

Дидковская М.В., Гранаткина Т.Д.

Учебно-научный комплекс «Институт прикладного системного анализа»,  
Национальный технический университет Украины  
«Киевский Политехнический Институт»

## ВЫЯВЛЕНИЕ ГРЯЗНЫХ ДАННЫХ В ХРАНИЛИЩАХ ДАННЫХ

### Аннотация

Исследована проблема выявления грязных данных в хранилищах данных. Проведен сравнительный анализ функций мер подобия записей. Рассмотрены подходы к выявлению грязных данных. Предложен алгоритм метода дедубликации данных.

**Ключевые слова:** хранилище данных, грязные данные, дедубликация, очистка данных, мера подобия, качество данных.

Didkovska M.V., Granatkina T.D.

Educational and Scientific Complex «Institute for Applied Systems Analysis»  
National University of Ukraine  
«Kyiv Polytechnic Institute»

## DIRTY DATA DETECTION IN DATA WAREHOUSES

### Summary

Dirty data detection problem in data warehouses was studied. Comparative analysis of similarity measure functions was performed. Methods of dirty data detection were reviewed. Algorithm for deduplication method was proposed.

**Keywords:** data warehouse, dirty data, deduplication, data cleansing, similarity measure, data quality.

УДК 004.932.72

## МЕТОД ЛОКАЛІЗАЦІЇ СТРУКТУРНИХ ОБ'ЄКТІВ НА ФОТОГРАФІЇ

Мельник В.В.

Національний технічний університет України  
«Київський політехнічний інститут»

Запропоновано метод локалізації структурних об'єктів на фотографіях з використанням нелінійної моделі. Метод застосовано для знаходження ключових точок тіла людини. Для даної задачі встановлено покращення метрик якості.

**Ключові слова:** структурні об'єкти, локалізація, нелінійна модель.

Одними з основних задач комп'ютерного зору є виявлення та локалізація об'єктів. Ці задачі слугують базовим блоком для алгоритмів сегментації, відстеження об'єктів та розпізнавання дії на відео.



Рис. 1. Структурна модель

Джерело: розроблено автором

Структурні об'єкти являються найбільш складним випадком для локалізації. Під структурним об'єктом ми будемо розуміти множину ключових точок та апріорні закони їх розміщення в тривимірному просторі. Дані точки являються візуально інваріантними, а їх кількість – фіксованою. Внаслідок рухливості ключових точок, їх відносне положення на площині фотографії не є фіксованим[3]. Процес проектування на площину також створює додаткові деформації розмірів частин об'єкту. Крім того, частина точок може бути перекрита іншими елементами сцени або об'єкта. Найбільш важливі приклади структурних об'єктів: кисть, обличчя, тіло людини. Знаходження елементів кисті необхідно для розпізнавання жестів людини. Локалізація ключових точок обличчя застосовується для обчислення «хешу» – індивідуальної інваріантної до проектування та поворотів характеристики, яку зручно використовувати для задачі ідентифікації. Виявлення положення частин тіла людини важливе для створення сучасних людино-машинних інтерфейсів, розпізнавання жестів. В даній роботі метод локалізації структурних об'єктів буде застосовано для пошуку ключових точок людини.

Структурною моделлю об'єкту будемо називати множину ключових точок та фіксований ациклічний однозв'язний граф, побудований на цих точках [1]. Даний граф задає взаємозв'язки між точками (Рис. 1).

Вимога ациклічності є необхідною для існування ефективної обчислювальної схеми. Граф, що задовольняє дані умови – дерево. Для задачі локалізації людини, в якості кореня дерева оберемо ключову точку, що відповідає маківці голови. Цей вибір не впливає на кінцевий результат, але від нього залежить порядок обчислень. Введемо орієнтацію ребер – від кореня до листів рис. 2.

