

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ  
“КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ІМЕНІ ІГОРЯ  
СІКОРСЬКОГО”**

**ФІЗИКО-ТЕХНІЧНИЙ ІНСТИТУТ  
КАФЕДРА ІНФОРМАЦІЙНОЇ БЕЗПЕКИ**

«До захисту допущено»

Завідувач кафедри

\_\_\_\_\_  
(підпис)  
М. В. Грайворонський

“ \_\_\_\_\_ ” \_\_\_\_\_ 2017 р.  
(ініціали, прізвище)

**Магістерська дисертація**

за спеціальністю 8.04030101 «Прикладна математика»

на тему «Розпізнавання обличчя з косметикою за допомогою

просторової реконструкції поверхні обличчя за одним або кількома фото»

Виконав студент 6 курсу групи ФІ-51м

Кригін Валерій Михайлович

Керівник к.т.н., Барановський Олексій Миколайович

Консультант т. к. “Люксофт”, Сдобніков Віктор Юрійович

\_\_\_\_\_  
(підпис)

\_\_\_\_\_  
(підпис)

\_\_\_\_\_  
(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 . . . . .                                   | 6  |
| 1 Теоретичні відомості . . . . .                  | 7  |
| 1.1 Попередні роботи. . . . .                     | 7  |
| 1.2 Породжувальна модель обличчя . . . . .        | 8  |
| 1.3 Задача . . . . .                              | 10 |
| 1.3.1 Початкові умови . . . . .                   | 10 |
| 1.3.2 Бассова задача розпізнавання . . . . .      | 12 |
| 1.3.3 Бінарна функція витрат . . . . .            | 13 |
| 1.3.4 Інтервальна функція витрат . . . . .        | 14 |
| 1.3.5 Різниця моделей . . . . .                   | 15 |
| 1.3.6 Різниця параметрів. . . . .                 | 18 |
| 1.4 Розв'язок . . . . .                           | 19 |
| 1.4.1 Дискретизація простору параметрів . . . . . | 19 |
| 1.4.2 Інтервальна функція витрат . . . . .        | 20 |
| 1.4.3 Різниця параметрів. . . . .                 | 25 |
| 1.5 Метод Монте-Карло . . . . .                   | 26 |
| 2 Практичні результати . . . . .                  | 31 |
| Висновки . . . . .                                | 32 |
| Перелік посилань . . . . .                        | 33 |

## ВСТУП

**Актуальність роботи.** Примітивні системи біометричної ідентифікації особистості за фотознімком для коректної роботи потребують певних умов освітлення, положення обличчя на знімку, виразу обличчя та інших параметрів фотозйомки. Наразі ці обмеження вдається обходити за допомогою різних підходів, серед яких є використання насичених навчальних виборок при машинному навчанні, просторова реконструкція обличчя та розпізнавання опорних точок обличчя (ніс, рот, очі). Наступним етапом є подолання таких складно передбачуваних перешкод, як наявність косметики, аксесуарів (сережки, окуляри), волосся на обличчі (чубчик, борода, вуса). Деякі аксесуари можна зняти перед розпізнаванням, проте бороду та косметику складно прибирати кожен раз лише для того, щоб бути розпізнаним.

*Об'єкт дослідження* — обличчя.

*Предмет дослідження* — розпізнавання обличчя з косметикою.

**Мета дослідження.** Розробка методу розпізнавання обличчя, на яке може бути нанесено косметику.

Завдання наступні:

- 1) Вивчити методи реконструкції просторової конфігурації обличчя за одним або кількома зображеннями;
- 2) Дослідити вплив косметики на результати реконструкції;
- 3) Розробити та перевірити метод розпізнавання облич, на які нанесено косметику.

**Практичне значення одержаних результатів.**

# 1 ТЕОРЕТИЧНІ ВІДОМОСТІ

## 1.1 Попередні роботи

Найпершою роботою зі створення тривимірної породжувальної моделі обличчя вважається стаття Фредеріка Парке 1972 року [1].

У даній роботі використовується породжувальна модель обличчя 3D Basel Face Model (BFM) розроблена командою Базельського університету [2]. Перший підхід, яких дозволяє отримати тривимірну модель обличчя лише за одним фото, було запропоновано Томасом Феттером та Фолькером Бланцом [3]. Вони використовували BFM та стохастичний градієнтний спуск [4], що дозволяло досягти добрих результатів за 40 хвилин з процесором Pentium III, 800MHz [5]. Згодом ті ж автори розробили та використали у своїй роботі стохастичний алгоритм Ньютона [6], яким досягли реконструкції просторової конфігурації обличчя за 4.5 хвилини на Pentium 4, 2GHz.

