

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
“КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ”
ФІЗИКО-ТЕХНІЧНИЙ ІНСТИТУТ
КАФЕДРА ІНФОРМАЦІЙНОЇ БЕЗПЕКИ

«До захисту допущено»

Завідувач кафедри

(підпис)
М. В. Грайворонський

“ _____ ” _____ 2017 р.
(ініціали, прізвище)

Дипломна робота

освітньо-кваліфікаційного рівня “магістр”

за спеціальністю 8.04030101 «Прикладна математика»

на тему «Тема»

Виконав студент 6 курсу групи ФІ-51м

Кригін Валерій Михайлович

Керівник к.т.н., Барановський Олексій Миколайович

Рецензент ,

(підпис)

(підпис)

(підпис)

Засвідчую, що у цій дипломній роботі
немає запозичень з праць інших авторів
без відповідних посилань.

Студент _____

РЕФЕРАТ

КЛЮЧЕВЫЕ СЛОВА

ABSTRACT

KEYWORDS

РЕФЕРАТ

СЛОВА

ЗМІСТ

Вступ	6
1 Теоретичні відомості	7
1.1 Задача	7
1.1.1 Бінарна функція витрат	9
1.1.2 Різниця моделей	10
1.1.3 Різниця параметрів.	12
1.2 Розв’язок	13
1.2.1 Бінарна функція витрат	13
1.2.2 Різниця параметрів з урахуванням їх гаусового розподілу	15
1.3 Метод Монте-Карло	17
2 Практичні результати	19
3 Охорона праці	20
Висновки	21
Перелік посилань	22

ВСТУП

Актуальність роботи.

Об'єкт дослідження —

Предмет дослідження —

Мета дослідження.

Завдання наступні:

- 1) Вивчити;
- 2) Розробити.

Практичне значення одержаних результатів.

1 ТЕОРЕТИЧНІ ВІДОМОСТІ

1.1 Задача

Позначимо множину T зображень. Кольори — відтінки сірого, що визначаються лише інтенсивністю від 0 до 1. Введемо множину I пікселів зображення. Зображення $t \in T$ є відображення з множини пікселів на множину їх значень

$$t : I \rightarrow [0; 1] .$$

Інтенсивність пікселя i в зображенні t позначимо як t_i .

Взагалі кажучи, I — множина індексів матриць однакового розміру

$$I = \{ \langle i, j \rangle \mid i = \overline{1..h}, j = \overline{1..w} \} .$$

Зазвичай зображення можуть бути розміром від $100 \times 100 = 10^4$ пікселів. Проте в середньому це значення досягає мільйона пікселів. Це означає, що при використанні $2^8 = 256$ градацій сірого маємо приблизно $10^{6 \cdot 256} = 10^{1536}$ різних зображень, тобто неймовірно багато.

Тривимірна модель обличчя визначається набором n дійсних параметрів. Множина всіх параметрів $X = \mathbb{R}^n$. Функцією, що перетворює набір параметрів на зображення, є відображення

$$f : X \rightarrow T .$$

Введемо позначення для зображення згенерованого з певним набором параме-

трів x

$$f(x) = t.$$

Інтенсивність i пікселя позначимо

$$f_i(x) = t_i.$$

Поставимо Баєсову задачу розпізнавання. Для цього потрібно визначитися з функцією витрат [1]

$$W : X \times X \rightarrow \mathbb{R}.$$

Введемо множину стратегій розпізнавання Q як функцій, які кожному $t \in T$ ставлять у відповідність параметри, з якими було згенеровано обличчя на даному зображенні

$$Q = X^T.$$

Стратегію $q \in Q$, яка для зображення t дає результат x , позначимо

$$q(t) = x.$$

Математичне очікування функції витрат для даного вирішального правила q як функції випадкової величини x за умови, що було пред'явлено зображення t , називається Баєсовим ризиком

