацією, визначає перетворення, зв'язує два різні положення фіксованою щодо камер системи координат.

Метод побудови 3D моделей з урахуванням епіполярних обмежень

Основною перевагою цього метода є уникнення перетворень та корекції стереозображень до початку етапу пошуку відповідних точок. Ця модель передбачає можливість пошуку відповідних точок безпосередньо на вихідних стереозображеннях. Побудова такої моделі пов'язана з двома принципово важливими аспектами проблеми. Епіполярні лінії на стереозображенні описуються рівняннями прямої, тому координати відповідних точок на них - дійсні числа. У той же час відповідні фрагменти на стереозображення характеризуються функціями розподілу яскравості на дискретній

множині відліків зображень і можуть бути визначені лише з точністю до величини міжпиксельної відстані. У зв'язку із зазначеною проблемою в рамках загальної моделі обліку епіполярних обмежень виникають такі завдання: **1)** розробка методу і чисельного алгоритму пошуку відповідних фрагментів з дискретними відліками яскравостей зображень на стереозображеннях з використанням безлічі заданих на епіполярних прямих точок, які мають дійсні координати; **2)** розробка і дослідження обчислювальної схеми пошуку відповідних фрагментів в деякій околиці епіполярної лінії шляхом урахування близькості цього фрагмента до неї. У першому випадку ми припускаємо реалізацію схеми, в якій спочатку задається деяка множина точок з дійсними координатами на епіполярних прямих, і для цих точок підбираються найбільш підходящі відповідні фрагменти зображень з використанням спеціально сформованих для цієї мети дескрипторів. У другому випадку найбільш «схожі» відповідні фрагменти підбираються в деякій околиці епіполярних ліній з урахуванням їх близькості до епіполярних ліній [4].

Використання даного методу реконструкції дозволяє уникнути застосування в явному вигляді ректифікації та пов'язаних з нею спотворень зображень. Разом з тим, епіполярні обмеження, що задають рішення в одній площині дозволяють отримати досить високу якість реконструкції сцени.

Метод побудови 3D моделей на основі карт диспаратності

Навідміну від методів, які використовують тільки ключові точки на зображеннях та визначаючи за ними спряжені точки, існують методи, які засновані на побудові карт диспаратності. Ці алгоритми використовують тоді, коли потрібно отримати модель з більш високою чіткістю. Вони дозволяють знайти відповідності для кожної точки зображення, або визначити, що такої точки немає. Вхідними даними для такого методу служать самі зображення. На початковому етапі роботи алгоритму виділяється множина точок на зображеннях, серед яких знаходяться ймовірні відповідності шляхом порівняння їх околиць. Окрім схожості околиць, знайдені пари точок повинні також задовольняти епіполярним обмеженням, для перевірки яких в ході алгоритму обчислюється та уточнюється фундаментальна матриця F . Фундаментальна матриця - це вироджена матриця 3×3 ($\text{rank} F = 2$), яка містить в собі всю інформацію про епіполярну геометрію між двома проєкціями. Зв'язані точки можуть задаватися вручну або відшукуватися автоматично за допомогою алгоритму автоматичного знаходження сполучених точок. Випрямлення зображень дозволяє отримати більш просту епіполярну геометрію:

$$F = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}. \quad (5)$$

Перед подальшим використанням до пар застосовується алгоритм випрямлення, в результаті якого епіполярні лінії стають паралельні осі x . Це досягається завдяки перенесенню епіполюса на нескінченність і мінімізації наступної суми

$$\sum_i d(Hx_i, Hx'_i)^2 \quad (6)$$

В результаті отримаємо, що зв'язані точки на зображеннях матимуть, крім рівних u -координат, також і близькі значення x -координат. На підставі отриманих точок можна побудувати карти диспаратності. Можна взяти лише основні точки зображення і інтерполювати їх, тим самим отримавши безліч точок, придатних для побудови карт диспаратності, але такий підхід не застосовується в реальному світі зважаючи на неточності кінцевого результату. В даний час розроблено ряд алгоритмів, які дозволяють знайти пари сполучених точок для практично всіх точок стереопари. Методи побудови щільної карти диспаратності можна розділити на дві групи: локальні методи, які працюють з невеликими околицями точок для знаходження відповідностей (наприклад, блоковий метод), і глобальні, які працюють з рядками зображення (динамічне програмування) або із цілим зображенням (метод максимального потоку в графі). Локальні методи можуть бути достатньо ефективні, проте вони чутливі до локально-неоднозначних регіонів (наприклад, регіонів з однорідною текстурою). Глобальні методи менш чутливі до таких регіонів, проте вони більш ресурсомісткі.

Найпростіший локальний алгоритм - блочний алгоритм. Він визначає диспаратність, порівнюючи невеликий регіон (блок) навколо точки першого зображення з послідовностей таких же регіонів на другому зображенні в деякій області пошуку. В якості міри схожості блоків зображень часто використовуються сума квадратів різниць інтенсивностей

$$\sum_{u,v} (I_1(u,v) - I_2(u+d,v))^2, \quad (7)$$

та нормована крос-кореляція:

$$\sum_{u,v} |I_1(u,v) - I_2(u+d,v)|, \quad (8)$$

де $I_k(u,v)$ - інтенсивність пікселя (u,v) на k -му зображенні, d - оцінюване значення диспаратності; підсумовування ведеться по прямокутному вікну.

