

Любчик, В.І. Стецюк, І.В. Файфура// Вісник Хмельницького національного університету. Технічні науки. – Хмельницький. – 2015. – №3. – С. 232-237.

5. Патент № 5392289 США. МПК G06F 11/00; H04L 12/00; H04L 7/00; H03M 13/00. Error rate measurement using a comparison of received and reconstructed PN sequences. / George R. Varian, Palo Alto, Calif заявлено 13.10.93; N136075; опубл. 21.02.95.

6. Sudhir Babu and K.V. Sambasiva Rao "Evaluation of BER for AWGN, Rayleigh and Rician Fading Channels under Various Modulation Schemes" International Journal of Computer Applications (0975 – 8887). – 2011. – Volume 26. – No. 9 July.

Рецензія/Peer review : 11.9.2015 р.

Надрукована/Printed :2.11.2015 р.

Рецензент: д.т.н., професор Філінюк М.А.

УДК 004.8

В.В. ГРИЦІК

Тернопільський національний університет ім. І.Пулюя

## ІНФОРМАЦІЙНО-АНАЛІТИЧНА СИСТЕМА СПРИЙНЯТТЯ ЗОВНІШНЬОГО СВІТУ

*Розробка інтелектуальної інформаційної технології доповненої реальності для осіб з обмеженими можливостями, зокрема перетворення візуальних образів у звукові і навпаки шляхом генерування єдиного поняття. Та використання в управляючому механізмі спільної бази знань, що зберігає образи понять.*

*Ключові слова: інтелектуальна інформаційно-аналітична система, реальний час опрацювання інформації, сприйняття зовнішнього світу, однорідні обчислювальні середовища.*

V.V. HRYTSYK

Ternopil's I.Pul'j National Technical University

### INFORMATION-ANALYTICAL SYSTEM OF PERCEPTION OF THE OUTSIDE WORLD

*Abstract - Development of intellectual information technology augmented reality for persons with disabilities, including converting visual images into sound and vice versa by generating a single concept are presented in article. Using the technology of the unified knowledge base for operating mechanism, which stores patterns of concepts, is developed in the article.*

*This approach allows best use computing resources by decreasing amounts of memory usage and computational elements. The system also reduces the number of applications and can be scaled to persons with hearing impairments and good eyesight for the purpose of data base common to all the tasks. In the case of the scaling proposed approach allows (unlike others) to establish communication between persons with disabilities to vote (silent) and persons with impaired vision (blind). The automatic adaptation to the environment and adaptive search of useful objects is an important feature of developed information technology.*

*The limitation of the system is the assumption that it contains some breaks (gaps of silence between words) that the system considers delimited words. Such systems need boarder between useful signal and background.*

*The optimization capabilities of neural network models for decomposition, synthesis and presentation of the images are obtained from different spectral source is considered in this work Their integration into a single information-analytical system is considered too. In this paper, The author continued research the concept of proposed model in which the neural network one output neuron responsible for term when various spectral data, including video and audio images are applied to the input.*

*Key words: Intellectual information-analytic system, real time of data processing, the perception of the outside world, homogeneous computing environment.*

### Аналіз стану та постановка проблеми

З мільйонами камер, спостереження у світі за тим, що відбувається на міських вулицях і місцях скупчення таких, як торгові центри та транспортні вузли [1], проблема автоматичної обробки даних, моніторингу та аналізу цих відеопотоків стала епічним завданням. Реалізацію розбиття обчислень до рівня однорідних середовищ для опрацювання та розпізнавання зорових образів і мови українськими науковцями описано в [2-7].

Європейці за допомогою різних методів опрацювання звуку будують цифрові міста [8, 21,22] та все більше уваги приділяється методам дистанційного відбору інформації [9]. Відтворюючи людську поведінку в робототехніці моделюють механізми людського навчання та сприйняття при реалізації методів комп'ютерного зору [10]. Роблячи аналіз провідних ІТ-компаній світу ми побачимо, що Google, Apple та Facebook застосовують нейронні мережі для розпізнавання звуків та пошуку об'єктів [11].

У новинах ІТ-індустрії Майкрософт реалізувала бета-тест миттєвого голосового перекладу наприкінці 2014 року [12], у той час, як Гугл заявила про те, що вони планують схоже рішення у 2015 [13].

