

Обробка, розпізнавання та розуміння сцен

Scene Processing, Recognition and Understanding

Особливості сприймання інформації за допомогою бінокулярного зору

Валерій Алексєєв

Інститут кібернетики АН України

Україна, 252207, Київ
просп. Академіка Глушкова, 40
Tel.: (044) 277-64-11

Викладається ряд ідей про особливості сприйняття обома очима інформації про об'єкти оточуючого середовища. Ці ідеї ґрунтуються на фактах, одержаних в експериментах, спостереженнях та з літературних джерел. Описані ідеї можуть бути використані при розробці систем технічного зору.

Описуються правила сприйняття зорової інформації за допомогою бінокулярного зору при участі текстурних структур і структур первинних елементів. Згідно з цими правилами, сприйняття зоровою системою інформації зі спільногом для обох очей поля зору здійснюється за допомогою правого ока (ПО) і лівого ока (ЛО) «плямами», тобто з невеликих ділянок цього поля зору, названих Інформаційними Мікрополями Зору (ІМЗ). Розміри, форма і місце знаходження ІМЗ не постійні і визначаються інформацією, яка одержується при текстурному аналізі на першому етапі акту сприйняття зорової інформації про сцену і при розв'язанні поведінкової задачі. Місце знаходження, форма і розмір ІМЗ для ПО і ЛО можуть збігатися, що має місце при сприйнятті за допомогою бінокулярного зору інформації про ділянки закриваючої і закритої поверхонь. З деяких ІМЗ інформація може не сприйматися одним оком (наприклад, неведучого), тобто бути «задавленою».

Наявність ефекту приросту-спадання текстури ділянок фонових поверхонь по краях закриваючої поверхні перед стоячим об'єктом говорить про ОБ'ЄМНІСТЬ СЦЕНИ, ЯКА СПРИЙМАЄТЬСЯ, але нічого не говорить про відстань до об'єкту. Відстань визначається за допомогою іншої структури мозку, яка відповідає за візуалізацію інформації.

Сукупність ІМЗ, «сформованих» на першому етапі акту сприйняття, складає Інформаційне Поле Зору (ІПЗ) бінокулярної системи зору, з якого суб'єктом потенційно може сприйматися інформація про сцену. При розв'язанні конкретної задачі в певний момент часу зорова інформація може сприйматися не з усього ІПЗ, а з деякої його частини. Величина ІПЗ буває не меншою за спільне для обох очей поле зору, але і не більшою двох таких спільніх полів зору.

На стиках ІМЗ в ІПЗ має місце перекручення зорової інформації, що порушує цілісність і неперервність зорової інформації про об'єкти сцени. При вертикальному розміщенні голови, при нахилі вперед чи назад перекручення інформації в місцях стиків ІМЗ ПО і ЛО в ІПЗ часто непомітні. А при нахилі голови вліво чи вправо перекручення інколи стають дуже помітними. Якщо на фоновій поверхні існують горизонтальні смуги і (або) лінії, то при нахилі

голови вправо фрагмент смуги (лінії), яка сприймається праворуч від закриваючого об'єкту *A* (об'єкт *A* такої ширини і знаходиться на такій відстані від спостерігача, що частина фонової поверхні не сприймається за допомогою ПО і ЛО), буде нижче ніж той, що сприймається зліва від *A* (при ведучому ПО). Сприйняття зсуву горизонтальної лінії (смуги) можливе при умові, що фонова поверхня справа від об'єкту *A* сприймається за допомогою ПО і зліва — за допомогою ЛО. Це не порушення правил введення інформації, оскільки такий зсув горизонтальної лінії вгору спостерігається тільки в тій області спільногополя зору для ЛО і ПО, в якій об'єкт *A* закриває фонову поверхню, що сприймається за допомогою ЛО, від сприймання за допомогою ПО.

Окреслимо границі IM3 ЛО, в яких відбувається «перекручення» зорової інформації про фонову поверхню, яка сприймається за допомогою ЛО, в порівнянні з зоровою інформацією, що сприймається за допомогою ПО. Правою границею цього IM3 є границя між лівою границею поверхні об'єкта *A* і фоновою поверхнею *B* при розгляданні фонової поверхні за допомогою ЛО. Лівою границею є границя між лівою границею поверхні *A* і фоновою поверхнею, яку видно правим оком. Фонова поверхня за цією областю сприймається ЛО. Зверху і знизу границі області, очевидно, проходять через найвищу і найнижчу точки границі поверхні об'єкта при сприйнятті за допомогою ПО (коли голова не нахиlena вліво чи вправо). В другому випадку границя нечітка (подвійна). Це пов'язано з тим, що для одного ока верхня границя малого об'єкту може бути вище, ніж для другого. Зліва від цієї області фонової поверхні, що сприймається за допомогою ЛО, знаходиться область фонової поверхні, що сприймається за допомогою ПО. У випадку досить простих сцен, довгих горизонтальних ліній і смуг в обох IM3 (справа і зліва від IM3 ЛО) горизонтальні лінії не зсунуті вгору чи вниз відносно тієї ж самої горизонтальної лінії в інших областях фонової поверхні, що сприймається за допомогою ПО.

