

ЗМІННА РОЗДІЛЬНА ЗДАТНІСТЬ У ЗОРОВОМУ СПРИЙНЯТТІ ТА ОБРОБЛЕННІ ЗОБРАЖЕНЬ

*Інститут проблем математичних машин і систем НАН України, м. Київ, Україна

Анотація. Вже існують методи для обробки зображень, що використовують різні значення роздільної здатності, відомі під загальною назвою «coarse-to-fine». Ідея цих методів полягає в тому, що початковий набір даних переглядається з різною роздільною здатністю для виключення невідповідних або нерелевантних областей зображення на ранніх стадіях обробки. Результат, отриманий при «грубій» роздільній здатності, використовується як початкове наближення на наступному кроці при збільшеній роздільній здатності. Тоді розрахунково-складну частину алгоритму буде застосовано на останньому кроці до меншого об'єму даних. Один із варіантів такого підходу – використання модифікації відомого алгоритму, яка реалізує підхід «coarse-to-fine» та придатна для розв'язання конкретного класу задач. У той же час велика кількість задач розпізнавання зображень, які мають складність NP чи взагалі не можуть бути розв'язані за допомогою традиційних методів, розв'язуються у зоровій системі людини миттєво, а задачі, пов'язані з обробкою відеоряду, у режимі реального часу. Зокрема, сегментація зображення відбувається на підсвідомому рівні, дуже швидко та без жодних відчутних зусиль, навіть при значній кількості завдань у полі зору. Тому природним буде звернути увагу на процеси, які супроводжують зорове сприйняття, а саме – на послідовну зміну роздільної здатності зорової системи під час зорового акту, від низької до максимально можливої. У статті запропоновано огляд сучасного стану за напрямками вивчення та застосування змінної роздільної здатності в кожній з означених дисциплін. Краще розуміння природи зорового сприйняття має стати наступним кроком до розроблення нових ефективних методів обробки візуальної інформації в інформаційних системах.

Ключові слова: змінна роздільна здатність, зорове сприйняття, зорова система, візуальна інформація, обробка зображень.

Аннотация. Уже существуют методы для обработки изображений, использующие разные значения разрешающей способности, известные под общим названием «coarse-to-fine». Идея этих методов заключается в том, что исходный набор данных просматривается при разной разрешающей способности для исключения неподходящих или нерелевантных областей изображения на ранних стадиях обработки. Результат, полученный при «грубой» разрешающей способности, используется как начальное приближение на следующем шаге при увеличенной разрешающей способности. Тогда вычислительно-затратная часть алгоритма будет применена на последнем шаге к меньшему объему данных. Один из вариантов такого подхода – использование модификации известного алгоритма, которая реализует подход «coarse-to-fine» и пригодна для решения конкретного класса задач. В то же время большое количество задач распознавания изображений, имеющих сложность NP или вообще не решаемых с помощью традиционных методов, решаются в зрительной системе моментально, а задачи, связанные с обработкой видеоряда, в режиме реального времени. В частности, сегментация изображения происходит на подсознательном уровне, мгновенно и без видимых усилий, даже при значительном количестве помех в поле зрения. Поэтому естественным будет обратить внимание на процессы, которые сопровождают зрительное восприятие, а именно – на последовательное изменение разрешающей способности в зрительной системе во время зрительного акта, от низкой до максимально возможной. В статье предложен обзор современного состояния по направлениям изучения и использования переменной разрешающей способности в каждой из указанных дисциплин. Лучшее понимание природы зрительного восприятия должно стать следующим шагом к разработке новых эффективных методов обработки визуальной информации в информационных системах.

Ключевые слова: переменная разрешающая способность, зрительное восприятие, зрительная система, визуальная информация, обработка изображений.

Abstract. *There exist image processing methods known as «coarse-to-fine» that use different image resolutions. The idea of these methods is lies in the fact that the initial data set considered with different resolution to exclude inappropriate or irrelevant parts of image on earlier stages of processing. Result that obtained for «coarse» resolution is used as initial approximation for processing with subsequent resolution. In that case the computationally-intensive part of algorithm will be applied to reduced amount of data on the final stage of processing. One of the variants of this approach is to use a coarse-to-fine modification of known algorithm, suitable for solving the particular class of problems. At the same time many image recognition tasks that have NP computational complexity or can't be addressed at all with traditional methods are solved by visual system in no moment and tasks of video stream processing are solved in real time. In particular image segmentation happens on subconscious level, in less than no moment, without tangible effort even with large amount of noise in field of view. So it is natural to pay attention to processes that accompany visual perception, namely on sequential resolution changes in visual system during visual act from lowest resolution to highest possible. In this article we propose a review of state of the art in studying and using of variable resolution in domains of scientific activity mentioned above. Better understanding of visual perception nature should be the next step towards developing of new effective methods for visual information processing in information systems.*

Keywords: *multiresolution, visual perception, visual system, visual information, image processing.*

1. Вступ

Під роздільною здатністю зображення (сигналу) розуміють кількість елементів растрового зображення (сигналу) на одиницю площі (одиницю довжини). Значення роздільної здатності, зазвичай, визначають інтуїтивно, зважаючи на задачу, яка має бути вирішена стосовно растрового зображення (сигналу).

Дослідження в нейрофізіології зорового сприйняття та створення інформаційних технологій для обробки візуальної інформації пов'язані областями діяльності людини. Дійсно, об'єктом дослідження для кожної з цих дисциплін є зорове сприйняття. Для фізіології зору це зорове сприйняття людини та тварин, у той час як для комп'ютерних наук це створення та налагодження засобів технічного (комп'ютерного) зору. Прогрес в одній з областей мав би впливати на прогрес в іншій. На жаль, обміну результатами, як такого, не відбувається. Можливо, через слабку взаємодію між спеціалістами, що є представниками різних наук, та через відмінність методів, що зазвичай ними використовуються.

Вперше змінна роздільна здатність (динамічні властивості рецептивних полів нейронів зорової системи [1]) була досліджена у зоровій системі вищих тварин (кішок).

У той же час є досить багато прикладів щодо оброблення зображень (сигналів) з використанням декількох значень роздільних здатностей при обробленні одного й того ж зображення.

Далі детально розглянуто змінну роздільну здатність зорового сприйняття вищих тварин (кішок), що принципово не відрізняється від людського, а також приклади оброблення зображень (сигналів), що використовують декілька значень роздільної здатності.

2. Змінна роздільна здатність у зоровій системі людини

2.1. Будова зорової системи

Зорова система [2, 3] містить в собі такі органи (рис. 1): очі (1), на сітківці (2), де проектується ліві і праві половини полів зору, зорові нерви (3), окорухийні нерви (4), хіазм (5), зоровий тракт (6), бічне колінчасте ядро (7), верхні бугри чотиригорбу (8), неспецифічний зоровий шлях (9), зорова кора головного мозку (10).

У сітківці кожного ока відбувається перетворення оптичних сигналів в електричні, а також первинна обробка візуальної інформації. У хіазмі відбувається поділ нервових шляхів, що зіставляються до лівої й правої частин поля зору кожного ока.

Бічне колінчасте ядро являє собою парний шаруватий орган. Для виконання функції керування окоруючими м'язами з метою суміщення сигналів зображень, сформованих в обох очах, у бічному колінчастому ядрі відбувається виділення певних ознак, наприклад, контрастних перепадів яскравості на зображенні. На підставі сигналів про ці перепади для різних очей у шарах бічного колінчастого ядра виробляються сигнали, що надходять у четверогорб, де, у свою чергу, формуються сигнали до очних м'язів.

