

# Ключові моменти становлення київської школи розпізнавання

До 50-річчя Інституту кібернетики ім. В.М. Глушкова НАН України.

Грудень 2007 року

М.І. Шлезінгер

Дослідження з розпізнавання образів розпочалися в Інституті кібернетики від самого моменту його заснування. Історія розвитку цього напрямку не є однорідною, і ми зазначимо кілька ключових моментів, які ознаменували завершення одних етапів і початок наступних. **Водночас/Втім** (ориг. однак), якби потрібно було неодмінно вказати характерну рису цього напрямку на всіх етапах його розвитку, то це була б єдність теоретичних і прикладних досліджень — одна з відомих трьох єдностей, визначених Глушковим. Саме така атмосфера панує поміж дослідників київської школи, і саме так її характеризують дослідники з інших шкіл, коли прагнуть виразити її найбільш визначну рису.

Ключовим моментом, що з самого початку визначив саме такий характер досліджень у київській школі, слід вважати роботу [1]. Із сучасної точки зору ця робота виглядає як комплекс, що складається з пристрою зчитування зображень (тобто, простіше кажучи, сканера), пристрою виведення зображень на екран (тобто дисплея) та універсальної обчислювальної машини «Київ». Проте, цей комплекс було створено задовго до появи самих слів «сканер» і «дисплей», тоді, коли теоретичні дослідження завершувалися у кращому разі перевіркою розроблених алгоритмів на обчислювальній машині на штучних, так званих модельних зображеннях, а практична перевірка розробок передбачала створення спеціалізованих технічних пристроїв, наприклад, автоматів для читання. Така технологія досліджень і розробок неминуче призводила до розриву між теоретичними й прикладними дослідженнями, оскільки вони розділялися як у часі, так і за складом виконавців. Створення комплексу [1] істотно зближувало теоретичні та прикладні дослідження. Дослідник-практик, що ставив за мету розв'язання практичної задачі, працюючи з комплексом, мимоволі починав мислити не в термінах електронно-променевих трубок чи фотопомножувачів, а в термінах масивів даних у пам'яті та операцій над ними, наближаючи таким чином своє мислення до теоретика. Дослідник-теоретик, працюючи з комплексом, на перших же етапах переконувався в нереалістичності певних уявлень, які йому здавалися очевидними — наприклад, що зображення **буквено-цифрового/абетково-цифрового** символу легко подати у вигляді двійкового коду, де 1 і 0 відповідають білим і чорним пікселям.

Єдність теоретичних і прикладних досліджень, про яку ми, слідом за Глушковим, говоримо, — це щось більше, ніж просто узгоджена робота теоретиків і практиків в одному колективі, або прагнення до того, щоб кожне теоретичне дослідження завершувалося практично значущим результатом, або розуміння того, що дослідно-конструкторській розробці має передувати серйозне теоретичне вивчення предмета. Єдність теоретичних і прикладних досліджень у розпізнаванні образів означає зникнення самого поділу дослідників на теоретиків і практиків. Кожен окремий дослідник на кожному етапі розв'язання своєї задачі виступає одночасно і теоретиком, і практиком. Такий стиль роботи у сфері розпізнавання образів був закладений у київській школі ще на етапі її зародження, і зараз, після багатьох років, можна стверджувати, що цей початковий імпульс виявився досить потужним, аби його вплив відчувався й досі.

Характер київської школи в значній мірі визначився ще одним потужним впливом на неї на початкових етапах її становлення. На своїх початкових стадіях розпізнавання образів, як і інші напрями в кібернетиці, а можливо, навіть більше за інші, характеризувалося великим «шумом» у найрізноманітніших значеннях цього слова. Це й гучні заяви, обіцянки на межі міфів, яким, як зараз зрозуміло, не судилося збутися, це й псевдонаукові роботи на рівні інформаційного шуму, коли точні твердження замінюються метафорами, а простіше кажучи — словесними маніпуляціями, міркуваннями на суто якісному рівні, які через свою неточності важко спростувати. Подібні шуми були характерні для багатьох напрямів у розпізнаванні образів, що тоді зароджувалися, але найбільш помітними вони були в перших публікаціях про перцептрон. Такі шуми, звісно, до певного часу пробачаються — подібно до того, як у період дитинства дозволено багато з того, що є неприйнятним у зрілому віці. Проте дитинство не повинно затягуватися, і тим паче ним не слід зловживати. У цьому сенсі важливу терапевтичну роль відіграли роботи [2,3], у яких було доведено, що багато «чудодійних» властивостей, анонсованих у перших публікаціях про перцептрон, такими залишилися лише на словах.