Якщо на сцені є кілька об'єктів, що закривають один одного і фонову поверхню, то має місце досить складне ІПЗ, в якому багаторазово чергуються області спільногополя зору, що сприймається ПО чи (i) ЛО. Тут «(i)» знаходиться не випадково, тому що в деяких випадках має місце сприйняття зорової інформації з однієї і тієї ж області спільногополя зору. Наприклад, при переміщенні об'єкта (поверхня якого виступає як фонова поверхня для не дуже широкого об'єкту, що стоїть попереду) на фоні досить віддалених об'єктів чи поверхонь одного з об'єктів сцени має місце багаторазове переключення сприйняття зорової інформації за допомогою ПО і ЛО (з відповідними перекрученнями на границях IM3 ПО і ЛО та повторним введенням інформації до 3/5 рухомого об'єкту для ПО и ЛО).

Таким чином, при сприйнятті інформації про реальний світ за допомогою бінокулярного зору в мозок подається досить складна і перекрученна зорова інформація. При візуалізації її на екрані дисплея повинно бути плямисте зображення з перекрученнями (наприклад, зсув горизонтальної лінії в IM3 ЛО і ПО), а також повторення ділянок зображення (наприклад, при наявності рухомого об'єкта частково закритого від спостерігача іншим об'єктом, який знаходиться між спостерігачем і рухомим об'єктом). Розпізнавання за шаблоном такої інформації про об'єкти стає дуже важким (особливо розпізнавання частково закритих об'єктів). Не буде адекватним і розпізнавання всієї сцени. На наш погляд, для роботи з такою зоровою інформацією найбільш придатне розпізнавання за допомогою зорової системи, що має знання про світ у вигляді динамічних моделей об'єктів і сцен. Динамічна модель об'єкту (сцени) містить в собі всі відомі суб'єктові знання про об'єкт. Для мозку «робота» з динамічною моделлю об'єкта рівноцінна «роботі» з реальним об'єктом.

При такій організації ІПЗ необхідно мати особливий механізм ідентифікації зорової інформації. Так, на другому етапі обробки зорової інформації для «роздільування» первинних елементів необхідна система детекторів (нейронних структур), які забезпечують виділення первинних елементів за правилами введення зорової інформації. Такою системою може бути структура, названа екраном (три екрани: детекторів смуг, відрізків різної ширини і кутів). Екран складається з набору детекторів для ПО і ЛО. Детектор виявляє певні первинні елементи в своєму полі зору. Поле зору детекторів ПО і ЛО, які мають тодіжні поля зору, працюють в залежності один від одного і від сигналів з відповідних управлюючих структур (наприклад, текстурних структур).



Методи оперативного відображення обстановки

Альберт Асланян, Георгій Горохов, Сергій Чумаченко

Київський інститут
військово-повітряних сил

Україна, 252186, Київ
Повітрофлотський проспект, 30
Тел.: (044) 271-62-29

Діяльність людини-оператора в контурі управління енергетичною системою базується на його взаємодії з інформаційною моделлю (ІМ), яка містить значення параметру процесу управління і відображається на екранах дисплейів. Забезпечення високої надійності функціонування енергетичної системи, зв'язаної з оглядом місцевості, досягається шляхом застосування ефективних засобів виявлення та розпізнавання об'єктів і елементів місцевості. Одним з найбільш актуальних напрямків забезпечення високої надійності виявлення та розпізнавання є об'єднання даних від сенсорів різних типів.

В системах без глибокої інтеграції дані, одержані від різних сенсорів, застосовувались незалежно одні від одних і в кращому разі забезпечували інформаційну надмірність та можливість взаємного підтвердження. Логічне об'єднання даних від багатьох сенсорів дозволяє одержати корисну інформацію, обсяг якої перевищує просту суму даних, що надійшли від кожного сенсора.

