

Вибір меж для шкал медико-біологічних сигналів в алгоритмах їх розпізнавання

Лагутін^f В. В., ORCID [0000-0003-3431-1686](https://orcid.org/0000-0003-3431-1686)

Сиротенко^f В. І., ORCID [0000-0002-7542-5885](https://orcid.org/0000-0002-7542-5885)

Шачиков^s А. Д., к.т.н., ORCID [0000-0003-2353-7647](https://orcid.org/0000-0003-2353-7647)

Шуляк^s О. П., к.т.н. доц., ORCID [0000-0002-6920-2212](https://orcid.org/0000-0002-6920-2212)

Кафедра електронної інженерії ee.kpi.ua

Національний технічний університет України

«Київський політехнічний інститут імені Ігоря Сікорського» kpi.ua

Київ, Україна

Анотація—Робота спрямована на вдосконалення процедур розпізнавання медико-біологічних сигналів в системах, що навчаються з учителем, в напрямку підвищення рівня правильності рішень, які приймаються.

Запропоновано шість модифікацій розпізнавальних алгоритмів з поданням сигналів в нормованих шкалах їх змін, в яких реалізовані спільний і роздільний варіанти вибору меж цих шкал на навчальних вибірках сигналів різних класів. Побудовано процедури порівняння сигналів з еталонами і прийняття рішень з підгонкою еталонів під сигнали і без неї.

Модифікації розпізнавальних процедур реалізовані і досліджені в середовищі MATLAB®. Підтверджена їх працездатність.

Спроможність розробок підтверджена розрахунками на контрольних вибірках в тестовому прикладі. Виявлено підвищення значень статистичних оцінок чутливості, специфічності і загальної валідності рішень з розпізнавання сигналів при переході до нормованих шкалам змін їх значень в порівнянні з варіантом використання звичайних шкал значень для опису процесів, що спостерігаються.

Ключові слова — медико-біологічні сигнали; розпізнавальні алгоритми; навчання з учителем; шкали змін значень.

I. ВСТУП

Результативність діагностики пацієнтів багато в чому пов'язана з правильною розпізнавання медико-біологічних сигналів [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]. В основі цього зв'язку лежить використання класифікації таких сигналів з інтерпретацією кожного класу, яка вказує на можливі стани організму пацієнтів, його окремих систем або на їх захворювання [1, 2, 3, 4, 5, 6, 7, 8, 9, 10].

Дослідження щодо підвищення результативності розпізнавання образів у вищезазначених цілях по розкриттю резервів у підвищенні точності та достовірності прийнятих рішень при обробці даних пацієнтів, оцінка таких резервів і їх реалізації шляхом створення і використання відповідних процедур обробки зазначених даних стимулюються медичною практикою [1, 2, 3, 4, 5, 6, 7, 8, 9, 10] і утворюють одне з важливих напрямків вдосконалення програмного інструментарію систем надання медичної допомоги [1, 2, 3, 4, 5, 6, 7, 8, 9, 10].

Проводиться робота з все більш повного здобуття інформації про сигнали при обстеженні пацієнтів, з пошуку різноманітних систем інформативних ознак станів організму і захворювань, з конструювання все більш чутливих критеріїв прийняття рішень при розпізнаванні сигналів [1, 2, 3, 4, 5, 6, 7, 8, 9, 10].

Дослідження [11, 12, 13, 14, 15, 16, 17, 18, 19, 20] [21] підтвердили продуктивність використання характеристик форми сигналів та критерію оцінки їх близькості у вигляді скалярного добутку, який виражає кореляцію форми сигналів [11].

Характеристика форми сигналів, представлених у вигляді послідовності відліків на рівнодискретній сітці часу, в найпростішому випадку являє собою [12, 13, 14, 15, 16, 17, 18, 19, 20, 21] ту ж саму послідовність відліків, позбавлену постійної складової, яка змінює співвідношення відліків сигналів і ускладнює їх розпізнавання, і додатково пронормовану по інтенсивності для виключення впливу енергії сигналів [11] на коректність їх порівняння.

Якщо x_i – відліки сигналу, $i = 1, \dots, N$, x_{\pm} – їх постійна складова, а y_i – відповідні відліки характеристики його форми, то вони пов'язані між собою співвідношеннями

$$y_i = (x_i - x_{\pm}) / \sqrt{\sum_{i=1}^N (x_i - x_{\pm})^2}$$

$$x_{\pm} = (1/N) \sum_{i=1}^N x_i$$

Питання, що розглядаються в роботі, пов'язані з однією особливістю цієї характеристики, яка стає



перешкодою в підвищенні правильності розпізнавання сигналів.

Зміна будь-якого відліку сигналу, за тою чи іншою причиною, наприклад, внаслідок варіабельності його форми або через її спотворення в процесі спостережень потребує перерахунку заново всіх відліків цієї характеристики. Це потрібно, навіть якщо інші відліки є ідеальними у порівнянні з еталонами.

І це дійсно так, оскільки в розрахунку будь-якого відліку характеристики форми сигналу беруть участь всі його відліки, в тому числі, згаданий змінений відлік.