З бічного колінчастого ядра сигнали про візуальну інформацію надходять у стриарну кору головного мозку, де відбувається виділення основних ознак, що ставляться до спостережуваного зображення – сегментація зображення на об'єкти.

Будову ока схематично представлено на рис. 2. Склера є зовнішньою оболонкою ока, всередині якої перебуває судинна оболонка. В задній частині ока, навпроти кришталика, розташована сітківка. Склера має білий колір з молочним відливом, крім передньої її частини, яка прозора й називається роговицею. Через роговицю світло надходить в око. Судинна оболонка – середній шар – містить кровоносні судини, по яких кров надходить для живлення ока. Прямо під роговицею судинна оболонка переходить у райдужну оболонку, яка й визначає колір очей. У центрі її перебуває зіниця. Функція цієї оболонки – обмежувати надходження світла в око при його високій яскравості. Це досягається за допомогою звуження зіниці при високій освітленості й розширення при низькій. За райдужною оболонкою розташований кришталік, схожий на двоопуклу лінзу. Світло, що проходить через зіницю, вловлюється й фокусується ним на сітківці.

Сітківка перетворює світло в сигнали нейронів у широкому діапазоні інтенсивності з діапазону хвиль «видимого» світла й забезпечує роздільну здатність, достатню, щоб вирізнити окремі об'єкти завтовшки з волос людини з кількох метрів [3]. Сітківка – це частина мозку, пов'язана з ним за допомогою зорового нерва. Подібно багатьом іншим структурам центральної нервової системи, сітківка має форму пластинки товщиною приблизно у чверть міліметра. Вона складається з трьох шарів тіл нервових клітин, розділених двома шарами синапсів, утворених аксонами й дендритами цих клітин.

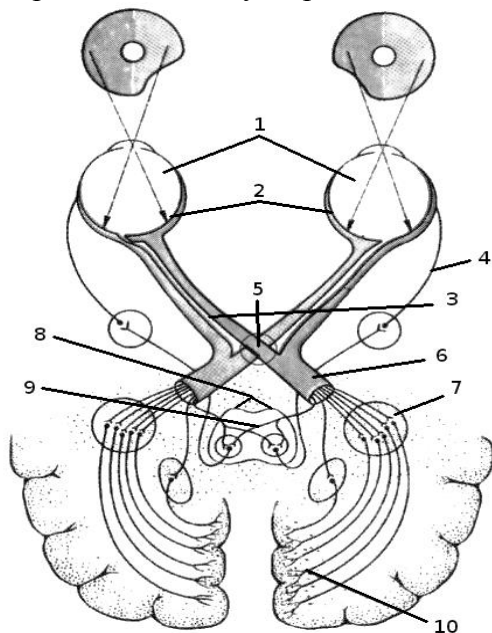


Рисунок 1 [2] – Зорова система

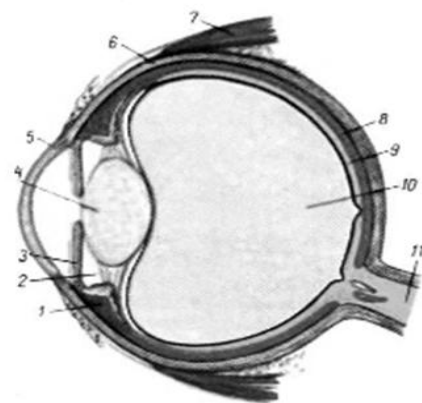


Рисунок 2 [2] – Схема будови ока: 1 – війкове тіло, 2 – задня камера, 3 – райдужна оболонка, 4 – кришталік, 5 – роговиця, 6 – війкове тіло, 7 – верхній прямий м'яз, 8 – судинна оболонка, 9 – сітківка, 10 – скловидне тіло, 11 – зоровий нерв

Шар клітин на задній поверхні сітківки містить світлочутливі рецептори – палички й колбочки. Чисельність паличок значно більша, ніж чисельність колбочок. Вони відповідальні за наше зорове сприйняття при слабкому освітленні й відключаються при яскраво-

му освітленні. Колбочки не реагують на слабе світло, але відповідальні за здатність бачити тонкі деталі та за сприйняття кольорів. У самому центрі сітківки, де здатність нашого зору розрізняти тонкі деталі максимальна, є тільки колбочки. Цю, позбавлену паличок зону діаметром приблизно півміліметра, називають центральною ямкою. Колбочки є по всій сітківці, але найбільш щільно упаковані саме в центральній ямці.

Оскільки палички й колбочки розташовані на задній поверхні сітківки (рис. 3), світло, що надходить до ока, проходить до них через два інших шари клітин. Можливою причиною такого розташування є те, що за рецепторами перебуває шар клітин із вміщенням чорного пігменту – меланіну. Меланін поглинає світло, яке пройшло через сітківку, не даючи йому відбиватися у зворотному напрямку і розсіюватися в середині ока, що створило б паразитне засвічення. Шари перед рецепторами досить прозорі і практично не погіршують чіткість. Однак у центральній частині сітківки два передні шари зміщені до периферії, оголивши центральні колбочки так, що вони виявилися на самій поверхні. маленьке поглиблення, що утворюється, і є центральна ямка.

2.2. Нейрони та їх рецептивні поля

Середній шар сітківки, що передує паличкам та колбочкам відносно падаючого світла, містить нейрони трьох типів: біполярні, горизонтальні та амакринові клітини. Біполярні клітини мають входи від рецепторів, як показано на рис. 3, і передають сигнали безпосередньо гангліозним клітинам. Горизонтальні клітини з'єднують рецептори й біполярні клітини порівняно довгими зв'язками, що йдуть паралельно шарам сітківки; подібним способом амакринові клітини зв'язують біполярні клітини з гангліозними.

Шар нейронів на передній стороні сітківки містить гангліозні клітини, аксони яких проходять по поверхні сітківки, збираються в пучок у сліпої плями й залишають око, утворюючи зоровий нерв. У кожному оці близько 125 мільйонів паличок і колбочок і близько 1 мільйона гангліозних клітин. Сигнали від фоторецепторів надходять до біполярних і далі до гангліозних клітин, між рецепторами й біполярними можуть бути включені ще й горизонтальні клітини, а між біполярними й гангліозними клітинами – амакринові клітини (рис. 3). Загальна площа, яку займають рецептори, пов'язані з однією гангліозною клітиною, становить усього близько квадратного міліметра. Ця зона є рецептивним полем гангліозної клітини – частиною сітківки, світлова стимуляція якої може впливати на стан і вихідні сигнали даної гангліозної клітини. Кожний рецептор може бути пов'язаний з декількома гангліозними клітинами. Рецептивні поля гангліозних клітин мають форму кола. Рецептивні поля сусідніх гангліозних клітин перетинаються [3].

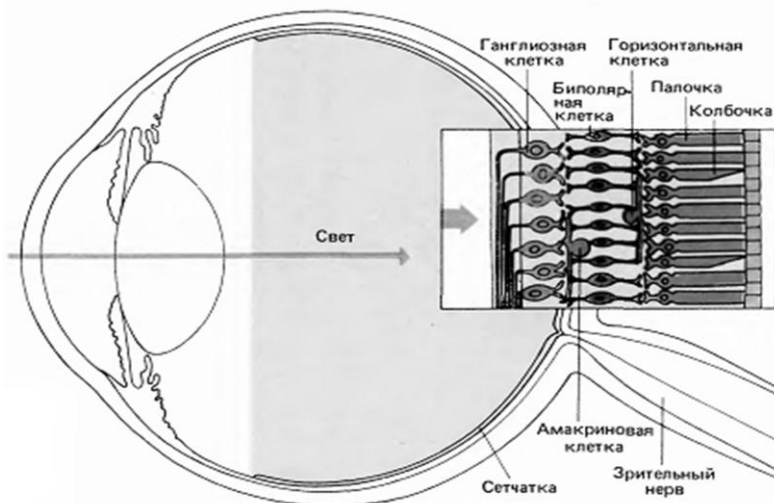


Рисунок 3 [3] – Відносне розташування трьох шарів сітківки в середині ока

Світло має пройти через шари гангліозних, біполярних та амакринових клітин перш ніж досягне паличок та колбочок.