Важливим елементом процесу об'єднання даних є кореляційна обробка сигнатур об'єктів-цилей та інших об'єктів. Звичайно, кореляція останніх для різних сенсорів не виконується, і тому помилкова тривога, що викликана відображенням від сторонніх об'єктів, одержана одним сенсором, не підтверджується даними другого. Якщо при цьому інформація про об'єкт-ціль, передана цими сенсорами корелюється, то система буде мати максимальну ефективність в режимі підтвердження даних.

Основоположним принципом об'єднання даних, які надходять від сенсорів різних типів є відношення правдоподібності $Q = \frac{P^*}{P}$, де P^* — ймовірність одержання інформації про об'єкт-ціль, P — ймовірність одержання інформації про відсутність об'єкта-цилі. При застосуванні декількох сенсорів загальне відношення правдоподібності визначається на основі відношення правдоподібності для кожного з сенсорів. Його конкретне формування залежить від умов інтеграції сенсорів і рівня кореляції даних, що від них надійшли.

Більш специфічним є метод логічного об'єднання даних, які надійшли від сенсорів, що працюють в різних діапазонах частот електромагнітного випромінювання. Цей метод ґрунтуються на припущеннях відсутності кореляції між перешкоджаючим відбиттям від підстелячої поверхні при роботі на різних

діапазонах довжин хвиль. Це означає, що ймовірність шумового сигналу в певному елементі зображення в будь-якому діапазоні не залежить від аналогічної ймовірності в іншому діапазоні. В цьому випадку можна використати множення цих сигналів для більш глибокого заглушення шумів.

Слід відзначити більшу стійкість цього методу об'єднання даних. Навіть при рівні кореляції 30 відсотків ще можна виділити траекторію об'єкту-цилі. Ефективність процесу розпізнавання об'єкту буде в цьому випадку залежати від рівня кореляції сигналів, які надходять двома і більше каналами [1].

Другою стороною цієї проблеми є безпомилковість та своєчасність розпізнавання відображеніх символів ІМ людиною-оператором і виконання ним необхідних дій. Підвищення надійності праці людини-оператора можливе на основі використання графічної ІМ, в якій наглядно здійснюється кодування окремих символів і яка враховує звичні асоціації людини при спостереженні зображень реальних об'єктів і елементів місцевості. Такий підхід до побудови умовних графічних символів дозволяє зменшити ймовірність переплутування символів, скоротити час сприймання ІМ та збільшити інтервал часу, відведений для прийняття рішення.

Методи вибору раціональної форми умовних графічних символів передбачають застосування моделі зорового сприйняття, розробленої на основі принципу ієархічної обробки інформації, який складається з декількох етапів завчасної обробки, стискання і узагальнення, виведення надмірної інформації і формування зорових образів [2].

Формалізований опис символів виконано за допомогою структурних методів, які забезпечують достатньо просту інтерпретацію людиною проваджуваних змін форми символів. Структура символа представлена графом, кількість вершин якого відповідає геометричним прикметам (інформаційним фрагментам точок фіксації погляду), а ребра характеризують взаємне просторове розміщення прикмет символу, який спостерігають. Кожному ребру приписують значення кутів і відстаней між вершинами.

Методи отримання мови опису зображень розглядаються в [3, 4]. Особливості процесу синтезу умовних графічних символів потребують введення до алфавіту мови додаткових засобів, які змальовують тип контурних ліній. При цьому підхіді символи представляються багатовимірним вектором, розмірність якого визначається множиною вершин графа, кількістю напрямків і типів контурних ліній між вершинами.

Ступінь наочності символа, яка характеризує рівень спрошення початкового повного зображення, ідентифікується як відстань між точками образу в багатовимірному просторі. Пропонується в формулу оцінки відстані за евклідовою метрикою ввести коефіцієнт, який враховує особливості зорового сприймання структурних елементів символу, які відрізняються кутовими розмірами.

Вибір форми умовних символів, які забезпечують найбільшу ймовірність розпізнавання, буде проводитись за результатом розв'язку екстремальної задачі при заданих обмеженнях на час спостереження ІМ. Визначається набір символів, для якого відстань між символами задовільняє вимогам максимуму по всій множині можливих спрощених форм символів.

Література

1. *Microwe JOURNAL*. — 1987. — № 9. — РР. 38, 40, 160, 162.
2. Глезер В.Д. Зрение и мышление. — Л. : Наука, 1985. — 246 с.
3. Завалишин Н.В., Мучник И.Б. Модели зрительного восприятия и алгоритмы анализа изображений. — М. : Наука, 1974. — 344 с.
4. Гинзбург В.М. Формирование и обработка изображений в реальном* времени. — М. : Радио и связь. 1986. — 232 с.