Пізніше студент Національного Технічного Університету України “Київський Політехнічний Інститут” запропонував та запатентував свій метод створення генеративної моделі обличчя та відновлення тривимірної поверхні обличчя за одним або кількома фото [7]. Основні відмінності полягають у тому, що задачу співставлення відповідних точок різних моделей було представлено та розв’язано як супермодулярну задачу розмітки [8], а нові моделі обличч генеруються як зважене середнє.

На момент написання дисертації одними з новіших робіт, де було використано модель Базельського університету, є методи відстеження обличчя [9] та переносу виразу обличчя однієї людини іншій [10].

## 1.2 Породжувальна модель обличчя

За допомогою високоточного 3D сканеру було зафіксовано 100 чоловічих і 100 жіночих облич. Це люди віком від 8 до 62 років (25 в середньому), вагою від 40 до 123 кілограмів (66 в середньому).

Щоб зважена сума облич теж була обличчям, потрібно знайти відповідність між вершинами моделей різних облич. Тобто, оскільки модель складається з впорядкованого набору вершин, точка, яка відповідає, наприклад, за горбинку на носі, повинна бути на одній і тій самій позиції в масиві кожної моделі. Для цього була використана модифікація ітеративного алгоритму найближчих точок ІСР [11]. Після цього достатньо відмітити опорні точки обличчя лише на одній моделі, щоб вони були відомі на всіх інших, у тому числі на похідних обличчях.

Розраховується середня модель обличчя як середнє арифметичне по всіх моделям

$$\bar{f}_v = \frac{\sum_{f \in F} f_v}{|F|}, \quad v \in V,$$

де  $V$  — множина вершин,  $F$  — множина облич,  $f_v$  — координати вершини  $v$  в обличчі  $f$ . Середня модель теж є обличчям завдяки попередньому кроку співставлення вершин.

Останній етап, який нас цікавить, це обробка даних за допомогою методу головних компонент [12]. На вході матриця, кожному стовбцю якої відповідає координата вершини, а стрічці — обличчя. На перетині  $f$  стрічки і  $v_i$  стовбця



знаходиться значення  $i$  компоненти координат вершини  $v$  обличчя  $f$

$$M = \begin{bmatrix} M_{v_x^1}^{f^1} & M_{v_y^1}^{f^1} & M_{v_z^1}^{f^1} & M_{v_x^2}^{f^1} & \dots & M_{v_y^m}^{f^1} & M_{v_z^m}^{f^1} \\ M_{v_x^1}^{f^2} & M_{v_y^1}^{f^2} & M_{v_z^1}^{f^2} & M_{v_x^2}^{f^2} & \dots & M_{v_y^m}^{f^2} & M_{v_z^m}^{f^2} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ M_{v_x^1}^{f^n} & M_{v_y^1}^{f^n} & M_{v_z^1}^{f^n} & M_{v_x^2}^{f^n} & \dots & M_{v_y^m}^{f^n} & M_{v_z^m}^{f^n} \end{bmatrix}.$$

На виході маємо дисперсії головних компонент і лінійне перетворення  $L$  коваріаційної матриці  $\widetilde{M}$ , яке робить  $L \cdot \widetilde{M} \cdot L^{-1}$  діагональною.

Оскільки маємо головні компоненти, нове обличчя отримується не як зважене середнє кількох облич, а як сума середнього обличчя та вектору параметрів  $x$  помноженого на матрицю  $L$

$$M(x) = \bar{f} + x \cdot L.$$

Переваги головних компонент:

- 1) для зменшення обчислювальних витрат можна обирати не всі параметри, а лише кілька перших, бо вони мають найбільшу дисперсію і несуть в собі більшу частину інформації;
- 2) згідно з центральною граничною теоремою розподіл параметрів вважається нормальним, що зручно при моделюванні та обчисленнях:
  - а) фактично відсутні обмеження на параметри моделі на відміну від зваженого середнього, де сума параметрів повинна дорівнювати 1;
  - б) за побудовою випадкові величини мають нульове середнє значення та діагональну коваріаційну матрицю;
  - в) якщо нормувати параметри (перенести їх дисперсію в матрицю  $L$ ),

отримаємо набір незалежних випадкових величин розподілених за стандартним нормальним законом.