По аналогії з гангліозними клітинами, які (опосередковано) зв'язані з рецепторами, сукупність сигналів, що підходять до нейрона від нейронів попереднього шару, називають

рецептивним полем даного нейрона. Зважаючи на те, що геометричні відношення відносно сітківки в наступних шарах бічного колінчастого ядра зберігаються, рецептивні поля нейронів бічного колінчастого ядра також мають форму кола. Центральна частина рецептивних полів називається зоною збудження, периферійна частина (кільце навкруги центральної частини) – зоною гальмування (рис. 4). При пред’явленні світлового стимулу на зону збудження рецептивного поля нейрона генерується певна кількість імпульсів. Чим більше стимул, тим більшу кількість імпульсів буде сгенеровано. Якщо розмір світлового стимулу перевищить діаметр зони збудження, співпадаючи з частиною зони гальмування, кількість імпульсів буде зменшуватися.

На рис. 5 ліворуч: чотири записи відповідей типової гангліозної клітини з он-центром (видає імпульси при впливі світлового стимулу на зону збудження рецептивного поля). Кожний запис отриманий при одиночному розгорненні променя осцилографу тривалістю 2,5 секунди. Через настільки повільне розгорнення висхідна й спадна фази імпульсу зливаються, так що кожний імпульс має вигляд однієї вертикальної лінії. Ліворуч показані стимули. Верхній запис – стан спокою (стимулу немає): імпульси виникають рідко й більш-менш випадково. Три нижні записи – реакції на невелику (оптимальної величини) пляму, на велику пляму, що покриває центр і периферію рецептивного поля, і на кільце, що покриває тільки периферію. Праворуч: реакції гангліозної клітини з off-центром (видає імпульси при впливі світлового стимулу на зону гальмування рецептивного поля) на такий же набір стимулів.

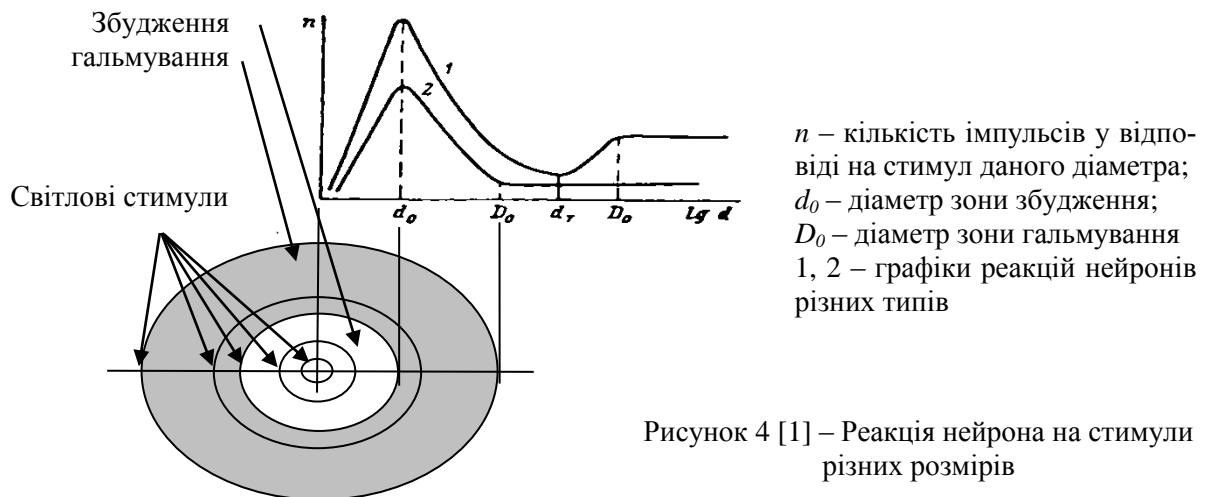


Рисунок 4 [1] – Реакція нейрона на стимули різних розмірів

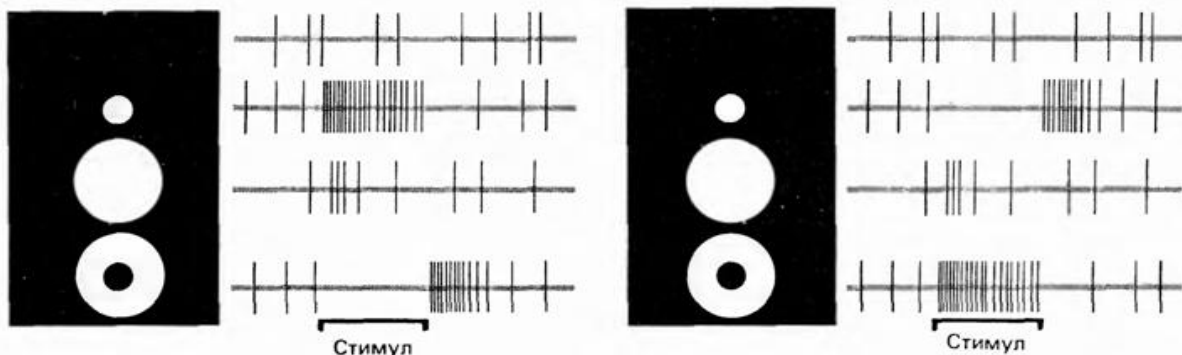


Рисунок 5 [3] – Реакції гангліозних клітин з on- і off-центрами на стимули

2.3. Зоровий акт та динамічні властивості рецептивних полів нейронів

Важливе відкриття було зроблено у 70-х роках минулого сторіччя – зміна зон збудження рецептивних полів під час зорового акту, що означає динамічну зміну роздільної здатності в зоровій системі.

Зорова система людини працює у режимі послідовних зорових актів. Зоровий акт триває приблизно 150 мс. Після чого відбувається саккада – посмикування ококорушійних м'язів, в результаті чого зображення на сітківці зміщується і починається наступний зоровий акт.

У результаті досліджень розмірів зон збудження рецептивних полів було встановлено [1], що їх розміри не залишаються постійними впродовж зорового акту. Пізніше це явище було знов досліджене та підтверджене вже на сучасному рівні [4]. Дослідження цих явищ виконувалися на кішках, зорова система яких у цій частині не суттєво відрізняється від зорової системи людини.

Один із методів визначення просторових характеристик рецептивного поля, що має форму кола за допомогою світлових плям із різною площею, полягає в тому, що, як відомо з [5], при збільшенні діаметра світлового стимулу, центр якого співпадає з центром рецептивного поля, відповідь нейрона – кількість імпульсів – спочатку збільшується (рис. 4), а потім зменшується. У відповідності з цим [6], діаметром d_0 зони збудження рецептивних полів, що мають форму кола, вважають найменший діаметр стимулу, при реакції на який кількість імпульсів у відповіді нейрона має максимальне значення. Діаметром рецептивного поля D_0 , якому ставиться у відповідність діаметр зони гальмування, вважають найменший діаметр стимулу, при збільшенні якого не відбуваються зміни у відповідях нейрона.

