



ПРИНЦИПИ ОРГАНІЗАЦІЇ СІТКІВКИ ОКА ЛЮДИНИ ТА ЇХНЄ ВИКОРИСТАННЯ В СИСТЕМАХ КОМП'ЮТЕРНОГО ЗОРУ

Анотація. Наведено короткий огляд принципів організації сітківки ока людини. Розглянуто принципи локальності під час взаємодії нейронів, кільцевої організації рецептивних полів з on- і off-центрами, спеціалізації шарів нейронів, організації зворотних зв'язків, адаптації до рівня освітлення і контрасту, стиснення даних у відеопотоці. Показано, що досконала організація сітківки ока людини дає змогу значно поліпшити технічні характеристики систем комп'ютерного зору в разі застосування сітківки як прототипу. Результати досліджень використано у процесі створення сім'ї інтелектуальних відеокамер і низки систем на їхній основі, а також під час побудови спеціалізованих нейронних мереж для первинного оброблення інформації безпосередньо на сенсорній матриці.

Ключові слова: сітківка; палички і колбочки; горизонтальні, біполярні, амакринові і гангліозні клітини; on- і off-центри; нейромережа; відеосенсор; керування параметрами зчитування інформації; інтелектуальні відеокамери; багатшарові матричні структури.

ВСТУП

Відеосистеми реального часу відіграють значну роль у системах автоматизації виробничих процесів, візуального контролю якості продукції, робототехніці, системах оборонного і військового призначення, системах автоматизації наукових та медико-біологічних досліджень тощо. Діапазон застосування цих систем та вимоги до них постійно розширюються. Це, зокрема, стосується відеосистем зі зворотним зв'язком, у яких результати оброблення інформації в реальному часі використовуються для керування процесом або інших дій. Такі системи мають задовольняти підвищені вимоги не лише до продуктивності обчислювальних засобів, але й до запізнювання інформації в контурі зворотного зв'язку, які не забезпечуються в межах традиційних підходів.

Зорова система людини вдосконалювалася впродовж мільйонів років і досягла надзвичайно високого рівня організації. Узагальнена модель зорової системи людини є багатофункціональною та складається з кількох десятків, а то й сотні, локальних моделей, які описують цілу низку структурних, фізичних, геометричних і психофізичних механізмів та процесів. Процес сприйняття візуальної інформації людиною є динамічним, з багатьма змінюваними параметрами і зворотними зв'язками. Ми не тільки дивимось, ми бачимо та реагуємо, тобто цей процес є активним. Тому феномен зору людини надає надзвичайно велику кількість різнопланових елегантних рішень для реалізації систем комп'ютерного зору.

Оскільки матеріал статті знаходиться на стику декількох спеціальностей, для кращого його розуміння надамо короткий опис організації сітківки ока людини, який стосується лише процесів сприйняття зображень та оброблення відеоінформації на шарах нейронів [1, 2].

ОРГАНІЗАЦІЯ СІТКІВКИ ЗОРОВОЇ СИСТЕМИ ЛЮДИНИ

Сітківка охоплює близько 200° внутрішньої поверхні очного яблука та являє собою надзвичайно складну мережу фоторецепторів і нервових клітин, які трансформують світлову енергію в нейрону активність. У багатьох аспектах поверхня сітківки є продовженням мозку. Сітківка складається з фоторецепторів двох видів — паличок і колбочок відповідно циліндричної та конічної форми. На периферії сітківки розміщено 120–130 мільйонів паличок, ширина яких становить близько 0,002 мм. Колбочки у кількості 6–8 мільйонів зосереджені переважно у центральній ямці (зона fovea) шириною близько 1 мм і мають діаметр 0,003–0,008 мм. На невеликій відстані від центральної ямки у бік носа розташований диск зорового нерва, з якого виходить зоровий нерв у мозок. На цій ділянці немає фоторецепторів (сліпа пляма) і відсутня будь-яка зорова реакція.

Зовнішня оболонка паличок і колбочок вкрита світлопоглинальним пігментом, який забезпечує умови їхнього збудження під час потрапляння на них світла. Сітківка являє собою складну десятишарову мережу нервових клітин різної спеціалізації з вертикальними і горизонтальними зв'язками.

Базовими елементами нервової системи людини (основними обробниками і передавачами інформації) є нейрони або нервові клітини, кількість яких становить близько 100 мільярдів. Як відомо, існують нейрони різних типів, при цьому кожний виконує специфічні функції. Деякі з цих нейронів пов'язані з рецепторними клітинами органів чуття (сенсорні нейрони), які сприймають певні види енергії (світло, звук, тиск, хімічну енергію). Ці клітини перетворюють сприйняту ними енергію в нервові імпульси та передають їх іншим нейронам нервової системи — мотонейронам для передавання імпульсів м'язам або інтернейронам для подальшого оброблення сприйнятої інформації. Більшість нейронів, незалежно від їхніх функцій, складається з трьох основних частин:

- тіла клітини, або соми, яка приймає та зберігає отриману інформацію, при цьому ядро клітини регулює хімічну активність нейрона;
- дендритів — розгалужених відростків, які приймають інформацію від інших нейронів;
- аксонів — довгих волокон, якими інформація від соми передається іншим нейронам.

Таким чином, дендрити сприймають інформацію від багатьох нейронів і передають її в тіло клітини. Сома залежно від сукупності прийнятої інформації (імпульси збудження і гальмування) та її спеціалізації видає через свої аксони відповідні сигнали іншим нейронам або м'язам для виконання відповідних дій. Розгалужений кінець аксона закінчується синаптичною головою, яка є місцем контакту з дендритом іншого нейрона. Залежно від типу сигналу, що надійшов аксоном (збудження або гальмування), у синаптичний контакт виділяється особлива хімічна речовина — нейротрансмітер, який збуджує або гальмує пов'язаний з цим аксоном нейрон. Завдяки використанню природою двох типів нейротрансмітерів забезпечується більш висока точність та скоординованість рухів м'язів.

Групи паличок і колбочок зв'язані по вертикалі з проміжними (біполярними) клітинами, які, в свою чергу, пов'язані з гангліозними клітинами, аксони яких входять до складу зорового нерва. Крім зв'язків з проміжними біполярними та гангліозними клітинами існують також два шари горизонтальних зв'язків: перша мережа створюється горизонтальними клітинами між фоторецепторами і біполярними клітинами, друга мережа бокових (латеральних) зв'язків формується амакриновими клітинами між біполярними і гангліозними клітинами. Однією з функцій цих бокових зв'язків є побудова такої взаємодії між сусідніми клітинами, за якої є можливою модифікація сигналів з фоторецепторів (латеральне гальмування).

Нервові клітини не обов'язково передають імпульси сусіднім нейронам щоразу, коли на них діє подразник. Нейрон має багато дендритів (до тисячі), пов'язаних з багатьма іншими нейронами. Ці дендрити доставляють не лише ту інформацію, яка збуджує відповідний нейрон, але й інформацію, що гальмує його роботу. По-

тенціал дії нейрона і передавання від нього імпульсу виникає тоді, коли досягається визначений поріг його стимулювання. Значення потенціалу дії не залежить від інтенсивності подразника, тобто є постійною величиною. Вплив інтенсивності подразника проявляється в кількості потенціалів дії та у часовому інтервалі між ними, тобто в частоті імпульсів потенціалу дії. Ця частота не може перевищувати 1000 імпульсів дії на секунду. Відчуття постійного за інтенсивністю подразника з часом зменшується, тобто відбувається адаптація до тривалого подразника. Хоча адаптація зазвичай свідчить про зниження чутливості, водночас вона приносить і значну користь, допомагаючи сприймати більш слабкі сигнали про зміни у довкіллі. Фізичним подразником зорової системи є світло: з одного боку, це електромагнітні коливання у вузькому діапазоні довжин хвиль (приблизно 380–760 нм для людини), а з іншого — це енергія, тобто неперервний потік частинок (квантів) енергії, який визначає інтенсивність світла. Психологічним ефектом від сприйняття електромагнітних хвиль різної довжини є кольори або відтінки (від фіолетового до червоного), а ефектом від інтенсивності світла, як фізичного подразника, є яскравість. Змінення інтенсивності світла не обов'язково призводить до пропорційної зміни сенсорного сприйняття, тобто яскравості.