Що ж до Європейського наукового простору, то ЄС через програму FP7 профінансував розробки складних технік для розуміння аудіовізуальних відображень впродовж життя для типових і нетипових популяцій [14].

Зараз все актуальнішими [15] стають структури, що масштабно відтворюють роботу людського мозку. Зокрема на цю тематику через програму FP7 з 2007 виділено 1,9 млрд. євро на 1200 проектів із 1200-ми учасниками. Вважається, що роботизовані системи вводу, які інтегровані в тулуб з руками і ногами є найближчим майбутнім комп'ютерного зору [16]. А це потребуватиме розробки моделей і методів синтезу

різнорідних даних. Отже, актуальною проблемою розробки штучного інтелекту є розробка принципів розуміння комп'ютером зовнішнього світу через інтеграцію не однотипних даних.

Узагальнюючи огляд досліджень та впровадженнь інтелектуальних ІТ в ЄС та США перейдемо до постановки задачі.

**Постановка задачі**

Ціль даної роботи – розробка концепції представлення зовнішнього світу за допомогою аудіовізуального сприйняття. Практичною стороною може бути технологія доповненої реальності для осіб з вадами зору або слуху [20].

Мета даної роботи – дослідження оптимізаційних можливостей інтегрованих систем аудіовізуального сприйняття для задач штучного інтелекту. Дослідження проводиться на частковій вибірці – цифр і тексту, що відповідає актуальності [11-13].

**Основна частина**

Концепція розробки аудіо-візуальної системи сприйняття, на перший погляд, є досить простою. Однак, між аудіо і відеопотоками є величезна різниця – кількість даних і чинників, які потрібно враховувати для прийняття рішення.

1.1. Система розпізнавання відеообразу.

В даний час найбільше поширені «інтелектуальні» системи, що з високим ступенем точності розпізнають більшість шрифтів. Деякі системи оптичного розпізнавання тексту здатні відновлювати початкове форматування тексту, включаючи зображення, колонки та інші нетекстові компоненти (рис. 1).