### 1.3 Задача

#### 1.3.1 Початкові умови

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

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

Зазвичай використовуються зображення розміром від  $100 \times 100 = 10^4$  пікселів. Проте сучасні камери на фотоапаратах, смартфонах та інших пристроях можуть зробити зображення площиною кілька мільйонів пікселів і більше. При використанні  $2^8 = 256$  градацій сірого маємо  $2^{8 \cdot 10^4} \approx 10^{24 \cdot 10^3}$  різних зображень розміром  $100 \times 100$ , 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

$$f : X \rightarrow T.$$

Вважаємо, що на вхідне зображення накладено шум  $\eta$ , що є вектором незалежних випадкових величин розподілених за центрованим нормальним законом та однаковою невідомою дисперсією  $\sigma^2$

$$t = f(x) + \eta.$$

Інтенсивність  $i$  пікселя

$$t_i = f_i(x) + \eta_i, \quad \eta_i \sim \mathcal{N}(0, \sigma^2), \quad i \in I.$$

Ймовірність того, що буде пред'явлено зображення  $t$  з параметрами  $x$ , є сумісною ймовірністю

$$\mathbb{P}(f(\xi) + \eta = t, \xi = x).$$

З визначення умовної ймовірності випливає

$$\mathbb{P}(f(\xi) + \eta = t, \xi = x) = \mathbb{P}(f(\xi) + \eta = t \mid \xi = x) \cdot \mathbb{P}(\xi = x).$$

Позбавимось умовної ймовірності, замінивши  $\xi$  на  $x$

$$\mathbb{P}(f(\xi) + \eta = t, \xi = x) = \mathbb{P}(\eta = t - f(x)) \cdot \mathbb{P}(\xi = x).$$

Це означає, що ймовірність отримати на вході зображення  $t$  згенероване з пара-

метрами  $x$  дорівнює добутку ймовірності отримати зображення з шумом  $t - f(x)$  на ймовірність того, що це зображення згенеровано з параметрами  $x$ .

### 1.3.2 Баєсова задача розпізнавання

Поставимо Баєсову задачу розпізнавання. Для цього потрібно визначитися з функцією витрат [13]

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

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

$$Q = X^T.$$

Стратегію  $q \in Q$

$$q : T \rightarrow X,$$

яка для зображення  $t$  дає результат  $x$ , позначимо

$$q(t) = x.$$

Математичне очікування функції витрат  $W$  як функції випадкової пари  $\langle t, x \rangle$  для даного вирішального правила  $q$  називається Баєсовим ризиком [14]

$$R(q) = \sum_{t \in T} \int_X W(x, q(t)) dF(t, x) = \mathbb{E}[W(x, q(t))],$$

де використовується інтеграл Лебега-Стільтєса, а  $F$  — функція сумісного розподілу зображення та параметрів зображеної моделі. Задача — знайти таке вирішальне правило  $q$ , за якого Баєсів ризик  $R(q)$  мінімальний для даної функції

штрафу  $W$  [15]

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

Оскільки зображення  $t$  не залежать одне від одного, мінімізація суми еквівалентна мінімізації кожного її елементу

$$q^*(t) = \arg \min_{q \in Q} \int_X W(x, q(t)) dF(t, x).$$

Можна переписати через умовне математичне очікування

$$q^*(t) = \arg \min_{q \in Q} \mathbb{E}[W(x, q(t)) | t] \cdot \mathbb{P}(t).$$

Константний множник не впливає на розв'язок

$$q^*(t) = \arg \min_{q \in Q} \mathbb{E}[W(x, q(t)) | t].$$

### 1.3.3 Бінарна функція витрат

Досить розповсюдженою, проте зазвичай неприродною є бінарна штрафна функція

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

Для неперервного випадку така функція витрат не підходить, бо Баєсів ризик

$$R(q) = \sum_{t \in T} \int_{x \in X} \mathbb{1}(x \neq q(t)) \cdot p(t, x) dx = 1, \quad \forall q \in Q.$$

Будь-яка стратегія дає невірну відповідь у неперервному випадку з точки зору бінарної функції витрат, тому для неї буде розглядатися лише дискретний випадок.