Структурні (конструктивні) особливості організації сітківки є такими:

- використання різних типів рецепторів з різними фізичними характеристиками для сприйняття світла (ахроматичний і колірний зір);
- розміщення рецепторів (переважно паличок і частково колбочок) у периферійній зоні для забезпечення широкого поля зору;
- щільне розміщення колбочок у зоні ямки (фовеа) для забезпечення гостроти зору;
- наявність кільцевої зони відносно осі ока зі специфічними алгоритмами оброблення відеоінформації;
- розміщення рецепторів на сферичній поверхні сітківки для компенсації нелінійних оптичних викривлень;
- інвертованість сітківки (розміщення рецепторів поза шарами нейронів);
- сприйняття світла боковою поверхнею рецепторів під гострим кутом;
- об'єднання багатьох рецепторів (паличок) сітківки у периферійній зоні на одній нервовій клітині для сумачії сигналів та підвищення чутливості (за рахунок зменшення просторової роздільної здатності);
- наявність спеціалізованих нейронів для передачі, оброблення сигналів та керування м'язами (біполярні, гангліозні, амакринові нейрони, мотонейрони тощо);
- наявність спеціалізованих зв'язків між нейронами сітківки для виділення геометричних ознак (лінія, кут нахилу до горизонту, кути між лініями, кривизна лінії, симетрія тощо).

Фізичні особливості сприйняття світла сітківкою є такими:

- розсіювання світла компонентами нейронів та перетворення його в шумоподібне зображення, що є найбільш ефективним способом адаптації до неоднорідності середовища. Що більш випадковим є зображення, то більшу кількість інформації воно містить. Це забезпечує краще пристосування до умов середовища, надає видам у процесі еволюції більше шансів на виживання. Наявність тремороподібних рухів у зоровому апараті в поєднанні з можливою часовою когерентністю світла дає змогу виділяти з рандомізованого зображення регулярні інтерференційні структури, за якими надалі можна відновити вхідне зображення. До того ж, завдяки цьому можна знизити диференційні рівні яскравості на фоторецепторах у результаті її розподілу за більшою площею фоторецепторів [3, 4];
- лінійне відносно простору і спектра перетворення світла на електричні сигнали;
- логарифмічне сприйняття яскравості;
- підвищена чутливість паличок для забезпечення ахроматичного «нічного» зору;
- різні спектральні характеристики колбочок, які забезпечують сприйняття кольорів;

- залежність сприйняття кольору від довжини хвилі світла, яке стимулює зорову систему. Лише випромінювання видимої області електромагнітного спектра з довжиною хвиль від 380 до 760 нм здатне викликати у людини колірні відчуття. Іншими словами, відчуття кольору — це суб'єктивний результат дії на нервову систему відбитого від предмета променя з певною довжиною хвилі у видимому діапазоні. Промені світла як такі не мають кольору, вони лише є результатом відбиття або поглинання променевої енергії у довкіллі, який інтерпретується зоровою системою. Навіть у мережі сітківки і зорової кори кольорів немає; вони з'являються тоді, коли відеоінформація отримує кінцеву інтерпретацію у свідомості спостерігача. Інакше кажучи, колір — це продукт діяльності зорової системи, а не невід'ємна властивість видимого спектру;

- між відчуттям кольору і фізичними параметрами світла існує тісна залежність, яку можна кількісно оцінити за трьома атрибутами світла: довжиною хвилі, інтенсивністю та спектральною чистотою. Цим фізичним параметрам відповідають свої психологічні аспекти відчуття кольору: колірний тон, яскравість і насиченість. Колірний тон, або просто колір, визначається довжиною хвилі світла, яскравість кольору визначається його інтенсивністю, а насиченість кольору є психологічним параметром, який відображає відносну кількість кольору поверхні предмета і пов'язаний з фізичним параметром спектральної чистоти.

Механізми і процеси, що відбуваються в сітківці, є такими:

- яскравісна адаптація до рівня освітлення (10 порядків: від порогу чутливості нічного зору, який забезпечується за рахунок збудження більш чутливих паличок сітківки, до порогу засліплювального блиску, який обмежується колбочками колірного зору сітківки). Суб'єктивна яскравість, яка сприймається зоровою системою людини, являє собою логарифмічну функцію від фізичної яскравості світла, що потрапило в око. Однак діапазон рівнів яскравості, які одночасно сприймаються оком, становить величину близько трьох порядків;

- збудження нейронів сітківки внаслідок появи у сфокусованому на ній зображенні перепадів яскравості або колірності (механізм уваги) використовується для швидкого пошуку в зображенні інформаційних ознак (сакади);

- рефлекторні тремороподібні рухи (мікрорухи) під час фіксації погляду на об'єкті, які забезпечують постійне збудження рецептивного поля; інакше образ починає розпливатися та зникає;

- адаптація до тривалого подразника, яка допомагає сприймати більш слабкі сигнали про зміни у навколишньому середовищі;

- блокування фону та виділення змін у зображенні.

Значна частина попереднього оброблення зорової інформації відбувається вже на рівні сітківки. Центральна ямка і периферія сітківки організовані за кільцевим принципом. Однак у центральній ямці кожна колбочка або невеликі їхні групи мають вихід на гангліозну клітину, а на периферійній сітківці палички об'єднуються в групи, підсумовуючи сигнали з більших ділянок рецептивного поля і забезпечуючи підвищення чутливості у разі недостатнього освітлення (в обмін на зменшення просторової роздільності) [1–3].

Центральна ямка спеціалізована на чіткий зір, а периферія сітківки — на високу чутливість. Горизонтальні та дифузні клітини є гальмівними на рівні сітківки. Амакринові та інтерплексіформні клітини відповідають за адаптацію розмірів рецептивних полів. Клітини центральної ямки та сітківки організовані за кільцевим принципом (on- і off-центри).

З огляду на досконалість і ефективність організації сітківки ока людини набуває актуальності задача побудови відеосистем комп'ютерного зору високої продуктивності з використанням зазначених принципів.

Метою роботи є підвищення продуктивності та ефективності систем комп'ютерного зору за рахунок використання принципів організації сітківки ока людини.

ОБҐРУНТУВАННЯ КОМП'ЮТЕРНОЇ МОДЕЛІ СІТКІВКИ

У роботі [3] запропоновано комп'ютерну модель сітківки на базі стільникових нейронних мереж гексагонального типу, однак вони не мають значних переваг порівняно з прямокутною організацією, крім того, вони не реалізуються на сучасних сенсорних матрицях. Частіше сітківку розглядають як структуру з логарифмічно-полярною організацією [4–7], у центрі якої розміщена центральна ямка (зона fovea), а навколо — периферійна сітківка. Це забезпечує широкий периферійний огляд і можливість детального сприйняття інформації в центральній ямці.