Окрім свого місцевого значення, а саме формального аналізу властивостей перцептрона, роботи [2,3] відіграли важливу методологічну роль. На прикладі перцептрона вони засвідчили, що розпізнавання образів перестало бути майданчиком для словесних вправ і умоглядних міркувань про те, як влаштоване людське мислення, а стало об'єктом наукового дослідження, що ґрунтується на однозначних визначеннях, строго доведених твердженнях, точних формулюваннях задач і обґрунтованих алгоритмах їх розв'язання. Першим прикладом саме такого розв'язання прикладної задачі розпізнавання був кореляційний метод розпізнавання зображень [4], що згодом став основою для цілого ряду прикладних розробок, серед яких найзначнішою була розробка першого в країні читаючого автомата для читання багатосторінкових **буквено-цифрового/абетково-цифрових** текстів [5].

У цій записці неможливо описати всі ті ключові моменти, що визначили формування розпізнавання образів як наукової дисципліни. Її формували багато

тисяч дослідників у всьому світі. Ми згадаємо тут лише ті моменти, що пов'язані з діяльністю київської школи.

У роботі [6], одночасно з працями інших дослідників, задача розпізнавання образів була сформульована як одна із задач статистичної теорії рішень. Таким чином, для задач розпізнавання було сформовано рамки таких формулювань, які можуть бути об'єктом формальних математичних досліджень. Як показав багаторічний розвиток теорії розпізнавання, ці рамки виявилися достатньо адекватними практичним задачам розпізнавання. Статистична теорія розпізнавання і зараз становить один з основних розділів науки про розпізнавання й є однією з найпомітніших сфер застосування математичної статистики та теорії статистичних рішень. Водночас результати, які розпізнавання запозичує з суміжних розділів математики, у даному випадку з математичної статистики, потрапляють у рамках розпізнавання в нетрадиційний для них контекст. Тут вони стикаються з результатами, що надійшли до розпізнавання з інших розділів математики, набувають дещо іншого відтінку, збагачуються й у такому більш насиченому вигляді, так би мовити, з відсотками, повертаються в ту галузь, з якої були запозичені. Найвиразнішими прикладами такого роду є статистична теорія навчання [7] і статистична теорія самонавчання [8], нині відома в світі під назвою EM-алгоритмів [9]. Статистична теорія розпізнавання таким чином демонструє свою зрілість: вона не лише засвоює результати, досягнуті в суміжних галузях, а й збагачує ці галузі.

Однією з ключових подій у розвитку теорії розпізнавання стало те, що під час розв'язання задач розпізнавання рядку тексту [10] та розпізнавання мовленнєвих сигналів [11,12] теорія статистичних рішень була змушена зіткнутися з теорією динамічного програмування. Значення цих робіт давно вже вийшло за межі задач розпізнавання текстів і мовленнєвих сигналів, які були початковим стимулом для їх виконання. Сьогодні важко перелічити всі прикладні задачі, розв'язання яких ґрунтується на методах, вперше описаних у [10,11]. Бібліографія цих робіт налічує багато тисяч публікацій. Роботи [10] і [11] справили потужний вплив на світову науку про розпізнавання й започаткували новий розділ цієї науки, який нині відомий як структурне розпізнавання.

У структурному розпізнаванні йдеться про аналіз об'єктів, що складаються з великої кількості взаємопов'язаних частин. Коли структура цих взаємозв'язків утворює нерозгалужений ланцюг (а саме таку структуру мають текстові рядки та мовленнєві сигнали), задачі розпізнавання зводяться до розв'язання задач динамічного програмування. Цей результат цілком природно узагальнюється і на випадок, коли структура взаємозв'язків не містить циклів [13]. Розпізнавання об'єктів зі складнішою структурою, ніж ланцюг або ациклічний граф, вимагало розробки нового математичного апарата, подібного до контекстно-вільних граматик Н. Хомського, але такого, що є їх двовимірним і багатовимірним узагальненням [13,14].

