

АЛГОРИТМ РОЗПІЗНАВАННЯ ОБЛИЧЧЯ ЛЮДЕЙ НА БАЗІ ЗГОРТКОВОЇ НЕЙРОННОЇ МЕРЕЖІ

Анотація: В статті розглянуто методи детектування та розпізнавання облич, проаналізована можливість їх комбінації з метою побудови комп'ютерної системи розпізнавання. Отриманий результат дозволив виділити переваги та недоліки при проектуванні системи розпізнавання у розглянутій області застосування. Розроблено алгоритм та реалізована програма ідентифікації обличчя особи.

Ключові слова: система ідентифікації, розпізнавання обличчя, ключові точки і ознаки, нейронна мережа, виділення області, помилкові спрацьовування.

Вступ

Вперше завдання розпізнавання обличчя виникло в криміналістиці. У зв'язку з цим були розроблені математичні методи і моделі як обличчя так і процедури розпізнавання особи, що можна назвати завданням верифікації. Добре відома величезна різноманітність факторів, що призводять до внутрішньокласової мінливості об'єкту розпізнавання: зміна освітлення, зміна положення, міміка, візит до перукаря – це все призводить до погіршення якості роботи детектора обличчя. Тим не менше, значна кількість досліджень, що відбуваються у лабораторіях усього світу протягом декількох десятиліть, так і не посприяла створенню вимог до реально працюючих систем комп'ютерного зору, здатних виявляти і розпізнавати людину в будь-яких умовах та позиціях. Сьогодні потреба в розпізнаванні загострюється в 5 напрямках: системи безпеки, банківська сфера (як для безпеки, так і при роботі з клієнтами), підтвердження особи, соціальні мережі і як особистий пароль.

Які можливі варіанти при розпізнаванні особи? Першу ситуацію можна описати як бажання об'єкта бути розпізнаним (наприклад на терміналі прикордонної служби або для визначення - чи це саме та людина, яка хоче зайти в приміщення, має на це дозвіл, а не якась інша). Інша назва такої ситуації: система з кооперацією – коли людина сама допомагає при такій процедурі.

Другий варіант ситуації – це відсутність кооперації суб'єкта з системою візуального моніторингу, наприклад камера на вулиці або прихована камера в банку. Цей варіант стикається з проблемою схожості людей один на одного. У найбільш складному випадку, при використанні системи виявлення та ідентифікації людини по зображенню його обличчя в обстановці, коли швидко змінюється сцена подій або загальний фон, з великим потоком вхідних даних (моніто-

ринг на міських вулицях з інтенсивним рухом, в метро, аеропортах і т.п.), вимагається використання максимально доступної додаткової інформації для досягнення задовільних результатів роботи алгоритму.

Постановка задачі

Метою роботи є розробка системи ідентифікації людини шляхом розпізнавання її обличчя в ситуації з кооперацією. Кооперація впливає з області застосування, наприклад, пропускний пункт охорони підприємства чи виявлення обличчя людини з вуличного потоку авто. Для виконання поставленої мети були розглянуті наступні завдання: проведення аналізу існуючих алгоритмів та підходів для розпізнавання облич; за отриманими результатами розробити більш ефективний метод та сформулювати вимоги до системи розпізнавання; реалізація та тестування розробленого методу.

Задача автоматичного розпізнавання особи, яка проходить крізь пункт пропуску за допомогою картки, враховує обмеження більшості підприємств – кількість осіб, що підлягатимуть розпізнаванню є постійним числом, що спрощує таке завдання і надає можливість зосередитись на найважливішому.

Загальний підхід до розробки системи розпізнавання

У загальному випадку алгоритм вирішення задачі виявлення та ідентифікації людини по зображенню його обличчя складається з наступної послідовності очевидних кроків: виявлення факту присутності людини на сцені, виділення фігури людини, виділення голови, визначення ракурсу спостереження голови (анфас, профіль), виділення особи, порівняння з еталонами кінцева ідентифікація.

Завдяки деяким алгоритмам можна відразу перейти до кроку визначення області на якій знаходиться обличчя людини. До таких відносять методи, що засновані на частинних ознаках: контурний аналіз, класифікацію по кольору шкіри, пошук/розпізнавання анатомічних ознак. Хоча слід відзначити: кращі результати можна отримати за допомогою методів, що засновані на узагальнених ознаках та використовують мета-алгоритми AdaBoost, RealBoost та «слабкі» класифікатори для побудови «сильних» класифікаторів, до них відносять: HAAR-Like features, LBP, Gaussians fields.