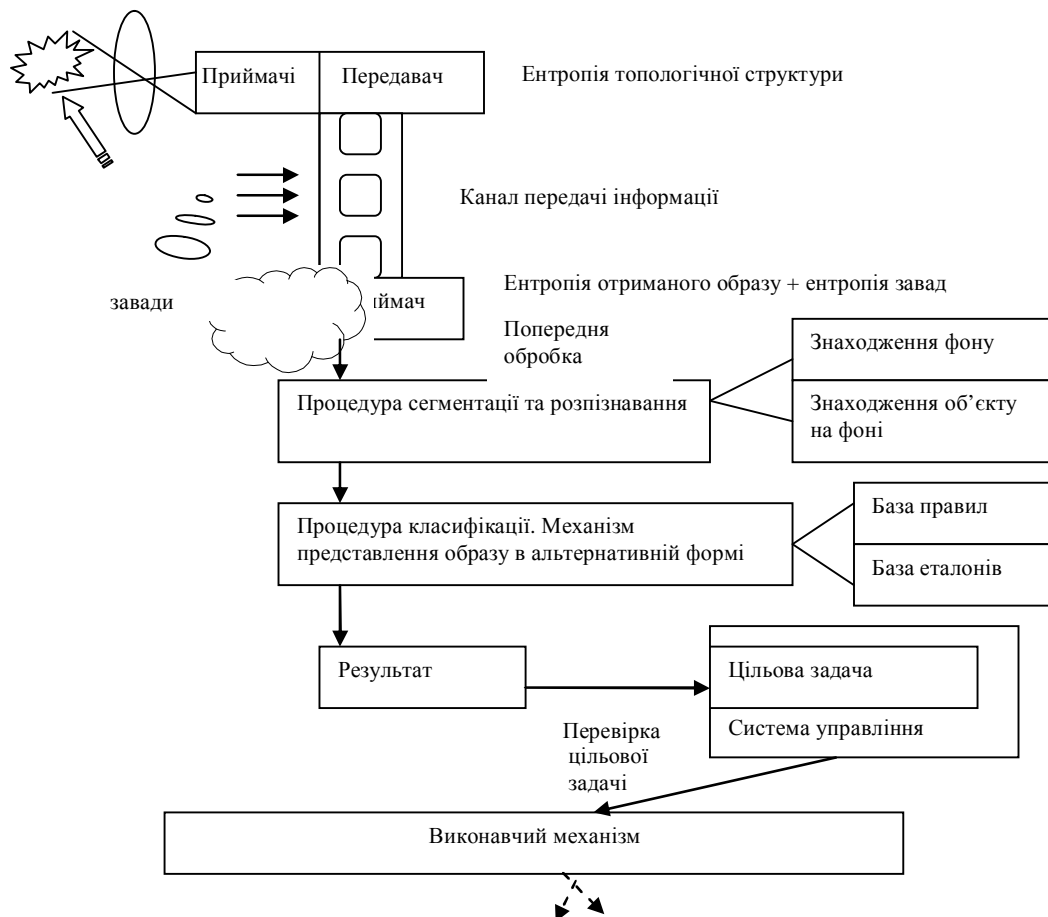


Рис. 1. Модель розпізнавання образів

Тому, перед обробкою зображення алгоритмами розпізнавання проводиться його попередня обробка, спрямована на поліпшення якості зображення. Вона включає фільтрацію зображення від шумів, підвищення різкості і контрастності зображення, вирівнювання і перетворення у використовуваний системою формат [4].

Виявлені фрагменти зображення подаються на вхід класифікатора, виходом якого є вектор можливості приналежності зображення до класу тієї чи іншої літери [5]. У запропонованому підході використовується класифікатор складовою архітектури, організований у вигляді дерева, листям якого є прості класифікатори, а внутрішні вузли відповідають операціям комбінування результатів нижніх рівнів (рис. 2).

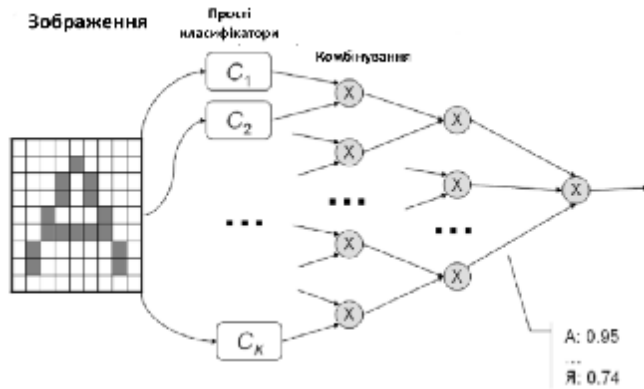


Рис. 2. Архітектура класифікатора

Робота простого класифікатора здійснюється в два кроки. Спочатку по вихідному зображенню обчислюються ознаки. Значення кожної ознаки є функцією від яскравості деякої підмножини пікселів зображення. У результаті виходить вектор значень ознак, який надходить на вхід нейронної мережі (рис. 3). Кожен вихід мережі відповідає одній з букв алфавіту, а одержуване на виході значення розглядається як рівень приналежності букви нечітких множин.

Завданням алгоритму комбінування є узагальнення інформації, що надходить у вигляді вхідних нечітких множин та обчислення на їх основі вихідного нечіткої підмножини множини символів, що розпізнаються. В якості алгоритмів комбінування використовуються операції теорії нечітких множин, вибір найбільш впевненого експерта.

Результатом роботи класифікатора є нечітка множина, отримана в результаті комбінування на верхньому рівні. На останньому етапі приймається рішення про найбільш правдоподібний варіант прочитання слова. Для цього

Рис. 3. Простий класифікатор

використовуються рівні можливості прочитання образів окремих букв.

Обробка зображення документа проводиться в три етапи: сегментація (із застосуванням операцій фільтрації і заливки), розпізнавання і орфографічна корекція [6]. Виділення цих етапів є функціональною декомпозицією і дозволяє проектувати три різних підсистеми, оптимізовані під алгоритми кожного етапу. Вироблені оцінки трудомісткості алгоритмів, що виконуються на різних етапах обробки, і надані рекомендації по машинно-залежним аспектам розробки паралельних алгоритмів.

### 1.2 Система розпізнавання голосу

Процес розпізнавання мови може бути розділений на дві основні фази: оцифрування і декодування. На першій фазі вхідний аудіосигнал записується і розбивається на фрагменти. На фазі декодування отримана інформація аналізується на основі використання різних моделей і алгоритмів. Алгоритми декодування можуть спиратися на зразки як цілих слів, так і окремих частин слів. Найменшою частиною слова є фонема, і будь-якій мові, зазвичай, достатньо 40-60 фонем, щоб описати вимову всіх слів.

