



БОЮН

Віталій Петрович — академік НАН України, завідувач відділу інтелектуальних відеосистем реального часу Інституту кібернетики ім. В.М. Глушкова НАН України

ВІДЕОСИСТЕМИ РЕАЛЬНОГО ЧАСУ З ЕЛЕМЕНТАМИ ШТУЧНОГО ІНТЕЛЕКТУ

Стенограма доповіді на засіданні Президії НАН України 13 жовтня 2021 року

У доповіді наведено найвагоміші результати робіт у галузі штучного інтелекту та його застосування для відеосистем реального часу як одного з пріоритетних напрямів діяльності Інституту кібернетики ім. В.М. Глушкова НАН України. Відзначено актуальність і високий рівень наукових досліджень, пов'язаних з розробленням теоретичних основ динамічної теорії інформації як бази для створення систем реального часу, інтелектуальних відеосистем з використанням принципів функціонування сітківки ока людини, а також важливість їх впровадження в галузі національної безпеки і оборони, у промисловості, робототехніці, медицині, на транспорті.

Вельмишановний Анатолію Глібовичу!

Шановні члени Президії! Шановні присутні!

Одним з важливих напрямів наукових досліджень Інституту кібернетики ім. В.М. Глушкова НАН України є розроблення інтелектуальних відеосистем реального часу для різних галузей застосування. Цей напрям є важливою складовою кібернетичної науки, зокрема розвитку кібернетичної техніки. На відміну від традиційних відеокамер, які сьогодні випускаються сотнями фірм у світі і розраховані на широке коло користувачів, ми орієнтуємо свої відеокамери на автоматичну роботу в системах реального часу, до яких належать системи комп'ютерного зору, робототехнічні комплекси, системи віртуальної реальності, системи військового й оборонного призначення, деякі системи медико-біологічного застосування тощо.

Сучасні традиційні відеокамери видають за секунду до гігабайта інформації, для оброблення якої вже зараз необхідні терафлопні комп'ютери. Для вирішення цієї проблеми світ обрав шлях створення все більш досконалих технологій, зокрема в галузі мікроелектроніки. Ми ж сповідуємо інший, інтелектуальний підхід з елементами штучного інтелекту, що уподібнює його до зорового аналізатора людини.

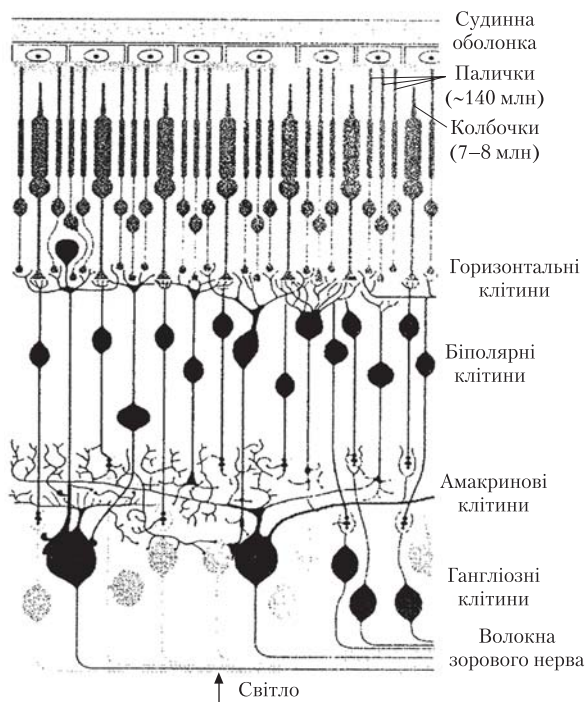


Рис. 1. Схема нейронних структур сітківки ока людини та їх взаємозв'язок (важливо, що фоторецептори обернені не в бік джерела світла, а «дивляться» вгору, на судинну оболонку)