Однак за такими відповідями безпосередньо можна визначити лише одне значення діаметра зони збудження рецептивного поля (d_0). Це значення не відображає змін розміру зони збудження під час дії стимулу. З метою отримання кількісного опису змін розмірів зони збудження рецептивного поля з плином часу, тобто знаходження d_0 у впорядкованій послідовності інтервалів часу $\Delta t = \{\Delta t_1, \Delta t_2, \dots, \Delta t_i, \dots\}$ від початку дії стимулу і визначення залежності $d_0 = f(t)$, був застосований метод часових зрізів [1]. Цей метод базується на доволі обґрунтованому припущенні, що якщо під час зорового акту діаметр зони збудження змінюється, то максимальна кількість імпульсів у відповіді нейрона припадає на той інтервал часу, для якого діаметр зони збудження та діаметр стимулу співпадають або відрізняються один від одного незначною мірою.

Метод часових зрізів для побудови залежності кількості імпульсів у відповіді нейрона під час дії світлового стимулу полягає у наступному.

Постстимульні гістограми (ПСГ) відповідей нейрона на стимули різної площі розбиваються на декілька послідовних часових інтервалів Δt (що дорівнювали 7,5 або 15 мс) та знаходять кількість імпульсів у кожному часовому інтервалі (рис. 6). У відповідності з означенням діаметр зони збудження дорівнює найменшому діаметру стимулу, при якому досягається максимум кількості відповідей нейрона. Метод часових зрізів дозволяє також визначити і номер часового інтервалу, на якому цей максимум досягається. Чим більший діаметр стимулу, тим раніше досягається максимум. Це дозволяє зробити висновок про те, що діаметр зони збудження рецептивного поля, яке має форму кола під час зорового акту, змінюється з плином часу, від максимального значення до мінімального.

Якщо на початку зорового акту розмір рецептивного поля має максимальне значення, то наприкінці зорового акту він зменшується до мінімально можливого. У випадку гангліозної клітини це 1–2 рецептори.

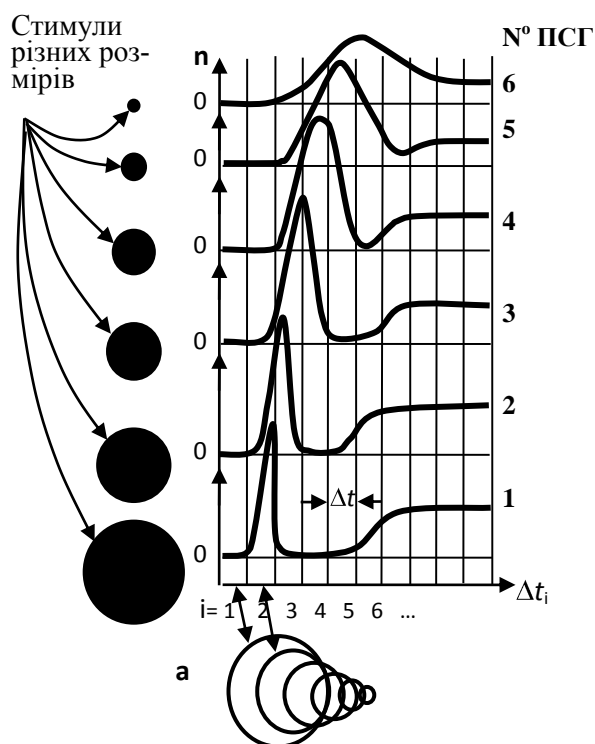


Рисунок 6 [1] – Обробка відповідей нейрона на стимули різних розмірів методом часових зрізів. На осі n для кожної ПСГ відкладена кількість імпульсів у відповідному часовому інтервалі Δt_i . Максимальна кількість імпульсів відповідає рівності розмірів зони збудження і стимулу; a – зміна розмірів зони збудження нейрона під час зорового акту

Залежність відповіді нейрона від часу дії стимулу та динамічні властивості рецептивного поля нейрона бічного колінчастого ядра кішки досліджено у [4]. Рецептивне поле подразнювалося за допомогою світлових стимулів упродовж 400–500 мс. Просторові параметри рецептивного поля оцінювалися за просторовими інтегральними кривими, що визначалися для послідовних часових інтервалів по 5мс на протязі всього часу пред'явлення стимулу. Таким чином, було можливо дослідити динамічні характеристики властивостей відповідей у подібних до природних умов станах та за допомогою методів, що не розглядають систему, яка моделюється, як лінійну.

Отримані результати показали виражені зміни у структурі рецептивного поля під час пред'явлення стимулу. На початку спостереження нейрони мали велику зону збудження рецептивного поля. Центральна частина рецептивного поля швидко зменшувалася до мінімального розміру, що спостерігалось приблизно через 70 мс від моменту пред'явлення стимулу, після чого спостерігалось незначне збільшення її розміру. Таким чином, максимальна просторова роздільна здатність досягалася за невеликий проміжок часу після пред'явлення стимулу. Автори наголошують на узгодженості їх результатів із гіпотезою про те, що під час пред'явлення світлового стимулу сильний початковий сигнал від нейронів містить просторово «грубе» повідомлення до зорової кори, який поступово змінюється слабкішим, але таким, що містить більш точну просторову інформацію.

Була перевірена гіпотеза про те, що стимул малого діаметра спочатку активує багато нейронів, але більшість з них – лише тимчасово (рис. 7). Моделювалися умови, коли стимул випадав із зон збудження рецептивних полів нейронів під час її зменшення. Для всіх трьох розташувань стимулу відносно центру рецептивного поля спостерігався початковий тимчасовий відгук, але наступна стійка частина відповіді, яка має виникати при знаходженні стимулу малого розміру у центрі зони збудження мінімального діаметра, у цьому випадку була відсутня.

Також у [4] було встановлено, що схожі зміни просторових характеристик спостерігалися і для рецептивних полів нейронів бічного колінчастого ядра.

Таким чином, можна вважати, що для зорової системи існує максимальна роздільна здатність, яка визначається кількістю рецепторів у центральній частині рецептивного поля і змінювана під час зорового акту змінна роздільна здатність, що визначається поточним значенням розміру зони збудження рецептивного поля нейрона.

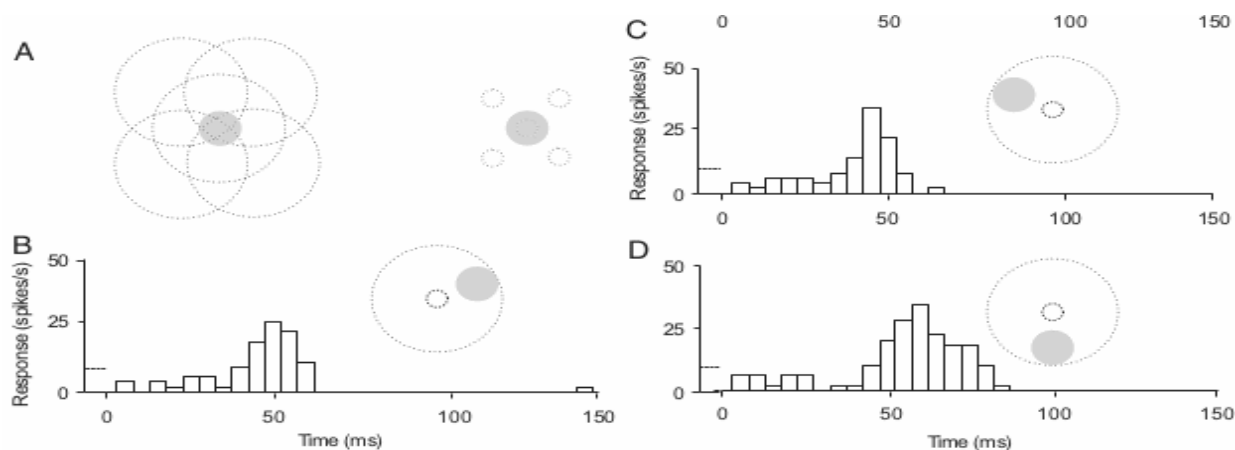


Рисунок 7 [4] – Великий початковий розмір зон збудження рецептивних полів передбачає, що кількість активованих нейронів зменшується під час моделювання.