Найбільш точними з точки зору розпізнавання є моделі, засновані на розпізнаванні слів цілком. Однак вони можуть використовуватися лише в системах зі словниками невеликого обсягу [7].

Моделі, засновані на фонемній структурі, є набагато більш універсальними і значною мірою вирішують проблему обсягу словника.

Для кращого розуміння особливостей задачі розпізнавання мови необхідно відзначити, що всі перераховані вище системи в принципі працюють однаково, використовуючи одні й ті ж методи та алгоритми. Різниця в типі диктування мови, розмірі словника тощо обумовлена лише специфікою задачі, наявними обмеженнями по швидкості обчислень і обсягом необхідної пам'яті. Спрощений процес розпізнавання може бути описаний у вигляді послідовності наступних основних кроків.

#### 1.2.1 Шумоочистка і відділення корисного сигналу.

Методи, що застосовуються для вирішення даного завдання, можна умовно розділити на чотири групи. Методи першої групи зазвичай зводяться або до виділення деяких інваріантних щодо шуму ознак, або до навчання в умовах шумів або модифікації еталонів розпізнавання з використанням оцінки рівня шумів. Як приклад методів другої групи можна назвати коефіцієнти лінійного передбачення, кепстральні коефіцієнти.

В якості прикладу елементи еталонів в даному випадку замість чисельних значень використовують імовірнісні розподілу (середнє + дисперсія).

Третя група методів пов'язана з цифровою обробкою сигналу. До методів цієї групи можна віднести методи маскування шумів (Чисельні значення, порівнянні з характеристиками шуму, ігноруються або використовуються з меншими ваговими коефіцієнтами) та методи шумозаглушення з використанням декількох мікрофонів (очищення від низькочастотних шумів з використанням двох мікрофонів на відстані

≈50 см і високочастотних - на відстані ≈5 см).

Четверта група методів, що застосовуються для відділення корисного сигналу від сторонніх шумів зазвичай пов'язана з використанням масивів мікрофонів, моделюють спрямований мікрофон зі змінним променем напрямку (найпростіший метод «затримки і підсумовування» або більш складний з модифікацією ваг мікрофонів методом найменших квадратів – Метод Фроста, який дозволяє пригнічувати і такий тип шумів, як спотворення).

### 1.2.2 Перетворення вхідного мовного сигналу в набір акустичних параметрів

Зазвичай звуковий сигнал розбивають на вікна однакової довжини і перетворюють в частотну область за допомогою дискретного перетворення Фур'є або більш складного перетворення, після чого частотні параметри факторизують з метою скорочення розмірності. За фізичним змістом частотні параметри найбільш близькі до тих, які використовує людина в процесі сприйняття мови.

### 1.2.3 Приведення акустичної форми сигналу до внутрішнього алфавітом еталонних елементів

Область значень акустичних параметрів мовлення розбивають на області згущення, які відповідають елементам фонем, однаковим для різних слів даної мови. Зазвичай таких областей для фіксованої мови налічується близько 1000, і, якщо словник системи розпізнавання містить більшу кількість слів, з метою економії пам'яті доцільно, як еталони системи розпізнавання, розглядати не слова, а відповідні фонемні елементи. Сукупність таких еталонних елементів утворює фонетичну кодову книгу.

Зразкові значення параметрів еталонних елементів для всіх дикторів даної мови відомі заздалегідь, і завдання початкового навчання полягає в уточненні значень цих параметрів. У цьому випадку в процесі розпізнавання мови за акустичними параметрами кожному вікну сигналу визначають найближчий до цього вікна еталонний елемент.