Так загальний порядок задач, що потрібно вирішити при створенні системи розпізнавання представлений на рис. 1.

При наявності зображення або відео потоку, розбитого на кадри, перше завдання яке виникає – де на зображенні знаходиться особа або особи, яких треба знайти. Визначення ключових точок і нормалізація обличчя (виділені прямокутником на рис. 1.) два надзвичайно важливих і пов'язаних між собою пункти, що мають найбільший вплив на якість і результати розпізнавання.

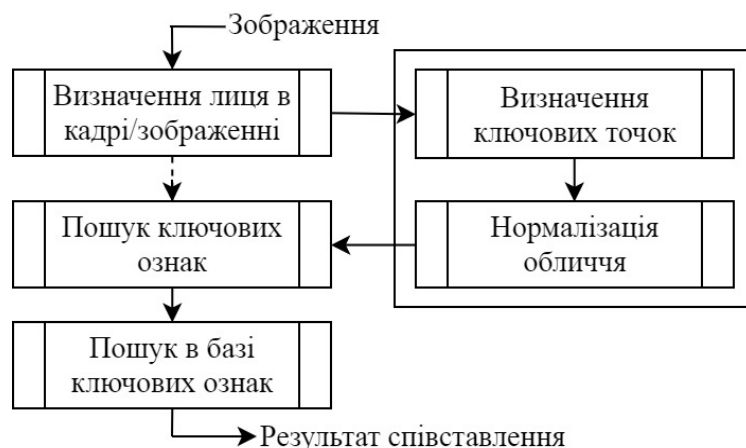


Рис. 1 – Послідовність завдань, що виникають при розпізнаванні обличчя

Аналіз існуючих методів ідентифікації людини

Локалізацію обличчя прийнято здійснювали детектором Віоли-Джонса, що в чистому вигляді будується за допомогою Machine Learning [1]. Він використовує каскад Хаара, який є набором примітивів, для яких розраховують їх згортку з зображенням. Використовуються найпростіші примітиви, що складаються з прямокутників і мають всього два рівні, +1 і -1. При стандартному розмірі примітива в 24x24 пікселя, можливі 162 тисячі [1] різних ознак і занадто довга процедура розрахунку. Тому в алгоритмі Віоли-Джонса використовується варіація алгоритму навчання AdaBoost, як для вибору ознак, так і для настройки класифікаторів [2]. Так, наприклад, очі будуть темніші, ніж область між ними, так само як область брів чи рота буде темнішою ніж лоб. Приклад таких примітивів зображений на рис. 2.