Було здійснено спроби відтворити таку фовеальну організацію візуального сприйняття в технічних та алгоритмічних моделях. Одним з напрямків є створення фовеальних сенсорів з радіальною та ієрархічною організацією рецептивного поля [8]. Недоліком радіальної організації є необхідність керування «поглядом», тобто напрямом оптичної осі сенсора, що потребує використання швидкісного приводу та системи керування ним. Ідеологія ієрархічного пірамідального сприйняття полягає в тому, що спочатку зчитується зображення з максимальною роздільною здатністю, а потім шляхом згладжування та проріджування (для цього найчастіше використовують піраміди Лапласа або Гаусса) формується наступний шар піраміди, який має у чотири рази менший розмір зображення. Ця процедура повторюється декілька разів до досягнення потрібного рівня загублення. Спочатку такий підхід був реалізований для стиснення і передавання зображення, при цьому зображення можна було відновити з попередньою високою роздільною здатністю. Такий підхід є ефективним в умовах обмеженої пропускної здатності каналів зв'язку. Згодом він був використаний у створенні пірамідальних систем машинного зору і моделюванні механізмів уваги [9]. Ідея полягала в тому, що на загубленому зображенні, побудованому на базі пірамідального сприйняття, значно простіше виявити об'єкт. Застосована після цього зворотна процедура дає змогу підвищити роздільну здатність подання об'єкта для його розпізнавання, вимірювання тощо. Проте такий підхід зовсім непридатний для оброблення інформації у відеосистемах реального часу. Крім того, він не відповідає принципам організації сітківки ока людини.

У роботах [10, 11] запропоновано протилежний до розглянутого вище підхід на базі механізму уваги зорового аналізатора, коли спочатку зчитується грубе зображення сцени, на якому за різними ознаками відшукується заданий об'єкт та оцінюються його габаритні розміри. Потім шляхом керування параметрами зчитування інформації з відеосенсора зчитується цей об'єкт (як фрагмент сцени) з високою роздільною здатністю для його розпізнавання, вимірювання або стеження за ним. Такий підхід добре узгоджується з можливостями технології «комплементарна структура метал-оксид-напівпровідник» (КМОН-технології) виробництва сучасних відеосенсорів, які мають декартову систему координат і забезпечують лінійне і просте сприйняття простору.

Враховуючи зазначене, в цій роботі запропоновано використання прямокутної (декартової) системи координат моделі сітківки ока людини.

Аналіз механізмів зорового аналізатора людини як прототипу для побудови систем комп'ютерного зору свідчить про необхідність організації цілої низки систем адаптації, регулювання та керування, які мають працювати в реальному часі. Для цього потрібно забезпечити можливість змінення параметрів зчитування інформації з відеосенсора та оброблення інформації в реальному часі. Обсяг інформації у відеопослідовності зазвичай оцінюють на основі амплітудно-просторової та часової роздільності:

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

де X і Y — розміри поля зображення; Z — координата яскравості зображення; $1/\Delta t = f$ — частота кадрів; Δx , Δy , δz , Δt — кроки дискретизації відповідних координат зображення.

Значення X , Y і Z у формулі зазвичай прийняті фіксованими і максимальними, фіксованими є також значення Δx , Δy , δz , Δt . Тому такий підхід дає оцінку зверху кількості інформації, що є дуже завищеною і не вказує шляхів скорочення надлишковості цифрового представлення зображень.

Можливість динамічної зміни параметрів у наведеній формулі дає змогу відеокамері адаптуватися до вимог і особливостей задачі, а також значно зменшити надлишковість подання зображень і відеопослідовностей. На базі цього підходу розроблено і досліджено динамічні моделі процесів панорамування та монтажу відеопослідовностей, автоматичного пошуку змін у системах кругового та секторного огляду, виділення змін (руху, кольору, розмірів) об'єктів та стеження за ними.

Нові інформаційні основи цифрового подання зображень і відеопослідовностей, які базуються на розробленій в Інституті кібернетики імені В.М. Глушкова динамічній теорії інформації [12, 13], дають змогу, виходячи з прикладної задачі, виділяти відповідну найбільш значущу та корисну (динамічну) інформацію. Наприклад, можна виділяти та зчитувати визначене місце в кадрі, змінювати його розміри та задавати потрібну роздільну здатність — просторову, часову та колірну, використовувати тільки одну колірну складову, адаптивно змінювати частоту кадрів залежно від динаміки процесів тощо. При цьому зменшення первинної надлишковості інформації з відеокамери доцільно здійснювати на сенсорному рівні, забезпечивши можливість програмування зазначених вище параметрів. КМОН-технологія виробництва сучасних сенсорів, яка забезпечує кожний елемент матриці пристроєм вибирання-зберігання і комутаційними засобами, що уможливають реалізацію різних законів зчитування зображення, практично повною мірою задовольняє ці вимоги. До того ж, сучасні відеосенсори містять у своєму складі до десятка спеціалізованих процесорів, організованих за конвеєрним принципом, призначених для первинного технологічного оброблення інформації, а також декілька сотень регістрів для керування процесами сприйняття і оброблення інформації. Враховуючи, що у разі достатнього рівня освітлення працює центральна ямка сітківки, яка забезпечує детальне сприйняття і гострий зір, а у разі недостатнього рівня освітлення працює периферійна сітківка, яка забезпечує грубе стосовно простору сприйняття інформації, можна вважати, що вони працюють не одночасно, а послідовно у часі. Використання цих можливостей розглянуто в наступному розділі.

ВИКОРИСТАННЯ ПРИНЦИПІВ ОРГАНІЗАЦІЇ СІТКІВКИ ОКА ЛЮДИНИ В СИСТЕМАХ КОМП'ЮТЕРНОГО ЗОРУ

Основна ідея роботи полягає не в точному копіюванні принципів організації сітківки ока людини, а в забезпеченні можливості використання найбільш важливих її принципів для створення досконаліших систем комп'ютерного зору.

Основні принципи організації сітківки, розглянуті і використані в роботі, є такими:

- широке поле зору;
- широкий динамічний діапазон сприйняття яскравості;
- гострота зору;
- адаптація до рівня освітлення і контрасту;
- робота в умовах недостатнього освітлення;
- суттєве стиснення інформації у відеопотоці.

Широке поле зору забезпечується, перш за все, шляхом вибору відповідної оптики та узгодженням її характеристик з розміром відеосенсора. Широкий динамічний діапазон забезпечується нелінійним (зокрема логарифмічним) сприйняттям яскравості, можливостями збільшення рецептивних полів з сумациєю сигналів з них, кільцевою організацією нейронів з можливістю нарощування кілець нейронів навколо центрального збуджувального елемента. Гострота зору забезпечується за рахунок зчитування зображення з максимальною фізичною роздільною здатністю та обчислення значень колірності в кожному пікселі за значеннями колірних складових сусідніх пікселів, а також додатково (за потреби) шляхом

інтерполяції значень яскравості для отримання субпіксельної точності. Як-то кажуть, око дивиться, мозок бачить. А бачить він за рахунок сприйняття перепадів яскравості, тобто контрасту. Контраст сприйняття забезпечується завдяки кільцевій організації нейронів сітківки і, додатково, шляхом нарощування кілець нейронів навколо центрального збуджувального елемента. Таким чином, гострота зору посилюється не лише завдяки збільшенню просторової роздільності, а й за рахунок підвищення контрасту сприйняття зображення. Адаптація до рівня освітлення і контрасту забезпечується шляхом керування (за допомогою амакринових клітин) нарощуванням кілець нейронів навколо центрального збуджувального елемента, зміненням коефіцієнта підсилювання маски кілець або зміненням порогу активації нейрона. Робота в умовах недостатнього освітлення забезпечується завдяки збільшенню рецептивних полів периферійної сітківки, кільцевій організації нейронів та нарощуванню їхніх кілець. Стиснення інформації у відеопотоці забезпечується виділенням інформативних ознак із зображення (точки, лінії, краю, контуру, плями тощо) за допомогою відповідних масок.