Таким чином, перетворення сигналу в характеристику його форми може призвести до спотворення всіх ідеальних відліків, що надійшли, через варіації одного єдиного (будь-якого) з них.

Для подолання цих труднощів порівняння сигналів з характерним їх видом для різних класів (з еталонами) запропоновано проводити, минаючи перерахунок сигналів в характеристики їх форми, підганяючи при цьому еталони під сигнали, що розпізнаються. У цьому полягають принцип і зміст модифікації процедури порівняння сигналу і еталонів при прийнятті рішень про тип сигналу, що розпізнається.

Головним в такому випадку стає питання про те, як забезпечити коректність порівняння сигналу з еталонами. В алгоритмах, що використовують для опису сигналів характеристики їх форми, питання вирішувалося їх нормуванням, завдяки якому зрівноважуються енергії процесів, що порівнюються, з оцінкою їх взаємної кореляції.

Тут же постає запитання про те, яким чином у такому разі можна організувати порівняння сигналу з еталонами для прийняття рішень про тип сигналів.

Принцип вирішення поставлених питань, розробка модифікацій розпізнавальних процедур відповідно до цього принципу представлені нижче.

Подібна обробка сигналів у роботі ілюструється на завданні розпізнавання N, A, V типів QRS-комплексів у півгодинному запису електрокардіограми пацієнта з бази даних в Internet [22].

Метою роботи в цілому є розробка програмного інструментарію для розкриття і дослідження можли-

вих резервів у підвищенні правильності розпізнавання медико-біологічних сигналів у діагностичних системах, що навчаються з учителем, за рахунок подання сигналів у нормованих шкалах змін їх значень та застосування спеціальних процедур підгонки еталонів під сигнали для збереження їх форми в перетвореннях під час підготовки до порівняння з еталонами різних класів у розпізнаванні спостережуваних процесів.

II. ДОСЛІДЖУВАНІ АЛГОРИТМИ І ВАРІАНТИ ВИБОРУ МЕЖ ДЛЯ ШКАЛ ЗМІНИ СИГНАЛІВ

Спочатку — про забезпечення коректності порівняння сигналу з еталонами. В алгоритмах, які розглянуті в цій роботі, вона досягається забезпеченням сумірності сигналів і еталонів на основі переходу до нормованих безрозмірних шкал при їх розгляді. Мається на увазі (рис. 1) перехід від шкал значень сигналів і еталонів до шкал їх змін і далі — до нормованих шкал змін і, в подальшому можливо, до шкал відхилень.

Вихідною для сигналу є шкала (рис. 2) його поточних значень в використовуваних одиницях вимірювання, наприклад, у мілівольтах при запису ЕКГ. Значення сигналу на цій шкалі відраховуються від нульового рівня, представленого на графіку віссю часу. Відзначено діапазон $[\min, \max]$ спостережуваних значень з межами їх зміни.

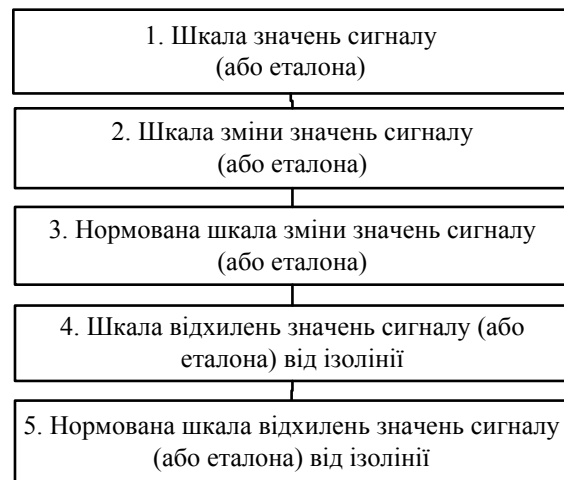


Рис. 1. Шкали для представлення сигналів і еталонів в розпізнавальних алгоритмах

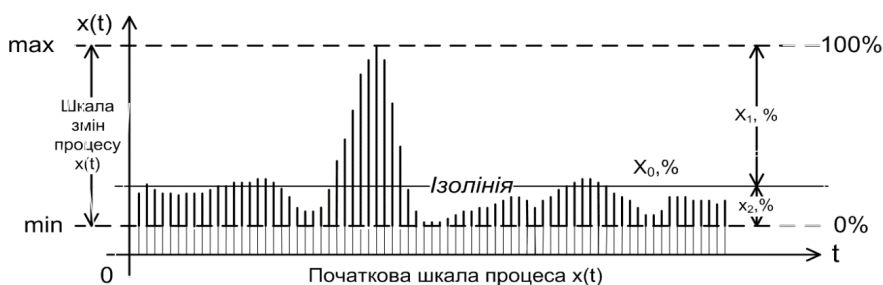


Рис. 2. Визначення та взаємозв'язок шкал представлення сигналів і еталонів в розпізнавальних алгоритмах



