УДК 628.321
© В Заяць[1], Г. Шокира[1], 2012

МОДЕЛІ ТА ПІДХОДИ ДО ПОБУДОВИ РОБОЧИХ СЛОВНИКІВ ОЗНАК ЗА ПРІОРИТЕТОМ ТА ІНФОРМАТИВНІСТЮ ПЕРВИННИХ ОЗНАК

Запропоновано імітаційну моделі для оптимізації робочого словника ознак на основі розробленого авторами підходу до встановлення пріоритету первинних ознак при створенні систем розпізнавання об'єктів та явищ. Визначено основні напрями прикладних застосувань отриманих результатів.

The simulation models for optimizing your dictionary features based on the authors developed by approach to priority primary signs when creating recognition systems of objects and phenomena is proposed. A main directions of applied applications of the results is determined.

1. ОПИС ІМІТАЦІЙНОЇ МОДЕЛІ

У роботі розглядаються моделі та підходи до побудови та оптимізації робочого словника ознак при побудові систем розпізнавання на основі встановленого пріоритету ознак. Розв'язання цієї задачі дає можливість оптимізувати робочий словник ознак, вибравши найбільш інформативні з них. У ролі моделі процесу розпізнавання виступають детерміновані ознаки, які приймають конкретні числові значення або лежать у певному інтервалі значень (детерміновані моделі) або/і вірогіднісні ознаки, в ролі яких виступають вірогідності віднесення об'єкту до того або іншого класу та умовні вірогідності приналежності ознаки даному об'єкту за умови його віднесення в певний клас (вірогіднісні моделі). Такий підхід дає можливість в умовах неминучих фінансових асигнувань забезпечити створення системи розпізнавання об'єктів чи явищ та їх ідентифікації з заданою точністю при неповній апріорній інформації та економії технічних засобів.

2. ВИБІР ПЕРВИННИХ ОЗНАК ДЛЯ РІЗНИХ КАТЕГОРІЙ ОБ'ЄКТІВ РОЗПІЗНАВАННЯ

Раціональний вибір параметрів, що характеризують образи розпізнавання, значною мірою визначає складність апаратури розпізнавання і рівень вірогідності правильного розпізнавання. Слід мати на увазі, що при автоматичному розпізнаванні образів принципово неможливо

[1] Національний університет «Львівська політехніка»

створити пристрій, що забезпечує абсолютне безпомилкове розпізнавання.

Завжди є деяка ймовірність помилки. Необхідно добитись того, щоб пристрій розпізнавання забезпечував найбільш можливу вірогідність правильної класифікації при вибраних параметрах, детерміновані або вірогіднісні властивості яких відомі і задані допустимі імовірності помилки.

Конструкція і технічна реалізація пристрою розпізнавання, особливо тієї його частини, яка служіть для сприйняття (вимірювання) параметрів, що характеризують образ, залежить від фізичної природи розпізнаваного об'єкта [1,2]. Один і той же об'єкт може бути охарактеризований, не тільки геометричними властивостями (висотою, шириною), а й механічними, електричними, хімічними, тепловими та іншими властивостями., тобто об'єкт можна оцінити з різних точок зору. Сукупність усіх ознак створює зображення (образ) об'єкта. При цьому, якщо по деякому ряду ознак об'єкти можна об'єднати в один загальний клас і розглядати як один образ, то за іншими ознаками ті ж об'єкти виявляються абсолютно відмінними один від одного і повинні бути віднесені до різних класів. Для прикладу, якщо характеризувати образ власною частотою коливань, то морський корабель і маятник належать до одного класу об'єктів, а якщо залучити такі ознаки, як геометричні розміри і форма, то – до різних.

Тому при розпізнаванні слід спочатку встановити, з якої точки зору необхідно розрізняти образи в даній задачі, і відповідно до цього вибрати фізичні чи інші параметри, що характеризують об'єкти.

В роботах [3-5] подані методи вибору ознак для різних прикладних застосувань. А в роботах [6,7] здійснюється аналіз вибраних ознак.

