



Теоретичні засади обробки та розпізнавання сигналів

Theoretical Basis of Signal Processing and Recognition

Багатокритеріальна оптимізація системи розпізнавання випадкових сигналів

Валерій Безрук, Олег Колесников, Віктор Омельченко

Харківський інститут радіоелектроніки

*Україна, 320127, Харків
просп. Леніна, 14*

Тел.: (057) 240-93-33

При розв'язанні прикладних задач розпізнавання в галузі радіолокації, технологічної та медичної діагностики інформація про образи, що розпізнаються, часто надходить у вигляді реалізацій випадкових сигналів. При цьому виникає необхідність побудови систем розпізнавання заданих класів випадкових сигналів. Вибір найкращої, в наперед встановленому розумінні, структури системи розпізнавання включає, в загальному випадку, три етапи: формування в явному або неявному вигляді множини припустимих структур (математичних моделей) системи, обґрунтування принципу оптимальності і розв'язання оптимізаційної задачі з метою одержання оптимальної структури системи.

В цій роботі пропонується і обговорюється методологія розв'язання задачі оптимізації системи розпізнавання випадкових сигналів у разі, коли якість систем характеризується не однією функцією мети, а сукупністю показників ефективності та витрат. Розглянуто особливості всіх етапів розв'язання такої задачі. Поставлено задачу розпізнавання заданих класів гаусівських випадкових сигналів, репрезентованих імовірнісною моделлю у вигляді ортогональних розкладів випадкових процесів. Введено сукупність показників якості синтезованої системи: надійність розпізнавання, швидкодія, а також витрати на проектування і реалізацію системи засобами обчислювальної техніки.

Синтез системи, оптимальної за векторним критерієм, полягає в тому, що знаходиться підмножина ефективних (оптимальних за Парето) припустимих альтернативних варіантів структури систем розпізнавання. При формуванні множини припустимих варіантів використано морфологічний підхід в умовах функціональної декомпозиції системи розпізнавання на три компоненти: датчик початкового опису, формувач інформативних ознак і класифікатор. Структурна множина альтернатив визначається припустимими комбінаціями можливих варіантів вказаних компонент системи. Розглянуто різні способи побудови інформативних ознак і правил прийняття рішень в межах обраної імовірнісної моделі випадкових сигналів.

Структурна множина альтернатив повинна утворюватись цілеспрямовано з урахуванням обмеженості ресурсів на проектування системи в морфологічному середовищі повних базових варіантів структури системи (парето-оптимальних,

або оптимальних за окремими показниками якості). З цією метою введено метризацію морфологічного простору і провадиться аналіз пошарової будови множини альтернатив.

Сформована множина припустимих альтернатив відображається в критеріальному просторі оцінок введених показників якості, де задаються відношення часткової переваги.

Використовуючи відношення строгої переваги, знаходять узгоджений оптимум за Парето показників якості, якому відповідає підмножина парето-оптимальних варіантів структури системи розпізнавання. Знаходження оптимума за Парето призводить до одержання багатовимірних потенціальних характеристик системи — потенціально найкращих значень кожного з показників при фіксованих значеннях решти показників. Парето-оптимізація лише звужує множину припустимих альтернатив за рахунок виключення безумовно гірших варіантів структури системи за відношенням строгої переваги. Для вибору єдиного варіанту системи повинно проводитись строге впорядкування парето-оптимальних варіантів з урахуванням додаткової інформації на заключному стані оптимізації. Розглянуто різні способи розв'язання цієї задачі, які зводяться до поновлення єдиної функції мети та її оптимізації або до знаходження спеціальних способів впорядкування парето-оптимальних варіантів.

Наводиться приклад багатокритеріальної оптимізації системи розпізнавання реальних сигналів, що ілюструє основні етапи розв'язання оптимізаційної задачі. Аналізуються особливості реалізації системи розпізнавання випадкових сигналів сучасними засобами обчислювальної техніки, зокрема з використанням паралельних обчислювальних структур.



Діалоговий пакет програм розпізнавання випадкових сигналів на персональній ЕОМ

Валерій Безрук, Віктор Омельченко, Едуард Шестак

Харківський інститут радіоелектроніки

*Україна, 310127, Харків
просп. Леніна, 14*

Тел.: (057) 240-93-33

Розглядаються особливості побудови та можливості пакету програм (ПП), призначеного для розпізнавання образів, при початковому опису у вигляді реалізації випадкових сигналів.

В ПП реалізовані алгоритми розпізнавання гаусівських випадкових сигналів, які базуються на математичній моделі сигналів у вигляді їх ортогональних розкладів. При цьому сигнали зображуються у традиційних та пристосованих до сигналів базисах.

Рішення приймається по одній реалізації та вибірках реалізацій за допомогою різних варіантів статистичних правил вибору розв'язків. Алгоритми дають можливість розпізнавати сигнали, які задані своїми навчальними класифікованими вибірками реалізацій.

За допомогою ПП може проводитись статистичний аналіз вибірок реальних сигналів, моделювання реалізацій гаусівських випадкових сигналів з типовими кореляційними функціями та енергетичними спектрами, реалізується робота різних алгоритмів розпізнавання сигналів у режимах навчання та екзамену, проводяться порівнювальні дослідження алгоритмів розпізнавання на вибірках

реальних та модельних сигналів з використанням сукупності показників ефективності та затрат.

До складу ПП входять чотири основні програмні модулі та бібліотека обслуговуючих підпрограм. Програмні модулі виконують такі функції:

- програма, що дає коротку інформацію про можливості та структуру ПП, зміст бібліотеки підпрограм, порядок роботи з ПП;
- програма моделювання вибірок реалізацій гаусівських сигналів із заданими статистичними характеристиками, а також проведення попереднього статистичного аналізу вибірок реальних та модельних сигналів;
- програма побудови різних типів інформативних ознак з метою подальшого скорочення розмірності зображення сигналів та оцінки параметрів правил прийняття рішень;
- програма, яка виконує розпізнавання екзаменаційних вибірок сигналів з використанням різних типів інформаційних ознак та правил прийняття рішень;
- тестова програма, яка проводить контроль роботи ПП на всіх етапах його роботи.

ПП має розвинутий сервіс користувача. Вибір режиму роботи ПП, структури та параметрів алгоритмів розпізнавання сигналів проводиться користувачем через систему вложений меню. Уся необхідна для роботи інформація виводиться на дисплей та друкуєчий пристрій. Результати роботи ПП представляються у вигляді таблиць та графіків. Передбачено заходи обходу аварійних ситуацій при невірному вводі інформації в процесі діалогу користувача з ЕОМ.

Допускається модифікація та розширення функціональних можливостей ПП. Для цього передбачено збереження початкових текстів програм та підпрограм. Є також текстовий приклад розпізнавання модельних сигналів, що дозволяє контролювати модифіковані варіанти ПП.

ПП орієнтований на ПЕОМ, сумісні з IBM PC із стандартним набором периферійного обладнання. При вводі реальних сигналів безпосередньо із фізичних датчиків початкового опису розпізнаваних образів повинно бути забезпечене стикування датчиків з ПЕОМ через аналого-цифровий перетворювач. Це дає можливість побудови експертної системи розпізнавання випадкових сигналів.

ПП може бути використаний для дослідження можливостей вирішення різних прикладних задач розпізнавання в галузях радіолокації та зв'язку, технічної та медичної діагностики, геології та сейсмології, в яких інформація про розпізнавані образи поступає у вигляді реалізацій випадкових сигналів.