Алгоритм, що використовує динамічне програмування для побудови щільної карти диспаратності працює з парами відповідних рядків на зображенні, розглядаючи їх незалежно від інших рядків, в результаті чого на карті диспаратності можуть з'явитися характерні горизонтальні штрихи. Поняття зображення простору диспаратності (DSI) - явне уявлення простору, в якому ведеться пошук відповідностей. Позначимо через $I_R(x,i)$ і $I_L(x,i)$ інтенсивність точки з координатами (x,i) на правому і лівому зображеннях відповідно, d - диспаратність. Тоді в загальному вигляді DSI можна представити функцією

$$DSI_L^L(x,d) = |I_L(x,i) - I_R(x-d,i)|, \quad (9)$$

де $-d_{\max} \leq d \leq d_{\max}$, $0 \leq x + d < N$, N – ширина зображення. Цей алгоритм дозволяє враховувати опорні точки при пошуку: допустимими в цьому випадку є тільки прямі, що проходять через всі опорні точки. Один з підходів до побудови карти диспаратності полягає в зведенні задачі до знаходження максимального потоку в графі. У побудованому спеціальним чином графі знаходиться мінімальний розріз, який і визначає шукані значення диспаратності. Глобальні алгоритми, мають дуже високу ресурсоемність. Обробка зображень великого дозволу вимагає значних тимчасових витрат. Звідси виникає необхідність пошуку способів оптимізації і прискорення виконання алгоритмів. Одним з підходів, що дають значні результати в збільшенні швидкості виконання алгоритмів, є розпаралелювання. Для паралельної реалізації необхідно зробити виділення незалежних блоків алгоритму, які можуть виконуватися на різних обчислювальних вузлах. Приміром для блочного способу при розпаралелюванні будуть виділятися частини вихідного зображення, і потім оброблятися на різних обчислювальних вузлах.

Порівняльний аналіз методів побудови 3D моделей

Для побудови тривимірних моделей за стереозображеннями існує велика кількість різноманітних способів, але не всі з них завжди доречні. В залежності від умов конкретної задачі, необхідно обирати відповідний алгоритм. Метод побудови за опорними точками добре застосовувати у випадках, коли об'єкти, які підлягають реконструкції, прості або ж припустима велика похибка при побудові. Це обумовлено тим, що відмінністю цього способу є обмеженість кількості точок, за якими будується модель, то ж деякі деталі можуть бути втрачені. Проте цей спосіб є швидким та не ресурсоемним. Метод заснований на побудові карт диспаратності дозволяє обійти обмеження кількості точок, але потребує великих затрат ресурсів та часу, особливо це стосується глобальних алгоритмів побудови карти диспаратності. Для використання цих способів в реальному світі доводиться застосовувати розпаралелення навантаження, проте результат виходить більш точним. Також до обох

зазначених вище способів можна застосувати епіпольярні обмеження. При цьому відсутня необхідність попередньої обробки вихідних зображень. Це дозволяє використовувати більш обширний спектр вхідних даних навіть не завжди дуже якісних, та все одно отримати достатньо точний результат.

Висновки

Достатньо часто на сьогоднішній день, в час, коли все стає комп'ютеризованим та уся інформація зберігається в цифровому вигляді, в різних сферах інформаційних технологій виникає необхідність швидкого та автоматичного розпізнавання та аналізу зображень. Тривимірне розпізнавання є набагато більш ефективним за двовимірне, тому що має більше вхідної інформації для аналізу. Існує велика кількість методів побудови тривимірних моделей об'єктів використовуючи інформацію з їх стереозображень. Маючи такі вхідні дані можна отримати достатньо точну та деталізовану модель об'єкту, яка в подальшому може бути використана у найрізноманітніших сферах.

Проаналізувавши різні методи побудови можна зробити висновок, що не існує єдиного універсального способу. Необхідно обирати підходящий метод в залежності від вхідних даних та необхідного результату. Від вибору алгоритму залежить кількість часу, необхідного для побудови моделі, та рівень деталізації та точності. Також існує можливість комбінувати різні способи для отримання ще кращого результату.

Список літератури

1. Prett U. *Digital processing of images: The lane with English*. U. Prett. – M: World, 1982. – Bookof 2. – 480 p.
2. Forsyte D., Pons. *Computer sight. Modern approach. The lane with English / Forsyte D., Pons.* – M: Williams publishing house, 2004. – 928 p.
3. Horn B.K.P. *Zreniye of robots: The lane with English / B.K.P. Horn.* – M: World, 1989. – 487 p.
4. Гошин С.В. *Модель реконструкції 3D-сцен з урахуванням епіпольярних обмежень [Текст] / С.В. Гошин // Молодий вчений.* – 2014. – № 12. – С. 71-73.

Надійшла до редколегії 21.07.2015

Рецензент: д-р фіз.-мат. наук, проф. С.В. Смеляков, Харківський університет Повітряних Сил ім. І. Кожедуба, Харків.

ИССЛЕДОВАНИЕ МЕТОДОВ ПОСТРОЕНИЯ 3D МОДЕЛЕЙ ДЛЯ РАСПОЗНАВАНИЯ СТЕРЕОИЗОБРАЖЕНИЙ

А.К. Гигоян, А.С. Чуприна

В статье рассматриваются различные методы и алгоритмы автоматического построения трехмерных моделей объектов по стереоизображениям. Проводится анализ и сравнительная характеристика их эффективности, точности и времени получения результатов. Приведено обоснование актуальности исследования данной темы и примеры практического применения описанных методов в различных сферах.

Ключевые слова: модель, метод, построение, стереоизображение, алгоритм, эффективность, точность

RESEARCH OF THE 3D MODELS BUILDING METHODS FOR STEREOIMAGES RECOGNITION

A.K. Gigoian, A.S. Chupryna

The article discusses different methods and algorithms of automatic construction of three-dimensional models of objects from stereo images. The article also provides the analysis and comparisons of their efficiency, accuracy and time of receiving the results, giving the substantiation of the relevance of the researched topic and examples of practical application of these methods in various fields.

Keywords: models, methods, construction, stereo algorithm, efficiency, accuracy