$$R(q) = \sum_{t \in T} \sum_{x \in X} \mathbb{P}(x, q(t)) \cdot W(x, q(t)).$$

Задача — знайти таке вирішальне правило q , яке мінімізує Баєсів ризик

$$q^* = \arg \min_{q \in Q} R.$$

1.1.1 Бінарна функція витрат

Досить розповсюдженою, проте зазвичай неприродною є бінарна функція штрафу

$$W(x, x') = \mathbb{1}(x \neq x').$$

Оберемо стратегію q^* , що мінімізує математичне очікування цієї функції витрат

$$\begin{aligned} q^*(t) &= \arg \min_{x'} \left\{ \sum_{x \in X} \mathbb{P}(x | t) \cdot \mathbb{1}(x \neq x') \right\} = \\ &= \arg \min_{x'} \left\{ \sum_{x \in X} \mathbb{P}(x | t) - \sum_{x \in X} \mathbb{P}(x | t) \cdot \mathbb{1}(x = x') \right\} = \\ &= \arg \min_{x'} \{1 - \mathbb{P}(x' | t)\}. \end{aligned}$$

В результаті

$$q^*(t) = \arg \max_x \mathbb{P}(x | t).$$

Отже, якщо використовується бінарна функція витрат, потрібно обирати найбільш ймовірний варіант. Така задача може бути розумною, коли є небагато різних варіантів відповіді. Проте в даному випадку відповідь — набір з сотень дійсних чисел. Аналітичного виразу для розрахування f немає, отже доведеться скористатися чисельними методами, які не дадуть точної відповіді.

Очевидно, що для неперервного випадку така функція зовсім не підходить. Тому що інтеграл

$$\int_{x \in X} \mathbb{1}(x \neq x') \cdot \mathbb{P}(x | t) dx = 1.$$

Тобто майже ніде немає вірної відповіді, бо майже всюди сплачується штраф.

1.1.2 Різниця моделей

Розглянемо більш природню функцію витрат — квадрат евклідової відстані між точками дійсної та обраної моделі.

Введемо множину вершин обличчя V . Кожна вершина має певні координати в тривимірному просторі \mathbb{R}^3 . Модель обличчя — відображення, яке кожній вершині v ставить у відповідність її координати

$$M : V \rightarrow \mathbb{R}^3.$$

Генеративна модель обличчя — відображення, яке кожному набору параметрів x ставить у відповідність модель m

$$G : X \rightarrow M.$$

Координати g вершини v моделі згенерованої з параметрами x позначимо

$$G_v(x) = g.$$

Координати кожної вершини v генеративної моделі отримуються шляхом перемноження компонент параметру x на відповідний коефіцієнт α^v отриманий шляхом методу головних компонент

$$G_v(x) = \sum_{i=1}^n \alpha_i^v \cdot x_i, \quad v \in V.$$

Функція витрат має вигляд

$$\begin{aligned}
 W(x, x') &= \|G(x) - G(x')\|^2 = \sum_{v \in V} [G_v(x) - G_v(x')]^2 = \\
 &= \sum_{v \in V} \sum_{p \in P} [\alpha_p^v \cdot (x_p - x'_p)]^2 = \sum_{p \in P} \left\{ (x_p - x'_p)^2 \cdot \sum_{v \in V} (\alpha_p^v)^2 \right\} = \\
 &= \left| \beta_p^2 = \sum_{v \in V} (\alpha_p^v)^2 \right| = \sum_{p \in P} \beta_p^2 \cdot (x_p - x'_p)^2.
 \end{aligned}$$

Оберемо стратегію q^* , що мінімізує математичне очікування цієї функції витрат

$$q^*(t) = \arg \min_{x'} \left\{ \sum_{x \in X} \left[\mathbb{P}(x | t) \cdot \sum_{i=1}^n \beta_i^2 \cdot (x'_i - x_i)^2 \right] \right\}.$$