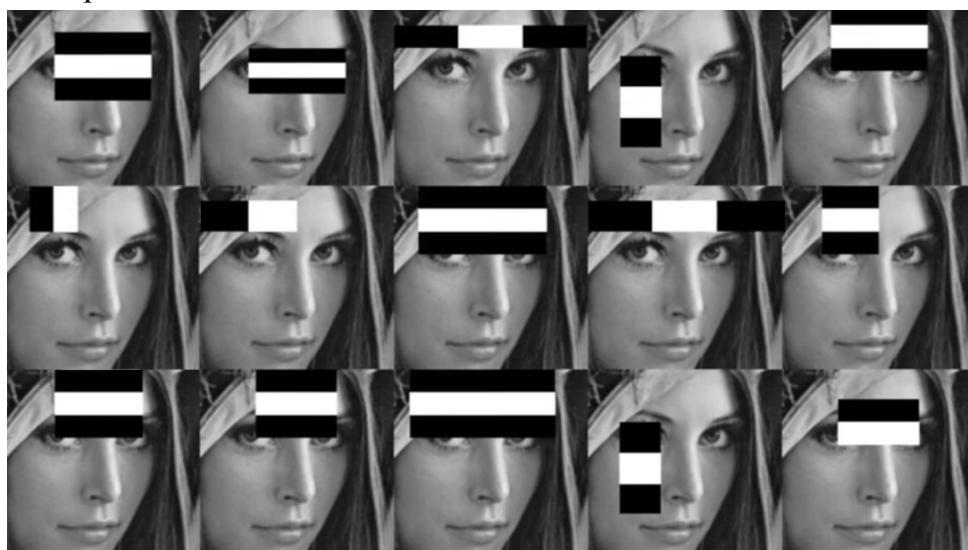


Рис. 2 – Примітиви Хаара

Обчислення згортки відбувається циклічно “ковзаючим вікном” чи вікном змінного розміру, або при постійному розмірі вікна змінюється розширення зображення. Важливим є уникання отримання помилкових оцінок, для цього помилка другого роду повинна бути менше 10^{-7} (бо при наявності 10^5 вікон буде вже 10 помилок, що є неприпустимим).

Наприклад, налаштований типовий дескриптор з відкритої бібліотеки OpenCV має 38 каскадів, що аналізує 6061 ознак. Коли на першому етапі використовуються лише дві ознаки, він відкидає 50% точок не обличчя при точності в 100%; другий етап – 10 ознак вже відкидає 80% не обличчя при точності близько 100%; на 3 і 4 етапах - по 25 ознак і т. д. [3].

Вибір алгоритму, який використовується для ідентифікації людини по зображенню його обличчя, також залежить від конкретних умов його застосування. Завдання виявлення конкретної людини в натовпі вимагає застосування зображень високої якості та витончених методів для зниження рівня помилкових спрацювань. При цьому потік інформації з відеокамери може значно збільшуватися, що може призводити до вимоги додаткових обчислювальних потужностей (наприклад, застосування комп'ютерного кластера) чи використання на першому етапі поточної ідентифікації фільтрованих зображень з меншою кількістю пікселів, а на другому – більш якісних.

Для наших умов розглянемо класичний та відносно новий методи ідентифікації.

Метод головних компонент. Для завдання розпізнавання має наступний вигляд: вхідні вектори – це відцентровані і приведені до єдиного масштабу зображення обличчя. Власні вектори, обчислені для всього набору зображень обличчя, називаються власними обличчями (eigenfaces). Вхідне зображення розкладається на набір лінійних коефіцієнтів, що називають головними компонентами. Сума N перших головних компонент, помножених на відповідні власні вектори, є апроксимацією зображення порядку N і зазвичай береться від 5 до 200 головних компонент. Інші компоненти кодують дрібні відмінності між особами і шум. Процес розпізнавання полягає в порівнянні головних компонент нерозпізаного зображення з компонентами всіх інших зображень. Для цього зазвичай застосовують будь-яку метрику (найпростіший випадок – евклідова відстань) [4].

Цей метод вимагає єдиних параметрів освітленості, нейтрального виразу обличчя, відсутності перешкод (окуляри чи борода). При недотриманні цих умов головні компоненти не будуть відображати міжкласові варіації. Наприклад, при різних умовах освітленості перші головні компоненти переважно відображають зміни освітлення, а порівняння видає зображення, що мають схожий рівень освітленості, і метод стає практично непридатним для роботи. Обчислення набору власних векторів відрізняється високою трудомісткістю. Тому цей метод можна використовувати лише в ідеальних умовах.

Згорткова нейронна мережа має деяку специфіку – вона полягає у перестановці згорткових шарів та шарів субдискретизації (пулінгу). Структура мережі багат шарова – однонаправлена (беззворотних зв'язків). Для навчання використовуються стандартні методи, частіше всього метод зворотного розповсюдження помилок. Функція активації нейронів (передаточна функція) – будь-яка, за вибором дослідника. Назву мережа отримала через наявність операцій згортки.

Після кількох проходжень згортки по зображенню і ущільнення за допомогою пулінгу система сама перебудовується від конкретної сітки пікселів з високою роздільною здатністю до більш абстрактних карток ознак. Як правило, на кожному наступному шарі збільшується число каналів і зменшується розмірність зображення в кожному каналі. Зрештою залишається великий набір каналів, що зберігають невелику кількість даних (навіть один параметр), які інтерпретуються як самі абстрактні поняття, виявлені з вихідного зображення.

Ці дані об'єднуються і передаються на звичайну повнозв'язну нейронну мережу (НМ), яка теж може складатися з декількох шарів. При цьому повнозв'язні шари вже втрачають просторову структуру пікселів і мають порівняно невелику розмірність [5].