### 1.2.4 Розпізнавання послідовності фонем і перетворення її до тексту слова

Після визначення ймовірної послідовності еталонних елементів у вхідному сигналі необхідно відновити по ній невідому послідовність фонем, що є транскрипцією одного із слів словника. Це завдання вирішується за допомогою динамічного програмування, коли в кожен момент часу визначається найбільш ймовірна передбачувана послідовність фонем у сигналі від початку слова до цього моменту часу. Якщо акустичні параметри перетворені в ймовірнісні з використанням кодової книги, а транскрипційні еталони слів задані у вигляді ймовірнісних автоматів, для розпізнавання зазвичай використовуються приховані ланцюги Маркова. У разі завдання еталонів слів у вигляді послідовності значень акустичних параметрів без застосування кодової книги (зазвичай для маленьких словників) застосовують іншу модель розпізнавання мови з використанням динамічного програмування, так звану динамічну деформацію часу.

На даний момент найскладнішими елементами при побудові системи розпізнавання мови є не алгоритми розпізнавання, а побудова акустичної моделі мови та початкове навчання еталонів слів словника. Як правило, для побудови достовірної моделі тієї чи іншої мови необхідне проведення багаторічної роботи зі збору та аналізу акустичних даних величезного числа носіїв цієї мови. Необхідно ретельно врахувати всі типи голосів і акцентів, і для кожної різновиди голосу і акценту отримати достовірну оцінку елементів кодової книги даної мови.

## Висновки

У роботі досліджено оптимізаційні можливості неймережевої моделі для декомпозиції, синтезу та представлення образів, що отримані з різних спектральних джерел та їх інтеграції в єдиній інформаційно-аналітичній системі.

У роботі автором продовжено дослідження концепції запропонованої моделі у якій в нейронній мережі один вихідний нейрон відповідає за поняття при подачі на вхід різних спектральних даних, зокрема відео та аудіо образів. Схематично, досліджену модель аудіо-візуального сприйняття нейронною мережею можна представити рисунком 4. Аналітична складова реалізована за рахунок обробки одним обчислювальним елементом інформації, що поступає з різних джерел. За рахунок використання одного обчислювального елемента для визначення поняття з різних джерел вдалося скоротити штучну нейронну мережу вдвічі, а також потенційно підвищити її точність.

Такий підхід дає змогу оптимально використати обчислювальні ресурси завдяки зменшенню об'ємів використовуваної пам'яті та обчислювальних елементів. Забезпечуючи реалізацію єдиної БЗ для усіх потреб система також зменшує кількість звернень та може бути масштабованою на осіб з вадами слуху і хорошим зором. У випадку масштабування запропонований підхід дозволяє (на відміну від інших) встановити зв'язок між особами з вадами голосу (німими) і особами з вадами зору (сліпим). Оскільки дозволяє (елемент навчання) навчити поняття і передавати його у довільному спектрі не змінюючи БЗ.

Дискретизація проводилася з частотою 44100 значень на секунду.

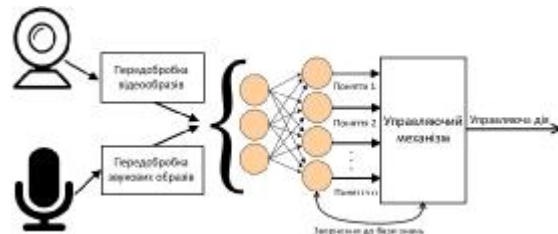


Рис. 4. Модель аудіо-візуального сприйняття

Важливою особливістю розробленої інформаційної технології є автоматичне пристосування до оточуючого середовища і адаптивний пошук корисних об'єктів. Обмеженнями системи є припущення, що мова містить у собі деякі паузи (проміжки тиші між словами), які система вважає роздільниками слів. Для таких систем для виділення корисного сигналу потрібен деякий поріг – значення вище якого є корисним сигналом, нижче – фоном.

У роботі розглянуто кілька варіантів:

- задання константою (є кілька варіантів задання порогу: метод експертних оцінок, метод пристосування);
- аналіз ентропії (ентропія – це міра невизначеності сили перепаду сигналу в рамках заданого фрейму).

Для того, щоб підрахувати ентропію конкретного фрейму, виконуються наступні дії:

- нормуємо сигнал так, щоб всі його значення лягли в діапазоні від -1 до 1;
- будуємо гістограму значень сигналу фрейму;
- розраховуємо ентропію за формулою 1:

$$E = \sum_{i=0}^{N-1} P[i] * \log_2(P[i]) \quad (1)$$

Маючи значення ентропії фрейму для відокремлення корисного сигналу від фону порівнюємо біжуче значення ентропії фрейму з середнім між максимальним і мінімальним значенням ентропії серед усіх фреймів. Оскільки ентропія величина відносно самостійна, то це дає змогу підібрати значення її порогу у вигляді константи.