Щоб мінімізувати неперервну функцію від параметрів x'_i , можна взяти по них похідну

$$\frac{\partial \sum_{x \in X} \left[\mathbb{P}(x | t) \cdot \sum_{i=1}^n \beta_i^2 \cdot (x'_i - x_i)^2 \right]}{\partial x'_i} = 2 \cdot \sum_{x \in X} \mathbb{P}(x | t) \cdot \beta_i^2 \cdot (x'_i - x_i), \quad i = 1..n$$

та прирівняти до нуля

$$\sum_{x \in X} \mathbb{P}(x | t) \cdot (x'_i - x_i) = 0, \quad i = 1..n.$$

Значення компоненти

$$x'_i = \frac{\sum_{x \in X} \mathbb{P}(x | t) \cdot x_i}{\sum_{x \in X} \mathbb{P}(x | t)} = \sum_{x \in X} \mathbb{P}(x | t) \cdot x_i, \quad i = 1..n.$$

Результуюча стратегія

$$q^*(t) = \sum_{x \in X} x \cdot \mathbb{P}(x | t).$$

У випадку неперервного розподілу ймовірностей

$$q^*(t) = \int_{x \in X} x \cdot \mathbb{P}(x | t) dx.$$

1.1.3 Різниця параметрів

Розглянемо більш просту функцію витрат — квадрат евклідової норми різниці між дійсними та обраними параметрами моделі зображеного обличчя

$$W(x, x') = \|x - x'\|^2 = \sum_{p \in P} (x_p - x'_p)^2.$$

Оберемо стратегію q^* , що мінімізує математичне очікування цієї функції витрат

$$q^*(t) = \arg \min_{x'} \left\{ \sum_{x \in X} \left[\mathbb{P}(x | t) \cdot \sum_{i=1}^n (x'_i - x_i)^2 \right] \right\}.$$

Маємо мінімізацію неперервної функції від параметрів x'_i , отже можемо взяти по них похідну

$$\frac{\partial \sum_{x \in X} \left[\mathbb{P}(x | t) \cdot \sum_{i=1}^n (x'_i - x_i)^2 \right]}{\partial x'_i} = 2 \cdot \sum_{x \in X} \mathbb{P}(x | t) \cdot (x'_i - x_i), \quad i = 1..n$$

та прирівняти до нуля

$$\sum_{x \in X} \mathbb{P}(x | t) \cdot (x'_i - x_i) = 0, \quad i = 1..n.$$

Значення компоненти

$$x'_i = \frac{\sum_{x \in X} \mathbb{P}(x | t) \cdot x_i}{\sum_{x \in X} \mathbb{P}(x | t)} = \sum_{x \in X} \mathbb{P}(x | t) \cdot x_i, \quad i = 1..n.$$

Результуюча стратегія

$$q^*(t) = \sum_{x \in X} x \cdot \mathbb{P}(x | t).$$

У випадку неперервного розподілу ймовірностей

$$q^*(t) = \int_{x \in X} x \cdot \mathbb{P}(x | t) dx.$$

Отримана та ж стратегія, що мінімізує математичне очікування суми квадратів різниць координат вершин дійсної та обраної моделі обличчя. Тобто це вирішувальне правило розв'язує обидві задачі.

1.2 Розв'язок

1.2.1 Бінарна функція витрат

Розглянемо функцію витрат, за якої вірним є лише один набір параметрів, а за всі інші сплачується штраф 1.

Вважаємо, що на даному зображенні t присутній нормальний шум з неві-

домою дисперсією σ_t^2 . Тоді ймовірність того, що дане зображення було отримано саме з параметрами x

$$\mathbb{P}_I(x | t) = \prod_{i \in I} \frac{\exp \left\{ -\frac{(t_i - f_i(x))^2}{2 \cdot \sigma_t^2} \right\}}{\sqrt{2 \cdot \pi \cdot \sigma_t^2}}.$$

Коли з контексту буде зрозуміло, якому саме зображенню належить дана дисперсія, індекс t не будемо використовувати.

