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

«До захисту допущено»
Завідувач кафедри
____
(підпис)
М. В. Грайворонський
(ініціали, прізвище)
“____” ____________ 2017 р.

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 ________ *(підпис)*
Керівник к.т.н., Барановський Олексій Миколайович ________ *(підпис)*
Рецензент , ________ *(підпис)*

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

Київ — 2017 року

РЕФЕРАТ

КЛЮЧЕВЫЕ СЛОВА

ABSTRACT

KEYWORDS

РЕФЕРАТ

СЛОВА

ЗМІСТ

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

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

Інтенсивність пікселя i в зображенні t позначимо як t_i.

Взагалі кажучи, I — множина індексів матриць однакового розміру

$$I = \left\{ \langle i, j \rangle \;\middle|\; i = \overline{1..h},\, j = \overline{1..w} \right\} .$$

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

$$f : X \to T.$$

Введемо позначення для зображення згенерованого з певним набором параметрів x

$$f\left(x\right)=t.$$

Інтенсивність i пікселя позначимо

$$f_i\left(x\right)=t_i.$$

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

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

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

$$q\left(t\right)=x.$$

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

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

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

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

1.1.1 Бінарна функція витрат

Досить росповсюдженою, 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

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

В результаті

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

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

$$\int\limits_{x \in X} \mathbb{1}(x \neq x') \cdot \mathbb{P}(x \mid t) \, dx = 1.$$

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

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

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

$$G : X \to M.$$

Координати g вершини v моделі згенерованої з параметрами x позначимо

$$G_v\left(x\right) = g.$$

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

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

Функція витрат має вигляд

$$W(x,x') = \|G(x) - G(x')\|^2 = \sum_{v\in V} \left[G_v(x) - G_v(x')\right]^2 =$$

$$= \sum_{v\in V}\sum_{p\in P}\left[\alpha_p^v \cdot \left(x_p - x'_p\right)\right]^2 = \sum_{p\in P}\left\{\left(x_p - x'_p\right)^2 \cdot \sum_{v\in V}\left(\alpha_p^v\right)^2\right\} =$$

$$= \left|\beta_p^2 = \sum_{v\in V}\left(\alpha_p^v\right)^2\right| = \sum_{p\in P}\beta_p^2 \cdot \left(x_p - x'_p\right)^2 .$$

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

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

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

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

та прирівняти до нуля

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

Значення компоненти

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

Результуюча стратегія

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

У випадку неперервного розподілу ймовірностей

$$q^*(t) = \int\limits_{x \in X} x \cdot \mathbb{P}(x \mid t) \, d\,x.$$

1.1.3 Різниця параметрів

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

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

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

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

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

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

та прирівняти до нуля

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

Значення компоненти

$$x'_i = \frac{\sum\limits_{x \in X} \mathbb{P}\left(x \mid t\right) \cdot x_i}{\sum\limits_{x \in X} \mathbb{P}\left(x \mid t\right)} = \sum_{x \in X} \mathbb{P}\left(x \mid t\right) \cdot x_i, \qquad i = 1..n.$$

Результуюча стратегія

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

У випадку неперервного розподілу ймовірностей

$$q^*\left(t\right) = \int\limits_{x \in X} x \cdot \mathbb{P}\left(x \mid t\right) d\,x.$$

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

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

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

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

Позбуваємося константного доданку та помножимо на подвоєну ненульову дисперсію

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

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 $f\left(x\right)$ належать обличчю. Тоді суму можна розбити на дві

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

Рахуємо вибіркову дисперсію для тієї частини зображення, де зображено обличчя

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

Екстраполюємо це значення на все зображення

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

Тоді вираз, який треба мінімізувати, можна представити як

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

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

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

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

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

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

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

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

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

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

Винесемо константи з добутків та позначимо

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

Ймовірність приймає вигляд

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

Стратегія розпізнавання

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

У неперервному випадку

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

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.

Треба підрахувати інтеграл

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

який є математичним очікуванням функції стандартного гаусового n-вимірного вектора

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

Внесемо все в експоненту

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

Для отримання оцінки інтегралу треба згенерувати N векторів з n незалежних випадкових величин, що мають стандартний нормальний розподіл

$$\xi^{(i)} = \left\langle \xi_1^{(i)}, \xi_2^{(i)}, \ldots, \xi_n^{(i)} \right\rangle, \qquad \xi_j^{(i)} \sim \mathcal{N}(0,1).$$

Для кожного вектору розрахувати значення підінтегрального виразу

$$Q^{(i)} = \frac{\xi^{(i)}}{\left|\xi^{(i)}\right|} \cdot \exp\left\{ \ln\left|\xi^{(i)}\right| - \frac{\left\|t - f\left(\xi^{(i)}\right)\right\|^2}{2 \cdot \sigma^2} - \frac{w \cdot h}{2} \cdot \left(\ln 2 + \ln \pi + \ln \sigma^2 \right) \right\}.$$