Зорова система людини легко справляється з такими великими потоками інформації завдяки своїй надзвичайно високій вибірковості, обробляючи змінну інформацію за допомогою фоторецепторів сітківки. Сітківка ока складається з шару фоторецепторів (близько 150 млн паличок і колбочок) та багатозарової мережі нейронів (100 млрд нейронів, 100 трлн зв'язків) з різними функціями (рис. 1). Палички, розміщені переважно на периферії сітківки, забезпечують сприйняття інтенсивності світла у відтінках сірого. Колбочки, більш щільно розташовані в центральній ямці (зона фовеа), забезпечують сприйняття світла у трьох діапазонах довжин хвиль, які інтерпретуються мозком як три базові кольори (R, G, B). Виходи шару паличок і колбочок дають набір окремих точок (~150 млн) різної інтенсивності і кольору, що є надлишковим і надзвичайно неінформативним для мозку.

Для вирішення цієї проблеми зоровий аналізатор має дві нейромережі. Перша знаходиться в сітківці ока і складається з кількох шарів нейронів — горизонтальні, біполярні, амакринові, гангліозні клітини тощо, які забезпечують селективний відбір інформації, на багато порядків зменшуючи її надлишковість. Друга нейромережа знаходиться в зоровій корі мозку і виконує функції розпізнавання об'єктів, аналізу сцен та керує сприйняттям інформації на сітківці. При цьому задіяно висхідні процеси, які передають інформацію від рецепторів, та низхідні, або процеси концептуалізації інформаційних даних, які ґрунтуються на отриманих раніше знаннях, попередньому досвіді, осмисленні та очікуванні. Ці процеси доповнюють один одного. Відома фраза «око дивиться, а мозок бачить» досить точно характеризує процес сприйняття, тобто осмислення стимуляції наших сенсорних рецепторів.

Можна виокремити такі основні принципи організації функціонування сітківки ока: розріджена периферійна сітківка і щільна центральна ямка забезпечують широке поле огляду з різною роздільною здатністю; вони організовані за принципом «центр—оточення», тобто центр — збуджуючий, оточення — гальмівне або навпаки. Висока пластичність нейронів завдяки зміні розмірів і форми цих центрів, кількості гальмівних шарів тощо дозволяє виділяти із зображення велику кількість різних інформативних ознак з різними масштабами для подальшого розпізнавання в зоровій корі. Крім того, сітківка має численні адаптаційні механізми для пристосування до умов сприйняття зображення з огляду на конкретні потреби.

Отже, для автоматичного виконання зорових функцій в інтелектуальній відеосистемі має бути закладено елементи інтелекту, подібні до властивостей зорового аналізатора людини, тобто такі відеосистеми повинні забезпечувати:

- ефективність роботи в різних режимах сприйняття інформації, а саме, виділяти корисну інформацію для кожного режиму;
- широке поле огляду з невисокою і керованою роздільною здатністю для ефективного

пошуку об'єкта та можливістю детального його представлення для розпізнавання об'єктів;

- значне зменшення обсягів інформації завдяки виділенню багатьох інформативних ознак з використанням спеціальних методів обробки інформації;
- реалізацію висхідних і низхідних процесів для керування процесом сприйняття відеоінформації;
- можливість адаптації відеокамери до умов використання для спрощення навчання штучної нейронмережі.

Оскільки для високодинамічних процесів і швидкісних рухомих об'єктів традиційні методи дискретизації сигналів, квантування та кодування, а також принципи обробки інформації в реальному часі не задовольняли вимог щодо запізнювання інформації в контурі зворотного зв'язку та продуктивності, в попередні роки було розроблено теоретичні основи кібернетичної техніки, зокрема динамічну теорію інформації, яка є базовою для систем реального часу. Мірою динамічної інформації є δ -ентропія як середнє значення похідної сигналу по модулю. Вона дозволяє виділяти і використовувати корисну (динамічну) інформацію з сигналів, ітераційних процесів, зображень, відео, просторових полів та значно зменшує обсяги інформації для оброблення. Отримані результати було узагальнено в 3 основоположних монографіях*.