Рис. 3. Зміст модифікації розпізнавальних процедур

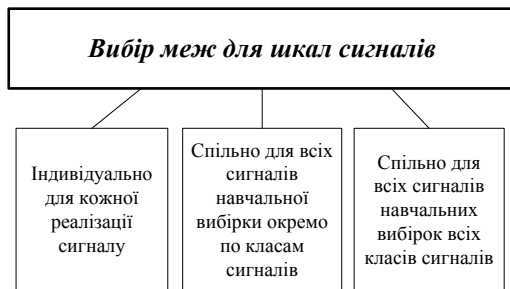


Рис. 4. Варіанти вибору меж для шкал сигналів

Під другою шкалою зміни значень сигналу, розуміється послідовність чисел, що отримується після того, як рівень \min приймається за нуль для відліків того ж самого процесу всередині шкали. Важливим, як виявляється в певних випадках, є та обставина, що всі відліки сигналу на цій шкалі є позитивними (хоча б невід'ємними).

Нормована шкала зміни значень сигналу виходить з попередньої шкали змін, якщо її пронормувати розміром цієї шкали. Значення сигналу в нормованій шкалі лежать в межах $[0, 1]$. Це безрозмірна шкала, в якій може бути показана форма сигналу, що розпізнається, і її можна порівнювати з еталоном, якщо він представлений подібним чином.

Ще одна шкала представлення значень сигналу, що представляє інтерес, виникає, якщо за нульовий рівень їх відліку прийняти рівень ізолінії цього сигналу, який розміщується на рівні його постійної складової. Тоді можна записати послідовність значень відхилень сигналу від цієї ізолінії, з'являється як позитивні, так і негативні значення сигналу в цій шкалі. І описаний таким чином спостережуваний процес буде трохи інакше виражати своїми відліками вихідну форму сигналу.

У роботі перехід до розгляду сигналів і еталонів у нормованих шкалах їх змін розглядається як спосіб забезпечення співмірності еталонів з сигналами для забезпечення коректності їх порівняння при прийнятті рішення про тип сигналу і, в той же час, як засіб поліпшення показників правильності розпізнавання сигналів заданої класифікації в системах, що навчаються з учителем.

В цілому, в роботі пропонується (рис. 3) реалізація розглянутого підходу для підвищення правильності розпізнавання сигналів у системах, що навчаються

з учителем, за рахунок приведення сигналів до шкал їх змін і введення запропонованого порядку їх порівняння в процесі модифікації розпізнавальних алгоритмів.

Таким чином, і сигнал, і еталони приводяться фактично до однієї і тієї ж шкали, чим забезпечується їх порівнянність і може бути виконано порівняння їх форми.

Графік кожного еталона повинен бути правильно накладений на графік сигналу і це було запропоновано зробити по положенню ізолінії еталона на графіку сигналу. Яким чином це організовується алгоритмічно, розглядається на конкретних прикладах процедур розпізнавання сигналів. Проте, загальний підхід полягає в наступному.

Еталон, який жорстко пов'язаний зі своєю ізолінією, переноситься на графік сигналу так, щоб його ізолінія зайняла на графіку сигналу точно таке ж положення (за рівнем у відсотках), як на власній шкалі еталона. Така дія виконується над кожним еталоном.

Положення ізоліній на графіках еталонів у загальному випадку різне. У середньому (за статистикою) найкраща відповідність забезпечується для сигналу з правильним еталоном, тобто з еталоном того класу, до якого належить сигнал, що розпізнається.

Після цього для прийняття рішень проводиться порівняння форми сигналу з різними еталонами. Рішення про тип сигналу приймається за максимумом їх схожості. На цьому побудований саме принцип розпізнавання сигналів.

Актуальним тут також є питання, що виникає додатково, – питання про порядок вибору меж розглянутих шкал. У попередніх дослідженнях було встановлено, що вплив порядку вибору меж для шкал змін сигналів на результативність застосування розглянутих процедур виявляється досить суттєвим. Мається на увазі (рис. 4), що такі межі можуть визначатися окремо для кожної реалізації сигналу відповідно до динаміки його зміни.

Вони можуть вибиратися також для кожної групи всіх сигналів навчальної вибірки в кожному класі сигналів роздільно або ж встановлюватися для всіх класів спільно. У той же час, в якості меж шкали для групи сигналів можуть бути обрані як середні, так і екстремальні значення.

Доцільність реалізації кожного з цих варіантів може бути визначена на підставі статистичних оцінок правильності прийнятих рішень у кожному з цих випадків за стандартними критеріями, в тому числі, по чутливості, специфічності, загальної валідності.

Очевидно, що використання індивідуальних меж шкал для різних реалізацій сигналів ускладнює їх обробку для прийняття рішень по ходу надходження сигналів, що розпізнаються. Загальні шкали, на противагу цьому, можуть бути отримані заздалегідь, на етапі навчання розпізнавальних алгоритмів, спрощуючи обробку поточних реалізацій. Крім того, в першому випадку буде проявляти себе варіабельність масштабу розгляду сигналів, умови їх розпізнавання

будуть більш складними. У другому випадку прояв цього фактору буде знижуватися.