A: схематичне подання гіпотези. B–D: стимул, що пред’являється в межах рецептивного поля, але за межами мінімального розміру зони збудження спочатку швидко видає відповідь, яка зникає при подальшому зменшенні зони збудження рецептивного поля до мінімальних розмірів. Побудовані лише відповіді для перших 150 мс з 500мс інтервалу моделювання. Пунктирна лінія перед початком пред’явлення стимулу показує рівень спонтанної активності. Ширина стовпчика гістограми відповідає інтервалу 5 мс. Для кожного розміщення було виконано 100 випробувань

3. Змінна роздільна здатність у системах технічного (комп’ютерного) зору

3.1. Приклади впливу роздільної здатності на результати обробки зображень

Зазвичай вважається [7], що вихідне зображення для подальшої обробки подано у аналоговій формі. Як правило, розглядається зображення, що обмежено прямокутником (полем зору), який має придатні для обробки розміри. Зображення має повністю вміщуватися у цьому прямокутнику та заповнювати його, якщо це можливо.

Обробка починається з дискретизації поля зору та квантування яскравості за рівнями, що відповідають кольорам у палітрі зображення (півтони, RGB та ін.). Наступні параметри системи комп’ютерного зору обираються із практичних міркувань: роздільна здатність – множина дискретних елементів зображення (пікселів, зазвичай мають квадратну форму), які можна розмістити в одиниці довжини (дюйми, сантиметри) та, відповідно, розмір зображення у пікселях; множина рівнів яскравості, що є областю визначення функції яскравості для представлення кожного з кольорів.

Вважається, що обрана роздільна здатність системи комп’ютерного зору якнайкраще підходить для класу зображень, які необхідно обробляти. Тобто непотрібні деталі не будуть відображені з надзвичайно великою роздільною здатністю (як-то спотворені завданнями контури об’єктів) та суттєві деталі об’єктів не зникнуть через дуже низьку роздільну здатність. До функції квантування та множини рівнів яскравості для подання кольорів зображення висуваються такі ж самі вимоги: можливість відобразити суттєві деталі об’єктів.

Таким чином, зображення може бути подане як двовимірний масив $V(M, N)$, що містить M рядків та N стовпчиків. Кожний елемент $v(m, n)$ цього масиву відповідає або яскравості пікселя з координатами (m, n) для півтонового зображення, або значенням яскравості базових кольорів (наприклад, червоного, зеленого, синього) для кольорового зображення. Цей масив може розглядатися як матриця або вектор у залежності від обраних математичних методів.

Розроблено та успішно застосовується велика кількість методів і алгоритмів для обробки зображень, що спираються на векторне або матричне представлення зображення.

Більшість цих методів може бути успішно застосована для обробки півтонових зображень, якщо для класу зображень, що розглядаються, були вдало обрані параметри дискретизації та квантування. Однак необхідно наголосити на обмеженнях такого підходу. Наприклад, для визначення міри схожості двох зображень, як $f(V_1, V_2)$ із використанням методів статистичного розпізнавання [8], використовуються значення яскравостей усіх пікселів, що належать до зображення. Але ж зазвичай пікселі зображення відносяться або до об'єкта, або до фону. Тобто результат розпізнавання залежить не тільки від значень яскравості пікселів об'єкта інтересу, а й від значень яскравості пікселів фону, що у багатьох випадках не прийнятно.

Розглянемо зображення, що містять довільний текст на однотонному фоні та на фоні деякої довільної решітки (рис. 8) [9]. Текст на рис. 8 а може бути розпізнаний як статистичними, так і структурними методами розпізнавання. Розпізнавання тексту на рис. 8 б є складнішою задачею. Якщо спробувати застосувати статистичні методи, результат обчислення схожості із еталонним зображенням буде спотворений через наявність пікселів зображення решітки із кольором об'єкта у полі фону. Також взаємне розташування тексту та решітки може змінюватися після застосування до зображення операцій дискретизації та квантування. При застосуванні до зображень, як на рис. 8 б, структурних методів замість контурів об'єктів будуть виділені контури клітин решітки. Схожу картину можна спостерігати, коли накладена на текст решітка має колір фону (рис. 9). У даному випадку при застосуванні статистичних методів розпізнавання результат обчислення схожості буде також спотворено через наявність в полі зображення пікселів, що належать до об'єкта, але мають колір фону. Знову ж, не гарантується однакове взаємне розташування тексту і решітки після накладення на зображення решітки та застосування до зображення операцій дискретизації й квантування. При спробі застосувати структурні методи розпізнавання до зображень на рис. 9, будуть отримані такі ж самі результати, як і для рис. 8 б: будуть виділені контури клітин решітки.

Задача розпізнавання тексту з накладеною на нього довільною решіткою може розглядатися як окремий випадок розпізнавання текстур. Але обчислювальна складність методів розпізнавання текстур зазвичай значно перевищує складність методів розпізнавання тексту. Тому їх застосування унеможливорюється або взагалі виключається для розв'язання вказаних вище задач.



Рисунок 8 [9] – Приклади зображень, що містять довільний текст: а) однорідний фон; б) фон містить прямокутну решітку

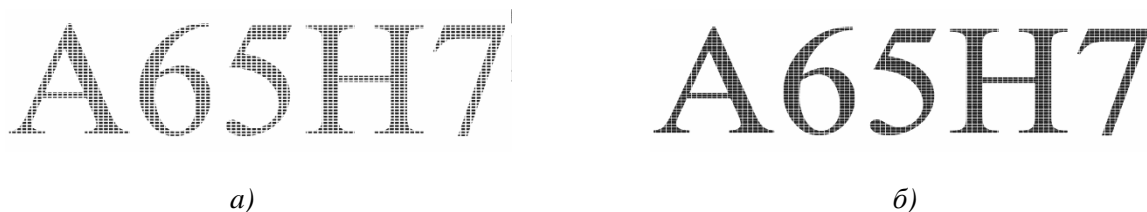


Рисунок 9 [9] – Решітка, що має колір фону, накладена випадковим образом на зображення тексту: а) лінії решітки на зображеннях; б) мають різну товщину

У зоровому сприйнятті людини всі чотири наведені задачі розпізнавання розв'язуються непомітно, без видимих зусиль, судячи з усього, на підсвідомому рівні, у

той час, як сучасні засоби обробки візуальної інформації впевнено справляються лише з розпізнаванням задач, подібних до рис. 8 *а*.

Як було показано вище, під час одного акту зорового сприйняття в зоровій системі людини зображення обробляється зі змінною роздільною здатністю. Можливо припустити, що саме спостереження послідовності зображень з різною роздільною здатністю на початку зорового акту робить можливим цілісне візуальне сприйняття символів на довільному фоні. Тоді можливо припустити, що якщо деяка програма розпізнавання текстів успішно впорається з розпізнаванням зображення на рис. 10 *а* зі зменшеною роздільною здатністю, то вона також успішно впорається з розпізнаванням зображень з такою ж самою, зменшеною, роздільною здатністю, що наведені на рис. 8 *б*, 9 *а*, 9 *б*. Для перевірки цього твердження можна провести простий експеримент із відомою програмою розпізнавання текстів FineReader.