Ідея динамічної інформації полягає у використанні замість повнорозрядних відліків сигналу, для отримання яких необхідно n тактів (n — розрядність), що приводило до збільшення кроку дискретизації та запізнювання на 1,5 кроку, тільки змін (приростів) сигналу, які кодуються величинами, кратними степеню 2, і визначаються за один такт. Це дозволило за-

* Малиновский Б.Н., Боюн В.П., Козлов Л.Г., Соловьев В.П. *Введение в кибернетическую технику: Обработка физической информации*. Киев: Наук. думка, 1979; Малиновский Б.Н., Боюн В.П., Козлов Л.Г. *Введение в кибернетическую технику. Параллельные структуры и методы*. Киев: Наук. думка, 1989; Боюн В.П. *Динамическая теория информации: Основы и приложения*. Киев: Наук. думка, 2001.

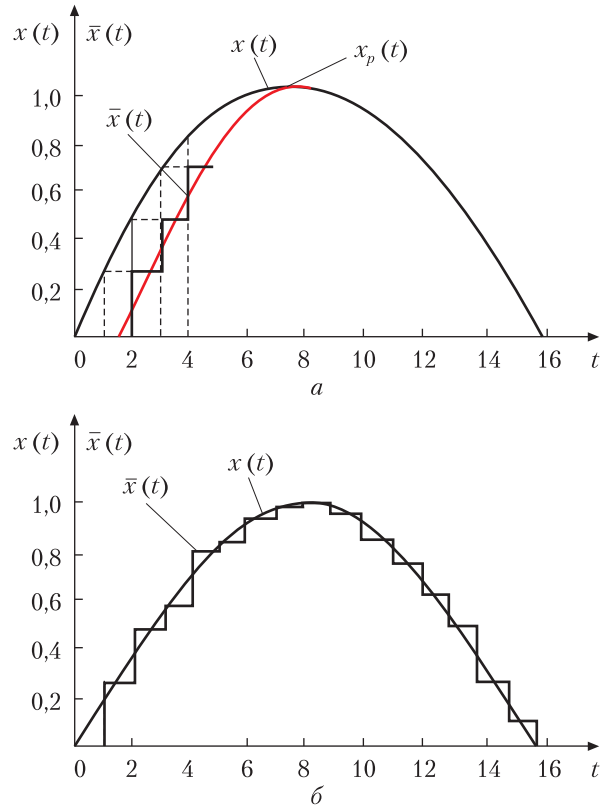


Рис. 2. Вимірювальна модель подання сигналів (а), яка вносить запізнювання на 1,5 кроку дискретизації, і слідкувальна модель (б) з компенсацією запізнювання шляхом збільшення порогів у 1,5 раза

мінити операцію множення операцією зсуву і додавання, що спростило обробку сигналів та на порядок підвищило частотний діапазон оброблюваних сигналів (рис. 2).

На цьому етапі було отримано такі основні результати:

— створено методи і алгоритми аналого-інкрементного перетворення й обробки інформації в реальному часі з мінімальним запізнюванням у контурі зворотного зв'язку, в яких операцію множення замінено операцією зсуву з додаванням, що дозволяє на 1–2 порядки підвищити ефективність використання обладнання;

— розроблено структури пристроїв цифрового функціонального перетворення, цифрової фільтрації, кореляційного та спектрального