Розпізнавання об'єктів і сцен на підставі тактильної інформації

Юрій Кондратенко

Миколаївський кораблебудівний інститут

Україна, 327025, Миколаїв
просп. Героїв Сталінграда, 9
Тел.: (051) 230-94-56

Дана доповідь присвячена питанням розробки перцептивних систем для роботів різного призначення, умови функціонування яких часто змінюються, коли повністю відсутня априорна інформація про параметри об'єкту маніпулювання, тобто об'єкту, з яким виконуватимуться маніпуляційні операції.

В доповіді розглядається адаптивний підхід до вирішення проблеми розпізнавання сцен в робочій зоні робота на підставі тактильної інформації. Тактильна інформація формується перцептивною системою робота. Детальні розглядаються методи розпізнавання об'єктів і сцен з використанням датчиків (*Sensors*) просковзування і датчиків оцінки конфігурації.

Датчики просковзування (*slip displacement sensors*) формують тактильну інформацію при просковзуванні об'єкту поміж пальцями захвату робота. Такі ситуації виникають в тих випадках, коли величина зажимного зусилля не відповідає величині маси об'єкту маніпулювання. Зростання зажимного зусилля буде забезпечуватись системою керування робота до того часу, поки об'єкт не буде надійно зафікований в захваті. На підставі тактильної інформації робот розпізнає об'єкти, що відрізняються між собою величиною маси. При необхідності, розпізнавання об'єктів може забезпечуватись за допомогою датчиків геометричних розмірів, розроблених на підставі тактильних принципів.

В доповіді детально розглянуті шляхи корекції алгоритмів керування при тактильному розпізнаванні, а також ряд нових технічних рішень по створенню тактильних датчиків для роботів, наведені приклади практичного використання розглянутих систем розпізнавання об'єктів і сцен.



Розпізнавання об'ємних тіл і розуміння просторових сцен

Віталій Рибак

Інститут кібернетики АН України

Україна, 252650, Київ
просп. Академіка Глушкова, 40
Тел.: (044) 266-45-89

Розглядаються питання, пов'язані з побудовою геометричної моделі об'ємних тіл та просторових сцен, процедурами одержання та обробки вихідних даних, порівняння результатів обробки з моделями і уточнення просторового положення об'єктів методами аналізу через синтез при вирішенні задач розпізнавання об'ємних тіл та розуміння просторових сцен.

Ключові слова: яскравістні та дальностні зображення, стереосприйняття, геометрична модель, аналіз зображень, синтез зображень, розпізнавання об'єктів, розуміння сцен.

Означення. Зрозуміти просторову сцену означає дати її опис у термінах, що дозволяють виразити властивості проблемної області «користувача». Універсальний опис сцени — опис положення власних систем координат (ВСК) моделей об'єктів сцени у системі координат (СК) моделі сцени і опис моделей об'єктів у ВСК останніх.

Розпізнати об'ємне тіло означає за зображеннями сцени побудувати його модель і знайти її відповідну з числа допустимих моделей класів, а також визначити просторове положення об'єкту.

Уточнити сцену означає виявити відхилення стану відомих об'єктів сцени від очікуваного стану її моделі.

Для досягнення прийнятних для практики результатів при створенні систем комп'ютерного аналізу сцен прикладного призначення в [1] запропоновано два основних шляхи скорочення перебору варіантів. Перший пов'язаний з використанням доступної «незорової» інформації про стан об'єктів сцени і переходом завдяки їй від задачі розпізнавання до задачі уточнення стану сцени. Другий шлях пов'язаний з підвищеннем достовірності результатів обробки вихідних даних за рахунок комплексного використання даних радіометричних та просторових вимірювань і зменшенням за рахунок цього кількості початкових варіантів кандидатів у структурні елементи моделей об'єктів.

У даній роботі розглядаються питання, пов'язані з реалізацією другого згаданого шляху. Запропоновано універсальну геометричну модель об'ємних тіл і просторових сцен. Вона дозволяє описати тіла довільної форми і складних сцен без зміни структури моделі і організації доступу до даних, допускає використання апріорних даних про предметну область і однаково придатна як для розпізнавання об'єктів, так і для синтезу їх зображень.