Максимізуємо логарифм ймовірності, тому що це зручніше, ніж максимізація добутку кількох десятків тисяч або мільйонів значень

$$\ln \mathbb{P}_I(x | t) = \sum_{i \in I} \left\{ -\frac{(t_i - f_i(x))^2}{2 \cdot \sigma^2} - \frac{\ln 2 + \ln \pi + 2 \cdot \ln \sigma}{2} \right\} \rightarrow \max.$$

Позбуваємося константного доданку та помножимо на подвоєну ненульову дисперсію

$$\sum_{i \in I} (t_i - f_i(x))^2 \rightarrow \min.$$

Бачимо, що мінімізація суми квадратів різниць між дійсним та згенерованим зображенням розв'язує задачу з бінарною функцією втрат та гаусовим шумом на зображенні без інших додаткових умов, які буде розглянуто далі.

Введемо множину пікселів $F \subset I$, які на зображенні $f(x)$ належать обличчю. Тоді суму можна розбити на дві

$$\sum_{i \in I} (t_i - f_i(x))^2 = \sum_{i \in F} (t_i - f_i(x))^2 + \sum_{i \in I \setminus F} (t_i - f_i(x))^2.$$

Рахуємо вибірковою дисперсію для тієї частини зображення, де зображено облич-

чя

$$\overline{\sigma_F^2} = \frac{\sum_{i \in F} (t_i - f_i(x))^2}{|F| - 1}.$$

Екстраполюємо це значення на все зображення

$$\overline{\sigma_F^2} \approx \frac{\sum_{i \in I} (t_i - f_i(x))^2}{|I| - 1}.$$

Тоді вираз, який треба мінімізувати, можна представити як

$$\sum_{i \in I} (t_i - f_i(x))^2 \approx (|I| - 1) \cdot \overline{\sigma_F^2} \rightarrow \min.$$

Оскільки розмір зображення фіксований і його зменшити не вийде, потрібно зменшити вибірккову дисперсію на пікселях обличчя

$$\overline{\sigma_F^2} = \frac{\sum_{i \in F} (t_i - f_i(x))^2}{|F| - 1} \rightarrow \min.$$

1.2.2 Різниця параметрів з урахуванням їх гаусового розподілу

Вважаємо, що на даному зображенні t присутній нормальний шум з невідомою дисперсією σ_t^2 . Ймовірність того, що зображення моделі з конкретними параметрами x є очищеним від шумів зображенням t

$$\mathbb{P}_I(x | t) = \prod_{i \in I} \frac{\exp \left\{ -\frac{(t_i - f_i(x))^2}{2 \cdot \sigma_t^2} \right\}}{\sqrt{2 \cdot \pi \cdot \sigma^2}}.$$

Відомо, що параметри x — набори коефіцієнтів нормованих головних компонент, тобто мають стандартний гаусовий розподіл. Ймовірність того, що було згенеровано саме такий набір

$$\mathbb{P}_X(x | t) = \prod_{i=1}^n \frac{\exp\left(-\frac{x_i^2}{2}\right)}{\sqrt{2 \cdot \pi}}.$$

Ймовірність того, що дане зображення t було отримано саме з параметрами x , є добутком розглянутих ймовірностей

$$\mathbb{P}(x | t) = \mathbb{P}_I(x | t) \cdot \mathbb{P}_X(x | t) = \prod_{i \in I} \frac{\exp\left\{-\frac{(t_i - f_i(x))^2}{2 \cdot \sigma_t^2}\right\}}{\sqrt{2 \cdot \pi \cdot \sigma_t^2}} \cdot \prod_{i=1}^n \frac{\exp\left(-\frac{x_i^2}{2}\right)}{\sqrt{2 \cdot \pi}}.$$

Коли з контексту буде зрозуміло, якому саме зображенню належить дана дисперсія, індекс t не будемо використовувати.