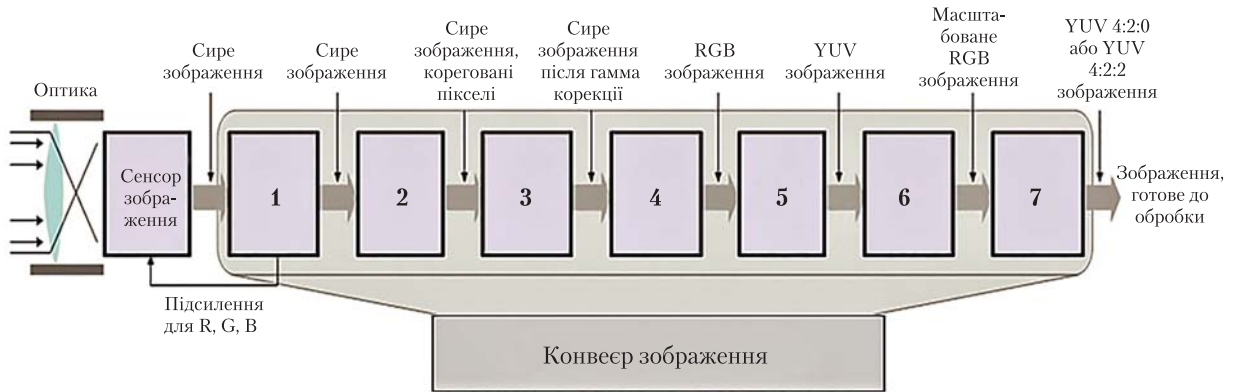


Рис. 3. Конвеєр введення і попередньої підготовки зображення: 1 – автоматичне підсилення, експозиція, баланс білого; 2 – коригування затемнення лінзи, дефектних пікселів; 3 – гамма-корекція; 4 – визначення відсутніх компонентів кольору за допомогою інтерполяції; 5 – перетворення кольорового зображення з формату RGB у формат YUV; 6 – масштабування, децимація або проріджування; 7 – кольорова субдискретизація

аналізу, розв’язання систем рівнянь та ін., на які отримано більш як 200 патентів;

– впроваджено понад 20 процесорів, комплексів і систем керування високодинамічними і складними процесами, зокрема 3 системи керування положенням і параметрами плазми в термоядерних установках типу ТОКАМАК.

У подальшому динамічну теорію інформації було допрацьовано і поширено на процеси сприйняття та обробки зображень і відеопослідовностей.

Обсяг інформації, що міститься у зображенні, визначається такою формулою:

$$C = \frac{x}{\Delta x} \cdot \frac{y}{\Delta y} \cdot \log_2 \left(\frac{z}{\delta z} + 1 \right) \cdot \frac{1}{\Delta t},$$

де: x і y – розміри поля зображення; z – координата яскравості (колірності) зображення; Δx , Δy , δz , Δt – дискретність подання відповідних координат зображення. Маємо, отже, три складові – кількість пікселів у зображенні, їх розрядність і частота кадрів. Звідси випливають шляхи скорочення надлишковості: змінення розмірів зчитуваного зображення; просторове загрублення зображення; зменшення розрядності в поданні інформації; змінення частоти кадрів.

У традиційних відеокамерах всі ці параметри в процесі роботи не змінюються, але якщо

забезпечити таку можливість, можна, керуючи параметрами зчитування інформації, в різних режимах виділяти тільки корисну інформацію, значно зменшивши її надлишковість.

Сучасний фотосенсор – це не лише матриця фотоелементів. Він має всередині до десятка спеціалізованих процесорів та близько двох сотень регістрів для налаштування його роботи і забезпечує на виході видачу інформації в стандартних форматах (JPEG, MPEG та ін.) для передачі або запам’ятовування. Схему конвеєра введення і попередньої підготовки зображення наведено на рис. 3. Традиційний конвеєр містить 7–8 спецпроцесорів та 250 регістрів для керування параметрами зчитування інформації з сенсора і орієнтований на забезпечення високої якості зображення та його комфортне сприйняття людиною.

Отже, системи реального часу працюють в автоматичному режимі і потребують забезпечення необхідного темпу сприйняття й обробки інформації та мінімальної її затримки в контурі зворотного зв’язку.