ІКДП-технологія автоматичного розпізнавання, розуміння та синтезу сигналів мовлення і порівняння її з іншими технологіями

Тарас Вінцюк

Інститут кібернетики АН України

*Україна, 252207, Київ
просп. Академіка Глушкова, 40*

Тел.: (044) 266-4356 Факс: (044) 266-1570

E-mail address: vintsiuk%golos.kiev.ua@relay.ussr.eu.net

Суть ІКДП-технології. Найбільш робастною системою розпізнавання мовлення була б така, яка запам'ятовувала б всі можливі сигнали мовлення як прототипи, а потім, при розпізнаванні, порівнювала б розпізнаваний сигнал з запам'ятованими прототипами. На превеликий жаль, із-за надзвичайного розмаїття мовних сигналів цей підхід до розпізнавання «не проходить»: немає та й не знайдеться

ні тепер, ні в майбутньому таких комп'ютерів, які будуть в змозі запам'ятати всі можливі мовні сигнали й, тим більше, їх порівняти. Але, на щастя, при всьому розмаїтті мовні сигнали є пов'язані сильними детермінованими залежностями, які найкраще описати за допомогою допустимих перетворень, які дозволяють перейти від однієї реалізації мовного образу до другої. Наприклад, всі реалізації одного й того ж слова при вимовлянні одним й тим же диктором відрізняються нелінійно змінюваним темпом вимовляння. Отже, розмаїття мовних сигналів, яке зумовлене змінюваним темпом вимовляння, задається економно описом допустимих перетворень осі часу зі зберіганням його прямого ходу. Задамо, наприклад, кожне слово якоюсь його окремою реалізацією, яку оголосимо модельною реалізацією цього слова або прототипом. Далі визначимо (опишемо) правила допустимих перетворень цих прототипів. За допомогою цих правил будемо утворювати (генерувати) різноманітні похідні модельні прототипи, наприклад, такі, що відрізняються нелінійно змінюваним темпом й інтенсивністю вимовляння.

Нехай нам вдалось якимсь чином побудувати модель, яка економно описує (або дозволяє генерувати) різноманітні модельні сигнали мовлення, які своєю сукупністю більш-менш якісно описують реальне розмаїття мовних сигналів. Якщо так, то таким способом вирішується проблема пам'яті. Далі спробуємо вирішити проблему обчислень — порівняння розпізнаваного сигналу з генерованими модельними сигналами. Ясно, що простий перебір модельних сигналів при порівнянні не підходить. Необхідно відшукати ефективні шляхи знаходження модельних сигналів, які згідно з тією чи іншою мірою є найбільш схожими на розпізнаваний сигнал. Отже, повинен бути забезпечений ефективний направлений пошук оптимальних рішень.

Підсумовуючи, модель (алгоритм) розпізнавання повинна задовольняти двом вимогам: бути й адекватною, й конструктивною. Остання вимога означає економність опису розмаїття мовних сигналів, а також направлений перебір варіантів при порівнянні сигналів.

В ІКДП-методі множини модельних сигналів мовлення складаються (композиція (К)) або генеруються ієрархічною (І) структурою стохастичних породжувальних автоматних граматики, а порівняння розпізнаваного сигналу мовлення з генерованими модельними (рівно ж й формування відповіді розпізнавання) реалізується направленим перебором варіантів і пошуком оптимального рішення за допомогою динамічного програмування (ДП). Відповідь розпізнавання визначається тим згенерованим модельним сигналом, який є найбільш схожим (в певному розумінні) на розпізнаваний сигнал. В ІКДП-методі генерація й пошук оптимального рішення виконуються направлено, але прийняте рішення еквівалентне повному перебору. Синтез мовлення в ІКДП-методі розглядається як процес в зворотньому зв'язку по відношенню до розпізнавання.

Коротка історична довідка. ІКДП-технологія була започаткована в 1966 році. Перша широко відома робота була опублікована в 1968 році [1]. Вона відноситься до розпізнавання окремо вимовлюваних слів. Кожне слово мовлення задавалось своїм початковим еталоном (прототипом) — послідовністю так званих еталонних елементів. З початкового еталону утворювались похідні прототипи: породжувальна граматика дозволяла, зберігаючи порядок слідування еталонних елементів в прототипі, повторювати окремі еталонні елементи. Процес породження похідних прототипів задавався автоматною породжувальною граматиною — графом, що синтезує похідні прототипи (послідовності з еталонних елементів довільної довжини). Розпізнаваний сигнал (послідовність спостережуваних елементів) порівнюється з синтезованими прототипами. Для цього вводились поняття елементарних і

інтегральних мір схожості, що мають ймовірнісну природу: спостережувані послідовності елементів розглядаються як спотворені послідовності еталонних елементів такої ж довжини. Далі обчислювалась ймовірність спостережуваного сигналу за умови, що вимовлялось дане слово. При цьому кожний спостережуваний елемент розглядався як спотворення відповідного еталонного елементу. Обчислення ймовірності зводилось до знаходження довжини найдовшого шляху на графі слова. Відповіддю розпізнавання було те слово, для якого спостережуваний сигнал мав найбільшу ймовірність. Отже, допустимі перетворення еталону слова задавались детерміновано — автоматною породжувальною граматику, а зв'язок похідних прототипів з розпізнаваним сигналом встановлювався як такий, що має ймовірнісний характер — незалежні спотворення еталонних елементів. Одночасно розв'язувались задачі навчання розпізнавати окремо вимовлювані слова — задачі формування початкового еталону (прототипу) слова на підставі навчальної вибірки (сукупності реалізацій) для цього слова [2, 3].

Узагальнення методу на розпізнавання зв'язного мовлення було зроблено в 1970 році [4]. Розглядався випадок вільного порядку слідування слів. Як і раніше, кожне слово задавалося автоматною породжувальною граматику — графом з детермінованою функцією переходу з одного стану в інший (ланцюжок станів). Початковий та кінцевий стани слова об'єднуються в один стан так, що утворюється одна «пелюстка». Всі «пелюстки»-слова своїми об'єднаними станами в свою чергу об'єднуються так, що утворюють граф-«квітку» злитого мовлення, зі спільним об'єднуючим станом у центрі: пелюстки квітки відповідають словам, пелюсток стільки, скільки слів у словнику, одна з пелюсток відповідає слову-паузі. Граф «квітка» породжує різні послідовності слів, різні прототипи зв'язного мовлення, що відрізняються темпом вимовляння, кількістю слів, довжиною пауз між словами (в тому числі й нульовою). Розпізнаваний сигнал порівнюється з генерованими модельними за допомогою динамічного програмування й для нього знаходиться найбільш правдоподібний модельний сигнал; далі визначається послідовність слів, похідні модельні сигнали яких утворюють цей найбільш правдоподібний модельний сигнал зв'язного мовлення. Ця послідовність слів оголошується відповіддю розпізнавання. Показано, що для формування відповіді розпізнавання достатньо для кожного поточного дискретного моменту часу обчислити потенціально-оптимальне слово, яке може закінчитись в цей момент, й потенціально-оптимальний момент його початку (індекс). Також було показано, що при розпізнаванні зв'язного мовлення об'єм обчислень несуттєво збільшується в порівнянні з розпізнаванням окремо вимовлюваних слів.

Подальше узагальнення методу враховує нелінійно змінювану інтенсивність вимовляння [5], а також бере до уваги обмеження на повторюваність еталонних елементів [6].

Наступні кроки починаючи з 1972 року пов'язані з урахуванням тої обставини, що сигнали мовлення (в тому числі й прототипи слів) складаються з невеликої кількості спільних еталонних елементів (просте пофонемне розпізнавання [7–10], або якихось інших структур, що відповідають фонемам (глибоке пофонемне розпізнавання [11–13]). Подальший розвиток ІКДП-технології ґрунтувався на врахуванні апріорної інформації про мовотворення, сприйняття мовлення, словотворення, синтаксичні, семантичні та прагматичні обмеження. В 1971 році [11, 12] запропоновано враховувати ймовірності слідування пар фонем при пофонемному розпізнаванні зв'язного мовлення. В 1974 році сформульовано алгоритм розпізнавання зв'язного мовлення, який бере до уваги обмеження на порядок слів, що задається автоматною граматику синтаксису [14]. З 1975 року ІКДП-метод узагальнюється

на смислову інтерпретацію (розуміння) зв'язного або квазізлитого мовлення [15–17, 13]. Синтаксичні та семантичні обмеження на порядок слів враховуються при розпізнаванні за допомогою так званих усномовної орієнтованої семантичної мережі, типів смислу та типів речень [13]. Запропоновані алгоритми розв'язання так званої задачі узагальненого розпізнавання зв'язного мовлення, коли видається N -ка ($N \geq 1$) найбільш ймовірних відповідей розпізнавання — послідовностей слів [18]. ІКДП-метод в сучасній та в узагальненій інтерпретації подано в [13, 17, 19–21].