Наскільки суттєвою буде різниця в результатах розпізнавання в таких випадках, має перевірятися порівняльними оцінками статистичним шляхом підрахунку кількості правильних і помилкових рішень на контрольних вибірках і розрахунку показників чутливості, специфічності, загальною валідності [1, 2, 3, 4, 5, 6, 7, 8, 9, 10].

У роботі досліджується і порівнюється результативність двох основних можливих варіантів вибору зазначених шкал за сукупністю сигналів навчальних вибірок – роздільному по класах і спільному для них.

Як приклади для отримання зазначених модифікацій розпізнавальних процедур шляхом приведення сигналів і еталонів до нормованих шкал їх змін були відібрані такі алгоритми (рис. 5).

Таких алгоритмів для дослідження вибору меж для шкал зміни сигналів, що розпізнаються, три. Умовно назвемо їх першим, другим і третім. Шкали нормовані, безрозмірні, для виразу значень, що спостерігаються в своїх межах у відсотках, як для сигналів, що розпізнаються, так і для еталонів.

Незалежно від відмінностей цих алгоритмів за змістом обробки даних при визначенні типу аналізованих сигналів, вони усі в цьому дослідженні використовують їх відліки в згаданих шкалах. Приведення вихідних даних про сигнали до такого виду, перетворення навчальних, контрольних вибірок і сформованих еталонів виконується відповідно до визначення цих шкал

Перший алгоритм можна вважати вихідним, базовим. Він заснований на формуванні для кожного класу сигналів єдиного еталона у вигляді усередненої характеристики форми реалізацій його навчальної вибірки. Спочатку розробка цього алгоритму не була розрахована [12, 13, 14, 15, 16, 17, 18, 19, 20, 21] на використання нормованих шкал змін сигналів. Тут цей алгоритм адаптований до такого варіанту і досліджується в нових умовах на оброблюваних даних зазначеного типу.

Блок-схема модифікацій першого алгоритму представлена на рис. 6.

Алгоритм розроблявся в двох варіантах в середовищі MATLAB® [23, 24, 25]. У першому, межі шкал вибиралися по мінімальному і максимальному значенням сигналів навчальних вибірок окремо за типами сигналів, а в другому – спільно.

Відповідно до цього в першому випадку для перетворення в нормовану шкалу змін сигналу, що розпізнається, використовувалося три варіанти меж, знайдених для еталонів. З усіма еталонами по типам сигналів порівнювалися всі три реалізації сигналу, які були перераховані в нормовані шкали. Таким чином з кожним еталоном виконувалося три порівняння з розрахунком скалярних добутків. З усіх їх відшукувався максимум, який вказував на правдоподібний варіант рішення.

У другому випадку межі для шкал змін вибиралися один раз за характеристиками форми всіх сигналів навчальних вибірок спільно. Для вхідного сигналу виходив один екземпляр, перетворений до нормованої шкали змін, який порівнювався з кожним з еталонів. Рішення приймалося за тим самим принципом.

Блок-схема модифікацій другого алгоритму представлена на рис. 7.

Алгоритм представлений в двох варіантах. У першому межі шкал для еталонів вибиралися по мінімальному і максимальному значенням показників форми сигналів навчальних вибірок окремо за типами сигналів, а в другому – спільно.

Для вибору меж шкал сигналів, що розпізнаються також використовувалися навчальні вибірки, але для початкового вигляду реалізацій, які надходять. Перерахунок сигналів у характеристики форми не виконувався.

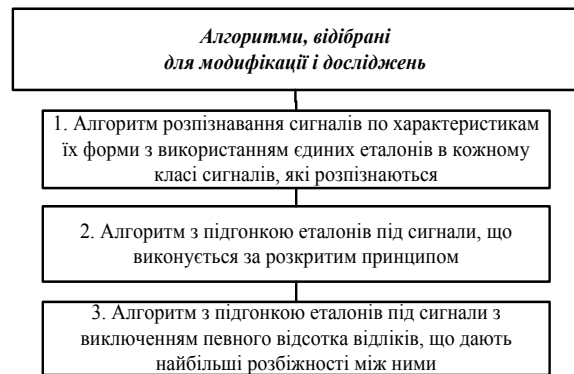


Рис. 5. Склад алгоритмів, відібраних для модифікації і досліджень



Рис. 6. Блок-схема модифікацій алгоритму розпізнавання сигналів за характеристиками їх форми