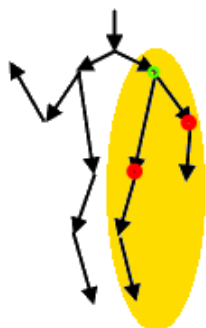


Рис. 2. Топологія структурної моделі

Джерело: розроблено автором

Підграф, що можна досягти рухаючись по орієнтованим ребрам з певної фіксованої вершини, назовемо піддеревом. Наприклад, підграф в жовтому еліпсі є піддеревом з коренем в «зеленій» вершині. Вершини, в які виходять ребра фіксованої вершини, назвемо дочірніми для неї. Наприклад, «червоні» вершини являються дочірніми для «зеленої» вершини.

Кожна ключова точка задається своїм положенням  $(x_i, y_i)$  та прихованим станом  $t_i$ . Прихований стан кодує інформацію про орієнтацію ключової точки (частини об'єкту) у просторі. Структуру, що задає усі ключові точки, будемо називати скелетом об'єкту.  $S = \{s_1, \dots, s_n\}$ ,  $s_i = ((x_i, y_i), t_i)$

Введемо функцію оцінки скелету, що визначає узгодженість скелета на зображенні з апріорним знанням про вигляд об'єкта на фотографіях.

$$F(I, S) = F(\vec{w}, \vec{\beta}, I, S) \rightarrow \mathbb{R}$$

Де  $I$  – зображення,  $S$  – скелет,  $\vec{w}, \vec{\beta}$  – параметри.

Оптимізуючи параметри функції ми прагнемо досягти наступних властивостей:

- Якщо скелет повністю відповідає об'єкту, значення більше 1
- Якщо скелет відповідає фону, значення менше -1
- В інших випадках – проміжні значення.

Представимо функцію в вигляді суми:

$$F(\vec{w}, \vec{\beta}, I, S) = \sum_{(i,j) \in Edge} w_{i,t_i,t_j} + \sum_{(i,j) \in Edge} d_{i,j}(s_i, s_j) + \sum_i w_i f(\vec{\beta}_i, I, s_i)$$

Перша група доданків штрафую скелет за вибір несумісних типів для сусідніх ключових точок. Кожен з доданків групи – константа вибрана з таблиць сумісності типів.

Друга група доданків оцінює відносно положення точок. Кожен з доданків групи – функція з параметрами. Параметри задають очікуване відносне положення точок. Якщо точки розташовані не ідеально, нараховується штраф пропорційний квадрату похибки.

$$d_{i,j}(s_i, s_j) = w_{t_i t_j A} (\Delta x - w_{t_i t_j B})^2 + w_{t_i t_j C} (\Delta y - w_{t_i t_j D})^2 =$$

$$= (\vec{w}, (\Delta x^2, \Delta x, \Delta y^2, \Delta y)^T)$$

Третя група доданків консолидує інформацію про відповідність вигляду точки на зображенні з її типом. Кожна функція  $f(\vec{\beta}_i, I, s_i)$  задається у вигляді нейрону з нелінійною функцією активації (гіперболічний тангенс). Кожній ключовій точці та кожному типу відповідає свій набір ваг.

$$f(\vec{\beta}_i, I, s_i) = \tanh(\vec{\beta}_{t_i}, \vec{x})$$

Де  $\vec{x}$  – об'єднання усіх каналів по всім пікселям зображення в околі ключової точки. Для спрощення обчислень окіл має квадратну форму.

Зображення у просторі RGB не є зручним для пошуку ключових точок людини оскільки люди мають різний колір шкіри, мають різнобарвний одяг та використовується різноманітне освітлення, але існують характерні людські форми та силуети. Тому необхідно перейти в ознаковий простір, який утилізує інформацію про кордони на зображенні. В алгоритмі використовується ознаковий простір на основі HOG[4].

Алгоритм перетворення зображення:

1. В кожному пікселі за кожним каналом підраховується градієнт.
2. Градієнтом пікселя вважається вектор з найбільшою довжиною.
3. Зображення розбивається на прямокутну сітку з сегментів.
4. В кожному сегменті будується гистограма на 18 інтервалів. Вектор потрапляє до одного з інтервалів в залежності від напрямку.
5. Кожен інтервал акумулює суму довжин векторів.
6. Для кожного сегменту підраховується ознаковий вектор з наступними групами ознак:
  - a. інтервальні суми – 18 ознак
  - b. сума протилежних інтервалів – 9 ознак
  - c. сума всіх довжин нормована різними способами – 4 ознаки
7. Локальна нормалізація ознак.