Порівняння відомих технологій. ІКДП-технологія отримала широке визнання та поширення в світі, вона має послідовників в країнах Балтії, СНД, Заходу, роботи українських вчених цитуються в США, Великобританії, Франції, Німеччині, Японії та інших країнах [22–27]. В той же час, особливо в останнє десятиріччя, популярності набула НММ (*Hidden Markov Model* — прихована марківська модель)-технологія. НММ-технологія була започаткована Джелінеком й вперше опублікована в 1976 році [28]. НММ-метод базується на використанні прихованої марківської моделі (однорідної марківської моделі) для опису реалізацій фонем. В моделі один початковий та один кінцевий стани; задані: кінцева кількість внутрішніх станів, ймовірності переходів з стану в стан в дискретному часі, ймовірності спостереження елементів (векторів або векторів-скалярів) при перебуванні в кожному стані. Ймовірність сегменту при умові фонем обчислюється як сума або максимум за всіма траекторіями (ланцюжками станів) ймовірностей сегменту при умові фонем й траекторії, кожна з останніх ймовірностей — це добуток, кожен зі співмножників якого є ймовірність переходу зі стану в стан помножена на ймовірність поточного спостережуваного елемента при умові поточного стану в траекторії.

Отже, на рівні моделей сегментів фонем є певні відмінності в ІКДП- і НММ-технологіях: в ІКДП-методі функції переходу зі стану в стан є детермінованими й такими, які визначають неоднорідний процес породження модельних сигналів й адекватно моделюють допустимі довжини сегментів фонем. В НММ-методі функції переходу зі стану в стан моделюються однорідним марківським процесом, який погано моделює реальні довжини сегментів фонем (ймовірність довжини сегменту монотонно спадає з ростом довжини). Щодо використання ієрархічної структури образів мовлення, прийняття оптимальних рішень (сумування чи максимізація) за допомогою динамічного програмування, то ці та інші властивості порівнюваних технологій є однаковими. Обидві технології, які є зараз пануючими, якщо не звертати увагу на окремі суттєві деталі, принципово не відрізняються. Обидві технології використовують ієрархію мовних образів, конструктивні композиційні прийоми економного задання різноманітних і змінюваних сигналів мовлення за допомогою стохастичних автоматних породжувальних граматик, базуються на теорії оптимальних рішень за допомогою динамічного програмування. Навіть графічне зображення моделей розпізнавання за допомогою ланцюжків станів зі стрілками-петлями й стрілками-переходами є аналогічними. Але не зважаючи на те, що багато дослідників вказують на пріоритет ІКДП-технології (започаткована в 1967 році, НММ — в 1976 році), спільність постановок і засобів часто приховується, а самі технології протиставляються [22–27, 29], ІКДП-метод і пріоритет Інституту кібернетики АН України інколи «забуваються».

Основи ІКДП-технології. Ознаками, які використовуються при описі мовних сигналів, є миттєві передавальна характеристика мовного тракту та характеристика джерел його збурення або різні їх еквіваленти — послідовності спосте-

режуваних векторів-елементів або елементів-імен (символів або скалярів). Світ модельних сигналів задається ієрархічною структурою автоматних породжувальних граматик. Приклади ієрархії мовних образів: мікрофонема—діфон—слово—речення(фраза)—передаваний смисл; фонема—слово—речення—передаваний смисл. Задаються стохастичні автоматні породжувальні граматики, що генерують (синтезують) допустимі модельні сигнали (послідовності елементів), які відповідають образам найнижчого рівня ієрархії (наприклад, мікрофонемам, фонемам або діфонам), й дозволяють обчислювати ймовірність спостережуваного сегменту (підпослідовності елементів) за умови образів найнижчого рівня. Найуживанішим способом задання цих стохастичних автоматних породжувальних граматик є граф з детермінованими функціями переходів зі стану в стан: виділяються початкові та кінцеві стани, визначаються правила стиковки графів образів найнижчого рівня з певними початковими станами наступного образу цього ж рівня; при переході в дискретний момент часу в певний стан з визначеною ймовірністю (залежною від стану й образу) генерується спостережуваний (модельний) елемент; генерація є статистично незалежною; залежним є детермінований порядок проходження станів. Всі старші образи задаються однією чи декількома транскрипціями, які виражаються в алфавіті образів на одиницю меншого рівня ієрархії (декілька фонетичних транскрипцій на слово). Образи найвищого рівня ієрархії — передаваний смисл — специфікуються особливими транскрипціями (орієнтованою семантичною мережею або типами допустимих речень). Ймовірність спостережуваного сегменту з фіксованими границями за умови образів певного рівня обчислюється як згортка (сумування або максимізація) добутоків (згідно транскрипції) ймовірностей сегментів образів нижчого (на одиницю) рівня за всіма варіованими границями між цими нижчими сегментами. Згортка виконується за допомогою багатоступеневого динамічного програмування. Багатоступеневість виникає тому, що ймовірність нижчого сегменту, в свою чергу, обчислюється як згортка за границями ще нижчих сегментів, аж поки не доберемось до сегментів найнижчих рівнів. В ІКДП-технології багатоступеневе динамічне програмування зводиться до ефективного одноступеневого. Так само використовуються багатозначні рішення на кожному рівні ієрархії. Обмежуючи кількість рівнів ієрархії, отримуємо розв'язання задач розпізнавання окремо вимовлюваних слів, зв'язного мовлення, що складається зі слів вибраного словника, ключових слів в потоці злитого мовлення, смислової інтерпретації квазізлитого й зв'язного мовлення. Навчання та самонавчання розпізнаванню використовується для формування автоматних породжувальних граматик образів найнижчого рівня за навчальними вибірками. Синтез мовлення є складовою частиною ієрархічної генеративної моделі розпізнавання.

Література

1. Винцюк Т.К. Распознавание слов устной речи методами динамического программирования // Кибернетика. — 1968. — № 1. — С. 81-88.
2. Винцюк Т.К. Обучение поэлементному распознаванию речи // Распознавание образов и конструирование читающих автоматов. — 1969. — Вып. 2. — С. 23-35.
3. Винцюк Т.К. Алгоритм определения эталонных элементов слова по совокупности его реализаций // Тр. Акуст. ин-та. — 1970. — Вып. 12. — С. 163-168.
4. Винцюк Т.К. Поэлементное распознавание непрерывной речи, составленной из слов выбранного словаря // Кибернетика. — 1971. — № 2. — С. 133-143.
5. Винцюк Т.К. Конструктивный подход к распознаванию речи. Применение методов математического программирования к распознаванию слуховых образов // Тр. IV Всесоюз. шк.-семинара «Автоматическое распознавание слуховых образов» (1968). — Киев, 1969. — С. 21-41.