Геометричну модель репрезентовано у вигляді семантичної мережі з ієрархічною структурою. Задання моделі уніфіковане зверху вниз до рівня примітивів включно. Самі примітиви в залежності від типу об'єктів репрезентовані двома способами. Універсальна репрезентація полягає в описі у ВСК примітиву осьової лінії та описі ліній перетину поверхні примітиву площинами, які перетинають вісь на заданій відстані одна від одної під кутом, що забезпечує мінімальну площину фігури, обмеженої лінією перетину. Окрема репрезентація примітиву полягає в описі структурних елементів (формуючих поверхонь і граничних ліній) та відношень між ними.

При побудові моделі об'ємного тіла шляхом його представлення системі визначення значень параметрів структурних елементів та відношень між ними здійснюється тими ж процедурами обробки даних, що й при розпізнаванні. Тим самим вже на етапі побудови моделі визначаються допустимі класи об'єктів, які можна розрізняти, використовуючи наявні процедури обробки даних.

Дані радіометричних та просторових вимірів (яскравістні та дальності зображення) отримуються за допомогою телевізійної системи стереосприйняття. Відповідні точки на стереопарі зображень визначаються шляхом виділення на кожному зображенні одноіменних елементарних утворень — зображень штучних або природних відміток на поверхні тіла або зображень світлових смуг структурованого підсвічування. При використанні для структурованого підсвічування транспаранта, що забезпечує одночасно підсвічування сцени набором паралельних світлових смуг, проблема виділення відповідних смуг на зображеннях стереопарі вирішується присвоєнням смугам на транспаранті імен (кодів), які визначаються шириною смуги, відстанню між сусідніми смугами та позицією відносно інших смуг. Для відтворення кодів на зображеннях запропоновано алгоритм, що базується на методі динамічного програмування.

Виділення однорідних у певному значенні областей зображень, що використовуються для знаходження кандидатів до структурних елементів, здійснюється шляхом знаходження на яскравістних та дальності зображеннях ліній порушення однорідностей та виділення областей, що примикають до них, шляхом аналізу просторової близькості точок вздовж світлових смуг або ліній перетину

сцени площинами та за допомогою алгоритму росту областей однорідності. Процедури, що використовуються, реалізують принцип обробки вихідних даних із збереженням семантичних зв'язків.

Значення параметрів однорідних областей, що використовуються при співставленні із значеннями параметрів структурних елементів моделей в процесі пошуку кандидатів у структурні елементи, одержують як результат апроксимації масивів точок однорідних областей рівняннями першого та другого порядків.

При розпізнаванні сцени за значеннями параметрів до одного структурного елементу моделі, що перевіряється на відповідність вихідним даним, може бути віднесена значна кількість кандидатів. Вибір відповідних здійснюється шляхом перевірки відповідності відношень між структурними елементами моделі відношенням між відповідними кандидатами. Відношення між структурними елементами, примітивами та більш складними частинами моделі описуються у термінах відстані, кутів, належності, які обчислюються за певними процедурами. Для прийняття рішення розробляються програми, що базуються на методі стохастичної розмітки, запропонованому в [2].

Похибки апроксимації, викликані похибками стереосистеми, доступністю для апроксимації тільки частини точок поверхні тіла та включенням до складу апроксимованих точок тих, що належать іншим поверхням, приводять до неоднозначної відповіді системи розпізнавання. Для подальшого уточнення використовується апарат аналізу через синтез. За даними моделі стереосистеми і моделі тіла з числа моделей допустимих класів, яка відповідає гіпотезі, що превірюється, синтезується яскравістне зображення моделі і перевірюється його відповідність зображеню, що сприймається реальною стереосистемою. При цьому здійснюється пошук найбільш придатного просторового положення ВСК моделі в околі значення, що відповідає гіпотезі. Цей же апарат використовується при розв'язанні задачі уточнення сцени.

Перевірка працездатності алгоритмів, що розробляються, здійснюється шляхом натурних експериментів та імітаційного моделювання. Розроблене математичне забезпечення використовується при створенні прикладних систем технічного зору різного призначення, в тому числі для реалізації інформаційного забезпечення автономних роботів.

Література

1. Rybak V.I. Recognition of 3-D Bodies and Scenes // Proc. First Intern. Conf. on Inform. Technol. for Image Processing and Pattern Recogn. — 1990. — Vol. 1. — P. 90–93.
2. Faugeras O.D., Price K.E. Semantic Description of Aerial Images Using Stochastic Labeling // IEEE Trans. on Pattern Recognition and Machine Intelligence. — 1991. — PAMI-3, № 6. — P. 633–642.