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

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

В результаті

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

Отже, якщо використовується бінарна функція витрат, потрібно обирати найбільш ймовірний набір параметрів. Аналітичний вираз для розрахування  $f$  досить складний, тому доведеться скористатися чисельними методами, які не завжди дають точну відповідь і можуть зупинитися у точці локального оптимуму, якщо такий є і не співпадає з глобальним.

### 1.3.4 Інтервальна функція витрат

Більш загальним варіантом бінарної штрафної функції є індикатор неналежності дійсного параметру  $x$  деякій  $\delta$ -околиці обраного параметру  $x'$

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

Оберемо стратегію, що мінімізує математичне очікування цієї функції витрат

$$\begin{aligned} q^*(t) &= \arg \min_{x'} \left\{ \int_X \mathbb{1}(x \notin \delta(x')) dF(t, x) \right\} = \\ &= \arg \min_{x'} \left\{ 1 - \int_X \mathbb{1}(x \in \delta(x')) dF(t, x) \right\} \end{aligned}$$

В результаті

$$q^*(t) = \arg \max_{x'} \int_{\delta(x')} dF(t, x) = \arg \max_{x'} \mathbb{P}(t, x \in \delta(x')).$$

Потрібно обрати таку точку, щоб ймовірність того, що параметри належать її  $\delta$ -околиці, була більшою, ніж для  $\delta$ -околиць інших точок при даному зображенні  $t$ .

### 1.3.5 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

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

Породжувальна 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  $v$  породжувальної моделі отримуються шляхом перемноження компонент параметру  $x$  на відповідний коефіцієнт  $\lambda^v$  отриманий за допомогою методу головних компонент та додавання результату до середнього положення поточної вершини

$$G_v(x) = g_0^v + \sum_{i=1}^n \lambda_i^v \cdot x_i, \quad v \in V.$$

Функція витрат має вигляд

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

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



витрат

$$q^*(t) = \arg \min_{x'} \left\{ \int_X \sum_{i=1}^n \beta_i^2 \cdot (x'_i - x_i)^2 dF(t, x) \right\}.$$

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

$$\frac{\partial \int_X \sum_{i=1}^n \beta_i^2 \cdot (x'_i - x_i)^2 dF(t, x)}{\partial x'_i} = 2 \cdot \int_X \beta_i^2 \cdot (x'_i - x_i) dF(t, x), \quad i = 1..n$$

та прирівняти до нуля

$$\sum_X (x'_i - x_i) dF(t, x) = 0, \quad i = 1..n.$$

Значення компоненти

$$x'_i = \frac{\int_X x_i dF(t, x)}{\int_X dF(t, x)} = \frac{\int_X x_i dF(t, x)}{\mathbb{P}(t)}$$

Результуюча стратегія

$$q^*(t) = \frac{\int_X x dF(t, x)}{\mathbb{P}(t)}.$$

У вигляді умовного математичного очікування

$$q^*(t) = \mathbb{E}[x \mid t].$$

Ділення на ймовірність появи зображення  $t$  можна сприймати як нормування при використанні чисельних методів розрахунку даного інтегралу.

### 1.3.6 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

$$q^*(t) = \arg \min_{x'} \left\{ \int_X \sum_{i=1}^n (x'_i - x_i)^2 dF(t, x) \right\}.$$

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

$$\frac{\partial \int_X \sum_{i=1}^n (x'_i - x_i)^2 dF(t, x)}{\partial x'_i} = 2 \cdot \int_{x \in X} dF(t, x) \cdot (x'_i - x_i), \quad i = 1..n$$

та прирівняти до нуля

$$\int_X (x'_i - x_i) dF(t, x) = 0, \quad i = 1..n.$$

Значення компоненти

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

Результуюча стратегія

$$q^*(t) = \frac{\int x dF(t, x)}{\mathbb{P}(t)}.$$

Оптимальна стратегія як умовне математичне очікування

$$q^*(t) = \mathbb{E}[x | t].$$

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.

Ділення на ймовірність появи зображення  $t$  можна сприймати як нормування при використанні чисельних методів розрахунку даного інтегралу, як і в минулому прикладі.