Наведемо алгоритм для локалізації у випадку відомого масштабу.

1. Конвертуємо зображення в ознаковий простір
2. Серед усіх можливих скелетів знаходимо скелети, для яких оціночна функція приймає значення більше порогу

3. Якщо два скелети перетинаються відповідними частинами, то відкидаємо скелет з меншою оцінкою

Алгоритм локалізації у випадку невідомого масштабу.

1. Будуємо піраміду зображень рис. 3. Піраміда складається з копій оригінального зображення з поступовим зменшенням масштабу.
2. На кожному масштабі, виділяємо скелети, для яких оціночна функція приймає значення більше порогу
3. Об'єднуємо множини скелетів з усіх масштабів
4. Якщо два скелети перетинаються відповідними частинами, то відкидаємо скелет з меншою оцінкою



Рис. 3. Піраміда зображень

Джерело: <http://goo.gl/TTdJXX>

Розглянемо наведений алгоритм детальніше. Нехай для певного зображення існує скелет з великим значенням оціночної функції (більше порогу). Якщо частину точок скелету перемістити на незначну відстань, то відносно положення майже не зміниться, а отже майже не зміниться оціночна функція. Таким чином, ми може побудувати велике сімейство скелетів з оцінкою більше порогу. Але на наступному кроці всі ці скелети будуть відкинуті і залишаться лише початковий. Це потенційна неефективність.

Для вирішення цієї проблеми було запропоновано наступну оптимізацію. Згрупуємо скелети по положенню маківки голови. Обчислимо в кожній групі скелет з найбільшою оцінкою:

$$S(s) = \underset{S, s_1=s}{\operatorname{argmax}} F(\vec{w}, \vec{\beta}, I, S)$$

Тоді результат роботи алгоритму можна представити в наступному вигляді:

$$\text{Result} = \{S(s) \mid \forall s, F(\vec{w}, \vec{\beta}, I, S(s)) \geq \text{threshold}\}$$

Очевидно, що ця оптимізація не змінює кінцевого результату.

Для побудови ефективної обчислювальної схеми для знаходження максимуму в групі скелетів переписемо оціночну функцію в зручному вигляді:

$$F(\vec{w}, \vec{\beta}, I, S) = \sum_{(i,j) \in \text{Edge}} (w_{i,j} + d_{i,j}(s_i, s_j)) + \sum_i w_i f(\vec{\beta}_i, I, s_i)$$

Функція розпадається на дві групи доданків. Перша група залежить від пар точок на кінцях ребер, а доданки другої – залежать тільки від однієї точки. Отже, ми можемо максимізувати цю функцію по частинах. Для цього введемо допоміжну функцію  $G(s_j)$ , що обчислює частковий максимум по піддереву.

$$G(s_k) = \max_{s_i \in \text{SubTree}_k} \left( \sum_{(i,j) \in \text{SubTree}_k} (w_{i,j} + d_{i,j}(s_i, s_j)) + \sum_{i \in \text{SubTree}_k} w_i f(\vec{\beta}_i, I, s_i) \right)$$

Тоді ми можемо позбутись від функції  $F(\vec{w}, \vec{\beta}, I, S)$ .

$$\max_{S, s_1=s} F(\vec{w}, \vec{\beta}, I, S) = G(s_1)_{s_1=s}$$

Переписемо  $G(s_j)$  в рекурентному вигляді:

$$G(s_j) = w_j f(\vec{\beta}_j, I, s_j) + \sum_{s_i \in \text{Child}_j} \max_{s_i} (w_{i,j} + d_{i,j}(s_i, s_j) + G(s_i))$$

Помітимо, що кожен з доданків можна максимізувати незалежно від інших. З даного запису випливає очевидний спосіб обчислення  $G(s_j)$  за допомогою динамічного програмування