Вихідний набір даних складається з чотирьох зображень (рис. 8 *а*, 8 *б*, 9 *а*, 9 *б*). Розмір кожного зображення складає 900x280 пікселів, при роздільній здатності 72x72 пікселі на дюйм. Довільна комбінація символів «А65Н7» була обрана спеціально, щоб виключити вплив на результат розпізнавання використання словників.

Текст на рис. 8 *а* був успішно розпізнаний. Розпізнавання інших трьох зображень дало відмову від розпізнавання через неможливість визначити об'єкт на зображенні. Зображення на рис. 10 відповідають зображенням на рис. 8, 9, що дискретизовані зі зменшеною у шість разів роздільною здатністю. На всіх цих зображеннях рядок тексту був успішно розпізнаний.

Проведений експеримент показує, що поняття оптимальної роздільної здатності може бути застосоване не лише до всього зображення, а й до кожного об'єкта на ньому. У цьому випадку достатньо було просто зменшити роздільну здатність зображень, щоб розв'язати задачу, яку інакше неможливо було вирішити традиційними методами. Тобто, при розглянутій зменшеній роздільній здатності зображень інформації про форму об'єктів у полі зору достатньо для прийняття рішення. У той самий час завади з великою просторовою частотою (решітки) майже непомітні і не впливають на прийняття рішення.



Рисунок 10 [9] – Зображення з рис. 8 (*а*, *б*) та рис. 9 (*в*, *г*) при зменшеній у 6 разів роздільній здатності при дискретизації

Природним чином розгляд зображень при різній роздільній здатності застосовується під час фотозйомки біологічних об'єктів із використанням звичайного світлового мікроскопа, тут дослідники стикаються з обмеженням глибини різкості, яке не дозволяє отримати чітке зображення на одному знімку. Якщо зразок товщий за максимальну глибину різкості, ділянки поверхні об'єкта за межами фокальної площини будуть розфокусовані. Для подолання цих обмежень зйомка зразка повторюється багато разів водночас зі зміною гли-

бини різкості вздовж оптичної вісі. Це дає послідовність зображень, де різні частини зразка з'являються і виходять із фокуса – різні частини зображення з'являються у полі спостереження з різним ступенем розмиття. Для отриманої послідовності зображень за допомогою вейвлет-перетворення виконується процедура відновлення, щоб отримати єдине різке зображення [10]. Таким чином, використання послідовності зображень з різним ступенем розмиття окремих областей (розгляд областей зображення при різній роздільній здатності) є ще одним прикладом задачі, що інакше не може бути розв'язана за допомогою існуючих методів.

3.2. Неявне використання різних роздільних здатностей у системах обробки зображень

В експерименті, який було описано вище, зменшення роздільної здатності під час дискретизації зображення призводить до розмиття кінцевого зображення. Система обробки зображень має обробляти зображення, спотворені завадами. У найпростішому випадку це зображення обробляється за допомогою фільтра Гауса з метою усунення небажаних деталей, алгоритм обробки застосовується до розмитого зображення. Саме така попередня обробка є, наприклад, складовою частиною широкоживаного [11–13] методу Канні [14] для виділення контурів на зображенні. Для вихідного зображення $V = \{v(m, n) | m = \overline{1, M}; n = \overline{1, N}\}$ виконують операцію розмиття зображення, використовуючи фільтр Гауса для зниження рівня завад, виключення небажаних деталей і елементів текстури зображення:

$$g(m, n) = G_{\sigma} * v(m, n),$$

де G_{σ} – фільтр Гауса для значення дисперсії σ , $g(m, n)$ – елемент «розмитого» зображення V_g .

Обчислюють часткові значення градієнтів для горизонтального $g_m(m, n)$ і вертикального $g_n(m, n)$ напрямків на розмитому зображенні $g(m, n)$, використовуючи, наприклад, оператор Собеля, щоб отримати значення загального градієнта $M(m, n)$ і його напрямку $\theta(m, n)$ як

$$M(m, n) = \sqrt{g_m^2(m, n) + g_n^2(m, n)},$$

$$\theta(m, n) = \arctg\left(\frac{g_n(m, n)}{g_m(m, n)}\right).$$

За значеннями $M(m, n)$, з використанням порога T , який має бути підібраний так, щоб усі контурні елементи були виділені, у той час, як більшість завад була б придушена, отримують значення $M_T(m, n)$:

$$M_T(m, n) = \begin{cases} M(m, n) & \text{if } M(m, n) > T \\ 0 & \text{if } M(m, n) \leq T \end{cases}.$$

Для підвищення якості роботи методу використовують два пороги T_1 і T_2 , причому $T_1 < T_2$. Якщо піксель $v(m, n)$ зі значенням $T_1 < M_T(m, n) < T_2$ має в напрямку градієнта $\theta(m, n)$ два сусідні пікселі, для кожного з яких $M_T(m, n) > T_2$, його значення як контурного елемента зберігається, а якщо ні, то прирівнюється до 0.

Усі ненульові елементи об'єднують, щоб вийшов замкнений контур об'єкта, для чого використовується спеціальний алгоритм.

У методі Канні змінна роздільна здатність використовується у неявному вигляді, оскільки оператор підбирає ступінь розмиття σ , але це робиться, виходячи з його суб'єктивних міркувань стосовно характеру завад.

Недоліки методу Канні:

– результатом роботи алгоритму є контур у вигляді послідовності пікселів, але піксель є двовимірним утворенням, у той час, як контур уявляється як лінія, зокрема, ламана лінія без товщини;

– оскільки модель зображення при застосуванні методу Канні не використовується, то метод не може автоматично розрізнити контури об'єктів та контури, що належать до текстури фону;

– результат роботи алгоритму Канні залежить від змінного параметра фільтра Гауса σ , який має значення дисперсії (рис. 11). У першоджерелі не надано жодних рекомендацій стосовно вибору значення цього параметра.

З метою придушення завад у вихідному сигналі можуть бути застосовані й інші фільтри, але результат буде також залежати від розміру апертури фільтра, який, у свою чергу, є невідомим параметром.

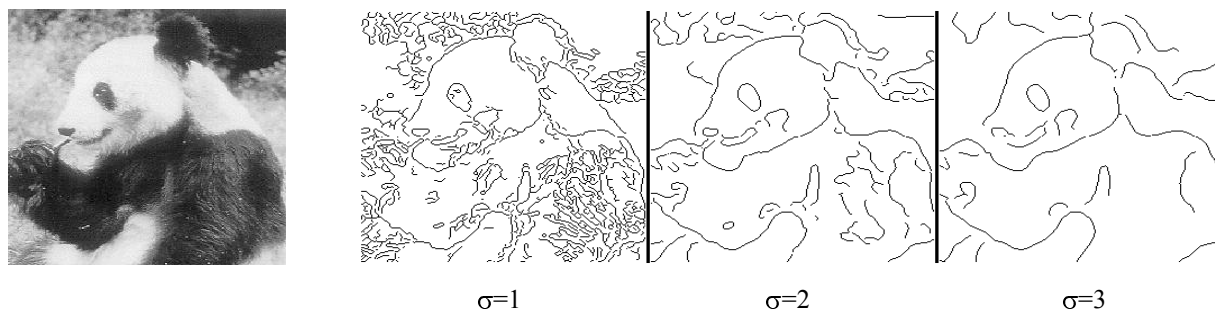


Рисунок 11 – Застосування методу Канні для виділення контурів на зображенні при різних значеннях параметра σ

3.3. Використання змінної роздільної здатності для зменшення обчислювальної складності алгоритмів

Зменшення роздільної здатності широко використовується для зменшення обчислювальної складності алгоритмів обробки, розпізнавання зображень та підвищення їх швидкодії.