Далі ми побудували динамічні моделі різних процесів, таких як панорамування, тобто зчитування тільки вертикальної або тільки горизонтальної смуги зображення, відновлення динамічного зображення з панорамного зображення в довільному темпі та за довільною траєкторією, використання панорамного

зображення для автоматичного пошуку змін між оборотами відеокамери, наприклад у системі відеоспостереження на блокпостах, виділення динамічної інформації під час стеження за об'єктом. При цьому досягнуто зменшення обсягів інформації для оброблення на 1–3 порядки та спрощення конструкції відповідних пристроїв.

Реалізація динамічних моделей на традиційних засобах відеокомп'ютерної техніки не може продемонструвати їх основні переваги. Тому ми розробили структуру нової керуючої відеосистеми реального часу, яка складається з інтелектуальної відеокамери — відеосенсора і процесора цифрової обробки зображень з багатьма каналами зворотного зв'язку для керування параметрами зчитування інформації, і має зв'язок з оператором та об'єктом. Суміщення процесів введення й обробки зображення дозволяє керувати параметрами зчитування з сенсора інформації наступного кадру з затримкою всього на кілька рядків, що особливо важливо для систем зі зворотними зв'язками.

У межах виконання Державної програми «Образний комп'ютер» було створено перше в Україні сімейство інтелектуальних відеокамер ІВК-1 (рис. 4), які побудовані на нових інформаційних засадах і вирізняються автоматичним виділенням, введенням динамічної частини зображення та його обробкою в реальному часі, системою керування параметрами зчитування відеоінформації, можливістю адаптації до зовнішніх умов, наявністю механізмів уваги та стеження за об'єктом тощо. Ці інтелектуальні відеокамери використано в низці розробок Інституту, зокрема у відеосистемах для визначення структурних змін в об'єктах, системі ідентифікації та контролю якості продукції за ознаками форми, розмірів, колірності, системі визначення концентрації біооб'єктів для потреб мікробіології.

Розроблено також метод визначення типу об'єкта за контурними ознаками, який дозволяє ідентифікувати об'єкт навіть в умовах завад, які можуть частково спотворювати його контур. З використанням цього методу створено пристрій стеження за рухомим об'єктом,



Рис. 4. Сімейство інтелектуальних відеокамер ІВК-1

коли оператор наводить курсор на довільну ділянку обраного об'єкта, а далі система наведення продовжує стежити за цією ділянкою до слушного для пострілу моменту. Цей пристрій уже передано для впровадження на підприємстві Укроборонпрому — ДП НВК «Фотоприлад» (м. Черкаси).

Крім того, розроблено цифровий оптичний капілярскоп, який дає змогу проводити дослідження статичних і динамічних параметрів мікроциркуляції крові людини. На базі цього цифрового оптичного капілярскопа та вдосконаленого нами доплерографа створено гемодинамічну лабораторію МакроМікроПотік для комплексного дослідження функціонального стану серцево-судинної системи на мікро- і макрорівнях.

Багаторічні мультидисциплінарні дослідження, спрямовані на вивчення структури і принципів функціонування сітківки ока людини, методів організації сприйняття, обробки інформації та зв'язків між нейронами, породили багато нових ідей для використання їх у перспективних системах комп'ютерного зору.

Так, для аналізу зображень у сітківці ока використовуються просторові частоти, для чого в її складі є спеціальні нейрони. Ми з цією метою використовуємо δ -ентропію по рядках і стовпчиках зображення, яка обчислюється