6. Винцюк Т.К., Пучкова Н.Г. Распознавание слов устной речи с помощью обучаемого алгоритма // Тр. VII Всесоюз. семинара «Автоматическое распознавание слуховых образов» (1972). — Алма-Ата, 1973. — С. 16–20.
7. Винцюк Т.К. Методы обучения, самообучения и распознавания речи, основанные на составлении эталонных сигналов из элементарных частей // Тр. VII Всесоюз. акуст. конф. (1973). Пленарные докл. — М., 1973. — С. 75–90.
8. Винцюк Т.К., Шинкаж А.Г. Фонемное распознавание слов устной речи: алгоритмы обучения, распознавания и экспериментальные результаты // Тр. VII Всесоюз. семинара «Автоматическое распознавание слуховых образов». — Львов, 1974. — Ч. 3. — С. 19–24.
9. Винцюк Т.К., Шинкаж А.Г. Автоматическое транскрибирование образов по обучающей выборке // Обработка и распознавание сигналов. — Киев, 1975. — С. 102–120.
10. Экспериментальная система фонемного распознавания слов и слитной речи / Т.К. Винцюк, О.Н. Гаврилюк, А.И. Куляс и др. // УСМ — 1982. — № 5. — С. 17–22.
11. Винцюк Т.К. Фонемне розпізнавання зв'язної мови. Вихідні передумови і постановка задачі // Автоматика. — 1972. — № 6. — С. 40–49.
12. Винцюк Т.К. Фонемне розпізнавання зв'язної мови. Алгоритми розпізнавання, навчання та самонавчання // Там же. — 1973. — № 1. — С. 63–72.
13. Винцюк Т.К. Распознавание и смысловая интерпретация речи // Кибернетика. — 1982. — № 5. — С. 101–111
14. Винцюк Т.К. Учет синтаксиса языка при распознавании речи // Обработка и распознавание сигналов. — Киев, 1975. — С. 86–102.
15. Винцюк Т.К. Проблема автоматического понимания речи // Распознавание образов. — Киев, 1977. — С. 28–34.
16. Винцюк Т.К. Автоматическое понимание речи при устном диалоге с ЭВМ // Распознавание графических и звуковых сигналов. — Киев, 1979. — С. 3–20.
17. Vintsiuk T.K. Phoneme-by-Phoneme Recognition and Semantic Interpretation of Multispeaker Speech (the HCDP-approach) // Proc. XI ICPHS (Eleventh Intern. Congr. of Phonetic Sci.). — Tallinn, 1987. — Vol. 5. — P. 407–410.
18. Винцюк Т.К. Обобщенная задача распознавания слитной речи // Автоматическое распознавание слуховых образов 1982. — Киев, 1982. — С. 345–348.
19. Винцюк Т.К. Анализ, распознавание и смысловая интерпретация речевых сигналов. — Киев : Наук. думка, 1987. — 264 с.
20. Винцюк Т.К. Сравнительный теоретический анализ ИКДП- и НММ методов распознавания речи // Всесоюз. семинар «Автоматическое распознавание слуховых образов». — Таллинн, 1989. — С. 18–24.
21. Vintsiuk T.K. HCDP-Technique for Automatic Analysis, Recognition and Understanding of Speech Signals // Proc. of First Intern. Conf. on Information Technology for Image Analysis and Pattern Recognition. — Lviv, 1990. — Vol. 1. — P. 108–112.
22. Lienard J.S. Le processus de la communication parlée. — Paris etc. : Masson, 1977. — 190 p.
23. Bridle J.S., Brown M.D., Chamberlain R.M. Continuous Connected Word Recognition using Whole Word Templates // The Radio and Electronic Eng. — 1983. — 53, № 4. — P. 167–175.
24. Nay H. Dynamic Programming as a Technique for Pattern Recognition // Proc. of 6th Intern. Conf. on Pattern Recognition. — Munich, 1982. — P. 1119–1125.
25. Levinson S.E. Structural Methodes in Automatic Speech Recognition // Proc. of the IEEE. — 1985. — 73, № 11. — P. 1625–1650.
26. Tscheschner W. Probleme der automatischen Sprachverarbeitung aus heutiger Sicht // Nachrichtentechnik, Elektronik. — 1979. — 29, № 1. — S. 26–29.
27. Nakagawa Sei-Ichi. A Machine Understanding System for Spoken Japanese Sentences // The Thesis for the Degree of Doctor of Eng. — Kyoto : University, 1976. — 174 p.
28. Jelinek F. Continuous Speech Recognition by Statistical Methods // Proc. IEEE, — 1976. — 64, № 4. — P. 532–556.
29. Jang B.H. On the Hidden Markov Model and Dynamic Time Warping for Speech Recognition. A Unified View // AT&T Bell Labs J. — 1984. — 63, № 7. — P. 1213–1260.



Нові результати в ІКДП-технології автоматичного розпізнавання, розуміння та синтезу сигналів мовлення

Тарас Вінцюк

Інститут кібернетики АН України

Україна, 252207, Київ
просп. Академіка Глушкова, 40
Тел.: (044) 266-43-56 Факс: (044) 266-15-70
E-mail address: vintsiuk%golos.kiev.ua@relay.ussr.eu.net

В останні п'ять років в ІКДП-технології [1] були зроблені певні поступи вперед — процеси обробки мовного сигналу стали більш адекватними, а процеси автоматичного розпізнавання та синтезу — більш узгодженими один з одним.

Синхронний з основним тоном аналіз мовного сигналу. Хоч й давно визнано, що універсальними ознаками мовного сигналу, поточні значення яких повинні обчислюватися на підставі спостережуваного сигналу, є миттєві передавальна характеристика мовного тракту та параметри джерел його збурення, конструктивних пропозицій щодо такого аналізу немає й понині. Задача аналізу в цій постановці є некоректною й такою, яка погано піддається формалізації. Тому пропонується при аналізі максимально враховувати структуру мовного сигналу, уникаючи різних «модельних привнесень». Враховуючи, що спостережуваний мовний сигнал є згорткою сигналів збурення з імпульсним відгуком мовного тракту, визначимо для ІКДП-технології такий робастний аналіз мовного сигналу, який ґрунтується на врахуванні квазіперіодичної структури згортки. Будемо виділяти поточний «квазіперіод» мовного сигналу: обчислювати поточні значення ознаки тон-шум, довжини «квазіперіоду», а також запам'ятовувати амплітудно-часову форму сигналу на довжині цього «квазіперіоду». Нагадаємо, що виділений поточний «квазіперіод» мовного сигналу є згорткою одного імпульсу голосових зв'язок з мовним трактом, якщо звук є дзвінким (ознака тон-шум має значення «тон»). Якщо ж поточний квазіперіод є шумним (ознака тон-шум має значення «шум»), то довжина квазіперіоду має умовне значення. Якщо ж звук є й дзвінким, й шумовим (фонема З й Ж, наприклад), то виділяють дві окремі частини: дзвінку — низькочастотну і шумову — високочастотну на всій довжині квазіперіоду. Далі вводяться міри схожості двох квазіперіодів. Якщо, наприклад, виявляється, що наступний квазіперіод по формі повторює поточний й відрізняється від нього тільки інтенсивністю, тобто ці сусідні квазіперіоди мають відносно велику міру схожості, то є сенс запам'ятати тільки один квазіперіод, а наступний період подати як повторення попереднього й для нього вказати тільки значення довжини квазіперіоду і відповідний множник інтенсивності.

Міри схожості квазіперіодів (елементарні міри схожості) мають різні вирази [1]. Але найбільш вживані міри схожості виражаються через амплітудні спектри, кепстри, автокореляційні функції, так звані *a*- чи *b*-предиктивні параметри, коефіцієнти відбиття, коди або інші описи порівнюваних квазіперіодів.

Отже, на виході аналізатора мовного сигналу маємо результат аналізу у формі часової послідовності спостережуваних елементів. Кожен елемент характеризується ознаками: 1) тон-шумом (елемент є тональним, шумним, комбінованим чи паузним); 2) довжиною квазіперіоду; 3) формою (амплітудно-часовою формою квазіперіоду або її еквівалентом: спектром, кепстром, автокореляційною функцією тощо, або формою попереднього квазіперіоду, взятою з певним множником інтенсивності).

Такий спосіб опису мовного сигналу максимально зберігає його структуру, а, отже, всі особливості, в тому числі й індивідуальні, його вимовляння.

Акустична база знань спільна як для автоматичного розпізнавання, так і для автоматичного синтезу мовлення. Сформулюємо єдину базу знань для даного диктора. Нехай заданою є сукупність сегментів (реалізацій) фонем для даного диктора. Сегментом (реалізацією) фонем будемо називати часову послідовність спостережуваних елементів, яка якимсь чином «вирізана» дослідником з експериментального матеріалу. Сформуємо навчальну вибірку з сегментів для певної фонем при фіксованому оточенні іншими фонемами (тою, що йде перед цією певною фонемою, й тою, яка йде слідом). В українському мовленні, наприклад, розрізнятимемо біля 70 різних фонем (серед них є наголошені й ненаголошені голосні) й, отже, буде всього 70^3 фонем-трійок. З усіх реалізацій навчальної вибірки фонем-трійки виберемо одну найкращу, яку оголоسیمо прототипом фонем-трійки й яка найкраще апроксимує всі інші реалізації навчальної вибірки [1]. Отже, користуємось мірою схожості сегментів-реалізацій. Ця міра є сумою елементарних мір схожості — мір схожості між двома порівнюваними елементами—«квазіперіодами». Оскільки порівнювані реалізації мають різну кількість елементів, тобто довжину, то при порівнянні їх довжини вирівнюються. Це робиться шляхом зміни кількості повторень тих форм-елементів прототипу, які відзначені аналізатором як такі, що можуть повторюватись. Загалом, в прототипі фонем-трійки максимально зберігається природа вимовляння фонем з урахуванням коартикуляції, допустимої зміни темпу та інтенсивності вимовляння, індивідуальних особливостей диктора. Однозначно визначено процес породження (генерації) різних модельних сегментів з прототипу фонем-трійки: зберігаючи порядок слідування форм-елементів, окремі форми-елементи повторюємо в допустимих границях, які вказані в прототипі фонем-трійки проти кожної її форми-елемента.