Інший напрям узагальнення результатів робіт [10,11] веде до значно складніших оптимізаційних задач [15,16], які нині стали відомі як задачі розмітки. Йдеться про

оптимізацію функції від великої кількості дискретних аргументів, яку можна подати у вигляді суми великої кількості доданків, кожен з яких залежить не від усіх аргументів, а лише від їх невеликої підмножини. Загальний алгоритм розв'язання таких задач, найімовірніше, неможливий, оскільки множина всіх можливих задач цього типу утворює NP-повний клас. Ці задачі споріднені з такими класично складними задачами, як задачі здійсненності [17] та псевдобулевої оптимізації [18].

Апарат розміток є адекватним засобом формулювання задач структурного розпізнавання, проте, з огляду на їхню фундаментальну складність, не є універсальним засобом їх розв'язання. Конкретне рішення, точне або наближене, для тієї чи іншої прикладної задачі знаходиться на основі аналізу специфічних особливостей цієї задачі, як це було зроблено при розробці апаратно-програмного комплексу «Теремки» для розпізнавання креслярсько-графічних зображень [19, 20, 21]. Вагомим результатом у цьому напрямі стала розробка апарату еквівалентних перетворень задач розмітки [22, 23] і виявлення на цій основі досить широких розв'язних підкласів цих задач [24]. До них належать деякі задачі розпізнавання великої прикладної значущості, такі як текстурна сегментація зображень і машинний стереозір [25]. Ми бачимо, таким чином, наскільки органічно поєднані прикладні й фундаментальні дослідження в межах розпізнавання образів і наскільки сильно це єдність відрізняється від гранично спрощеного уявлення про те, що розв'язання тієї чи іншої прикладної задачі досягається шляхом застосування відомого математичного засобу. Найважливіші результати в розпізнаванні досягаються тоді, коли розпізнавання образів слугує не лише випробувальним полігоном для відомих математичних засобів, а й середовищем, у якому вони вдосконалюються і, зміцнівши від зіткнення з реальністю, повертаються у свої материнські галузі. Вочевидь, і подальший прогрес залежатиме від того, чи зможе розпізнавання образів саме в такому стилі взаємодіяти з іншими розділами прикладної математики. Судячи з роботи [26], яка містить огляд досліджень київської школи, можна сподіватися, що київська школа розпізнавання, як і раніше, не лише черпатиме результати із суміжних розділів, а й збагачуватиме їх.

Минулих пів століття можна вважати етапом становлення науки про розпізнавання образів, і з певним задоволенням можна констатувати, що київська школа не залишалася осторонь цього процесу. Останні десять років дослідження з розпізнавання образів зосереджені в Міжнародному науково-навчальному центрі інформаційних технологій і систем, що входить до складу Кібернетичного центру. Це десятиріччя ознаменувало становлення розпізнавання образів не лише в суто науковому, але й, що не менш важливо, у науково-організаційному сенсі цього слова. Дослідження проблем ведуться в межах Державної науково-технічної програми «Образний комп'ютер» [27], що свідчить про усвідомлення їх державної важливості. Розпізнавання образів ніколи не страждало від недооцінки свого значення та браку уваги до себе. Розпізнавання образів завжди вважалося важливою складовою теоретичної кібернетики, обчислювальної техніки, автоматизованого проєктування, автоматизації наукових досліджень у біології, матеріалознавстві, космічних дослідженнях, криміналістиці. Важко перелічити всі

можливі галузі застосування розпізнавання образів. Ще важче назвати такі галузі науки й техніки, яким розпізнавання образів не потрібне. Проте в період зрілості для розпізнавання образів уже стало дещо незручним становище, коли його цінність визнається не сама по собі, а лише через його корисність для чогось іншого, нехай і дуже важливого. Формування ДНТП «Образний комп'ютер» якраз і завершує організаційне становлення розпізнавання образів як самостійної наукової дисципліни.

І так само, як і пів століття тому, дослідники київської школи не можуть провести чітку межу між теоретичними та прикладними дослідженнями, що виконуються в межах програми «Образний комп'ютер». Це дослідження машинного стереозору [28,29], текстурної сегментації зображень земної поверхні [30], ідентифікації особи за зображенням її обличчя [31], структурного аналізу зображень документів зі складною ієрархічною структурою [32], побудови просторових моделей будівель і міських територій за їх зображеннями [33] та, звісно ж, інші. Розв'язання всіх цих практичних задач нерозривно пов'язане з їх теоретичним аналізом, який приводить до усвідомлення того, що розпізнавання образів є лише однією складовою в надзвичайно складному механізмі обробки інформації, який можна було б назвати образним мисленням [34]. І так само, як і пів століття тому, київська школа відчуває себе лише на початку дослідження цих механізмів. Що ж, це теж ознака зрілості — бачити не лише точку свого поточного становища, а й той шлях, який ще слід пройти і щодо якого поточне становище є лише початком.