## 1.4 Розв'язок

### 1.4.1 Дискретизація простору параметрів

Дискретизуємо простір параметрів  $X$ . Введемо невелике  $\delta x$  таке, щоб ймовірність того, що значення потрапляє в рамки гіперкубу  $\left[x - \frac{\delta x}{2}; x + \frac{\delta x}{2}\right]$ , приблизно дорівнювала добутку щільності розподілу в його центрі на його об'єм

$$\mathbb{P}\left(\xi \in \left[x - \frac{\delta x}{2}; x + \frac{\delta x}{2}\right]\right) \approx p(x) \cdot \delta x^n,$$

де

$$\left(x \pm \frac{\delta x}{2}\right)_i = x_i \pm \frac{\delta x}{2}, \quad i = 1..n.$$

Далі треба розбити простір на гіперкуби зі стороною  $\delta x$  і вважати, що коли величина потрапляє на територію певного гіперкубу, то вона дорівнює значенню в його центрі.

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

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

Введемо зручне позначення для шуму в пікселі  $i$

$$y_i = t_i - f_i(x).$$

Розглядається одне зображення  $t$ . З контексту буде зрозуміло, яке значення має  $x$ . Ймовірність того, що обличчя на даному зображенні отримано з параметрами  $x$

$$\mathbb{P}_Y(t | x) = \delta x^{w \cdot h} \cdot \prod_{i \in I} \frac{\exp\left\{-\frac{y_i^2}{2 \cdot \sigma_t^2}\right\}}{\sqrt{2 \cdot \pi \cdot \sigma_t^2}}.$$

Сумісна ймовірність зображення  $t$  та параметрів  $x$

$$\mathbb{P}(x, t) = \mathbb{P}_I(x, t) \cdot \mathbb{P}_X(x, t)$$

#### 1.4.2 Інтервальна функція витрат

Розглянемо функцію витрат, за якої вірним є будь-який набір параметрів, що знаходиться в  $\delta$ -околі дійсного набору  $x$ , а за всі інші сплачується штраф 1.

Для максимізації розглянемо логарифми ймовірностей, тому що це зручніше, ніж максимізація добутку кількох десятків тисяч або мільйонів значень

$$\mathbb{P}_X(x) = \ln \delta x^n + \sum_{i=1}^n \left\{ -\frac{x_i^2}{2} - \frac{\ln 2 + \ln \pi}{2} \right\},$$

$$\mathbb{P}_Y(t | x) = \ln \delta x^{w \cdot h} + \sum_{i \in I} \left\{ -\frac{y_i^2}{2 \cdot \sigma^2} - \frac{\ln 2 + \ln \pi + \ln \sigma^2}{2} \right\}.$$

Приберемо константні доданки, що не впливають на результат максимізації ймовірності появи даного зображення  $t$  з даними параметрами  $x$

$$\sum_{i \in I} \left\{ -\frac{y_i^2}{2 \cdot \sigma^2} - \frac{\ln \sigma^2}{2} \right\} - \sum_{i=1}^n \frac{x_i^2}{2} \rightarrow \max_{x \in X}.$$

Помножимо на  $-2$  та винесемо логарифм дисперсії за знак суми

$$w \cdot h \cdot \ln \sigma^2 + \sum_{i \in I} \frac{y_i^2}{\sigma^2} + \sum_{i=1}^n x_i^2 \rightarrow \min_{x \in X}. \quad (1.1)$$

Відомо, що  $n$  набагато менше за  $w \cdot h$ . Визначимо, за яких умов можна знехтувати ймовірністю набору параметрів  $x$ . Для деякого  $\varepsilon$  повинна виконуватись нерівність

$$\left| \frac{\sum_{i=1}^n x_i^2}{w \cdot h \cdot \ln \sigma^2 + \sum_{i \in I} \frac{y_i^2}{\sigma^2}} \right| < \varepsilon.$$

Розв'яжемо цю нерівність відносно дисперсії шуму

$$\frac{\sum_{i=1}^n x_i^2}{\varepsilon \cdot w \cdot h} < \left| \ln \sigma^2 + \frac{\sum_{i \in I} y_i^2}{\sigma^2 \cdot w \cdot h} \right|.$$