Ентропія може просідати по середині слова (на голосних), а може раптово зростати через невеликий шум. Для того, щоб вирішувати першу проблему вводиться поняття «мінімальної відстані між словами» та виконується об'єднання сусідніх наборів фреймів, що розділені просіданням. Друга проблема вирішується використанням «мінімальної довжини слова» і відсіканням усіх кандидатів, які не пройшли відбір.

При роботі з фреймами, оптимальною довжиною одного фрейма є 10 мс, накладання 50% (при розбитті вхідних даних по невеликих часових проміжках – фреймах – вони повинні йти не строго одним за одним, а накладатися) і з урахуванням того, що середня довжина слова становить 500 мс – такий крок дає приблизно 100 фреймів на слово:

$$500 / (10 * 0.5) = 100.$$

### Література

1. Software: running commentary for smarter surveillance?. – research\*eu results supplement. - №24. – May 2010. – 29-30p.
2. Шупкина О.В. Модифицированная вероятностная нейронная сеть с комбинированным обучением в задачах обработки текстовой информации / О.В. Шупкина, С.В. Бодянский // Материалы конференции - ISCDMCI'2012. –С. 433 – 435.
3. Параллельная обработка информации: в 5 т. Т.2. Параллельные методы и средства распознавания образов / Под. Ред. А.Н. Свенсона. – Киев: Наукова думка, 1985. – 280 с.
4. Путятин Е.П. Обработка изображений в робототехнике / Путятин Е.П., Аверин С.И. – М.: Машиностроение, 1990. –320с.
5. Методы корреляционного обнаружения объектов / Путятин Е.П., А.В. Гиренко, В.В. Ляшенко, В.П. Машталир, В.Д. Рыжиков. – Харьков: Изд-во АО «Бизнес-Информ», 1996. – 111 с.
6. Путятин Е. П. Методы та алгоритми комп'ютерного зору: навч. посібник / Е.П. Путятин, В.О. Гороховський, О.О. Матат. – Х: СМІТ, 2006. 236 с.
7. Putyatin E., Matat E. Information systems technology. Image processing and pattern recognition. Kharkiv National University of Radio Electronics. Kharkiv, 2003. 105 p.
8. Using sound to picture the world in a new way // Research\*eu Results Magazine, №38 - December 2014 p. 34.
9. Tacticledis plays of the future to feel information // Research\*eu Results Magazine, №38 - December 2014 p. 36.
10. A glimpse into the future of robotic technology// Research\*eu Results Magazine, №28 - December 2013 p. 33.
11. Geoffrey Hinton, Li Deng. Deep Neural Networks for Acoustic Modeling in Speech Recognition. URL: <http://research.google.com/pubs/pub38131.html> (Дата звернення 23.03.2015)
12. Gurdeep Pall. Skype Translator Preview – An Exciting Journey to a New Chapter in Communication // Skype blog. URL: <http://blogs.skype.com/2014/12/15/skype-translator-preview-an-exciting-journey-to-a-new-chapter-in-communication/> (Дата звернення 22.03.2015)
13. Quentin Hardy. Language Translation Tech Startsto Deliver on Its Promise // NYT imes Bitsblog. URL: [http://bits.blogs.nytimes.com/2015/01/11/language-translation-tech-starting-to-deliver-on-its-promise/?\\_r=0](http://bits.blogs.nytimes.com/2015/01/11/language-translation-tech-starting-to-deliver-on-its-promise/?_r=0) (Дата звернення 22.03.2015).
14. "EU study tackles vision-mapping language" // Research\*euResultsMagazine, №19 – February 2013 p.