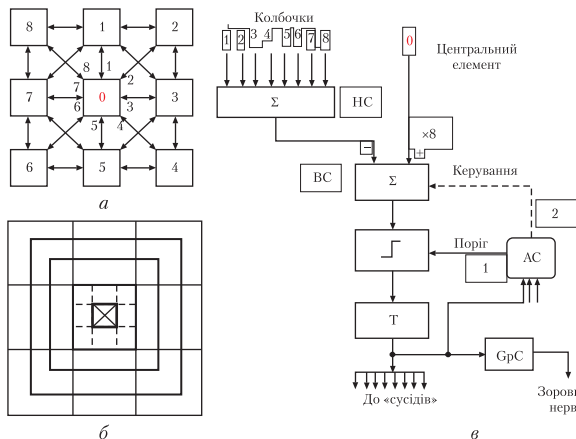


Рис. 5. Організація нейронів центральної ямки сітківки ока людини з нарощуванням кілець для виділення ознак: *а* – кільцева організація нейронів у центральній ямці (*on*-центр); елемент 0 у центрі – збуджувальний, елементи 1–8 навколо – гальмівні; *б* – нарощування кілець навколо центрального елемента для підвищення чутливості до сприйняття контрасту; *в* – реалізація *on*-центру розміром 3×3; центральна ямка сітківки організована на колбочках, горизонтальних (НС), біполярних (ВС) і гангліозних клітинах Р-типу (GrC) за кільцевим принципом; горизонтальні клітини є гальмівними, зворотні зв'язки через амакринові (АС) клітини керують сприйняттям контрасту внаслідок зміни порогу або нарощування шарів нейронів навколо центрального елемента

набагато простіше порівняно з фур'є-аналізом. У виокремлених текстурах δ -ентропія дозволяє визначити регулярність текстури, розміри зерен, контраст, орієнтацію, дефекти регулярності або колірності. Зараз такі системи комп'ютерного зору застосовують для мобільного відеонагляду за дорожньою і придорожною обстановкою, у системах розпізнавання номерів автомобілів, що рухаються, в системах розпізнавання складних символів тощо.

У сітківці ока людини дуже важливим є забезпечення принципу кільцевої організації рецептивних полів, оскільки він однаково ефективно працює як на моделі периферичної сітківки, так і на моделі центральної ямки. За принципом центр–оточення на моделі сітківки можна виділити безліч інформативних ознак різного рівня – плями або дрібні деталі, які відрізняються від фону за яскравістю, кольором,

орієнтацією, динамічними характеристиками, наявністю руху у відеопослідовності тощо. Ці ознаки забезпечують як пошук об'єктів у сцені, так і детальне їх розпізнавання.

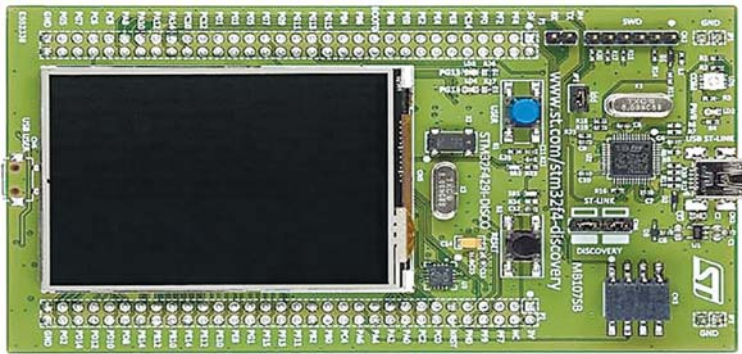
На рис. 5 показано кільцеву організацію нейронів центральної ямки, нарощування шарів та їх реалізацію. Адаптація до контрасту також здійснюється за принципом кільцевої організації. При високому рівні освітлення можливе детальне дослідження об'єкта за допомогою колбочок центральної ямки, організованих за кільцевим принципом. При цьому на вибір порогу впливає величина контрасту. Високий рівень контрасту при реалізації *on*- і *off*-центрів для виділення інформативних ознак вказує на необхідність збільшення порогу бінаризації, тоді як невисокий рівень контрасту потребує зменшення порогу. Прикладом реалізації *on*-центру може бути маска Лапласа розміром 3×3.

Загрублення зображення на периферичній ділянці сітківки дозволяє виділяти більш габаритні ознаки і значно зменшує обсяг обчислень. Сумація сигналів з паличок та їх кільцева організація дає змогу підсилити чутливість в умовах недостатнього освітлення в обмін на зменшення роздільної здатності.