#### Література:

1. Глушков В.М., В.А.Ковалевский, В.И.Рыбак. Универсальная установка для исследования алгоритмов распознавания изображений // Принципы построения самообучающихся систем.- Киев: Гостехиздат УССР, К., 1962. – С. 63-72.
2. Глушков В.М. Теория обучения одного класса дискретных перцептронов // Журн. Вычислительной математики и мат. физики.- 1962.- № 2.- С. 317-335.
3. Глушков В.М. К вопросу о самообучении в перцептроне // Там же. - № 6.- С. 1102-1110.
4. Ковалевский В.А. Корреляционный метод распознавания изображений // Журн. Вычислительной математики и мат. физики.- 1962.- 2, № 4.- С. 684-689.
5. Барашко А.С., Ковалевский В.А. и др. Корреляционный читающий автомат со сдвиговым регистром ЧАРС // Читающие автоматы. – Киев: Наук. думка, 1965.- С. 184-207.
6. Ковалевский В.А. Задача распознавания образов с точки зрения математической статистики // Читающие автоматы. – Киев: Наук. думка, 1965.- С.3-41.
7. Вапник В.Н., Червоненкис А.Я. Теория распознавания образов (статистические проблемы обучения). М.: Наука, 1974. 415 с.
8. Шлезингер М.И. Взаимосвязь обучения и самообучения в распознавании образов // Кибернетика.- 1968.- № 2.- С. 81-88.
9. Demster A., Laird N., Rubin D. Maximum likelihood from incomplete data via the EM algorithm // Journal of the Royal Statistic Society.- 1977.- Vol. B39.- P. 1-38.
10. Ковалевский В.А. Алгоритм разделения машинописной строки на знаки при отсутствии пробелов // III Всесоюзная конференция по информационно-

- поисковым системам и автоматизированной обработке научно-технической информации.- Т. 3.- М.: Всесоюз. ин-т научной и техн. Информации, 1967.- С. 156-164.
11. Винцюк Т.К. Распознавание устной речи методами динамического программирования // Кибернетика.- 1968. - № 1. – С. 81-88.
  12. Винцюк Т.К. Анализ, распознавание и интерпретация речевых сигналов.- Киев: Наук. думка, 1987.- 262 с.
  13. Шлезингер М.И., Главач В. Десять лекций по статистическому и структурному распознаванию. – Киев: Наук. думка, 2004, 546 с.
  14. Шлезингер М.И. Двумерное обобщение контекстно-свободных языков и грамматик // Методы и средства информатики речи.- Киев: Ин-т кибернетики НАН Украины, 1991.- С. 36-47.
  15. Шлезингер М.И. Синтаксический анализ двумерных зрительных сигналов в условиях помех // Кибернетика.- 1976.- № 4.- С. 113-130.
  16. Коваль В.К., Шлезингер М.И. Двумерное программирование в задачах анализа изображений // Автоматика и телемеханика.- 1976.- № 8.- С. 149-168.
  17. Jeavons P.G., Cohen D.A., Cooper M.C. Constraints, consistency and closure. *Artificial Intelligence*, 101(12):251-265, 1998.
  18. Boros E., Hammer P.L. Pseudo-Boolean Optimization // *Discrete Applied Mathematics*.- 2002.- Vol. 123, 1-3, (November 2002).- P. 155-225.
  19. Мацелло В.В., Провалов Ю.Л., Шарыпанов В.М., Шлезингер М.И. Система ввода и обработки графической информации // УСиМ.—1986. – №. 6. – С. 66-60.
  20. Мацелло В.В. Система автоматизированного кодирования эскизов принципиальных схем. // УСиМ.—1989. – №. 2. – С. 108-111.
  21. Кийко В.М. Автоматическое кодирование изображений слоев топологии печатных плат // УСиМ.—1988. – №. 5. – С. 86-89.
  22. Шлезингер М.И., Гигиняк В.В. Решение (max,+)-задач структурного распознавания с помощью их эквивалентных преобразований I // УСиМ.—2007. – №. 1. – С. 3-15.
  23. Шлезингер М.И., Гигиняк В.В. Решение (max,+)-задач структурного распознавания с помощью их эквивалентных преобразований II // УСиМ.—2007. – №. 2. – С. 5-17,
  24. Schlesinger M. I., Flach B. Some solvable subclasses of structural recognition problems. // *Czech Pattern Recognition Workshop*. – Praha. – 2000. – P. 55-62.
  25. Schlesinger M. I., Flach B. Analysis of optimal labelling problems and their application to image segmentation and binocular stereovision // *Proceedings East-West-Vision 2002 (EWV'02)*, Franz Leberl, Andrej Ferko (eds.), International Workshop & Project Festival on Computer Vision, Computer Graphics, New Media.- P. 55-60.
  26. Werner T. A Linear Programming Approach to Max-sum Problem: // A review. <ftp://cmp.felk.cvut.cz/pub/cmp/articles/werner/Werner-TR-2005-25.pdf>. (Будет опубликована в *Pattern Analysis and Machine Intelligence* в мае 2007).
  27. Винцюк Т.К. Генеративна модель образного комп'ютера // Шоста Всеукраїнська міжнародна конференція (УкрОБРАЗ'2002): Праці.- Київ: Укр. Асоціація з оброблення інформації та розпізнавання образів, 2002.- С. 7-14.