Наприклад, у [15] модель еталонів, яка складається з окремих частин, поєднаних не жорсткими зв'язками, розглядається з різною роздільною здатністю. Сформульовано алгоритм переходу від низької роздільної здатності до високої. Метод обробки, що пропонується, заснований на спостереженні, що пошук відповідностей частини зображення еталона є найбільш обчислювально-витратною операцією у порівнянні з виявленням суттєвих частин та розрахунком їх оптимальної конфігурації. Мінімізація кількості операцій порівняння частин еталонів із зображенням тягне за собою прискорення операції виявлення. Починаючи з найменшої роздільної здатності, еталони порівнюються з зображенням. Відбираються лише найбільш ймовірні їх розташування. Потім знайдені локально оптимальні розташування рекурсивно розповсюджуються на частини моделі з більшою роздільною здатністю. За допомогою рекурсивного видалення невідповідних розташувань з простору пошуку множина можливих розташувань зменшується так, що при максимальній роздільній здатності потрібно виконати лише декілька порівнянь еталонів із зображенням. Запропонований метод дозволяє отримати десятикратне прискорення обчислень у порівнянні зі стандартним методом динамічного програмування.

Алгоритм, що розглядається у [16], використовує подібну ідею виключення великих областей із простору гіпотез на ранніх стадіях розпізнавання, однак для кожної роздільної здатності використовується послідовність детекторів об'єктів. Результатом роботи детектора є кількісна оцінка області, що розглядалася. Рішення щодо застосування наступних детекторів з послідовності до цієї області приймається на підставі порівняння отриманої кількісної оцінки з певним порогом. Область буде розглянуто при наступній роздільній здатності, якщо її кількісні оцінки від кожного детектора перевищили відповідні порогові значення. Всі пороги встановлюються автоматично, на підставі вірогіднісних оцінок.

У [17] розглядається застосування стратегії coarse-to-fine до задачі кластеризації траєкторій руху транспортних засобів. Вихідні траєкторії об'єднуються у «грубі» кластери. До кожного «грубого» кластера входять траєкторії з приблизно однаковим напрямом, але з різними характеристиками розташування. Для подальшої точної кластеризації використовується перерахування множини точок траєкторії з застосуванням евклідової відстані як міри близькості.

При розпізнаванні облич процедура «coarse-to-fine» може бути реалізована шляхом застосування різних методів розпізнавання для скорочення кількості кандидатів на кожному кроці.

У [18] процес прийняття рішення має декілька стадій: 1) оцінка приналежності одному з усіх можливих класів (one-against-all of SVM); 2) визначення приналежності кожного кандидата одному з пари класів (one-against-one of SVM); 3) алгоритм Eigenface; 4) метод RANSAC. Стадії 1) та 2) використовують характеристики усього зображення обличчя, отримані з дискретного косинусного перетворення. На стадії 3) розглядаються проєкції зображень облич у простір характеристик. Простір облич визначається власними векторами множини облич та на підставі інформації про інтенсивність зображення обличчя. Метод RANSAC застосовується на останній стадії, де просторова інформація, отримана за допомогою методів епіпольярної геометрії із зображення, що перевіряється, порівнюється з двома еталонними зображеннями і обирається зображення з найбільшим значенням подібності та з найкоротшою відстанню до відповідних характерних точок.

Задача встановлення відповідності між пікселями двох зображень облич людини (знаходження розмітки) [19] ефективно розв'язується за допомогою побудови «каскадів» розміток. В одному каскаді розміри всіх зображень зменшуються вдвічі і будується нова розмітка. Після цього визначається початкове наближення для вихідної розмітки на підставі нової розмітки й виконується пошук поля руху відносно цього початкового наближення, але з меншою кількістю міток. При використанні одного «каскаду» алгоритм розв'язання задачі виконується у вісім разів швидше, зберігаючи при цьому точність знаходження поля руху для двох зображень. Хоч автор і описує цей метод як певний інженерний прийом, але слід зауважити, що фактично тут використовується обробка зображення зі змінною роздільною здатністю, оскільки в межах одного «каскаду» зображення обличчя розглядається зі зменшеною вдвічі роздільною здатністю, а отримана для зображень зі зменшеною роздільною здатністю розмітка використовується як початкове наближення при пошуку розмітки для зображень зі збільшеною роздільною здатністю.

Динамічне програмування дуже часто використовується у задачах розпізнавання мови, розпізнавання символів, пошуку відповідності за шаблоном для об'єктів, що можуть деформуватися, відстеження дороги. Однак постановки таких задач дуже часто призводять до простору станів величезних розмірів, через що розрахунки можуть стати нездійсненними, навіть із використанням динамічного програмування.

Для подолання такого роду перешкод у [20] пропонується використовувати динамічне програмування coarse-to-fine (CFDP). Основна ідея цього підходу – сформувати послідовність грубих наближень вихідного графа динамічного програмування шляхом об'єднання станів графа у «над стани». Для кожної грубої апроксимації оптимальний шлях

розраховується з «оптимістичними» вагами дужок між надстанами. Надстани вздовж цього оптимального шляху переглядаються, і процес повторюється знову, доки не буде знайдено доказово оптимальний глобальний шлях. У багатьох випадках глобальний оптимум досягається зі значно меншими витратами на розрахунки, ніж при прямому використанні динамічного програмування. Запропонований алгоритм особливо добре підходить для задач з великим простором станів. Відповідно до [21], швидкість алгоритму CFDP залежить від структури об'єднання і від природи задачі. У найліпшому випадку CFDP дозволяє отримати значне скорочення обчислень у порівнянні зі звичайним методом динамічного програмування; у найгіршому випадку він фактично буде працювати повільніше.

Метою застосування методів зі змінною роздільною здатністю у розглянутих вище випадках є визначення частин вихідного зображення або частин вихідного набору даних, що містять інформацію, яка здається корисною для розв'язання поставленої задачі. Складні розрахунки виконуються лише над цими частинами. При цьому природі механізму зміни роздільної здатності, що використовується у кожному конкретному випадку, не надається значення.

Зауважимо, що велика кількість задач розпізнавання зображень, які мають складність NP чи взагалі не можуть бути розв'язані за допомогою традиційних методів, розв'язуються у зоровій системі людини ментально, а задачі, що пов'язані з обробкою відеоряду, у режимі реального часу. Тому природним буде звернення до результатів вивчення процесів у зоровій системі людини, що отримані у нейрофізіології для створення нових методів та алгоритмів обробки візуальної інформації.

У попередні роки дослідники вже намагалися просунутися у цьому напрямі, використовуючи актуальні на той час результати з нейрофізіології зору.