Переваги згорткової НМ:

- Невелика кількість вагових коефіцієнтів, оскільки одне ядро використовується для всього зображення, замість того, щоб робити для кожного пікселя вхідного зображення свої вагові коефіцієнти. При навчанні відбувається узагальнення демонстрованої інформації, а не запам'ятовування кожної показаної картини у вигляді міриад вагових коефіцієнтів.

- Відносна стійкість до повороту і зсуву зображення.
- Зручне розпаралелювання обчислень, а, отже, можливість реалізації алгоритмів роботи і навчання мережі на графічних процесорах [6].

Експериментально-дослідне вирішення задачі

Для вирішення поставленого завдання ідентифікації особи за зображенням її обличчя та виявлення недоліків, було розроблено декілька версій ПЗ, що реалізує описані вище методи, а саме: локалізація обличчя здійснювалася алгоритмом Віоли-Джонса, а безпосередньо для розпізнавання використовувався метод Eigenfaces (головних компонент) або згорткова НМ.

На рис. 3. а-д зображена робота детектора Віоли-Джонса, з чого видно, що фронтальні обличчя чудово відшукуються, деякі труднощі виникають при повороті голови або зміні ракурсу. Також використання перешкод під очима та засобів маскування по типу капелюха, окулярів чи навіть пов'язки на око не змогли «приховати» обличчя.

Навіть при кількості у сотню головних компонент результати отримані при розпізнаванні вже не настільки успішні (рис. 3. *е-ж*). Тестування показало, що враховуючи, що в базі обличч була лише одна особа, точність для слабко зашумлених зображень з навчаючої вибірки склала 94-100%, а ось для звичайних фото точність розпізнавання варіювалася від 48% для фото в «профіль» і до 0,6% з повернутим обличчям на $10-25^{\circ}$.

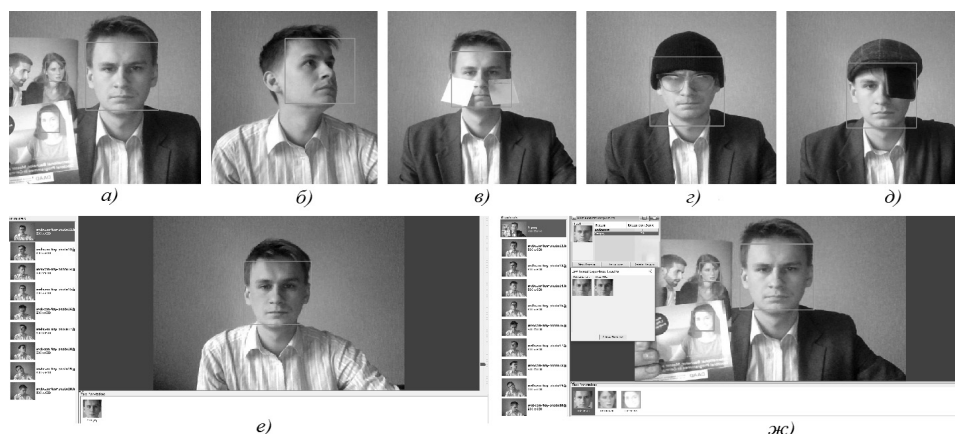


Рис. 3 – Результати знаходження та розпізнавання обличчя

Додаткова точність досягалася при використанні отриманих областей методом Віюлі-Джонса для більш швидкого заповнення бази осіб. Враховуючи дані результати програма буде часто давати хибні результати, і, незважаючи на її простоту, вона підкреслює недоліки класичних систем ідентифікації особи за зображенням.

Комбінування методу Віюлі-Джонса та згорткової НМ дало вже набагато кращі результати, що зображено на рис. 4.. Враховуючи те, що нейронна мережа проходила навчання на зображеннях фіксованого розміру (кількість пікселів має бути постійною, оскільки вона і є кількістю входів мережі), виникала необхідність нормувати зображення з камери і, таким чином, приводити його до розмірів, на яких навчалася мережа.