Горизонтальні клітини сітківки мають тисячі дендритів, але навряд чи вони забезпечують виділення якоїсь однієї ознаки. Вони, напевно, використовуються для організації виділення різних інформативних ознак. Цьому сприяють адаптаційні механізми, які, управляючи розмірами і формою рецептивних полів (сумациєю сигналів з колбочок і паличок та кільцевою організацією нейронів з нарощуванням кілець навколо нейронів), забезпечують можливість виділення більш глобальних і локальних ознак із зображення. Ці принципи будуть розглянуті далі більш детально.

Паралельний характер отримання зображення сенсорною матрицею погано узгоджується з подальшими послідовними процесами аналого-цифрового перетворення та оброблення інформації, зумовлює потребу в детальному програмуванні цих процесів. Хоча під час реалізації паралельного оброблення має місце значна надлишковість у представленні інформації, однак за рахунок однорідного оброблення інформації, зменшення пересилань та запам'ятовування, а також завдяки відсутності детального програмування послідовних процесів є можливим значне підвищення продуктивності та ефективності оброблення.

Нейромережева організація обчислень є надзвичайно ефективною. Однак традиційні нейромережеві обчислення, незважаючи на існування десятків варіантів для розв'язання різних задач, є надзвичайно складними для апаратної реалізації і потребують складного налаштування на задачу [14]. Як прототип запропоновано використовувати методи організації сприйняття, оброблення інформації та зв'язків між нейронами сітківки ока людини [1, 2].

На організацію сприйняття відеоінформації значною мірою впливають рівень освітлення сцени та величина контрасту. Характеристикою рівня освітлення є середнє значення яскравості, а контраст визначається як середнє значення модулів приростів яскравості за модулем (δ -ентропія) [11, 12]. Цією інформацією можна скористатися для попереднього налаштування параметрів сприйняття відеоінформації. За умови високого рівня освітлення є можливим детальне дослідження об'єкта за допомогою колбочок центральної ямки, організованих за кільцевим принципом. При цьому на вибір порогу впливає величина контрасту. Високий рівень контрасту під час реалізації on- і off-центрів для виділення інформативних ознак зумовлює потребу в збільшенні порогу бінаризації, тоді як невисокий рівень контрасту потребує зменшення порога бінаризації. Прикладом реалізації on-центра може бути маска Лапласа [15] розміром 3×3

$$\text{виду 1 } \begin{bmatrix} & -1 & \\ -1 & 4 & -1 \\ & -1 & \end{bmatrix} \text{ або виду 2 } \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

У масці для реалізації off-центра знаки при коефіцієнтах змінюються на протилежні.

Зовсім малий контраст для забезпечення виділення інформативної ознаки потребує нарощування гальмівних кілець навколо центрального збуджувального елемента і відповідного підвищення його ваги, тобто здійснення переходу від однокільцевої маски 3×3 до двокільцевих масок 5×5 :

0	0	-1	0	0
0	-1	-2	-1	0
-1	-2	16	-2	-1
0	-1	-2	-1	0
0	0	-1	0	0

-1	-1	-1	-1	-1
-1	-2	-2	-2	0
-1	-2	32	-2	-1
-1	-2	-2	-2	-1
-1	-1	-1	-1	-1

При цьому ступінь підсилення для масок Лапласа змінюється від 4, 8 для масок 3×3 до 16 та 32 для масок 5×5 відповідно.

Центральна ямка сітківки організована на колбочках, горизонтальних (англ. HC — horizontal cells), біполярних (англ. BC — bipolar cells) і гангліозних клітинах Р-типу (англ. GrC — P-type ganglion cells) за кільцевим принципом (on- і off-центри) (рис. 1). Горизонтальні клітини є гальмівними. Зворотні зв'язки за допомогою амакринових (англ. AC — amacrine cells) клітин керують сприйняттям контрасту шляхом зміни порогу або нарощуванням шарів нейронів навколо центрального нейрона.

Така організація нейромережі добре узгоджується з напрацьованим арсеналом методів виділення різних інформативних ознак за допомогою масок Лапласа, Превіта, Робертса та ін. [15], завдяки чому можна ефективно підкреслювати контури, виділяти ознаки краю, виявляти інформативні точки, лінії, їхню орієнтацію, обчислювати градієнти тощо. Розглянута кільцева організація нейронів центральної ямки забезпечує підвищення контрасту, який збільшується з нарощуванням шарів нейронів навколо центрального елемента. Організація зв'язків між нейронами є досить універсальною і забезпечує можливість реалізації на такій структурі згорток з матрицями 3×3 , 5×5 , ... за рахунок збільшення часу обчислення. Усі обчислення на структурі виконуються пара-

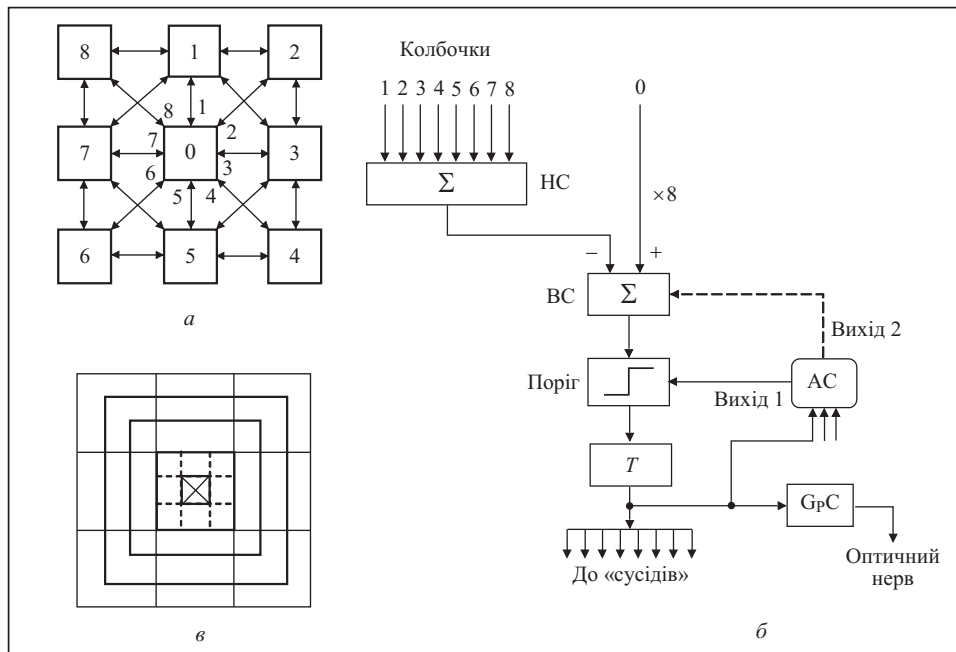


Рис. 1. Горизонтальні та біполярні клітини з нарощуванням кілець: *a* — кільцева організація нейронів у центральній ямці (on-центр), елемент 0 у центрі — збуджувальний, елементи 1–8 навколо — гальмівні; *b* — реалізація on-центра маскою Лапласа 3×3 ; *в* — нарощування кілець навколо центрального елемента для підвищення чутливості до сприйняття контрасту