Винесемо константи з добутків та позначимо

$$c_t = (2 \cdot \pi)^{\frac{w \cdot h + n}{2}} \cdot \sigma_t^{w \cdot h}.$$

Ймовірність приймає вигляд

$$\mathbb{P}(x | t) = c_t \cdot \exp\left\{-\frac{\|t - f(x)\|^2}{2 \cdot \sigma^2}\right\} \cdot \exp\left\{-\frac{\|x\|^2}{2}\right\}.$$

Стратегія розпізнавання

$$q^*(t) = c_t \cdot \sum_{x \in X} x \cdot \exp\left\{-\frac{\|t - f(x)\|^2}{2 \cdot \sigma^2}\right\} \cdot \exp\left\{-\frac{\|x\|^2}{2}\right\}.$$

У неперервному випадку

$$q^*(t) = c_t \cdot \int_{x \in X} x \cdot \exp \left\{ -\frac{\|t - f(x)\|^2}{2 \cdot \sigma^2} \right\} \cdot \exp \left\{ -\frac{\|x\|^2}{2} \right\} dx.$$

Оскільки σ^2 невідома, її можна оцінити за відомою формулою оптимальної оцінки дисперсії

$$\overline{\sigma^2} = \sum_{x \in X} \frac{\|f(x) - t\|^2}{w \cdot h \cdot n - 1}.$$

1.3 Метод Монте-Карло

Оскільки функція f , що дає зображення t для даного набору параметрів x , є складно влаштованою, вона не представлена в явному вигляді. Інтегрування виразу, що містить цю функцію, не може бути проведено аналітично. Метод Монте-Карло — чисельний метод, що підходить до розв'язку даної задачі.

Маємо інтеграл

$$q^*(t) = c_t \cdot \int_{x \in X} x \cdot \exp \left\{ -\frac{\|t - f(x)\|^2}{2 \cdot \sigma^2} \right\} \cdot \exp \left\{ -\frac{\|x\|^2}{2} \right\} dx,$$

який є математичним очікуванням функції стандартного гаусового n -вимірного вектора

$$q^*(t) = (2 \cdot \pi \cdot \sigma^2)^{\frac{w \cdot h}{2}} \cdot M_{\xi} \left[\xi \cdot \exp \left\{ -\frac{\|t - f(\xi)\|^2}{2 \cdot \sigma^2} \right\} \right], \quad \xi \sim \mathcal{N}(0, 1).$$

Внесемо все в експоненту

$$q^*(t) = M_\xi \left[\exp \left\{ \frac{w \cdot h}{2} \cdot (\ln 2 + \ln \pi + \ln \sigma^2) + \ln \xi - \frac{\|t - f(\xi)\|^2}{2 \cdot \sigma^2} \right\} \right].$$

Щоб його розрахувати, потрібно згенерувати набори з n незалежних випадкових величин, що мають стандартний нормальний розподіл, для кожного набору розрахувати значення виразу, математичне очікування якого потрібно визначити, просумувати та поділити на кількість спроб. Перед цим потрібно оцінити дисперсію шуму на зображенні t .

$$z_\gamma^2 = 2.575^2$$

$$\varepsilon^2 = 0.01^2$$

$$M(N) = \sum_{i=1}^N \left(\hat{Q}^{(i)}(x) \right)^2$$

$$S(N) = \sum_{i=1}^N \left(\hat{Q}^{(i)}(x) \right)$$

$$\hat{V}_N^2 = \frac{1}{N-1} \cdot \left[M(N) - \frac{1}{N} \cdot (S(N)) \right]$$

$$\hat{Q}_N^2 = \left(\frac{1}{N} \cdot S(N) \right)$$

$$N^* = \min \left\{ N \mid N > \frac{z_\gamma^2 \cdot \hat{V}_N^2}{\varepsilon^2 \cdot \hat{Q}_N^2} \right\}$$

2 ПРАКТИЧНІ РЕЗУЛЬТАТИ

3 ОХОРОНА ПРАЦІ

ВИСНОВКИ

В результаті виконання роботи вдалося.

ПЕРЕЛІК ПОСИЛАНЬ

1. Berger, J.O. Statistical Decision Theory: Foundations, Concepts, and Methods / J.O. Berger // Springer Series in Statistics. — Springer New York, 1980.