Права частина містить суму, яка дорівнює оцінці дисперсії ділень на реальну дисперсію шуму. Вважаємо, що зображення має достатньо пікселів, щоб сума була приблизно рівною одиниці

$$\frac{\sum_{i=1}^n x_i^2}{\varepsilon \cdot w \cdot h} < |\ln \sigma^2 + 1|.$$

Потрібно позбутися модуля. Якщо  $\ln \sigma^2 + 1 > 0$ , то

$$\sigma^2 > e^{-1} \approx 0.37.$$

Це завеликий шум, який сильно спотворить зображення, тому не будемо розглядати цей випадок. Отримали нерівність

$$-\frac{\sum_{i=1}^n x_i^2}{\varepsilon \cdot w \cdot h} - 1 > \ln \sigma^2.$$

Візьмемо експоненту від обох частин

$$\sigma^2 < \exp \left\{ -\frac{\sum_{i=1}^n x_i^2}{\varepsilon \cdot w \cdot h} - 1 \right\}.$$

Перепишемо суму квадратів параметрів моделі як їх середнє помножене на кількість

$$\sigma^2 < \exp \left\{ -\frac{\|x\|^2}{\varepsilon \cdot w \cdot h} - 1 \right\}.$$

Якщо в нерівність підставити максимально можливе значення для  $\|x\|^2$ , отримає-

мо більш жорсткі обмеження на дисперсію шуму. Такий прийом лише збільшить точність результату.

$$\sigma^2 < \exp \left\{ -\frac{n \cdot x_{max}^2}{\varepsilon \cdot w \cdot h} - 1 \right\}.$$

Нехай  $\varepsilon$  лінійно залежить від максимально можливого значення параметрів  $x$ .  
Тоді

$$\sigma^2 < \exp \left\{ -\frac{n}{\varepsilon' \cdot w \cdot h} - 1 \right\}, \quad \varepsilon = \varepsilon' \cdot x_{max}^2.$$

Параметри  $x$  мають стандартний нормальний розподіл, отже за правилом трьох сігма не варто брати значення, що за модулем перевищують 3. Поріг  $\varepsilon$  достатньо взяти 0.1 — це буде означати, що доданком, який вдесятеро менший за інший доданок, можна знехтувати. Тоді  $\varepsilon' \approx 0.01$ .

Якщо вважати шум з дисперсією  $e^{-1.2} \approx 0.3$  завеликим (таким, що не зустріється на реальних зображеннях), отримаємо залежність максимальної кількості параметрів від розміру зображення

$$n < 0.2 \cdot \varepsilon' \cdot w \cdot h \leq \frac{w \cdot h}{500} \quad \varepsilon' \leq 0.01. \quad (1.2)$$

Тоді буде виконуватись умова

$$\sigma^2 < e^{-1.2}.$$

Повернемося до (1.1) та замінимо дисперсію її оцінкою [16]

$$w \cdot h \cdot \ln \frac{\sum_{i \in I} y_i^2}{w \cdot h - 1} + w \cdot h - 1 + \sum_{i=1}^n x_i^2 \rightarrow \min_{x \in X}.$$

Позбавимось константних доданків

$$w \cdot h \cdot \ln \frac{\sum_{i \in I} y_i^2}{w \cdot h - 1} + \sum_{i=1}^n x_i^2 \rightarrow \min_{x \in X}.$$

В загальному випадку зображення містить не тільки обличчя. Генерація усіх можливих фонів для кожної моделі є недоцільною. Потрібно отримати формулу, яка не прив'язана до кожного пікселя вхідного зображення  $t$ . Для цього перепишемо доданки як середньоквадратичні відхилення

$$w \cdot h \cdot \ln \tilde{y}^2 + (n - 1) \cdot \tilde{x}^2 \rightarrow \min_{x \in X}.$$

Треба брати середньоквадратичне відхилення шуму по тим пікселям, де на згенерованому зображенні  $f(x)$  знаходиться обличчя. За умовою накладений шум є набором незалежних випадкових однаково розподілених величин, тому квадрат середньоквадратичного відхилення будь-якої їх сукупності буде оптимальною оцінкою їх дисперсії. Це означає, що квадрат середньоквадратичного відхилення шуму на обличчі буде тим точнішою оцінкою дисперсії шуму, чим більше місця займає обличчя на зображенні.

Знехтуємо ймовірністю параметрів  $x$

$$\tilde{y}^2 \rightarrow \min_{x \in X}.$$

Помножимо на  $(w \cdot h - 1)$ . Отримали мінімізацію відомої цільової функції — суми квадратів відхилень

$$\sum_{i \in I} y_i^2 = \sum_{i \in I} (t_i - f_i(x))^2 \rightarrow \min_{x \in X}.$$

Мінімізація суми квадратів різниць між дійсним та згенерованим зображенням розв'язує задачу з бінарною або інтервальною функцією штрафу та гаусовим шумом на зображенні без додаткових умов на кшталт розподілу параметрів  $x$ . Проте це знехтування доцільне при виконанні умов (1.2).