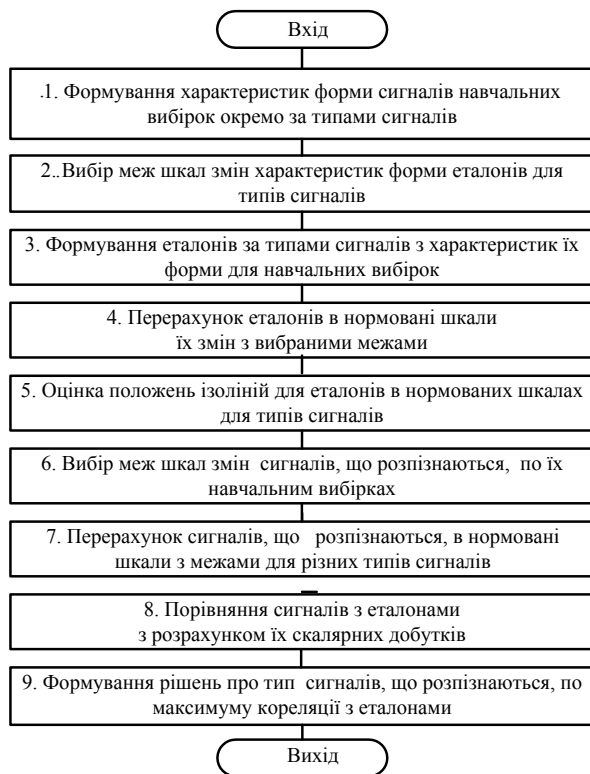


Рис. 7. Блок-схема модифікацій алгоритму з підгонкою еталонів під сигнали за розкритим принципом



Рис. 8. Блок-схема модифікацій алгоритму з підгонкою еталонів під сигнали з виключенням окремих відліків

Для перетворення еталонів у нормовані шкали змін у типах сигналів для першого варіанту використовувалося три пари меж – за типами сигналів. У другому – пара меж для еталонів різних типів була єдиною.

Для перетворення сигналів, що розпізнаються, в нормовані шкали змін у типах сигналів для першого варіанту використовувалося три пари меж. У другому – одна пара меж, знайдена для початкового вигляду реалізацій по навчальних вибірках, що розглядаються спільно.

У першому випадку три варіанти подання вхідного сигналу порівнювалися з кожним з трьох еталонів. У другому – сигнал, що розпізнається, був представлений в одному вигляді і порівнянь виявлялося три замість дев'яти в попередньому варіанті. Принцип прийняття рішень залишався незмінним.

Третій алгоритм відрізняється від попереднього в основному тим, що додатково відкидається з розгляду частина відліків з усієї їх послідовності при прийнятті рішень. Для кожного типу сигналів тут виключаються відліки сигналів і еталонів, відповідність між якими найгірша. Перед фінальним блоком порівняння вставляється зазначений блок відбору відліків.

Блок-схема модифікацій третього алгоритму представлена на рис. 8.

При проведенні досліджень з розгляду виключалося в кожному класі 20% гірших відліків. Для різних класів сигналів їх склад у загальному випадку різний. Таким чином, підгонка в цьому разі виявлялася більш глибокою і це є одним з питань, які треба було перевірити при дослідженнях представлених модифікацій розроблених розпізнавальних процедур.

Розглянуті модифікації процедур розпізнавання сигналів програмно реалізовані в середовищі MATLAB® [23, 24, 25] і були досліджені на тестовому прикладі.

III. ПОРЯДОК І РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ АЛГОРИТМІВ

- 1) Підтвердження працездатності розроблених модифікацій розпізнавальних процедур, які використовують нормовані шкали змін сигналів і еталонів, для різних варіантів вибору меж цих шкал.
- 2) Оцінка рівня правильності розпізнавання сигналів заданої класифікації за стандартними показниками чутливості, специфічності, загальною валідністю в разі використання нормованих шкал змін сигналів і еталонів.
- 3) Оцінка результативності переходу до варіанту порівняння сигналів за підгонкою еталонів під сигнали при прийнятті рішення про їхній тип.
- 4) Порівняльна оцінка роздільного і групового варіантів вибору меж для нормованих шкал змін сигналів і еталонів по навчальним вибіркам для сигналів різних типів.

Дослідження виконувалися на тестовій задачі з вичерпною класифікацією сигналів з трьох класів. Розпізнавані сигнали – QRS-комплекси типів N, A, V у півгодинному запису електрокардіограми пацієнта з бази даних в Internet [22].

У такому записі для кожного комплексу вказано положення R-піку, що дозволяє аналізувати форму всіх комплексів в єдиному вікні їх аналізу з однаковим положенням таких піків. Еталони синхронізовані з сигналами по положенню в цьому вікні, що забезпечує зручність їх порівняння. Для кожного QRS-комплексу в зазначених даних відзначений його тип. У цілому, в наявності є необхідні дані для розгляду розроблених алгоритмів при їх навчанні з учителем. Примірники QRS-комплексів можуть бути відібрані в навчальні та контрольні вибірки для тестування розпізнавальних процедур.

При виконанні роботи були проведені статистичні дослідження правильності рішень запропонований алгоритм з підрахунком кількості вірних рішень та помилкових із записом виду помилок і підрахунком їх кількості. За цими даними оцінювалися чутливість, специфічність, загальна валідність [1, 2, 3, 4, 5, 6, 7, 8, 9, 10] розпізнавання таких сигналів кожної досліджуваної процедури. Використані стандартні формули розрахунку цих показників. Результати обробки даних випробування зазначених процедур узагальнювались в звичні для такого випадку таблиці рішень [1, 2, 3, 4, 5, 6, 7, 8, 9, 10].