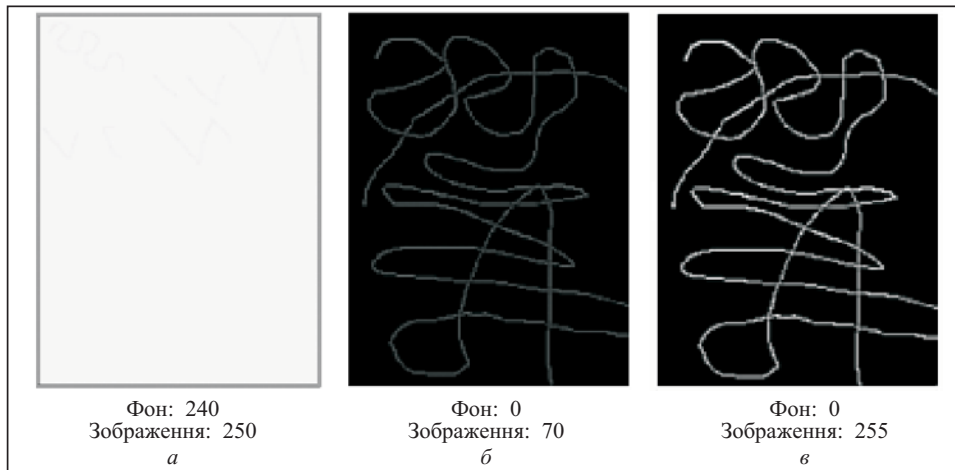


Рис. 2. Підсилення контрасту на колбочках центральної ямки за рахунок кільцевої організації нейронів: вхідне зображення (*a*); результат застосування маски 3×3 (*б*); результат застосування маски 5×5 (*в*)

тельно. Слід також зауважити, що обчислення сум з виходів кільцевих нейронів та обчислення різниці між збуджувальними та гальмівними нейронами може здійснюватися послідовним кодом, що дає змогу реалізувати на такій схемі не тільки бітові коди після порогового обмеження, але і повнорозрядні коди. При цьому час обчислень збільшується практично пропорційно розрядності повнорозрядних кодів. Коефіцієнти матриць майже завжди можна прийняти числами, пропорційними степеню двійки (0, 1, 2, 4, 8, ...), що виключає необхідність виконання операції множення.

Для дослідження принципів виділення інформативних ознак із зображення (точка, лінія, край, контур тощо) і наочного представлення відповідних результатів розроблено спеціальний інструментарій, за допомогою якого можна задавати різні об'єкти (зокрема, кольорові) та побачити результати використання різних масок як у цифровому вигляді, так і в графічному та тривимірному просторі.

Результати підсилення контрасту у випадку використання кільцевої організації нейронів з одним і двома кільцями навколо центрального елемента наведено на рис. 2.

Корекцію параметрів сприйняття контрасту, як і для функцій амакринових клітин (яких дослідники нараховують до 35 видів), запропоновано здійснювати шляхом оцінювання кількості колбочок (підключених до збуджувального входу біполярної клітини) на визначеній ділянці рецептивного поля, що здолали поріг дискретизації. За їхньої відсутності або невеликої кількості поріг знижується, поки не досягне заданого рівня, наприклад 10 % загальної кількості колбочок цього рецептивного поля (див. рис. 1, вихід 1). Більш радикально діє вихід 2 амакринової клітини, який керує нарощуванням кілець навколо центрального елемента.

У разі використання повнорозрядних кодів, тобто без порогового елемента, амакринові клітини, керуючи коефіцієнтом підсилення маски і нарощуванням кілець навколо збуджувального елемента, можуть підтримувати контраст для змін освітлення в заданому діапазоні, наприклад 128–255 одиниць яскравості.

Розглянемо деякі особливості роботи on- і off-центрів. On-центри забезпечують виділення більш світлих об'єктів на темному фоні (показують внутрішню границю об'єкта), а off-центри, навпаки, виділяють більш темні об'єкти на світлому фоні (показують зовнішню границю об'єкта). Результати роботи on- і off-центрів щодо світлих і темних об'єктів показано на рис. 3.

Для більш глибокого розуміння процесів, які відбуваються під час проходження маски зображенням, моделювання здійснюється у повнорозрядних кодах, тобто без порогового обмеження. Проведені дослідження дали змогу виявити

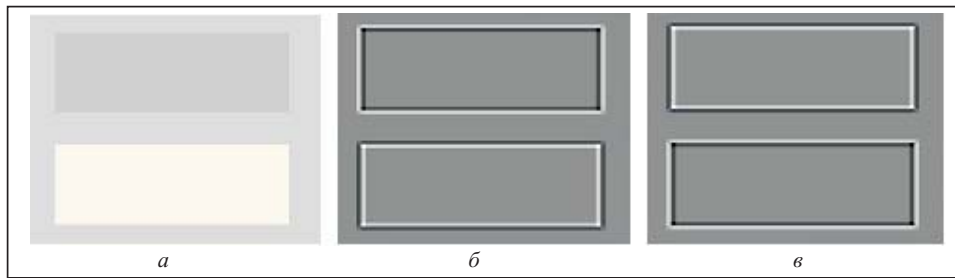


Рис. 3. Результати роботи он- і off-центрів під час виділення світлих (темних) об'єктів на темному (світлому) фоні: вхідні зображення (а); із застосуванням он-центра (б); із застосуванням off-центра (в) (результати подано на шкалі ± 128)

низку цікавих і корисних особливостей оброблення зображень масками. Наприклад, зазвичай для виділення контурів або виділення інформативних ознак здійснюється проходження зображенням однієї маски, наприклад з он- або off-центром. Однак у зображеннях є елементи більш світлого кольору на фоні більш темного і навпаки; а вони виділяються різними масками.

Запропоновано адаптивно використовувати обидва центри (рис. 4, рис. 5). Для цього, виходячи з різниці між збуджувальним входом біполярної клітини та сумою гальмівних входів з горизонтальної клітини, використовується той або інший центр (можливо, цим процесом управляє амакринова або якась інша клітина). Якщо ця різниця додатна, то процес іде по он-центру, в іншому разі — по off-центру. При цьому від'ємні результати використання маски обнуляються. Результатом застосування такого підходу є можливість відображати сцену з більш точними розмірами об'єктів незалежно від яскравості фону та об'єктів.