Використання акустичної бази знань при розпізнаванні індивідуального мовлення. Почнемо з розпізнавання окремо вимовлюваних слів. Нехай дано словник. Кожне слово задане своїм орфографічним текстом. Від орфографічного тексту слова переходимо до однієї чи декількох фонетичних транскрипцій слова. Далі, виходячи з фонетичних транскрипцій слова, складаємо прототипи слова шляхом об'єднання у послідовності відповідних прототипів фонем-трійок. При розпізнаванні пред'явлена реалізація слова порівнюється з перетворюваними прототипами слова. При перетвореннях прототипів слова зберігається порядок слідування форм елементів прототипа та варіюється в дозволених межах повторюваність форм-елементів. Процес порівняння та пошук найкращої міри схожості реалізується методами динамічного програмування (ДП). Розпізнавана реалізація відноситься до того слова, перетворений прототип якого дав найбільшу інтегральну схожість на розпізнаваний сигнал [1].

Аналогічно розглядаємо орфографічні тексти фраз і речень, їх фонетичні еквіваленти, а також згідно ієрархії (І) творимо (складаємо (композиція (К))) прототипи фраз і речень з прототипів слів і знову ж розглядаємо їх допустимі перетворення. Як і у випадку розпізнавання окремо вимовлюваних слів, процес перебору допустимих фраз і можливих границь між словами досягається методами ДП. Відповіддю розпізнавання є та допустима фраза або речення, перетворений прототип якої виявився найбільш схожим на пред'явлений для розпізнавання сигнал. Процеси розпізнавання ускладнюються, коли окрім лексичних накладаються синтаксичні та семантичні обмеження на порядок слів. Але порядок обробки інформації на всіх рівнях ієрархії, за винятком розглянутого тут фонемного рівня, є стандартним для ІКДП-технології [1]. Найбільш складною

є задача смислової інтерпретації зв'язного мовлення. На найвищому рівні ієрархії задаються такі структури-транскрипції, які генерують всі можливі, допустимі в межах предметної області, речення та тексти, що виражають один й той самий смисл. І це робиться для всіх смислів, що допускається передавати в межах предметної області. Далі, опускаючись на нижчі рівні ієрархії й звертаючись до акустичної бази знань для кожного з допустимих речень, згідно фонетичних текстів та транскрипцій слів, синтезуються допустимі модельні сигнали зв'язного мовлення. Ці сигнали складаються з модельних сегментів фонем-трійок. Синтезовані модельні сигнали порівнюються з розпізнаваним сигналом, результат порівняння використовується як зворотний зв'язок для генерації та направлено пошуку перетворених модельних сигналів, які є найбільш схожими на розпізнаваний сигнал. Ці найбільш схожі модельні сигнали далі аналізуються: з'ясовується, яким послідовностям слів вони відповідають й заодно виявляється, який смисл ці послідовності передають. Базовою технікою для обчислень залишається те ж ДП, яке стає багатоступеневим, ієрархічним [1]. Реалізація цієї техніки потребує солідної комп'ютерної підтримки: сотень мегабайтів оперативної пам'яті й сотень мегафлорсів швидкодії.

Використання акустичної бази знань при синтезі індивідуального мовлення. Для моделювання синтезу індивідуального мовлення за текстом скористаємося індивідуальною акустичною базою знань про фонем-трійки, сформованою за навчальними вибірками диктора в режимі навчання розпізнаванню мовлення. Вхідний орфографічний текст розмічається наголосами в словах та синтагматичними розділовими знаками. Далі цей орфографічний текст за допомогою автоматичного транскриптора трансформується в фонемний текст й розбивається на синтагми, а кожна синтагма на ритмогрупи. При цьому, наприклад, прийменники приєднуються до слів й утворюють фонемні слова. Синтагма — це фонемний текст між двома сусідніми розділовими знаками. Ритмогрупа — підпослідовність з фонемних слів, серед яких тільки одне слово має сильний (а є ще слабкий) наголос. Поняття синтагм та ритмогруп використовуються модулями ритміки та інтонування. Модуль ритміки обчислює коефіцієнти-множники збільшення-зменшення «стандартної» довжини-тривалості кожної фонем-трійки, яка використовується в даному фонемному контексті. В модулі інтонування обчислюються поточні значення періоду основного тону (для дзвінких фонем) з врахуванням «рекомендованого» модулем ритміки збільшення-зменшення тривалості прототипа фонем-трійки в даному контексті [2, 3].

Результати роботи фонетичного транскриптора, модулів ритміки та інтонування «апелюють» до акустичної бази знань про прототипи фонем-трійок певного диктора, голос котрого вирішено моделювати.

Синтез мовлення ведеться посинтагменно, ритмогрупа за ритмогрупою, фонема за фонемою. Для кожної поточної фонемі та її оточення вибирається відповідний прототип фонем-трійки. Кожна поточна форма-елемент цього прототипу зчитується з бази знань, визначається кількість повторень цієї форми-елемента (згідно «вказівок» модуля ритміки), обчислюється тривалість (періоду основного тону — для дзвінких звуків) повторення цієї ж форми-елемента (згідно корекцій модуля інтонування) і далі утворювані таким способом «повторювані» й скореговані по тривалості форми-елементи прототипу фонем-трійки один за одним, в стик, передаються через цифро-аналоговий перетворювач на динамік, де фізично озвучуються.

Проблема багатодикторності. Запропонована технологія орієнтована на використання індивідуальної акустичної бази знань щодо фонем-трійок, яка формується в режимі навчання-самонавчання розпізнаванню мовлення разово для даного диктора. Щодо створення багатодикторних систем розпізнавання, то видається, що ця проблема не буде вирішена доти, доки не будуть встановлені закономірності перетворення мовного сигналу одного диктора в сигнали іншого. При синтезі мовлення, аналогічно, фонетичний транскриптор, модулі ритміки та інтонування є загальними для розглядуваної мови-мовлення, індивідуальною ж залишається акустична база знань про фонемі-трійки. Але саме останнє дозволяє моделювати бажану індивідуальність мовлення. Отже, при автоматичному розпізнаванні та синтезі мовлення наявною є діалектика цілого й частинного, індивідуального й загального.

Література

1. Винцюк Т.К. Анализ, распознавание и интерпретация речевых сигналов. — Киев : Наук. думка, 1987. — 264 с.
2. Винцюк Т.К., Лобанов Б.М., Шинкаж А.Г. Система распознавания речи и система устного диалога СРД «Речь-1» на основе микроЭВМ // Автоматическое распознавание слуховых образов 1982. — Киев, 1982. — С. 516–521.
3. Final Report on the UNESCO Contract SC/RP 261060.8 «Development of the Multilingual (including English, Russian languages) Speech Dialogue System for Micro-Computer». — Kyjiv : Institute of Cybernetics, 1988. — 97 p.



Адаптивна обробка сигналів в задачі цифрової фільтрації авторегресійного випадкового процесу з невідомими параметрами

Леонід Житецький, Микола Проценко, Леонід Яковенко

Інститут кібернетики АН України

Україна, 252207, Київ
просп. Академіка Глушкова, 40
Тел.: (044) 267-60-22

На сьогодні відома значна кількість робіт в галузі цифрової фільтрації. Переважна більшість публікацій цього напрямку стосується методів й алгоритмів побудови так званих частотно-вибіркових фільтрів. В той же час методи й алгоритми цифрової фільтрації, орієнтовані на обробку випадкових сигналів певних класів, не набули широкого поширення.