Без шкоди до змістовного боку досліджень, з метою спрощення виконуваних розрахунків, обсяги навчальних вибірок для сигналів кожного типу були відібрані однаковими. Однаковими за обсягом були також контрольні вибірки того ж розміру. Навчальні та контрольні вибірки за своїм складом не перетиналися. У таких умовах обсяги зазначених вибірок призначалися максимально можливими відповідно до конкретного вмісту запису ЕКГ конкретного взятого для прикладу пацієнта.

У тестовому прикладі обсяг кожної навчальної та контрольної вибірки для будь-якого з трьох типів сигналів, що розпізнавалися, склав 62 екземпляри. Всі сигнали вибірок були розглянуті в єдиному вікні аналізу. Кожен алгоритм у двох різновидах досліджувався по повній програмі відповідно до поставлених цілей. Таблиці рішень в кожному випадку були поєднані зі зведених розрахунків показників правильності рішень, що були отримані на контрольних вибірках.

Отримані результати представлені у табл.1, де наводяться оцінки результативності двох варіантів групового вибору меж шкал змін сигналів і еталонів трьох обраних для дослідження алгоритмів розпізнавання сигналів.

Табл 1. розкриває отримані при статистичному дослідженні оцінки результативності варіантів вибору меж шкал змін сигналів і еталонів на тестовому прикладі для досліджуваних алгоритмів розпізнавання сигналів.

КОМПЛЕКСНА ТАБЛИЦЯ 1 – ОЦІНКИ РЕЗУЛЬТАТИВНОСТІ ВАРІАНТІВ ВИБОРУ МЕЖ ШКАЛ ДЛЯ ДОСЛІДЖУВАНИХ АЛГОРИТМІВ

РОЗДІЛЬНИЙ ВИБІР MIN/MAX					СУМІСНИЙ ВИБІР MIN/MAX				
<i>Базовий алгоритм</i>					<i>Базовий алгоритм</i>				
	A	N	V	Спец.		A	N	V	Спец.
A	27	0	12	69,23	A	29	0	6	82,86
N	35	62	2	62,63	N	33	62	1	64,58
V	0	0	48	100,00	V	0	0	55	100,00
Чутлив.	43,55	100,00	77,42	73,66%	Чутлив.	46,77	100,00	88,71	78,49%
АЛГОРИТМ З ПІДГОНКОЮ					АЛГОРИТМ З ПІДГОНКОЮ				
	A	N	V	Спец.		A	N	V	Спец.
A	24	0	8	75,00	A	35	0	4	89,74
N	38	62	2	60,78	N	27	62	1	68,89
V	0	0	52	100,00	V	0	0	57	100,00
Чутлив.	38,71	100,00	83,87	74,19%	Чутлив.	56,45	100,00	91,94	82,80%
АЛГОРИТМ З ПІДГОНКОЮ (мінус 20% відліків)					АЛГОРИТМ З ПІДГОНКОЮ (мінус 20% відліків)				
	A	N	V	Спец.		A	N	V	Спец.
A	24	0	8	75,00	A	38	0	4	90,48
N	38	62	2	60,78	N	24	62	1	71,26
V	0	0	52	100,00	V	0	0	57	100,00
Чутлив.	38,71	100,00	83,87	74,19%	Чутлив.	61,29	100,00	91,94	84,41%



Дані наведених таблиць є підставою для таких висновків.

- 1) Наведені результати підтверджують працездатність розроблених модифікацій розпізнавальних процедур, які використовують нормовані шкали змін сигналів і еталонів, для двох розглянутих групових варіантів вибору меж цих шкал.
- 2) Рівень правильності розпізнавання сигналів заданої класифікації тестового прикладу за стандартними показниками чутливості, специфічності, загальною валідністю в разі використання нормованих шкал змін сигналів і еталонів є високим.
- 3) Використання наведених шкал для опису сигналів і еталонів з метою визначення типів сигналів виявилось результативним.
- 4) Більш результативним для сигналів розглянутого типу виявився вибір кордонів шкал збільшень сигналів при спільному розгляді навчальних вибірок одночасно всі трьох типів сигналів їх використовуваної класифікації.
- 5) Підгонка еталонів під сигнали при забезпеченні їх порівняння для прийняття рішень про тип сигналів у порівнянні зі звичайною оцінкою близькості сигналів (новоприбулих і еталонних) по їх скалярним добуткам дає більш високі результати правильності прийнятих рішень.
- 6) Підгонка еталонів під сигнали з відкиданням певного відсотка гірших відліків призводить до додаткового підвищення правильності рішень розпізнавального алгоритму.

ВИСНОВКИ

При виконанні роботи отримані такі результати.