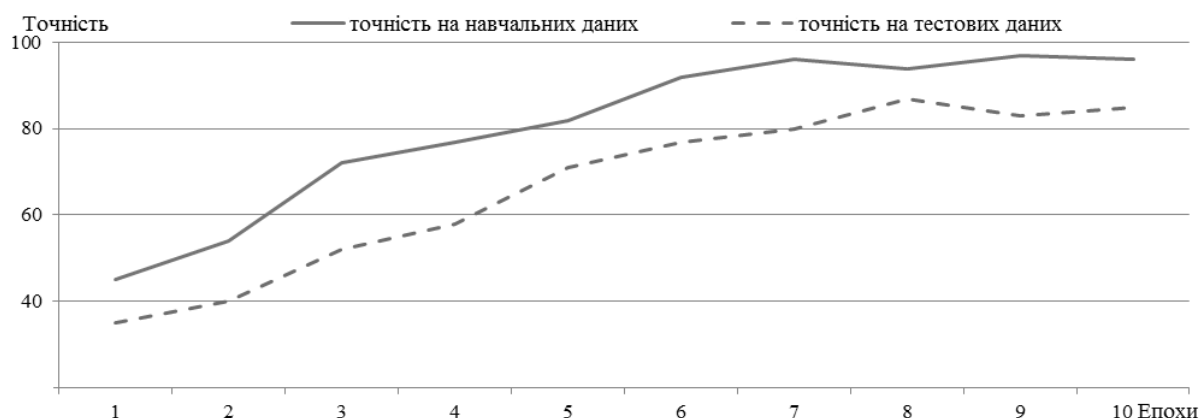


Рис. 4 – Точність розпізнавання при навчанні мережі

Нормалізація здійснювалася за допомогою детектору Віюлі-Джонса - виділена область стискалась до необхідних розмірів за допомогою методів бібліотеки OpenCV, далі зображення передавалося на вхід згорткової НМ, а на виході активувався нейрон, що відповідав за певну особу (алгоритм комбінованої обробки за допомогою згорткової НМ або алгоритму Eigenface зображений на рис.5). Внаслідок неминучих спотворень при нормалізації точність падала, і складала 80-85%. Враховуючи малий розмір навчаючої вибірки у 50 зображень та після 10 епох навчання з підвибірками по 10 зображень, отримані результати наочно підтверджують ефективність та перспективність використання згорткових НМ.