15. From electronic brains to artificial vision // research\*eu results magazine. - №30. - March 2014. - p.25-26.
16. A project applies neuro science to robot vision. – research\*eu results magazine. - №6. – October 2011. – 9p.
17. Audio-visual answer to modern computing. – research\*eu result supplement. - №26. – July/August 2010. – 31-32p.
18. Kunanets N. Digital Libraries for Disabled Persons / V. Pasichnyk, N. Kunanets, O. Malynovskyi // Intelligent data acquisition and advanced computing systems (IDAACS) : proceedings of the 2013 IEEE 7th International conference, September 12-14, 2013, Berlin, Germany. - Ternopil, 2013. – Vol. 1. – P. 212–215. (каф. НДР). – Berlin, 2013. – P. 241-261. (SCOPUS)
19. Грицик В.В. Модель аудіо-візуального сприйняття. – Intellectual systems for decision making and problems of computational intelligence / В.В. Грицик, А.З. Грондзаль. – МК ISDMCI'2015. – Kherson: KNTU. – 2015. – 51-53p.
20. Volodymyr Hrytsyk, Andrii Grondzal, Andrij Bilenyk. Augmented reality for people with disabilities. – IEEE: Computer science and information technologies. – 14-17 September 2015. – Lviv. – 188-191p.
21. Bucki R., Information Linguistic Systems, Network Integrators Associates, Parkland, Florida, USA, 2007, p. 102, ISBN 9780978860646.
22. Bucki R., Thorough Analysis of the Technological Case Control, Management & Informatics, Network Integrators Associates, Parkland, Florida, Volume 1, No. 1, 2007, pp. 68-112, ISSN 1939-4187.

Рецензія/Peer review : 8.9.2015 р. Надрукована/Printed :3.11.2015 р.  
Стаття рецензована редакційною колегією

УДК 621.382

Л.Б. ЛІЩИНСЬКА

Вінницький торговельно-економічний інститут  
Київського національного торговельно-економічного університету

## КЛАСИФІКАЦІЯ ІНФОРМАЦІЙНИХ ПРИСТРОЇВ НА УЗАГАЛЬНЕНИХ ПЕРЕТВОРЮВАЧАХ ІМІТАНСУ

*Розглянута класифікація інформаційних пристроїв комп'ютерних та інформаційно-вимірювальних систем, які будуються на основі однокристальних багатопараметричних узагальнених перетворювачів імітансу, надана їх характеристика.*

*Ключові слова: інформаційна система, інформаційний пристрій, узагальнений перетворювач імітансу, імітансні логічні елементи, оптоімітансні логічні елементи, радіочастотні датчики, частотно-вибіркові пристрої, напіпровідникові множувачі індуктивності.*

L.B. LISHCHYNSKA

Vinnitsa Trade and Economic Institute of Kyiv National Trade and Economic University

## CLASSIFICATION OF INFORMATIVE DEVICES BASED ON GENERALIZED CONVERTER IMMITANCE

*Abstract - Classification of informative devices of computer and information systems which is built on the basis of the one-chip multi-parameters generalized immitance were considered, their description is given.*

*The conducted researches had showed wide functional possibilities of one-chip multi-parameters generalized immitance, perspective of their application at the construction of the radio frequency informative systems.*

*This classification had showed possibility of realization of high-efficiency managers of elements, frequency selective built on, radio frequency sensors, semiconductor multipliers of inductance, radio frequency, immitance and opto-immitance logical elements.*

*Key words: information system, information device, generalized immitance converter, radio frequency sensors, semiconductor multipliers of inductance, radio frequency, immitance and opto-immitance logical elements.*

### Постановка проблеми

Прогрес розвитку сучасного суспільства у значній мірі визначається успіхами у розвитку обчислювальної техніки, основою побудови якої є кодування та обробка інформації відеоімпульсними сигналами. Такі обчислювальні системи мають низку суттєвих переваг, зокрема високу надійність і швидкодію, відносно невисоку вартість тощо. Але основою глобальних інформаційних систем є радіочастотні канали зв'язку, об'єднання яких з відеоімпульсними системами здійснюється шляхом використання прийнятно-передавальної апаратури, аналогово-цифрових і цифро-аналогових перетворювачів, це може привести до часткової втрати інформації і погіршення швидкодії таких систем, що є особливо актуальним при створенні інформаційних систем передачі, прийому та обробки інформації щодо рухливих об'єктів спеціального призначення.

Одним із шляхів вирішення цієї проблеми є використання інформаційних систем, побудованих на сучасних інформаційних пристроях, що здійснюють безпосередню обробку сигналів на носійній частоті. Використання таких інформаційних систем і пристроїв дозволяє підвищити завадостійкість систем передачі та обробки інформації, значно знизити необхідний енергетичний рівень інформаційного сигналу, зменшити