Труднощі, що виникають при розв'язуванні задачі цифрової фільтрації випадкових процесів, полягають у тому, що в ряді прикладних задач параметри цих сигналів априорі невідомі. У таких випадках доцільним здається адаптивний підхід [1, 2]. На жаль, відомі публікації в галузі адаптивної фільтрації випадкових процесів обмежені головним чином теоретичними дослідженнями і не доведені до інженерних розробок.

У даній роботі розглядається можливість застосування теорії адаптивних систем обробки сигналів для побудови конкретних алгоритмів і програм цифрової фільтрації стаціонарних випадкових процесів авторегресійного типу в умовах неповної априорної інформації щодо параметрів цих процесів.

Задача адаптивної фільтрації формулюється так. Нехай деякий сигнал u_n , який доступний для виміру в кожний n -й дискретний момент часу, є сумою корисного сигналу x_n і шуму v_n , тобто

$$u_n = x_n + v_n. \quad (1)$$

Припускається, що послідовність $\{x_n\} = x_1, x_2, \dots$ є реалізацією стаціонарного випадкового процесу, який описується різницеvim рівнянням авторегресії l -го порядку виду

$$x_n + a_1 x_{n-1} + \dots + a_l x_{n-l} = \xi_n \dots \quad (2)$$

В цьому рівнянні a_1, \dots, a_l — постійні коефіцієнти, а $\{\xi_n\} = \xi_1, \xi_2, \dots$ — послідовність стохастично незалежних величин (дискретний білий шум) з одною і тою ж густиною розподілу ймовірностей, причому

$$\lim_{n \rightarrow \infty} \frac{1}{k} \sum_{k=1}^n \xi_k = 0, \quad (3)$$

$$\lim_{n \rightarrow \infty} \frac{1}{k} \sum_{k=1}^n \xi_k^2 = \bar{\sigma}^2. \quad (4)$$

На відміну від [2] далі вважається, що коефіцієнти a_i ($i = 1, 2, \dots, l$) в рівнянні (2) невідомі; відомо лише, що всі корені відповідного характеристичного рівняння

$$\lambda^l + a_1 \lambda^{l-1} + \dots + a_l = 0$$

по модулю менші 1, тобто модель (2) описує стійкий процес.

Вважається, що v_n — дискретний білий шум з нульовим середнім значенням і відомою дисперсією $\bar{\sigma}_n^2$.

Задача полягає в тому, щоб побудувати алгоритм адаптивної фільтрації послідовності y_n з метою одержання асимптотичної оцінки \hat{x}_n корисного сигналу x_n , що визначається класичною умовою мінімуму середнього квадрату відхилення \hat{x}_n від x_n у вигляді

$$\lim_{n \rightarrow \infty} M \{ (x_n - \hat{x}_n)^2 \} = \min \quad (5)$$

при виконанні припущень (1) – (4).

При побудові алгоритму адаптивної фільтрації в даній роботі суттєво використані результати, які раніше були одержані та теоретично обгрунтовані Фомінін [2]. В рамках запропонованого підходу був синтезований алгоритм параметричної ідентифікації авторегресійного процесу (2), що є рекурентною процедурою методу найменших квадратів.

В роботі наводяться конкретні алгоритми та програми адаптивних фільтрів для цифрової обробки випадкових процесів типу (2) при різних значеннях l , а також результати моделювання на ПЕОМ процесів адаптації та фільтрації. Розроблені алгоритми практично використовувались при обробці сигналів, що надходили від датчика зовнішнього діаметра скляних труб в системі керування технологічним процесом виготовлення скляних труб.

Література

1. Цыпкин Я.З. Основы теории обучающихся систем. — М. : Наука, 1970. — 252 с.
2. Фомин В.Н. Рекуррентное оценивание и адаптивная фильтрация. — М. : Наука, 1984. — 288 с.



Параметричні моделі стохастичних коливань та методи їх статистичної обробки

Віктор Михайлишин, Ігор Яворський
Фізико-механічний інститут АН України

Україна, 290601, Львів
вул. Наукова, 5
Тел.: (032) 265-45-60

Нерегулярність і повторюваність — визначальні риси стохастичних коливань. В ймовірносних моделях періодично корельованих випадкових процесів (ПКВП) та їх узагальнень — бі-, полі- та майже ПКВП — повторюваність властивостей описується певними типами часових змін характеристик другого порядку: математичного сподівання, кореляційної функції, спектральної густини [1]. При цьому детермінованим зображенням коливань у вигляді періодичних та майже періодичних функцій, якими повністю ігнорується нерегулярність часових змін, та зображенням у вигляді стаціонарних випадкових процесів, які в свою чергу вже не припускають ніякого збереження в часі тої чи іншої фазової структури, належать дві крайні вироджені позиції. Поєднання опису повторюваності та стохастичності досягається врахуванням стохастичної амплітудної та фазової модуляції сукупності гармонічних коливань з певним набором частот. Для ПКВП множина таких частот задається арифметичною прогресією $\omega_k = k \omega_0$:

$$\xi(t) = \xi_0(t) + \sum_{k \in \mathbb{N}} [\xi_k^c(t) \cos k \omega_0 t + \xi_k^s(t) \sin k \omega_0 t] \quad (1)$$

Математичні сподівання випадкових компонентів $E \xi_k^c(t) = m_k^c$, $E \xi_k^s(t) = m_k^s$ визначають Фур'є-компоненти математичного сподівання ПКВП, а їх авто- і взаємні кореляційні функції — кореляційні компоненти:

$$m(t) = E \xi(t) = m_0 + \sum_{k \in \mathbb{N}} (m_k^c \cos k \omega_0 t + m_k^s \sin k \omega_0 t), \quad (2)$$

$$b(t, u) = E \xi(t) \xi(t+u) = B_0(u) + \sum_{k \in \mathbb{N}} [B_k^c(u) \cos k \omega_0 t + B_k^s(u) \sin k \omega_0 t], \quad (3)$$

при цьому

$$B_0(u) = \frac{1}{4} \sum_{l \in \mathbb{Z}} [(D_l^{cc}(u) + D_l^{ss}(u)) \cos l \omega_0 u + (D_l^{cs}(u) - D_l^{sc}(u)) \sin l \omega_0 u], \quad (4)$$

$$B_k^c(u) = \frac{1}{2} \sum_{l \in \mathbb{Z}} [(D_{l+k,l}^{cc}(u) + D_{l+k,l}^{ss}(u)) \cos l \omega_0 u + (D_{l+k,l}^{cs}(u) - D_{l+k,l}^{sc}(u)) \sin l \omega_0 u], \quad (5)$$

$$B_k^s(u) = \frac{1}{2} \sum_{l \in \mathbb{Z}} [(D_{l+k,l}^{cc}(u) + D_{l+k,l}^{ss}(u)) \sin l \omega_0 u + (D_{l+k,l}^{cs}(u) - D_{l+k,l}^{sc}(u)) \cos l \omega_0 u]. \quad (6)$$

Нульовий кореляційний компонент є парною функцією зсуву $B_0(-u) = B_0(u)$, а вищі задовольняють умовам

$$B_k^c(-u) = B_k^c(u) \cos k \omega_0 u - B_k^s(u) \sin k \omega_0 u,$$

$$B_k^s(-u) = B_k^s(u) \cos k \omega_0 u + B_k^c(u) \sin k \omega_0 u.$$

Спектральні характеристики коливань пов'язані з кореляційними співвідношеннями

$$\begin{aligned} \operatorname{Re} f(\omega, t) &= \frac{1}{2\pi} \int_{R^+} [b(t, u) + b(t, -u)] \cos \omega u \, du, \\ \operatorname{Im} f(\omega, t) &= \frac{1}{2\pi} \int_{R^+} [b(t, u) - b(t, -u)] \sin \omega u \, du, \\ \operatorname{Re} f_k(\omega) &= \frac{1}{4\pi} \int_R [B_k^c(u) \cos \omega u - B_k^s(u) \sin \omega u] \, du, \\ \operatorname{Im} f_k(\omega) &= \frac{1}{4\pi} \int_R [B_k^s(u) \cos \omega u + B_k^c(u) \sin \omega u] \, du, \\ f_0(\omega) &= \frac{1}{\pi} \int_{R^+} B_0(u) \cos \omega u \, du. \end{aligned} \quad (7)$$

Використання всієї сукупності ймовірносних характеристик першого й другого порядку ПКВП-моделі дозволяє описати суттєво нові закономірності коливних процесів, а саме, повторюваність в часі властивостей флуктуаційних змін. В арсеналі детермінованої моделі — тільки функція $m(t)$ та компоненти m_0 , m_k^c , m_k^s , а у випадку стаціонарного зображення — величини m_0 , $B_0(u)$, $f_0(\omega)$, що характеризують, відповідно, середній рівень регулярних коливань, усереднені кореляційні зв'язки та спектральний склад.