### 1.4.3 Різниця параметрів

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

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

Винесемо константи з добутоків

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

В даному випадку не можна замінити дисперсію на її оцінку, бо для кожного  $x$  вона буде різною, через що інтеграл ймовірностей не буде дорівнювати одиниці, що порушує умову, якій повинна задовольняти ймовірнісна міра [17]. Введемо константу

$$c_t = \frac{(2 \cdot \pi)^{-\frac{n+w \cdot h}{2}} \cdot (\sigma_t^2)^{-\frac{w \cdot h}{2}}}{\mathbb{P}(t)}$$

Ймовірність певного набору параметрів  $x$  на даному зображенні  $t$

$$\mathbb{P}(x | t) = c_t \cdot \delta x^{n+w \cdot h} \cdot \exp \left\{ -\frac{\|t - f(x)\|^2}{2 \cdot \sigma_t^2} \right\} \cdot \exp \left\{ -\frac{\|x\|^2}{2} \right\}.$$

Стратегія розпізнавання

$$q^*(t) = c_t \cdot \delta x^{n+w \cdot h} \cdot \sum_{x \in X} x \cdot \exp \left\{ -\frac{\|t - f(x)\|^2}{2 \cdot \sigma_t^2} \right\} \cdot \exp \left\{ -\frac{\|x\|^2}{2} \right\}.$$

Щільність розподілу набору параметрів  $x$  на даному зображенні  $t$

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

Стратегія розпізнавання у неперервному випадку

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

## 1.5 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

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

який є математичним очікуванням функції стандартного гаусового  $n$ -вимірного вектора

$$q^*(t) = c_t \cdot \mathbb{E}_\xi \left[ \xi \cdot \exp \left\{ -\frac{\|t - f(\xi)\|^2}{2 \cdot \sigma^2} \right\} \right], \quad \xi_i \sim \mathcal{N}(0, 1).$$

Внесемо все в експоненту для полегшення машинних обчислень

$$q^*(t) = \mathbb{E}_\xi \left[ \frac{\xi}{|\xi|} \cdot \exp \left\{ \ln |\xi| - \frac{\|t - f(\xi)\|^2}{2 \cdot \sigma^2} + \ln c_t \right\} \right].$$

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

$$\xi^{(i)} = \langle \xi_1^{(i)}, \xi_2^{(i)}, \dots, \xi_n^{(i)} \rangle, \quad \xi_j^{(i)} \sim \mathcal{N}(0, 1).$$

Для кожного вектору розрахувати значення підінтегрального виразу

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

Значення константи  $c_t$  визначимо з умови нормування

$$c_t \cdot \sum_{i=1}^N \exp \left\{ \frac{\|t - f(\xi^{(i)})\|^2}{2 \cdot \sigma^2} + \frac{\|\xi^{(i)}\|^2}{2} \right\} = 1.$$

Явний вигляд

$$c_t = \frac{1}{\sum_{i=1}^N \exp \left\{ \frac{\|t - f(\xi^{(i)})\|^2}{2 \cdot \sigma^2} + \frac{\|\xi^{(i)}\|^2}{2} \right\}}.$$

Оцінкою інтегралу буде середнє значення

$$\bar{Q}_N = \frac{\sum_{i=1}^N \hat{Q}^{(i)}}{N}.$$

Визначимо оптимальне значення  $N$ . Згадаємо центральну граничну теоре-

му

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

Для певного  $z_\gamma > 0$

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

де  $\Phi$  — інтегральна функція стандартного гаусового розподілу. Нас цікавить ймовірність

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

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

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

Поділимо обидві частини нерівності на модуль середнього значення виборки

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

Позначимо відносну похибку

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

та перепишемо вираз

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

Перенесемо все окрім кореню від кількості спроб в праву частину нерівності

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

та піднесемо обидві частини нерівності до квадрату

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

Перепозначимо отриману ймовірність

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

Щоб виразити  $z_\gamma$ , потрібно скористатися квантилем нормального розподілу

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

Щоб відносна похибка обчислення математичного очікування методом Монте-Карло не перевищувала  $\varepsilon$  з ймовірністю  $\gamma$  і вище, необхідна та достатня кіль-

кість ітерацій рахується за формулою

$$N = \left\lceil \frac{z_\gamma^2 \cdot D(Q)}{\varepsilon^2 \cdot \overline{Q}_N^2} \right\rceil, \quad z_\gamma = \Phi^{-1} \left( \frac{\gamma + 1}{2} \right). \quad (1.3)$$

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

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

Формула (1.3) 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**

## **ВИСНОВКИ**

В результаті виконання роботи вдалося.