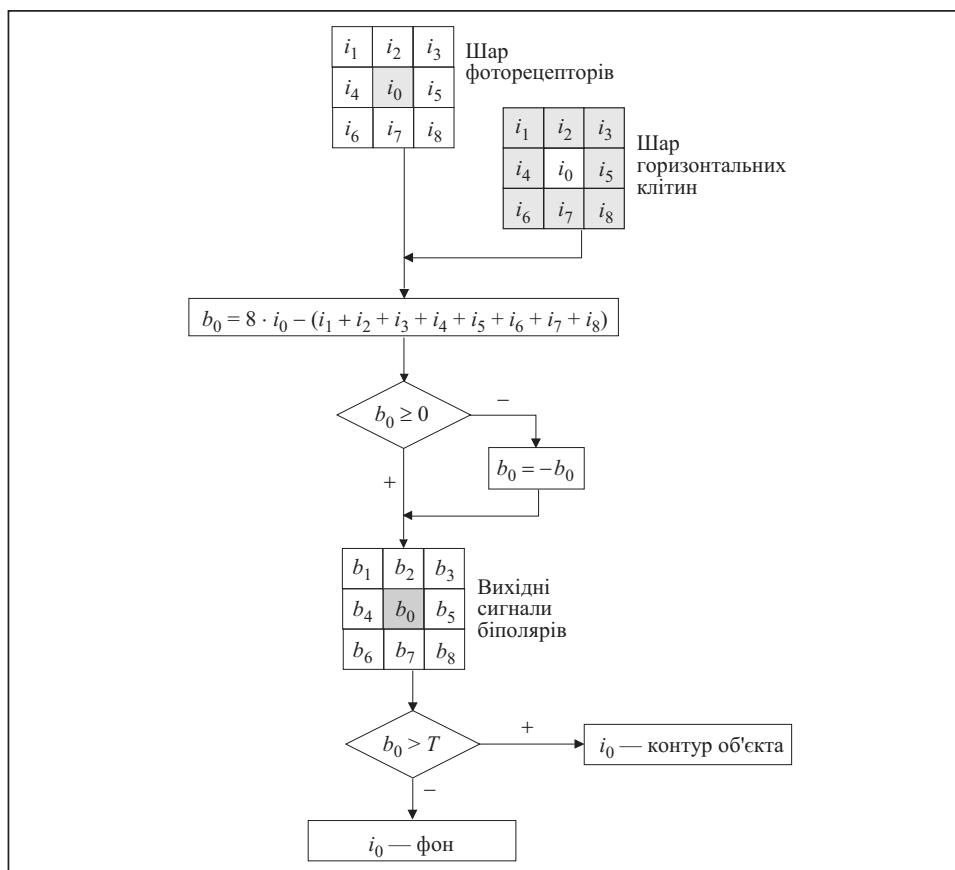


Рис. 4. Блок-схема адаптивного алгоритму виділення контурів

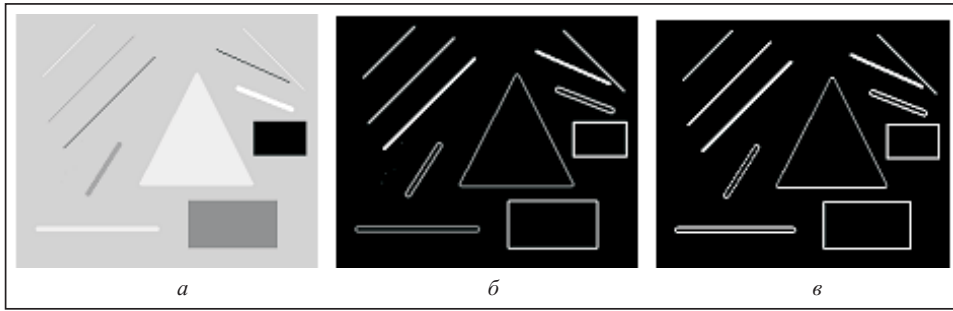


Рис. 5. Адаптивне використання рецептивних полів біполярних клітин з on- і off-центрами: вхідне зображення (а); результат роботи алгоритму (б); контури об'єктів після порогового перетворення (в)

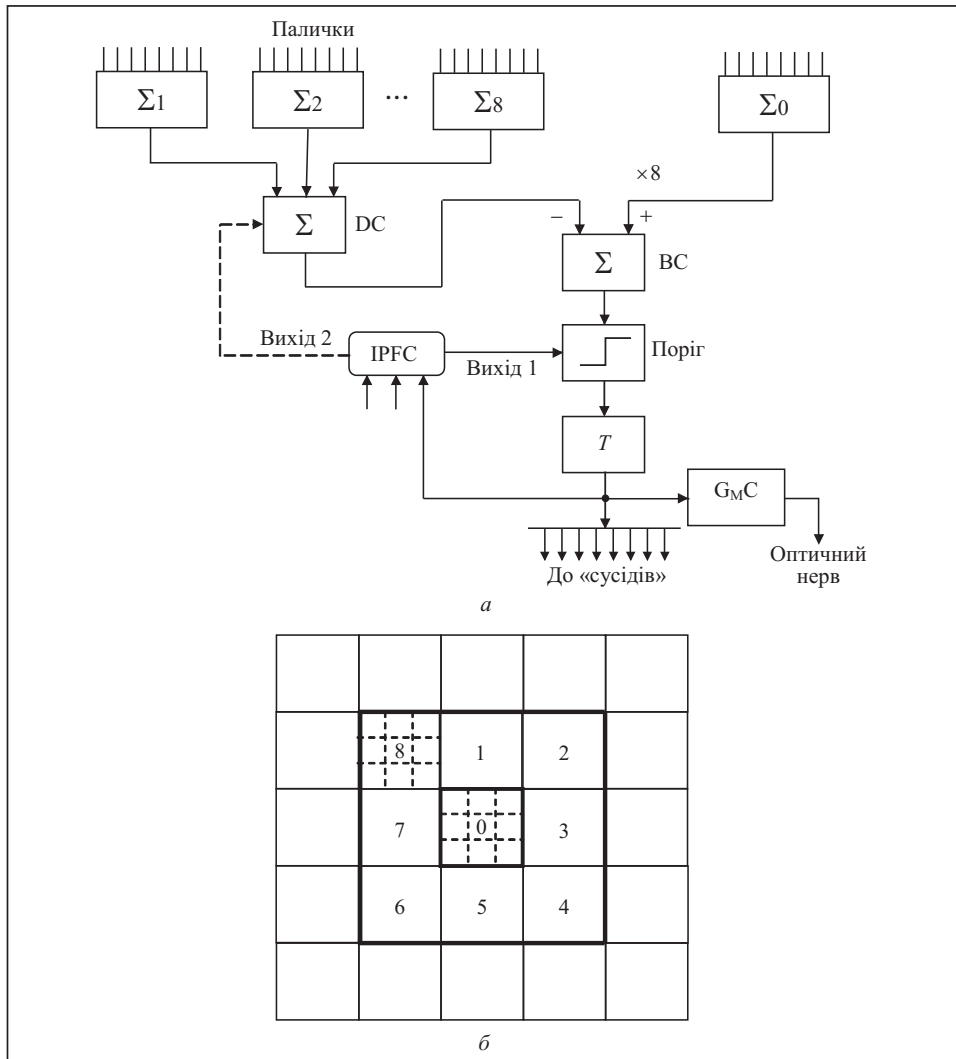


Рис. 6. Дифузні та біполярні клітини з нарощуванням кілець: реалізація кільцевої організації нейронів периферійної сітківки (а); нарощування кілець для підвищення чутливості в умовах недостатнього освітлення (б)

В умовах недостатнього рівня освітлення ефективно працюють палички сітківки за рахунок підсумовування сигналів з більших рецепторних полів, обмінюючи просторову роздільність на підвищення чутливості. Великий контраст у цих умовах не потребує значного розширення рецепторних полів, тобто достатньо ско-

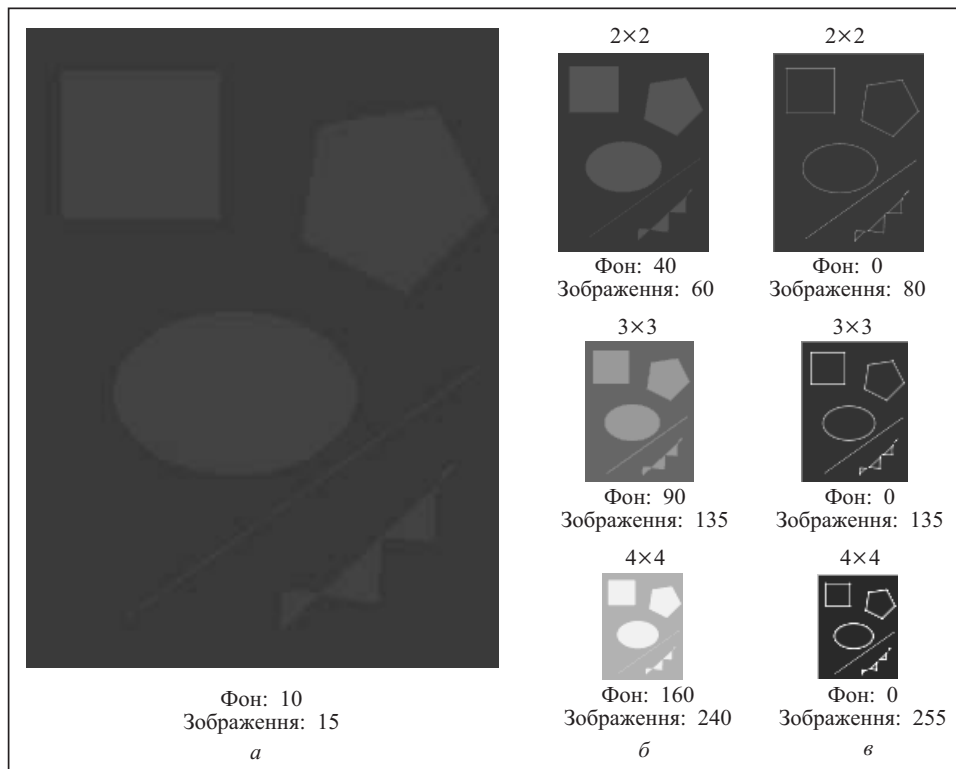


Рис. 7. Підсилення чутливості периферійної сітківки шляхом сумачії сигналів з паличок та їхньої кільцевої організації: вхідне зображення (а); результат підсилення за рахунок сумачії сигналів полів розміром 2×2, 3×3 або 4×4 (б); результат підсилення за рахунок кільцевої організації з маскою Лапласа 3×3 (в)