- 1) У роботі запропоновані та перевірені на тестовому прикладі обробки медико-біологічних сигналів конкретного типу з трьома їх різновидами у вичерпній класифікації шість модифікацій розпізнавальних алгоритмів для систем, які навчаються з учителем, з поданням сигналів і еталонів у нормованих шкалах їх змін.
- 2) У модифікаціях реалізовано два варіанти спільного вибору меж зазначених шкал – спільний і роздільний для навчальних вибірок сигналів різних класів. Процедури порівняння сигналів з еталонами і прийняття рішень побудовані з підгонкою еталонів під сигнали і без неї.
- 3) Модифікації розпізнавальних процедур реалізовані та досліджені в середовищі MATLAB®. Результати досліджень представлені й узагальнені у відповідних таблицях рішень алгоритмів, в які вбудовані значення оцінок показників їх чутливості, специфічності і загальної валідності при обробці контрольних вибірок.

Основними можна вважати такі висновки.

- 1) Перехід до нормованих шкал змін представлення сигналів для їх опису при розпізнаванні є спроможним. Сумірність сигналів і еталонів, як впливає з аналізу результатів проведених розрахунків, досягається.
- 2) Підгонка, як спосіб порівняння сигналів і еталонів та прийняття рішень за результатами перевірки на тестовому прикладі коректна і стає більш результативною у разі її поглиблення за рахунок відкидання однієї і тієї ж кількості погано погоджених відліків сигналів і еталонів з розгляду за абсолютним значенням різниці між ними.
- 3) Перехід до нормованих шкал змін сигналів і еталонів можна розглядати як один із способів, що сприяють підвищенню правильності розпізнавання сигналів.
- 4) Більш продуктивним є вибір меж для шкал змін сигналів при спільному використанні реалізацій навчальних вибірок всіх класів сигналів використовуваної класифікації.

ПЕРЕЛІК ПОСИЛАНЬ

- [1] A. A. Genkin, *New informational technology of medical data analysis (OMIS software package)*, Saint Petersburg: Politekhnik, 1999, p. 191, ISBN: 5-7325-0326-9.
- [2] V. A. Djuk and V. L. Jemanujel', *Information technology in biomedical studies*, Saint Petersburg: Piter, 2003, p. 528, ISBN: 5947235013.
- [3] M. U. Antomonov, *Mathematical processing and analysis of biomedical data*, Moscow: Nauka, 2005, p. 558, ISBN: 966-7865-70-3.
- [4] V. I. Vasil'ev, *Recognition systems. Reference book*, Kiev: Naukova Dumka, 1983, p. 421.
- [5] O. Mincer, Y. Voronenko and V. Vlasov, *Information technology in health and practical medicine: Educ. textbook*, Kiev: Vyshcha shkola, 2003, p. 350, ISBN: 966-642-160-7.
- [6] N. I. Yabluchanskij and A. V. Marty'nenko, *Heart rate variability. To help the practitioner*, Khar'kov, 2010, p. 131.
- [7] V. A. Gel'man, *Medical informatics*, Saint Petersburg: Piter, 2001, p. 480, ISBN: 5-272-00312-8.
- [8] R. M. Baevskij, *Prediction of conditions on the verge of normalcy and pathology*, Moscow: Meditsina, 1979, p. 295, ISBN: 978-5-458-39258-7.
- [9] R. M. Baevskij, O. I. Kirillov and S. Z. Kleczkin, *Mathematical analysis of changes in heart rate during stress*, Moscow: Nauka, 1984, p. 220.
- [10] R. M. Baevskij and R. E. Moty'lyanskaya, *Heart rhythm in athletes*, Moscow: Fizkul'tura i sport, 1986, p. 142.
- [11] A. A. Korostelev, N. F. Klyuev, Y. A. Mel'nik, A. A. Vegetyagin, V. A. Gubin and V. E. Dulevich, *Theoretical foundations of radar: textbook for universities*, Moscow: Sov. radio, 1978, p. 608.
- [12] A. P. Shulyak and Shachykov A. D., "Criteria and procedures for estimation the informativity and feature selection in biomedical signals for their recognition," *«Visnyk NTUU KPI» Seriya – Radiotekhnika. Radioaparaturbuduvannya*, no. 66, p. 86, 2016, DOI: [10.20535/RADAP.2016.66.79-86](https://doi.org/10.20535/RADAP.2016.66.79-86).
- [13] A. D. Shachykov and A. P. Shulyak, "Analysis of the Biomedical Signals' of the Problem of Recognition," in *2014 IEEE XXXIV International Scientific Conference Electronics and Nanotechnology (ELNANO)*, Kyiv, 2014, DOI: [10.1109/ELNANO.2014.6873982](https://doi.org/10.1109/ELNANO.2014.6873982).