## ПЕРЕЛІК ПОСИЛАНЬ

1. Parke, Frederick I. Computer Generated Animation of Faces / Frederick I. Parke // *Proceedings of the ACM Annual Conference - Volume 1.* — 1972. — Pp. 451–457.
2. A 3D Face Model for Pose and Illumination Invariant Face Recognition / P. Paysan, R. Knothe, B. Amberg et al. // *Proceedings of the 6th IEEE International Conference on Advanced Video and Signal based Surveillance (AVSS) for Security, Safety and Monitoring in Smart Environments.* — 2009.
3. Blanz, V. A morphable model for the synthesis of 3D faces / V. Blanz, T. Vetter // *Proceedings of the 26th annual conference on Computer graphics and interactive techniques.* — 1999. — Pp. 187–194.
4. Jones, M. J. Multidimensional morphable models / M. J. Jones, T. Poggio. — 1998. — Jan. — Pp. 683–688.
5. Blanz, V. Face Identification across Different Poses and Illuminations with a 3D Morphable Model / V. Blanz, S. Romdhani, T. Vetter // *Automatic Face and Gesture Recognition, 2002. Proceedings. Fifth IEEE International Conference on.* — 2002. — Pp. 192–197.
6. Blanz, V. Face recognition based on fitting a 3D morphable model / V. Blanz, T. Vetter // *IEEE Transactions on Pattern Analysis and Machine Intelligence.* — 2003. — Sept. — Vol. 25, no. 9. — Pp. 1063–1074.
7. Tyshchenko, M. 3D Reconstruction of Human Face Based on Single or Several Images / M. Tyshchenko // *Управляющие системы и машины.* — 2011. — № 2. — С. 79–85.
8. Rossi, Francesca. Handbook of Constraint Programming (Foundations of

- Artificial Intelligence) / Francesca Rossi, Peter van Beek, Toby Walsh. — New York, NY, USA: Elsevier Science Inc., 2006. — P. 978.
9. Saito, Shunsuke. Real-Time Facial Segmentation and Performance Capture from RGB Input / Shunsuke Saito, Tianye Li, Hao Li // *Computer Vision – ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part VIII*. — 2016. — Pp. 244–261.
  10. Face2Face: Real-time Face Capture and Reenactment of RGB Videos / J. Thies, M. Zollhöfer, M. Stamminger et al. // *Proc. Computer Vision and Pattern Recognition (CVPR), IEEE*. — 2016.
  11. Amberg, Brian. Optimal Step Nonrigid ICP Algorithms for Surface Registration. / Brian Amberg, Sami Romdhani, Thomas Vetter // *2007 IEEE Conference on Computer Vision and Pattern Recognition*. — 2007.
  12. Айвазян, С.А. Прикладная статистика: Классификация и снижение размерности : Справочное издание / С.А. Айвазян. — М.: Финансы и статистика, 1989. — 606 с.
  13. Berger, J.O. Statistical Decision Theory: Foundations, Concepts, and Methods / J.O. Berger // *Springer Series in Statistics*. — Springer New York, 1980. — P. 428.
  14. Schlesinger, M.I. Ten Lectures on Statistical and Structural Pattern Recognition / M.I. Schlesinger, V. Hlavác // *Computational Imaging and Vision*. — Springer Netherlands, 2002. — P. 522.
  15. Wald, A. Selected Papers in Statistics and Probability / A. Wald, I.M. Statistics // No. 10. — McGraw-Hill, 1955. — P. 702.
  16. Боровков, А. А. Математическая статистика / А. А. Боровков // *Учебники*

для ВУЗов. Специальная литература. — 4 изд. — Санкт-Петербург: Лань, 2010. — С. 705.

17. Дороговцев, А. Я. Элементы общей теории меры и интеграла / А. Я. Дороговцев. — К.: Выща школа, 1989. — С. 152.