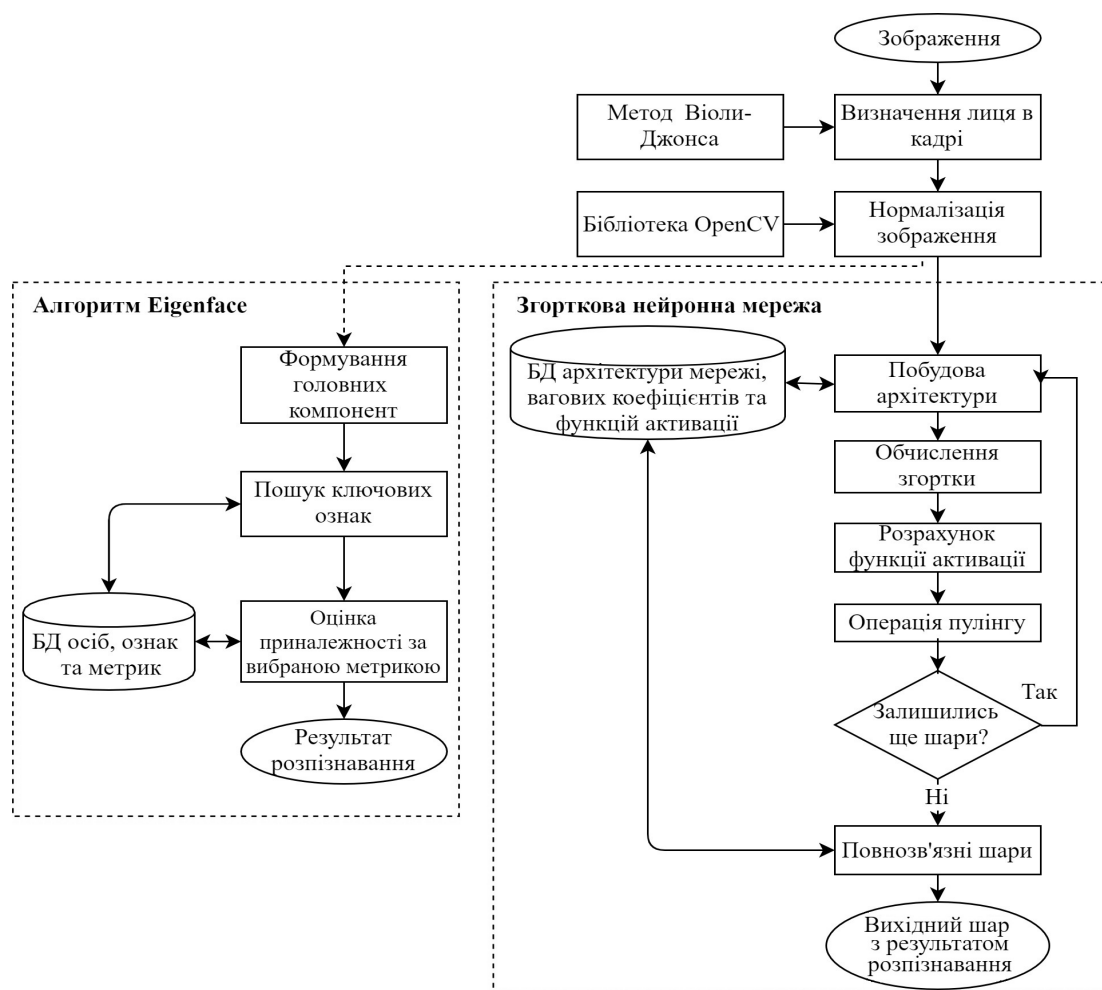


Рис. 5. Структура алгоритму комбінованої обробки зображень

Використовуючи таку систему слід пам'ятати, що результат є ймовірністю, перенавчання мережі внаслідок додавання нового члена колективу може займати багато часу, тому виникає потреба у резервному копіюванні хоча б вагових коефіцієнтів. Для нейронних мереж, у яких база осіб по факту є значеннями їх

вагових коефіцієнтів, швидкість обробки зменшується при додаванні нових осіб, що обумовлює перенавчання мережі, а такий процес надзвичайно ресурсномісткий. Велику проблему для систем безпеки складають також помилкові спрацювання.

Для розглянутої задачі можна підібрати метод та побудувати систему розпізнавання, що даватиме задовільні результати. Теоретично можна збільшити швидкодію шляхом передачі нормалізованого зображення на вхід мережі після успішного проходження лише половини каскадів Хаара.

Висновки

Отримані результати показали перспективність використання комбінації двох методів – за допомогою детектору Віоли-Джонса та згорткової НМ. Розглянуті та реалізовані вище методи підкреслюють обмеженість використання методів розпізнавання з використанням комп'ютерного зору. Так обмеження, що накладені на детектори обличчя створюють парадокс: швидкі алгоритми, що добре визначають область з обличчям, але знаходять його добре лише у фронтальному положенні з мінімальною інваріантністю, а методи, що знаходять обличчя у будь-яких умовах з високою точністю вимагають великих обчислювальних витрат. До того ж необхідний швидкий і точний пошук по базі даних осіб, так як накладаються обмеження на виділення вектору ознак і при декількох невірно визначених ознаках метод може взагалі не спрацювати.

Тому питання про створення класу універсальних швидкодіючих, інваріантних до перешкод та точних алгоритмів ідентифікації особи за зображенням її обличчя залишається відкритим.

Список використаних джерел

1. P. Viola and M.J. Jones, «Rapid Object Detection using a Boosted Cascade of Simple Features», proceedings IEEE Conf. on Computer Vision and Pattern Recognition (CVPR 2001), 2001
2. P. Viola and M.J. Jones, «Robust real-time face detection», International Journal of Computer Vision, vol. 57, no. 2, 2004., pp.137–154
3. Bradsky G., Kaehler A. Learning OpenCV — O'Reilly, 2008. — P. 580 — ISBN 978-0-596-51613-0
4. M. Turk; A. Pentland (1991). "Face recognition using eigenfaces". Proc. IEEE Conference on Computer Vision and Pattern Recognition. pp. 586–591.
5. Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard and L. D. Jackel: Backpropagation Applied to Handwritten Zip Code Recognition, Neural Computation, 1(4):541-551, Winter 1989.
6. Matusugu, Masakazu (2003). «Subject independent facial expression recognition with robust face detection using a convolutional neural network». Neural Networks 16 (5): 555–559. DOI:10.1016/S0893-6080(03)00115-1.