Для прикладу, у роботах з діагностики хвороб [8] в якості ознак хвороб використовувалися, крім її симптомів, дані лабораторних аналізів і обчислювалися ймовірності наявності тих чи інших ознак за даних хвороб. Це служило вихідним статистичним матеріалом для їх розпізнавання. Як показали експерименти, ймовірність правильного розпізнавання хвороб за допомогою формального алгоритму на електронних обчислювальних машинах істотно більше, ніж вірогідність правильного розпізнавання тих же хвороб лікарем.

Методами теорії розпізнавання вирішуються також найбільш складні завдання класифікації корисних копалин за результатами геофізичних вимірів: так, наприклад, в [9] описано, як проводився процес розпізнавання з досить малою ймовірністю помилок водоносних і нафтоносних пластів в свердловинах.

Використання зібраного статистичного матеріалу і застосування оптимальних алгоритмів розпізнавання та встановлення пріоритету

первинних ознак дозволяють зробити це з більшою ймовірністю прийняття вірних рішень.

Таким чином, вибір в якості параметрів розпізнавання тих чи інших фізичних величин визначається:

а) інваріантністю параметрів до всіляких перетворень,

б) простотою пристрою вимірювання обраних параметрів і пристрою, що перетворює параметри в код, зручний для використання в пристрої прийняття рішення,

в) чутливістю параметрів до протидій, тобто завадостійкістю параметрів;

г) складністю навчання та самонавчання систем.

У кожній конкретній задачі розпізнавання найважливішим є оптимальний вибір комплексу найбільш інформативних параметрів розпізнаваного образу, тобто встановлення пріоритету кожної із ознак в процесі розпізнавання. Такий вибір зумовить і кращу конструкцію пристрою і більш широкі перспективи його використання [10, 11].

3. ПІДХОДИ ДО ПОБУДОВИ РОБОЧОГО СЛОВНИКА ОЗНАК В УМОВАХ ОБМЕЖЕНИХ РЕСУРСІВ

При розробці робочого словника ознак доводиться зустрічатися з наступними обмеженнями:

1) у словник включають тільки ознаки, щодо яких може бути отримана апріорна інформація, необхідна для опису класів;

2) деякі з ознак недоцільно включати в апріорний словник з огляду на те, що вони малоінформативні;

3) деякі з ознак, як правило, найбільш інформативні, не можуть бути визначені зважаючи на відсутність відповідних вимірюючи пристроїв, а ресурси, асигновані на створений цієї системи розпізнавання, обмежені.

Саме тому апріорний словник ознак в загальному випадку може бути застосований лише як основа для побудови реально використовуваного в системі розпізнавання робочого словника, що включає лише ознаки, які, з одного боку, найбільш інформативні і, з іншого боку, можуть визначатися наявною або спеціально створеною вимірювальною апаратурою.

Таким чином, завдання розробки робочого словника ознак системи розпізнавання в загальному випадку зводиться до того, щоб в межах виділених ресурсів визначити перелік технічних засобів, спостережень, створення яких необхідно забезпечити для ознак з найвищим пріоритетом. Цей пріоритет може бути встановлено на основі методів та підходів, запропонованих в роботі [12].

На основі аналізу цих робіт, реальних технічних обмежень та результатів імітаційного моделювання розробленої авторами комп'ютерної системи розпізнавання детермінованих об'єктів формульовано підхід до побудови робочого словника ознак. Основними результатами застосування даного підходу є:

– процес розпізнавання буде точніші у тому випадку, коли об'єкти, що належать кожному з заданих класів в просторі ознак компактно розташовані, а відстань між класами є достатньо велика.

– при збільшенні числа класів для зменшення середнього ризику необхідно включати до складу робочого словника такі ознаки, які мають менший розкид.

– зниження ефективності розпізнавання за рахунок збільшення числа класів може бути компенсовано збільшенням розмірності вектора ознак.