ристання полями для підсумовування розміром 2×2. Це забезпечує підвищення чутливості сприйняття контрасту з використанням лапласіана 3×3 у $8 \times 4 = 32$ рази. Якщо контраст є малим, потрібно збільшити рецептивні поля для підсумовування до 3×3, 4×4, що забезпечить ступінь підсилення у $8 \times 9 = 72$ та $8 \times 16 = 128$ разів. При цьому розміри зображення зменшуються відповідно у 4, 9 та 16 разів.

Периферія сітківки спеціалізована на високу чутливість і організована на паличках сітківки, дифузних (англ. DC — diffuse cells), біполярних (BC) і гангліозних клітинах М-типу (англ. G_MC — M-type ganglion cells) також за кільцевим принципом (рис. 6). Підвищення чутливості у разі недостатнього освітлення забезпечується завдяки сумачії сигналів від значної кількості паличок і дії гальмівних дифузних клітин. При цьому відповідно зменшується роздільна здатність щодо простору (рис. 7). Зворотні зв'язки за допомогою інтерплексіформних клітин (англ. IPFC — interplexiform cells) керують порогами або розмірами рецептивних полів (рис. 6).

Корекцію параметрів сприйняття зображення в умовах недостатнього освітлення за аналогією до функцій інтерплексіформних клітин запропоновано здійснювати шляхом оцінювання кількості груп підсумованих паличок на певній ділянці рецепторного поля, що здолали поріг дискретизації. За їхньої відсутності або невеликої кількості вихід 2 інтерплексіформної клітини керує нарощуванням кільця навколо центрального елемента.

Описані підходи реалізовано у процесі створення інтелектуальних відеокамер на базі відеосенсорів з конвеєром для попереднього технологічного оброблення зображень і можливістю керування параметрами зчитування інформації та мікроконтролера для подальшого оброблення зображення в реальному часі [13, 16]. Такі інтелектуальні відеокамери покладено в основу створення низки відеосистем для контролю:

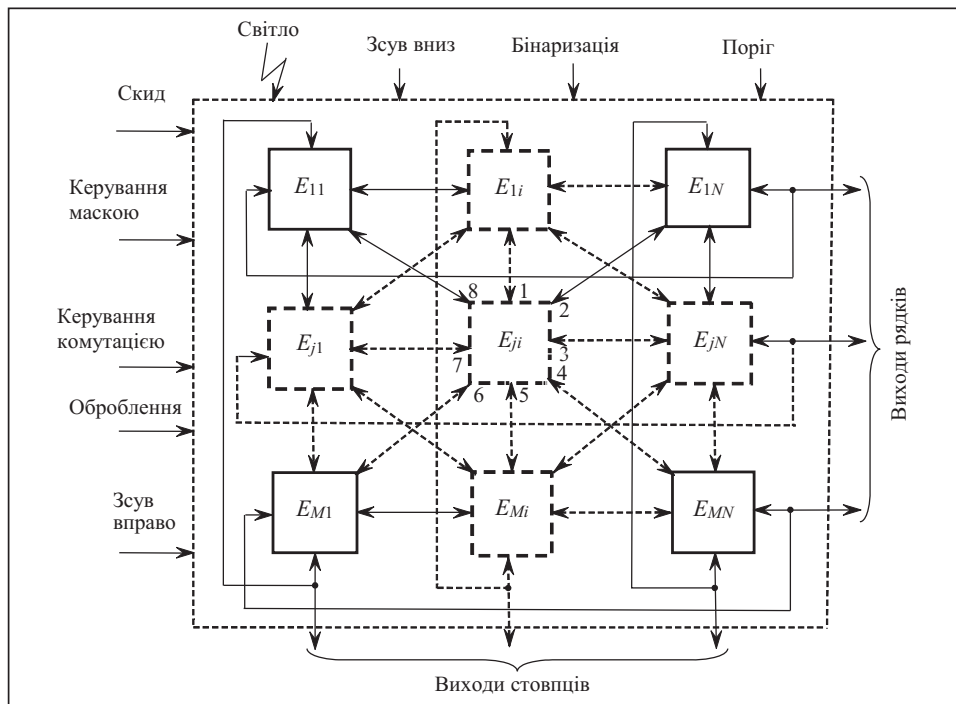


Рис. 8. Структура сенсорної матриці з обробленням зображень

- якості продукції за ознаками кольору, розмірів, форми та текстури;
- параметрів фізичних, хімічних та біологічних об'єктів;
- статичних і динамічних параметрів мікрокапілярів тощо.

До того ж, запропоновано узагальнену структуру неймережі (рис. 8, виходи стовпців та рядків сенсорної матриці, або додаткові шари нейронів, використовуються для обчислення моментів, визначення місцеположення об'єктів тощо), яка після переналаштування на одному або кількох шарах нейронів може працювати в режимі центральної ямки або периферійної сітківки.

Кожний елемент матриці має зв'язки лише з сусідніми елементами матриці (принцип локальності), тобто за аналогією до зорового аналізатора людини може реалізувати on- і off-центри з кільцевою організацією.

Елемент матриці забезпечує сприйняття світла фотодетектором, бінарзацію, виконання логічного оброблення інформації, комутацію входів і виходів елемента матриці та кільцеві зсуви інформації вздовж рядків і стовпців матриці, а також запам'ятовування спочатку бінаризованої вхідної інформації, потім результатів логічного оброблення або результатів зсуву.

Крім того, підключення на виходи рядків і стовпців сенсорної матриці додаткових пристроїв з мікроконтролером для управління процесами дає змогу значно розширити її функціональні можливості. Кільцеві зв'язки в рядках і стовпцях матриці забезпечують не лише зсув інформації в них, але й записування в матрицю проміжних результатів обчислень. Це дає змогу застосовувати матрицю декілька разів для послідовних процесів обчислення. При цьому може бути використаний цілий арсенал напрацьованих методів на базі масок для придушення шумів і покращення якості зображення, виділення країв, підкреслення або виділення контурів, виділення орієнтаційних ознак, обчислення градієнтів. Коефіцієнти масок задаються комплексно (тобто вказується тільки тип маски), і відповідні процедури реалізуються без операції множення.

Середні значення яскравості та контрасту для попереднього налаштування матриці можна обчислити на запропонованій матриці шляхом підключення накопичувальних суматорів на виходах рядків або стовпців. Зсуваючи зображення, на-

приклад, вправо і підсумовуючи значення яскравостей пікселів рядків, отримуємо середні значення за рядками, а потім, зсуваючи їх по вертикалі і накопичуючи в мікроконтролері, отримуємо середнє значення яскравості всього зображення. Для розмірів зображення, кратних степеню двійки, легко реалізується усереднення шляхом зчитування відповідних старших розрядів накопичувальних суматорів. Аналогічно можна обчислити і середнє значення контрасту. При цьому можуть бути реалізовані підходи до адаптації налаштувань шляхом обчислення середніх значень яскравості та контрасту для окремих ділянок зображення.