Наприклад, чутливі до напрямку руху стимулу клітини зорової системи приматів показують наявність певного діапазону просторових розмірів, зокрема, якщо розміри рецептивних полів порівнюються між різними ділянками кори, такими як первинна зорова кора та середня частина скроневої частки [22]. З урахуванням цього, у [23] досліджено, як інтегрування інформації про рух об'єкта по всіх просторових масштабах може допомогти покращити оцінку оптичного потоку. Було запропоновано адаптивний, багатомасштабний метод, де масштаб дискретизації обирається локально, відповідно до оцінки відносної помилки визначення швидкості щодо властивостей зображення. Було показано, що запропонований метод дає суттєво кращі оцінки оптичного потоку, ніж традиційні алгоритми, при незначному збільшенні витрат на розрахунки. На думку авторів, це є важливим з огляду на велику кількість ітерацій, необхідних для релаксаційних алгоритмів та дивовижної швидкості, з якою людина може надійно оцінити швидкість руху. На засадах цього підходу у [24] представлено дворівневу багатомасштабну адаптивну нейромережеву модель обрахування параметрів руху в середній частині скроневої частки приматів. На першому етапі локальні швидкості вимірюються при багатьох просторових роздільних здатностях, після чого поле оптичного потоку обраховується мережею чутливих до напрямку нейронів при багатьох просторових роздільних здатностях. При виникненні конфліктів між сигналами від клітин при різних роздільних здатностях застосовується схема ветування coarse-to-fine, згідно з якою перевагу мають сигнали від клітин на «грубіших» роздільних здатностях. Подальші експерименти з моделювання властивостей неklasичного рецептивного поля виявилися цілком узгодженими з результатами, що були отримані у нейрофізіології. Також було запропоноване нове пояснення феномена захоплення руху з застосуванням стратегії розв'язання конфліктів coarse-to-fine при розгляді інформації з різних вхідних каналів.

4. Висновки

Таким чином, у результаті проведеного аналізу поточного стану речей в областях нейрофізіології та створення систем комп'ютерного зору встановлено, що в системах комп'ютерного зору обробка візуальної інформації з використанням різних роздільних здатностей використовується спонтанно, як певний інженерний прийом. У той самий час нейрофізіологічні дослідження свідчать про наявність механізмів обробки візуальної інформації зі змінною роздільною здатністю в зоровій системі людини.

З розглянутих прикладів видно, що велика кількість задач розпізнавання зображень, які мають складність NP чи взагалі не можуть бути розв'язані за допомогою традиційних методів, розв'язуються у зоровій системі людини моментально, а задачі, що пов'язані з обробкою відеоряду – у режимі реального часу. При цьому вибір відповідної роздільної здатності суттєво впливає на результат розпізнавання зображень.

Тому існує необхідність розробки нових методів та алгоритмів, що поєднують отримані у нейрофізіології результати вивчення процесів, які пов'язані зі зміною роздільної здатності у зоровій системі людини, з існуючими методами і алгоритмами обробки зображень для їх подальшого використання в задачах розпізнавання зображень на основі сегментації спотворених завадами зображень та сигналів при відсутності апріорної інформації про характер завад.

СПИСОК ДЖЕРЕЛ

1. Подвигин Н.Ф. Динамические свойства нейронных структур зрительной системы. Ленинград: Наука, 1979. 158 с.
2. Ерошевский Т.И. и др. Глазные болезни: учебное пособие. М.: «Лидер М», 2008. 316 с.
3. Hubel H.D. Eye, brain and vision. New York: Scientific American Library, Distributed by W.H. Freeman, 1988. 240 p.
4. Ruksenas O., Bulatov A., Heggelund P. Dynamics of Spatial Resolution of Single Units in the Lateral Geniculate Nucleus of Cat During Brief Visual Stimulation. *J Neurophysiol.* 2007. N 97. P. 1445–1456.
5. Barlow H.B. Summation and inhibition in the frog's retina. *J. Physiol.* 1953. Vol. 119. P. 69–88.
6. Глезер В.Д. Механизмы опознания зрительных образов. Л.: Наука, 1966. 204 с.
7. Павлидис Т. Алгоритмы машинной графики и обработки изображений. М.: Радио и связь, 1986. 400 с.
8. Шлезингер М., Главач В. Десять лекций по статистическому и структурному распознаванию. Київ: Наукова думка, 2004. 535 с.
9. Sharypanov A., Antoniouk A., Kalmykov V. Joint study of visual perception mechanism and computer vision systems that use coarse-to-fine approach for data processing. *Information content & processing.* 2014. Vol. 1, N 3. P. 287–300.
10. Forster B., Van De Ville D., Berent J., Sage D., Unser M.. Complex Wavelets for Extended Depth-of-Field: A New Method for the Fusion of Multichannel Microscopy Images. *Microsc. Res. Tech.* September 2004. Vol. 65, N 1–2. P. 33–42.
11. Grigorescu C. et al. Contour Detection Based on Nonclassical Receptive Field Inhibition. *IEEE Transactions On Image Processing.* 2003. Vol. 12, N 7. P. 729–739.
12. Ярмілко А.В. Використання методів візуалізації у діагностиці та управлінні електронно-променевим мікрообробленням оптичних матеріалів. *Вісник Чернігівського державного технологічного університету.* 2013. № 2 (65). С. 156–161.
13. Костенко В.В., Лавренюк І.В., Пономарьов В.М. Дослідження можливостей ефективного використання методів розпізнавання для обробки фотозображень автомобільних номерних знаків. *Вісник АМСУ. Технічні науки.* 2014. N 1 (51). С. 48–56.
14. Canny J.F. A computational approach to edge detection. *IEEE Trans. Pattern Anal. Machine Intell.* 1986. Vol. PAMI-8, N 6. P. 679–698.
15. Pedersoli M., Vedaldi A., Gonzalez J. A Coarse-to-fine approach for fast deformable object detection. *CVPR.* 2011. June. P. 1353–1360.

16. Moreels P., Perona P. Probabilistic Coarse-To-Fine Object Recognition. *Technical report*. Pasadena: California Institute of Technology, 2005. 49 p.
17. Li X., Hu W., Hu W. A Coarse-to-Fine Strategy for Vehicle Motion Trajectory Clustering. *Proceedings of the 18th International Conference on Pattern Recognition (ICPR'06) Year. 2006*. Vol. 1. P. 591–594.
18. Lee J.-D., Kuo C.-H. A Multi-Stage Classifier for Face Recognition Undertaken by Coarse-to-fine Strategy, State of the Art in Face Recognition. *Tech.* 2009. URL: http://www.intechopen.com/books/state_of_the_art_in_face_recognition/a_multistage_classifier_for_face_recognition_undertaken_by_coarse-to-fine_strategy.
19. Тищенко М.А. Тривимірна реконструкція людського обличчя в задачах ідентифікації особи: дис. ... канд. техн. наук. Київ: Міжнародний науково-навчальний центр інформаційних технологій та систем, 2012. 120 с.
20. Raphael C. Coarse-to-Fine Dynamic Programming. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2001. Vol. 23. P. 1379–1390.
21. Lucena O.B. Dynamic Programming, Tree-width and Computation on Graphical Models: PhD thesis. Division of Applied Mathematics. Providence: Brown University, 2002. 85 p.
22. Maunsell J.H.R., Van Essen D.C. Functional properties of neurons in middle temporal visual area of the macaque monkey. I. Selectivity for stimulus direction, speed and orientation. *J. Neurophysiol.* 1983. Vol. 49. P. 1127–1147.
23. Battiti R., Amaldi E., Koch C. Computing Optical Flow Across Multiple Scales: An Adaptive Coarse-to-Fine Strategy. *International Journal of Computer Vision*. 1991. Vol. 6, N 2. P. 133–145.
24. Wang H.T., Mathur B., Koch C. A Multiscale Adaptive Network Model of Motion Computation in Primates. *Advances in Neural Information Processing Systems*. 1990. Vol 3. P. 349–355.

Стаття надійшла до редакції 17.05.2018