– вирішуючи задачу підвищення ефективності систем розпізнавання за рахунок збільшення розмірності вектора ознак, необхідно звернути увагу на те, що вказане збільшення – це в основному зростання числа технічних засобів вимірювань, кожен з яких забезпечує визначення одного або групи ознак. Означає при цьому зростають витрати на побудову систем розпізнавання. А ресурси часто обмежені.

В умовах обмежених ресурсів на створення системи розпізнавання (обмежені фінансові ресурси, відсутні всі необхідні технічні засоби вимірювання, наявність малоінформативних, непріоритетних ознак) тільки деякий компроміс між розмірами алфавіту класів і об'ємом робочого словника ознак забезпечує рішення задачі розпізнавання та коректної ідентифікації об'єктів і явищ оптимальним чином.

4. ВИСНОВКИ

Застосування запропонованої моделі до побудови робочих словників ознак після встановлення їх пріоритету дозволяє оптимізувати процедуру формування робочого словника і в умовах неповної інформаційної визначеності і ефективно здійснити процедуру розпізнавання з заданою точністю.

1. Пересада В. П. Автоматическое распознавание образов./ В. П. Пересада. – Л. : «Энергия», 1970. – 92 с. 2. Заяць В. М. Методи розпізнавання образів. / В. М. Заяць, Р. М. Камінський. – Навчальний посібник. – Львів : „Львівська політехніка", 2004.- 176 с. 3. Аникина Г. О. Непараметрические алгоритмы распознавания образов в пространстве дискретных признаков. / Г. О. Аникина. – Информатика и системы управления: межвуз. сб. науч. тр. Красноярск: ГУ НИИ ИПУ КГТУ, 2003. – Вып. 9. – С. 55-64. 4. Барило К. В. Алгоритм оптимізації рівня селекції координат еталонних векторів класів розпізнавання./ К. В. Барило. – Тези доповідей Третьої міжнародної науково-практичної кон-

ференції (Суми, 2-4 листопада 2011 року) Інтелектуальні системи в промисловості і освіті (ІСПО). – Суми «Видавництво СумДУ», 2011. – Том І – С. 12-13. 5. Тарасов Е. М. Выбор признаков и решающих правил при распознавании состояний рельсовых линий сортировочных горок. / Е. М.Тарасов, М. В. Трошина. – Уфа : УГАТУ, 2012. – Т. 16. № 3 (48). – С. 194-197. 6. Aeberhard S. Comparative analysis of statistical pattern recognition methods in high dimensional settings. / S. Aeberhard, D. Coomans, O. de Vel. – Pattern Recognition, 1994. – Vol.27. № 8 – P. 1065-1077. 7. Ball G. H. A novel Method of Data Analysis and Pattern Classification. / G. H. Ball, D. J. Hall. – Menlo Park, CA: Stanford Research Institute, 1995. 8. Кравченко А. И. О решении на ЭЦВМ задачи распознавания логичного образа, заданного дискретной информацией. / А. И. Кравченко, М. И. Кравченко. – Известия ВУЗов. "Электромеханика", 1996.– №4. 9. Вапник В.М. Системы обучения распознаванию образов при помощи обобщенных портретов. / В. М. Вапник. – Известия АН СССР, "Техническая кибернетика", 1965 – №1. 10. Горелик А. Л. Об одном методе решения задач классификации объектов или явлений / А. Л. Горелик, В. А. Скрипник. – Техническая кибернетика, 1965. -№1. – С. 10-19 с. 11. Заяць В. М. Доцільність встановлення пріоритету первинних ознак при побудові систем розпізнавання та ідентифікації об'єктів і процесів на основі детермінованих та імовірнісних методів. / В. М. Заяць – Радіоелектроніка. Інформатика. Управління, 2009. – № 1.– С. 92-98. 12. Заяць В. М. Алгоритм корекції пріоритету первинних ознак при побудові систем розпізнавання. / В. М. Заяць, Шокира Г. Я. – Збірник наукових праць. Інститут проблем моделювання в енергетиці. Київ, 2012. – Вис. 63 – С. 105-111.