Припустимо, що для всіх дочірніх вузлів  $G(s_j)$  вже пораховано, а їх кількість константа. Тоді асимптотика часу підрахунку одного доданку –  $O(ST)$ , де  $S$  – площа зображення,  $T$  – кількість типів для точки. Отже, час підрахунку  $G(s_j)$  для всіх положень одні-

єї точки  $O((ST)^2)$ . Але застосувавши алгоритм узагальненого перетворення відстані, ми можемо для всіх  $s_j$  знаходити  $\max_{s_i} (w_{i,j} + d_{i,j}(s_i, s_j) + G(s_i))$  за  $O(ST^2)$ .

Використовуючи вищезгаданий алгоритм, сумарний час роботи алгоритму локалізації складає  $O(NST^2)$ , де  $N$  – кількість ключових точок.

Для знаходження коефіцієнтів оціночної функції застосовуються методи машинного навчання. Навчання моделі зводиться до складної оптимізаційної задачі. Якщо застосувати класичний оптимізаційний алгоритм, то внаслідок величезної кількості параметрів модель не зможе ефективно навчитись. Найімовірніше трапиться перенавчання, за якого модель ідеально запам'ятовує тренувальну вибірку, але не виявляє закономірності і помиляється на перевіірній вибірці. Тому було запропоновано ітеративний алгоритм:

1. «Розмноження» тренувальних прикладів шляхом афінних перетворень.
2. Виявлення прихованого стану точок за допомогою кластеризації
3. Запуск алгоритму детектування на фонових зображеннях
4. Додавання знайдених скелетів до «активної» множини
5. Оптимізація параметрів  $\vec{w}$  (при фіксованих  $\vec{\beta}$ )
6. Відкидання скелетів з «активної» множини
7. Оптимізація параметрів  $\vec{\beta}$  (при фіксованих  $\vec{w}$ )
8. Якщо не виконано критерій зупинки – перейти на крок №3

#### Тестування

Для оцінки ефективності методу було проведено тестування на датасеті PARSE. Даний датасет складається з 305 зображень: 100 – тренувальна вибірка, 205 – перевірна вибірка. Кожне зображення містить один основний образ людини. Для нього існує розмітка на 14 ключових точок. Частина фотографій на задньому фоні містить інших людей, ключові точки яких не мають бути знайдені. Виходячи з особливостей датасету, алгоритм має повертати не більше одного скелету. Застосовується стандартна метрика якості – доля правильно розпізнаних ключових точок [2].

Результати тестування занесемо до таблиці (Табл. 1) для порівняння.

**Висновки.** Запропонований метод для задачі локалізації людини дозволяє отримати покращення результатів. Час роботи методу пропорційний до площі зображення, але константа більша ніж в аналогічних методах. Алгоритм чудово локалізує голову та тулуб, але має певні складнощі при знаходженні кисті.

Таблиця 1

Порівняння досягнутих результатів

| %                   | Голова | Плечі | Лікті | Кисть | Таз  | Коліна | Щикол. | $\Sigma$ |
|---------------------|--------|-------|-------|-------|------|--------|--------|----------|
| Д. Раманан          | 90.2   | 85.4  | 68.   | 47.1  | 77.1 | 75.4   | 67.1   | 72.9     |
| Отриманий результат | 95.1   | 89.8  | 68.8  | 47.8  | 78.5 | 80.    | 69.8   | 75.686   |

#### Список літератури:

1. P. Felzenszwalb and D. Huttenlocher, «Pictorial structures for object recognition», International Journal of Computer Vision, vol. 61, 2005.
2. Y. Yang and D. Ramanan, «Articulated pose estimation with flexible mixtures-of-parts», in IEEE Conference on Computer Vision and Pattern Recognition, 2011
3. L. Bourdev and J. Malik, «Poselets: Body part detectors trained using 3d human pose annotations», in IEEE Conference on Computer Vision and Pattern Recognition, 2009.
4. N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. in CVPR, 2005.