Оцінкою інтегралу буде середнє значення

$$\overline{Q}_N = \frac{\sum\limits_{i=1}^{N} \hat{Q}^{(i)}}{N}.$$

Визначимо оптимальне значення N. Згадаємо центральну граничну теорему

$$\sqrt{N} \cdot \frac{\overline{Q}_N - M(Q)}{\sqrt{D(Q)}} \Rightarrow \mathcal{N}(0,1).$$

Для певного $z_\gamma > 0$

$$\mathbb{P}\left(\sqrt{N} \cdot \frac{\overline{Q}_N - M(Q)}{\sqrt{D(Q)}} \le z_\gamma \right) = \Phi(z_\gamma),$$

де Φ — інтегральна функція стандартного гаусового розподілу. Нас цікавить ймовірність

$$\mathbb{P}\left(-z_\gamma \leq \sqrt{N} \cdot \frac{\overline{Q}_N - M\left(Q\right)}{\sqrt{D\left(Q\right)}} \leq z_\gamma\right) = \Phi\left(z_\gamma\right) - \Phi\left(-z_\gamma\right).$$

Використаємо модуль для перевірки входження виразу в довірчий інтервал та скористаємося симетричністю гаусового розподілу

$$\mathbb{P}\left(\left|\sqrt{N} \cdot \frac{\overline{Q}_N - M\left(Q\right)}{\sqrt{D\left(Q\right)}}\right| \leq z_\gamma\right) = 2 \cdot \Phi\left(z_\gamma\right) - 1.$$

Поділимо обидві частини нерівності на модуль середнього значення виборки

$$\mathbb{P}\left(\left|\sqrt{N} \cdot \frac{\overline{Q}_N - M\left(Q\right)}{\overline{Q}_N \cdot \sqrt{D\left(Q\right)}}\right| \leq \frac{z_\gamma}{\left|\overline{Q}_N\right|}\right) = 2 \cdot \Phi\left(z_\gamma\right) - 1.$$

Позначимо відносну похибку

$$\varepsilon = \left|\frac{\overline{Q}_N - M\left(Q\right)}{\overline{Q}_N}\right|$$

та перепишемо вираз

$$\mathbb{P}\left(\left|\sqrt{N} \cdot \frac{\varepsilon}{\sqrt{D\left(Q\right)}}\right| \leq \frac{z_\gamma}{\left|\overline{Q}_N\right|}\right) = 2 \cdot \Phi\left(z_\gamma\right) - 1.$$

Перенесемо все окрім кореню від кількості спроб в праву частину нерівності

$$\mathbb{P}\left(\left|\sqrt{N}\right| \leq \left|\frac{z_\gamma \cdot \sqrt{D\left(Q\right)}}{\varepsilon \cdot \overline{Q}_N}\right|\right) = 2 \cdot \Phi\left(z_\gamma\right) - 1$$

та піднесемо обидві частини нерівності до квадрату

$$\mathbb{P}\left(N \leq \frac{z_\gamma^2 \cdot D\left(Q\right)}{\varepsilon^2 \cdot \overline{Q}_N^2}\right) = 2 \cdot \Phi\left(z_\gamma\right) - 1.$$

Перепозначимо отриману ймовірність

$$\mathbb{P}\left(N \leq \frac{z_\gamma^2 \cdot D\left(Q\right)}{\varepsilon^2 \cdot \overline{Q}_N^2}\right) = \gamma.$$

Щоб виразити z_γ, потрібно скористатися квантилєм нормального розподілу

$$\Phi^{-1}\left(\frac{\gamma+1}{2}\right) = z_\gamma.$$

Щоб відносна похибка обчислення математичного очікування методом Монте-Карло не перевищувала ε з ймовірністю γ і вище, необхідна та достатня кількість ітерацій рахується за формулою

$$N = \left\lceil \frac{z_\gamma^2 \cdot D\left(Q\right)}{\varepsilon^2 \cdot \overline{Q}_N^2} \right\rceil, \qquad z_\gamma = \Phi^{-1}\left(\frac{\gamma+1}{2}\right). \tag{1.1}$$

Коли дисперсія отриманої випадкової величини Q невідома, використовується її незміщена оцінка

$$\widetilde{Q}_N = \frac{\sum\limits_{i=1}^{N} \left(\hat{Q}^{(i)}\left(x\right) \right)^2 - \overline{Q}_N^2}{N-1}.$$

Формула (1.1) використовується лише для перевірки того, чи дає наявна вибірка з N елементів відповідь з вказаною похибкою. Спочатку необхідно дати алгоритму “розігрітись” і не порівнювати значення N з необхідною кількістю кроків. Також немає необхідності у перевірці відповідності значення N на кожній ітерації, особливо якщо це займає багато часу — достатньо робити перерахунок кожні 100 або 1000 кроки в залежності від інтегралу.

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.