Періодичність характеристик другого порядку є наслідком кореляцій між модулюючими процесами в зображенні (1). Ця кореляція накладає певні умови на квадратурні складові кожного компонента $\xi_k^c(t)$ і $\xi_k^s(t)$, оскільки величина

$$D_{k,-k}(u) = \frac{1}{4} [D_{kk}^{cc}(u) - D_{kk}^{ss}(u) - i [D_{kk}^{cs}(u) + D_{kk}^{sc}(u)]]$$

буде відмінною від нуля або у випадку, коли квадратурні складові не корелюють між собою, або коли їхні кореляційні функції не є однаковими. Якщо обидві умови не виконуються, то приходимо до адитивної моделі $\xi(t) = \eta(t) + f(t)$, де

$$R_\eta(u) = D_{00}(u) + \sum_{l \in \mathbb{N}} D_{ll}^f(u) \cos l \omega_0 u.$$

Якісно такий же спектральний склад має і мультиплікативний процес $\xi(t) = \eta(t)f(t)$. Тому встановити, яка з цих двох моделей відповідає тому чи іншому коливанню можна, вже залучивши до аналізу величини, що характеризують нестационарність другого порядку — вищі кореляційні та спектральні компоненти.

Мультиплікативна модель $\xi(t) = \eta(t)f(t)$ описує тільки амплітудну модуляцію коливань. Найпростішою моделлю ПКВП, що враховує фазову модуляцію, є зображення у вигляді квадратурних складових

$$\xi(t) = \xi^c(t) \cos \omega_0 t + \xi^s(t) \sin \omega_0 t,$$

де $D_c^c(u) = D_s(u)$ і $D_{cs}(u) = 0$. Така модель може бути в першому наближенні використана для дослідження коливань морської поверхні, геомагнітних пульсацій, іоносферного короткохвильового радіосигналу, добового ходу температури повітря й води.

Якщо у виразах (4)—(6) прийняти до уваги тільки кореляційні зв'язки між нульовим та вищими компонентами, то приходимо до формул

$$B_0(u) = D_{00}(u), \quad B_k^{\cos}(u) = D_k^{\cos}(u) + D_k^{\sin}(u) \cos k \omega_0 u + D_k^{\sin}(u) \sin k \omega_0 u, \\ B_k^{\sin}(u) = D_k^{\cos}(u) + D_k^{\sin}(u) \sin k \omega_0 u + D_k^{\cos}(u) \cos k \omega_0 u.$$

Такий тип ПКВП описує сезонну ритміку температури й солоності морської води, коливань рівня моря, льодовитості, стоку рік.

Параметричні представлення певних типів ПКВП можуть бути одержані з використанням виразів (2)—(7) на основі апроксимацій авто- та взаємних кореляційних функцій стаціонарних компонентів, що формують їх. Як показують результати аналізу обширного експериментального матеріалу, такі апроксимації можуть бути вибрані у вигляді згасаючих експонент чи експоненціально згасаючих косинусоїдальних функцій. Конкретні значення параметрів апроксимацій відображають специфічні особливості того чи іншого коливного процесу, і на їх основі можуть бути сформовані ті вимоги до довжини реалізації та інтервалу дискретизації, що дозволяють досягти певної якості результатів при статистичній обробці натурних даних.

В доповіді наводяться результати досліджень якості оцінювання всього комплексу характеристик другого порядку основних типів ПКВП. При цьому розглядаються оцінки, що знаходяться за допомогою як когерентного, так і компонентного методів. Перший з них базується на усередненні відліків, що вибираються через період корельованості, а другий виходить з розкладу ймовірносних характеристик в ряд Фур'є: оцінки характеристик будуються на основі оцінок їх Фур'є-компонентів. Порівняльний аналіз компонентних і когерентних статистик показує, що при швидкому згасанні кореляційних зв'язків та невеликій кількості компонентів достовірність перших може значно перевищувати достовірність других, але при збільшенні кількості компонентів відмінність в якості оцінювання пропадає.

І когерентні, і компонентні оцінки можна розглядати як результат перетворення реалізацій ПКВП за допомогою лінійного інваріантного фільтру. Оптимальні оцінки характеристик (з мінімальною дисперсією) одержуються вже в класі лінійних неінваріантних фільтрів. Структура таких оптимальних оцінок залежить від тих параметрів, які визначають певний тип ПКВП. Слід зауважити, що клас цих оцінок співпадає з класом оптимальних компонентних оцінок гаусових ПКВП, що одержуються за допомогою методу максимальної правдоподібності. Метод максимальної правдоподібності може бути використаний і для оцінок параметрів, що визначають кореляційну та спектральну структуру ПКВП.

Література

1. Драган Я.П., Рожков В.А., Яворский И.Н. Методы вероятностного анализа ритмики океанологических процессов. — Л. : Гидрометеониздат, 1987. — 320 с.



Розвиток моделей стохастично періодичних процесів і полів

Микола Приймак

Тернопільський приладобудівний інститут

Україна, 282001, Тернопіль
вул. Руська, 56
Тел.: (035) 222-4133

Широке коло задач в багатьох областях науки і техніки (акустика, енергетика, медицина тощо) пов'язане з дослідженням стохастично періодичних процесів і полів, для яких певні ймовірнісні характеристики є періодичними. Практика свідчить, що для успішного аналізу таких процесів і полів першочерговою задачею є вибір або побудова їх нових моделей, тому що в залежності від того, в якому об'ємі і яким чином модель враховує фактичну періодичність процесів (полів), настільки достовірними і повними будуть результати їх статистичного аналізу.

В ряді випадків для опису стохастично періодичних процесів (полів) використовують адитивне зображення, одною із складових якого є періодична функція (поле), а другою — стаціонарний процес (однорідне поле). Але поскільки таке зображення враховує періодичність тільки першого моменту — математичного сподівання процесу (поля), то в ряді випадків для більш адекватного описування стохастично періодичних процесів (полів) використовують періодичні або періодично корельовані випадкові процеси (поля), які вже дозволять враховувати періодичність їх вищих ймовірнісних характеристик.

В останній час все більш широке застосування при описі і статистичному аналізі стохастично періодичних процесів і полів знаходять лінійні випадкові процеси і поля. Перевага цих моделей в тому, що вони дозволяють розділити причини виникнення стохастичної періодичності процесів і полів і, крім того, дають можливість враховувати цю періодичність в рамках всіх моментних функцій, тобто в рамках характеристичних функцій. Це в кінцевому результаті дає можливість вдосконалювати методи досліджень вказаних процесів і полів.

В доповіді приведені моделі лінійних випадкових процесів і полів, сформульовані теореми, при виконанні умов яких вказані процеси і поля будуть періодичними. Вказується на конкретні явища, для опису і статистичного аналізу яких ці моделі успішно використовуються. Це в першу чергу стосується шумів кавітації в гідроакустиці, навантажень енергосистем в енергетиці тощо.



Асимптотика похибки апроксимації сигналів в базисі многочленів Якобі

Ярослав П'янило

Інститут прикладних проблем механіки і математики АН України

Україна, 290053, Львів
вул. Наукова, 36
Тел.: (032) 265-19-48

Серед багатьох способів розв'язування прикладних задач, зокрема задач обробки сигналів, спектральні методи виділились в окремий напрямок. В даний час інтенсивно проводяться дослідження спектральних методів як в існуючих ортогональних базисах, так і при побудові нових базисів, оптимальних для розв'язування певного класу задач. Оскільки задачі обробки сигналів є в основному некоректними за Тихоновим, то важливе місце серед досліджень займає вивчення похибки вихідних результатів та використання апріорної інформації для побудови

регуляризованих алгоритмів. В даній роботі розглядаються питання оцінки похибки апроксимації сигналів та один із шляхів побудови регуляризованих алгоритмів.