З використанням зазначених методів розроблено низку спеціалізованих технічних рішень, захищених патентами на винаходи, для поєднання процесів сприйняття відеоінформації з процесом її оброблення безпосередньо на сенсорній матриці. Зокрема, це сенсорні матриці з паралельною бінаризацією зображення і визначенням місцеположення та параметрів об'єкта [17, 18], обчисленням перших і других моментів інерції бінаризованого зображення для рядків, стовпців і всього зображення [19], з морфологічною обробкою бінаризованого зображення [20], з паралельним аналого-цифровим перетворенням та можливістю нелінійного сприйняття яскравості [18].

ВИСНОВКИ

Використання низки принципів організації нейронів зорового аналізатора людини на рівні сітківки дає змогу розробляти ефективні інтелектуальні відеокамери реального часу на базі звичайних відеосенсорів і засобів оброблення інформації, а також багат шарові сенсорні матриці з нейромережу високої продуктивності для відбору і первинного оброблення інформації в системах реального часу безпосередньо на зазначених матрицях.

СПИСОК ЛІТЕРАТУРИ

1. Харви Р. Ш. Ощущение и восприятие. Санкт-Петербург: Питер, 2003. 928 с.
2. Боюн В.П. Зоровий аналізатор людини як прототип для побудови сімейства проблемно-орієнтованих систем технічного зору. Матеріали міжнародної науково-технічної конференції «Искусственный интеллект. Интеллектуальные системы ИИ-2010». Донець: ИПИИ «Наука і освіта», 2010. Т. 1. С. 21–26.
3. Huang C.-H., Lin C.-T. A bio-inspired computer fovea model based on hexagonal-type cellular neural networks. *IEEE Transactions on Circuits and Systems I: Regular Papers*. 2007. Vol. 54, Iss. 1. P. 35–47.
4. Shah S., Levine M.D. Visual information processing in primate cone pathways. I. A model. *IEEE Transactions on Systems, Man, and Cybernetics. Part B: Cybernetics*. 1996. Vol. 26, N 2. P. 259–274.
5. Benoit A., Caplier A., Durette B., Hérault J. Using human visual system modeling for bio-inspired low level image processing. *Computer Vision and Image Understanding*. 2010. Vol. 114, Iss. 7. P. 758–773.
6. Андерсон Д. Когнитивная психология. Санкт-Петербург: Питер, 2002. 496 с.
7. Kolb H. How the retina works: much of the construction of an image takes place in the retina itself through the use of specialized neural circuits. *American Scientist*. 2003. Vol. 91, N 1. P. 28–35.
8. Шелепин Ю.Е., Бондарко В.М., Данилова М.В. Конструкция фовеолы и модель пирамидальной организации зрительной системы. *Сенсорные системы*. 1995. № 1. С. 87–97.
9. Burt P.J. Smart sensing within a pyramid vision machine. *Proceedings of the IEEE*. 1988. Vol. 76, Iss. 8. P. 1006–1015.
10. Боюн В.П. Інтелектуальне вибіркоче сприйняття візуальної інформації. Інформаційні аспекти. *Штучний інтелект*. 2011. № 3. С. 16–24.
11. Boyun V. Intelligent selective perception of visual information in vision systems. *Proc. 6-th IEEE Intern. Conf. on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Application (IDAACS'2011)*. (Czech Republic, Prague, 15–17 September 2011). 2011. Vol. 1. P. 412–416.
12. Боюн В.П. Динамическая теория информации. Основы и приложения. Киев: Институт кибернетики им. В.М. Глушкова, 2001. 326 с.
13. Boyun V. Directions of development of intelligent real time video systems. *Application and Theory of Computer Technology*, [S. l.]. 2017. Vol. 2, N 3. P. 48–66. <https://doi.org/10.22496/atct.v2i3.65>.
14. Руденко О.Г., Бодянский Е.В. Искусственные нейронные сети. Харьков: Компания «СМИТ», 2005. 407 с.
15. Гонсалес Р., Вудс Р. Цифровая обработка изображений. Москва: Техносфера, 2005. 1072 с.

16. Boyun V. Bioinspired approaches to the selection and processing of video information. *Proc. IEEE Second International Conference on Data Stream Mining & Processing (DSMP)*. 2018. P. 498–502. <https://doi.org/10.1109/DSMP.2018.8478541>.
17. Боюн В.П. Пристрій для визначення місцеположення та параметрів об'єкта в зображенні. Патент України № 76597. Опубл. БВ № 6, 10.01.2013.
18. Боюн В.П. Пристрій для визначення місцеположення та центра ваги об'єкта. Патент України № 106292. Опубл. БВ № 12, 11.08.2014.
19. Боюн В.П. Сенсорний пристрій для визначення місцеположення та моментів інерції об'єкта в зображенні. Патент України № 106301. Опубл. БВ № 15, 11.08.2014.
20. Боюн В.П. Сенсорна матриця з обробкою зображень. Патент України № 109335. Опубл. БВ № 6, 10.08.2015.

Надійшла до редакції 03.01.2019

В.П. Боюн, Л.А. Возненко, И.Ф. Малкуш

ПРИНЦИПЫ ОРГАНИЗАЦИИ СЕТЧАТКИ ГЛАЗА ЧЕЛОВЕКА И ИХ ИСПОЛЬЗОВАНИЕ В СИСТЕМАХ КОМПЬЮТЕРНОГО ЗРЕНИЯ

Аннотация. Приведен краткий обзор принципов организации сетчатки глаза человека. Рассмотрены принципы локальности при взаимодействии нейронов, кольцевой организации рецептивных полей с on- и off-центрами, специализации слоев нейронов, организации обратных связей, адаптации к уровню освещения и контрасту, сокращения объема данных в видеопотоке. Показано, что совершенная организация сетчатки глаза человека позволяет значительно улучшить технические характеристики систем компьютерного зрения при использовании сетчатки в качестве прототипа. Результаты работы использованы при создании семейства интеллектуальных видеокамер и ряда систем на их основе, а также при построении специализированных нейросетей для первичной обработки информации непосредственно на сенсорной матрице.

Ключевые слова: сетчатка; палочки и колбочки; горизонтальные, биполярные, амакриновые и ганглиозные клетки; on- и off-центры; нейросеть; видеосенсор; управление параметрами считывания информации; интеллектуальные видеокамеры; многослойные матричные структуры.

V.P. Boyun, L.O. Voznenko, I.F. Malkush

THE PRINCIPLES OF HUMAN RETINA ORGANIZATION AND THEIR USE IN COMPUTER VISION

Abstract. The paper provides a summary of the principles of human retina's organization. The paper explores the principles of: locality in the interaction of neurons; receptive fields ring organization with on- and off-centers (center-surround organization); neuron layers specialization; feedbacks organization; adaptation to light and contrast levels; data reduction in a video stream. The perfect organization of the human retina makes it possible to significantly improve the technical characteristics of computer vision systems when using the retina as a prototype. The results of this research were used in creation of a family of intelligent video cameras and systems based on them, as well as in building a number of specialized neural networks for primary information processing directly on the sensor matrix.

Keywords: retina, rods and cones, horizontal, bipolar, amacrine and ganglion cells, on- and off-centers, neural network, video sensor, information reading parameters control, intelligent video cameras, multilayer matrix structures.

Боюн Віталій Петрович,

чл.-кор. НАН України, доктор техн. наук, професор, завідувач відділу Інституту кібернетики ім. В.М. Глушкова НАН України, Київ, e-mail: vboyun@gmail.com.

Возненко Людмила Олександрівна,

науковий співробітник Інституту кібернетики ім. В.М. Глушкова НАН України, Київ, e-mail: l.voznenko@gmail.com.

Малкуш Ірина Феліксівна,

науковий співробітник Інституту кібернетики ім. В.М. Глушкова НАН України, Київ, e-mail: mif@gmail.com.