На основі зазначених вище властивостей було розроблено кілька типів відеокамер різної продуктивності і вартості для різних застосувань (на базі мікроконтролера, мікрокомп'ютера, програмовних логічних інтегральних схем) (рис. 6). Деякі з них використовуються в інноваційному проекті «Система мобільного відеонагляду дорожньої і придорожньої обстановки», який реалізується у співпраці з ТОВ «Магнітприлад».

Розроблено також низку методів визначення аномальних і конфліктних ситуацій у натовпі.

З метою значного підвищення продуктивності (на 3–4 порядки) розроблено низку багат шарових структур, які поєднують паралельне зчитування, аналого-цифрове перетворення і обробку інформації. Виходи рядків і стовпчиків матриці можна використовувати для підключення додаткових пристроїв, призначених, наприклад, для обчислення низки



а



б



в

Рис. 6. Інтелектуальні відеокамери нового покоління: а – на базі мікроконтролера; б – на базі мікрокомп'ютера; в – на базі програмовних логічних інтегральних схем

інваріантних моментів, визначення місцеположення об'єктів тощо.

За результатами цих робіт отримано чотири патенти на винаходи, проте в Україні, на жаль, немає технологій для їх практичної реалізації.

Отже, використання динамічної теорії інформації, ідей та методів, що ґрунтуються на механізмах зорового аналізатора людини, і запропонованих способів їх реалізації дозволило створити низку інтелектуальних відеокамер і систем різного призначення, розробити узагальнену динамічну модель процесів пошуку об'єктів у зображенні, стеження за ними у відеопослідовності, класифікації об'єктів та їх розпізнавання. Ці підходи дають змогу на 2–3 порядки підвищити продуктивність та ефективність систем комп'ютерного зору і розширити їх функціональні можливості.

Уже понад 10 років Інститут має тісні зв'язки з Китайською Народною Республікою в рамках Союзу науково-технічного співробітництва з країнами СНД. Інтерес до наших розробок великий, проте, на жаль, обмеження з китайського боку щодо роботи з державними установами не дозволяють укласти прямі договори. Однак нещодавно було підписано Угоду про співробітництво з ТОВ «Сянчу Китайсько-українські ядерно-енергетичні технології», яка передбачає розроблення інтелектуального обладнання для робототехніки ядерного класу з метою подовження терміну експлуатації АЕС, їх безпечної роботи та виведення з експлуатації ядерних об'єктів.

Результати наших робіт неодноразово доповідалися на засіданнях РНБО України, на круглих столах в Міністерстві оборони Укра-

їни тощо, ухвалювалися протокольні рішення про доцільність продовження таких робіт, але на заваді їх виконанню стають заборони на фінансування дослідно-конструкторських робіт. Головна проблема полягає в забезпеченні можливості тиражування і впровадження роз-

робок, що потребує додаткових людських і фінансових ресурсів.

Дякую за увагу!

*За матеріалами засідання
підготувала О.О. Мележик*

Vitaliy P. Boyun

V.M. Glushkov Institute of Cybernetics of the National Academy of Sciences of Ukraine, Kyiv, Ukraine

ORCID: <https://orcid.org/0000-0002-3932-3558>

REAL-TIME VIDEO SYSTEMS WITH ELEMENTS OF ARTIFICIAL INTELLIGENCE

Transcript of the report at the meeting of the Presidium of NAS of Ukraine, October 13, 2021

The report presents the most important results of work in the field of artificial intelligence and its application for real-time video systems as one of the priority scientific fields of the V.M. Glushkov Institute of Cybernetics of the National Academy of Sciences of Ukraine. The relevance and high level of research, which relate to the development of theoretical foundations of dynamic information theory as a basis for creating real-time systems, intelligent video systems using the principles of human retina functioning, is noted as well as the importance of their implementation in national security and defense, industry, robotics, medicine, transport.

Keywords: real-time video system, human visual analyzer, dynamic information theory, intelligent video system.