Сигнал, що задається своїми дискретними значеннями, апроксимується рядом Фур'є-Якобі

$$f(x) = \omega(x) \sum_{n=0}^{\infty} \frac{f_n}{r_n} P_n^{(\alpha, \beta)}(x). \quad (1)$$

В формулі (1) позначено: $P_n^{(\alpha, \beta)}(x)$ ($\alpha > -1$, $\beta > -1$) — многочлени Якобі,

$\omega(x) = (1-x)^\alpha (1+x)^\beta$ — вагова функція,

$$r_n = 2^{\alpha+\beta+1} \frac{\Gamma(n+\alpha+1) \Gamma(n+\beta+1)}{n! (2n+\alpha+\beta+1) \Gamma(n+\alpha+\beta+1)}, \quad (2)$$

$$f_n = \int_{-1}^1 f(x) P_n^{(\alpha, \beta)}(x) dx. \quad (3)$$

Не зменшуючи загальності, можна вважати, що сигнал задається на проміжку ортогональності многочленів Якобі $[-1, 1]$, оскільки в протилежному випадку довільний проміжок $[a, b]$ лінійною заміною змінних зводиться до $[-1, 1]$. Оскільки відомі дискретні відліки сигналу, то для обчислення коефіцієнтів f_n доцільно користуватись такою оптимальною в класі L_2 квадратурною формулою

$$f_n^* \approx \sum_{j=0}^N p_{j,N}^2 P_n^{(\alpha, \beta)}(x_j) \frac{f(x_j)}{\omega(x_j)}, \quad (4)$$

де x_j — нулі многочлена $P_{N+1}^{(\alpha+\beta)}(x)$, а

$$p_{j,N}^2 = 2 \lambda_N \frac{(N+\alpha+1)(N+\beta+1)}{(2N+\alpha+\beta+2)} \left[P_n^{(\alpha, \beta)}(x_j) \right]^2 / (1-x_j^2), \quad (5)$$

$$\lambda_N = \frac{(N+1)! \Gamma(N+\alpha+\beta+2)}{2^{(\alpha, \beta)} (2N+\alpha+\beta+2) \Gamma(N+\alpha+1) \Gamma(N+\beta+1)}. \quad (6)$$

На практиці для апроксимації сигналу використовується N -а часткова сума ряду (1). Крім цього, формула (4) є наближеною. Все це приводить до виникнення деякої похибки при відновленні сигналу рядом (1).

Якщо підставити формулу (4) в N -у часткову суму ряду (1), то після тотожних перетворень отримаємо формулу

$$f_N(x) \approx \omega(x) \sum_{n=0}^N \frac{1}{W_{N,j}} \frac{f(x_j)}{\omega(x_j)} \left[P_N^{(\alpha, \beta)}(x) \right]^{-2} \times \sum_{n=0}^N \frac{1}{r_n} P_n^{(\alpha, \beta)}(x) P_n^{(\alpha, \beta)}(x_j), \quad (7)$$

де

$$W_{N,j} = 2 \lambda_N \frac{(N+\alpha+1)(N+\beta+1)}{(2N+\alpha+\beta+2)} / (1-x_j^2).$$

Якщо тепер в рівності (7) покласти $x = \cos \Theta$ й використати формули Крістофеля-Дарбу та Ейлера для сумування інтегралів, то головний член асимптотики $f(x)$ при великих N буде мати вигляд

$$f_N(\cos \Theta) \approx \frac{1}{\pi} f(\cos \Theta) \{ Si [N_1 (\Theta - \Theta_0)] + Si [N_1 (\Theta_N - \Theta)] \}. \quad (8)$$

В формулі (8) позначено: $Si(x)$ — інтегральний синус, $N_1 = N + 1 + (\alpha + \beta)/2$, $\Theta_0 = \arccos x_N$, $\Theta_N = \arccos x_0$.

Оскільки $Si(x) \rightarrow \frac{\pi}{2}$ при $x \rightarrow \infty$, то $f_N(\cos \Theta) \rightarrow f(\cos \Theta)$ при $N \rightarrow \infty$. З формули (8) випливає, що головний член відносної похибки $\Delta f_N(\cos \Theta)$, яка виникає за рахунок використання N -ої часткової суми ряду (1) та квадратурної формули (4), має вигляд

$$\Delta f_N(\cos \Theta) \sim \frac{1}{\pi} \{ Si [N_1(\Theta - \Theta_0)] + Si [N_1(\Theta_N - \Theta)] \}.$$

Теоретично при збільшенні N в формулі (4) $f_n^H \rightarrow f_n$. Але за рахунок втрати точності при виконанні алгебраїчних операцій завжди існує таке N_{\max} , що при $N > N_{\max}$ формула (4) дає неточні результати. Крім цього, кількість n_0 коефіцієнтів f_n , обчислених за формулою (4), менша $N (n_0 < N)$ і не завжди дозволяє відновити сигнал з необхідною точністю. Для подолання цих труднощів та побудови регуляризуючого алгоритму на основі певної апріорної інформації про відновлюваний сигнал побудовано асимптотичні формули для коефіцієнтів f_n при великих n та асимптотику N -го залишку ряду (1). Оцінено похибки, які виникають за рахунок використання наближеної формули (4) та асимптотичних формул для обчислення коефіцієнтів. Якщо відома необхідна апріорна інформація, то одержані результати дозволяють побудувати регуляризуючий алгоритм, який дає можливість відновити сигнал з наперед заданою точністю. При цьому n_0 та N виступають регуляризуючими параметрами даного методу. Слід зауважити, що побудований алгоритм відновлення спрацьовує тоді, коли задана апріорна інформація дозволяє «зшити» з необхідною точністю коефіцієнти, обчислені за наближеними та асимптотичними формулами з їх відповідними точними значеннями. Якщо апріорної інформації недостатньо для «зшивання», то виникає невизначеність і даний алгоритм не спрацьовує.



Проективні перетворення та зображення сигналів

Богдан Яворський

Тернопільський приладобудівний інститут

Україна, 282001, Тернопіль
вул. Руська, 56
Тел.: (035) 222-41-33

Ефективність комп'ютерних технологій наукових досліджень залежить від рівня інтегрованості бази концептуальних знань. Виникає необхідність розробки потужних концепцій чи більш повного використання можливостей відомих концепцій. В теорії сигналів широко використовується концепція лінійності, яка приводить до розгляду зображень сигналу у вигляді розкладів (ортогональних), отримуваних відповідними перетвореннями. Але всі розклади коректні в конкретному випадку обробки при звичайних рамках концепції лінійності. Це є причиною втрати загальності та однорідності алгоритмів, що зменшує ефективність комп'ютерних технологій дослідження сигналів.

У доповіді розглянуто трактовку зображення сигналу як проективного перетворення, що ширше звичайного розуміння в теорії сигналів зображення (яке в термінах проективних перетворень формулюється як ортогональна колінеація). Таке розуміння зображення сигналу веде до його погодження з операторами обробки.

Для сигналу природньо вводиться відношення порядку на множині простору його завдання, що дозволяє розглядати його підпростори як вектори при умові

існування базису. Базис вводиться наступним чином. Розклад лінійного оператора на проектори приводить до характеристичних функцій підпросторів і до розбиття на ці підпростори простору. При виборі базису у випадку евклідового простору виняткову роль відіграє символ Кронекера. Останній приймається за базис зображення сигналу. Розглянуто можливості отримання інших базисів і способи розкладів при відомому операторі обробки. Показано еволюцію понять та методів при переході від евклідового простору до гільбертового.

Приведені приклади зображення сигналу у випадку дискретного лінійного оператора другого порядку при обробці тонального сигналу та періодично нестационарного випадкового сигналу. Обговорюється оцінка ефективності дослідження структури сигналу у випадку інтеграції концепції в програмне середовище. Отримано підвищення ефективності за рахунок збільшення інформативності завдяки можливості послідовного переходу від проективного до лінійного зображення.