- 28.Рябокoнь Д.И. Прoстрaнствeннaя рeкoнстpукция пoверхнoстей пo стeрeопaрe изoбpaжeний с пoмoщью алгoритмoв пoискa минимaльнoгo сeчeния нa гpaфe // УСиМ.—2004. – №. 3. – С. 47-51.
- 29.Рябокoнь Д.И. Вoсстaнoвлeниe прoстрaнствeннoй кoнфигуpации oбъeктoв и сцeн пo их стeрeoизoбpaжeниям // УСиМ.—2005. – №. 1. – С. 22-31.
- 30.Кoвтун И.В. Тeкстурнaя сeгмeнтaция изoбpaжeний нa oснoвaнии мaркoвскиx случaйннoх пoлeй// УСиМ.—2003. – №. 4. – С. 46-55.
- 31.Кийкo В.М., Кийкo К.В., Мaцeллo В.В., Прoвaлoв Ю.Л., Шaрипaнoв В.М. Систeмa дoступу дo приміщeння нa oснoві рoзпiзнaвaння людськиx oблич // Вoсьмa Вceукрaїнськa мiжнaрoднa кoнфeрeнцiя (УкрOБPАЗ'2006), Киiв, 28-31 сeрпня 2006: Пpaci.- Киiв: Мiжнaрoдний нaуковo-нaвчaльний цeнтp iнфoрмaцiйних тeхнoлoгiй тa систeм, 2006.- С. 123-126.
- 32.Шлeзiнгeр М.І., Сaвчинський Б.Д., Aнoхiнa М.О. Синтaкcичний aнaлiз тa рoзпiзнaвaння дpукoвaних нoтних тeкстiв // Упpавляючiєся систeми и мaшини. – 2003.- № 4.- С. 30-38.
- 33.Рябокoнь Д.І. Ствoрeння тpивимiрних мoдeлeй oб'єктiв зa стeрeo пaрaми зoбpaжeнь для eлeктpoнних мyзeiв // Eлeктpoннi зoбpaжeння тa вiзуaльнi мистeцтвa EVA 2002, Киiв, 22-24 тpавня 2002 рoкy: Пp. Кoнф.- Киiв 2002.- С. 61-68.
- 34.Рoзрoбкa кoнцeпцiї oбpaзнoгo мислeння в iнтeлeктуaльних iнфoрмaцiйних тeхнoлoгiях: Звiт пpо НДP, викoнaну в рaмкaх пpогpaми ДНТП „Oбpaзний кoмп'ютeр” / Мiжнaрoдний нaуковo-нaвчaльний цeнтp iнфoрмaцiйних тeхнoлoгiй тa систeм; Нaуковий кeрiвник Шлeзiнгeр М.І.- № ДP 0104U008470; iнв. № 0205U004231.- К., 2004.- 231 с.