- [14] A. D. Shachykov and A. P. Shulyak, "Decomposition of Biomedical Signals on Mutually Orthogonal Components in the Diagnosis of Diseases," in *2014 IEEE XXXIV International Scientific Conference Electronics and Nanotechnology (ELNANO)*, Kyiv, 2014, DOI: [10.1109/ELNANO.2014.6873914](https://doi.org/10.1109/ELNANO.2014.6873914).
- [15] A. Shulyak and A. Shachykov, "Working off the analysis principles of cyclic biomedical signals' structure for their detection, recognition and classification," *Bulletin of National Technical University of Ukraine "Kyiv Polytechnic Institute". Series Instrument Making*, no. 49(1), 2015, DOI: [10.20535/1970.49\(1\).2015.47101](https://doi.org/10.20535/1970.49(1).2015.47101).
- [16] A. D. Shachykov and A. P. Shulyak, "About the Impact of Informative Features Selection in the Mutually Orthogonal Decompositions of Biomedical Signals for their Recognition," in *2016 IEEE 36th International Conference ELECTRONICS and NANOTECHNOLOGY (ELNANO)*, Kyiv, 2016, DOI: [10.1109/ELNANO.2016.7493054](https://doi.org/10.1109/ELNANO.2016.7493054).
- [17] A. D. Shachykov, V. A. Shelofast and O. P. Shulyak, "Modifications of the procedures for feature selection of cyclic biomedical signals for their recognition," *Bulletin of NTUU "KPI". Series Instrument Making*, no. 53(1), 2017, DOI: [10.20535/1970.53\(1\).2017.106810](https://doi.org/10.20535/1970.53(1).2017.106810).
- [18] A. Shulyak and Shachykov A.D., "Features of using the shape characteristics of biomedical signals during their recognition," *Bulletin of NTUU "KPI". Series Instrument Making*, no. 51(1), 2016, DOI: [10.20535/1970.51\(1\).2016.78217](https://doi.org/10.20535/1970.51(1).2016.78217).
- [19] I. D. Mandel, *Cluster analysis*, Moscow: Finance and Statistics, 1988, p. 176.
- [20] A. Kotov and N. Krasilnikov, *Data clustering*, 2006.
- [21] B. Duran and P. Odell, *Cluster analysis*, Moscow: Statistics, 1977, p. 128.
- [22] "PhysioNet. St.-Petersburg Institute of Cardiological Technics 12-lead Arrhythmia Database," 2008. [Online]. Available: <https://physionet.org/content/incartdb/1.0.0/>.
- [23] V. Potemkin, *Introduction to MATLAB*, Moscow: Dialog-MIFI, 2000.
- [24] V. P. D'yakonov, *MATLAB. Complete tutorial*, Moscow: DMK Press, 2012, p. 768, ISBN: 978-5-94074-652-2.
- [25] A. B. Sergienko, *Digital signal processing*, Saint Petersburg: Piter, 2002, p. 608, ISBN: 5-318-00666-3.

Надійшла до редакції 28 лютого 2020 року

The Choice of Boundaries for Scales of Biomedical Signals in the Recognition Algorithms

V. V. Lahutin^f, ORCID [0000-0003-3431-1686](https://orcid.org/0000-0003-3431-1686)

V. I. Sirotenko^f, ORCID [0000-0002-7542-5885](https://orcid.org/0000-0002-7542-5885)

A. D. Shachykov^s, PhD, ORCID [0000-0003-2353-7647](https://orcid.org/0000-0003-2353-7647)

O. P. Shulyak^s, PhD Assoc.Prof., ORCID [0000-0002-6920-2212](https://orcid.org/0000-0002-6920-2212)

Electronic Engineering Department ee.kpi.ua

National technical university of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute» kpi.ua

Kyiv, Ukraine

Abstract—The work is aimed at improving the recognition procedures of biomedical signals in supervised learning systems in the direction of increasing the fidelity level of decisions. The prerequisites of the work is the increasing data flow during medical diagnostics and deepening search of informative features of signals of patient's state. The results of the work are applied to the field of medical diagnostic software.

The work is based on using shape characteristic of a signal and scalar product as criteria of several signals' similarity. The reasoning for the work is to provide signal shape characteristic algorithms with the methods to manipulate the standard for more tailored supervised learning and effective recognition. While simultaneously maintaining correctness of comparing incoming signals with the heavily modified standards with custom composition of features.

Six modifications of recognition algorithms with the representation of signals in normalized measurement scales are proposed, in which joint and separate options for choosing the boundaries of these scales on training samples of signals of different classes are implemented. Procedures have been built for comparing signals with standards and making decisions with and without fitting standards to signals.

Modifications to recognition procedures are implemented and explored in the MATLAB[®] environment. Their performance has been confirmed. The half-hour ECG recording from an open access database has been used for signal recognition for the normal and two types of abnormal cardiac beats.

The viability of the development has been confirmed with the control samples in test example. An increase was found in the values of statistical estimates of the sensitivity, specificity, and general validity of solutions for recognizing signals when switching to normalized measurement scales in comparison with the option of using ordinary scales of measurements in describing the observed processes. The modification with variable boundaries of the scales has been found to be more effective for the used signal types while learning all three types of signals simultaneously.

Additional modifications with variable standards and discarding worse features allowed further increase in correct decisions rate. As the result, the statistical values of recognition has been calculated for all six algorithms with variable calculation complexity and effectiveness.

Special consideration has been paid to the consistency of signals and their standards for the cases of variable scales, boundaries, and composition during learning and recognition phases.

Keywords — *biomedical signals; recognition algorithms; supervised learning; values' change scales.*