**Мельник В.В.**

Национальный технический университет Украины «Киевский политехнический институт»

## МЕТОД ЛОКАЛИЗАЦИИ СТРУКТУРНЫХ ОБЪЕКТОВ НА ФОТОГРАФИИ

### Аннотация

Предложен метод локализации структурных объектов на фотографиях с использованием нелинейной модели. Метод применен для нахождения ключевых точек тела человека. Для данной задачи установлено улучшение метрик качества.

**Ключевые слова:** структурные объекты, локализация, нелинейная модель.

**Melnyk V.V.**

National Technical University of Ukraine «Kyiv Polytechnic Institute»

## THE METHOD OF LOCALIZATION OF STRUCTURAL OBJECTS IN PHOTOS

### Summary

A method of localizing structural objects in photos using a nonlinear model was proposed. The method was used to find the key points of the human body. For this task it improves quality metrics.

**Keywords:** structural models, localization, nonlinear model.

УДК 004.4

## РАЗРАБОТКА И РЕАЛИЗАЦИЯ ПОДХОДА К ОБУЧЕНИЮ С ПОМОЩЬЮ МОБИЛЬНЫХ УСТРОЙСТВ

**Туник В.С.**

Харьковский национальный университет радиоэлектроники

В статье рассмотрены основные современные технологии и тенденции, которые могут помочь улучшить качество образовательных процессов. Ключевой идеей статьи является выявление и описание роли и преимущества использования мобильных устройств в современном образовании, а также формулирование на основании представленного анализа подхода к организации обучения с помощью мобильных устройств. Реализацией подхода представляет собой мобильное Android приложение, которое может быть использовано преподавателями высших учебных заведений для проведения лекций и презентаций, а также представления электронных материалов студентам. Разработанное приложение может помочь повысить эффективность процесса проведения занятий, а также будет способствовать повышению степени вовлеченности студентов в процесс проведения занятия.

**Ключевые слова:** мобильные устройства, образование, презентации, приложение.

**Введение.** На сегодняшний день разработка системы для организации обучения через мобильные устройства является актуальным и перспективным направлением научной и практической деятельности. Во-первых, распространенность и удобство мобильных устройств не могут и далее оставаться незамеченными для современных процессов обучения. Образовательная система требует модернизации, для того чтобы быть более приспособленной к новым тенденциям и способам распространения и представления информации. Во-вторых, в связи с распространенностью мобильных устройств уже зародившееся мобильное обучение с каждым днем набирает всё большую популярность среди преподавателей школ и высших учебных заведений. Мобильные технологии трансформируют баланс между процессом обучением и участием обучающегося. Именно поэтому мобильное обучение – это новая форма обучения, отличная от дистанционного или смешанного, характеризующая новый виток развития и информатизации человеческого общества [1].

Преимущества и возможности мобильного обучения вызывают стремительный рост количества обучающих мобильных систем и приложений для различных возрастных категорий. Именно поэтому разработка подхода к организации обучения с по-

мощью мобильных устройств является особенно актуальной задачей. Предметом разработки описываемого подхода является анализ процессов обучения в высших учебных заведениях, как процесса взаимодействия преподавателей и студентов в процессе проведения занятий. При реализации подхода к обучению, в основном делается акцент на возможность улучшения процесса взаимодействия преподавателей и студентов высших учебных заведений, улучшение качества проведения занятия, увеличение количества студентов, которые участвуют в проведении занятия. Новизна предлагаемого подхода состоит в том, что на сегодняшний день он не имеет абсолютных аналогов в сфере систем мобильного обучения.

**Анализ предметной области.** Сегодня учебные мобильные приложения стали практически обязательной частью любого курса на Западе [2], и в целом всё больше примеров можно найти подтверждению того факта, что высшие учебные заведения Запады стремятся к внедрению мобильного обучения для студентов и преподавателей. Существует целое множество тем, и связанных с ними уже реализованных приложений, которые могут использоваться в образовательном процессе высших учебных заведений. Это приложения с расписанием занятий, приложения для контроля успеваемости и