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 Попередні роботи

Найпершою в створенні тривимірної породжувальної моделі обличчя вважається робота 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 Униерситету 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 в термінах задачі розмітки [8], а нові моделі облич генеруються як зважене середнє.

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 ICP [9]. 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 арифметичнє по всім моделям

$$\overline{f}_v = \frac{\sum\limits_{f \in F} f_v}{|F|}, \qquad v \in V,$$

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 [10]. 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 робіть $L \cdot \widetilde{M} \cdot L^{-1}$ 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

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

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

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

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 $10^{6 \cdot 256} = 10^{1536}$ різних зображень розміром 1000×1000, 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

$$f : X \to T.$$

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

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

Інтенсивність i пікселя

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

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

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

З визначення умовної ймовірності випливає

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

Позбавимося умовної ймовірності, замінивши ξ на x

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

Це означає, що ймовірність отримати на вході зображення t згенероване з параметрами x дорівнює добутку ймовірності отримати зображення з шумом $t - f\left(x\right)$ 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 [11]

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

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

$$q(t) = x.$$

Математичне очікування функції витрат W як функції випадкової пари $\langle t, x \rangle$ для даного вирішального правила q називається Баєсовим ризиком [12]

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

де використовується інтеграл Лебега-Стілтьєса, а F — функція сумісного розподілу зображення та параметрів зображеної моделі. Задача — знайти таке вирішальне правило q, яке мінімізує Баєсів ризик [13]

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

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

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

1.3.3 Бінарна функція витрат

Досить росповсюдженою, 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

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

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

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

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

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 отптимуму, 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 δ-околиці обраного параметру x'

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

Оберемо стратегію, що мінімізує математичне очікування цієї функції витрат

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

В результаті

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

Потрібно обрати таку точку, щоб ймовірність того, що параметри належать її δ-околиці, була більшою, ніж для δ-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

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

Породжувальна 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

$$G_v(x) = g.$$

Координати кожної вершини v породжувальної моделі отримуються шляхом перемноження компонент параметру x на відповідний коефіцієнт λ^v отриманий за допомогою методу головних компонент та додавання результату до середнього положення поточної вершини

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

Функція витрат має вигляд

$$W\left(x,x'\right)=\left\|G\left(x\right)-G\left(x'\right)\right\|^{2}=\sum_{v\in V}\left[G_{v}\left(x\right)-G_{v}\left(x'\right)\right]^{2}=$$
$$=\sum_{v\in V}\sum_{p\in P}\left[\lambda_{p}^{v}\cdot\left(x_{p}-x'_{p}\right)\right]^{2}=\sum_{p\in P}\left\{\left(x_{p}-x'_{p}\right)^{2}\cdot\sum_{v\in V}\left(\lambda_{p}^{v}\right)^{2}\right\}=$$
$$=\left|\beta_{p}^{2}=\sum_{v\in V}\left(\lambda_{p}^{v}\right)^{2}\right|=\sum_{p\in P}\beta_{p}^{2}\cdot\left(x_{p}-x'_{p}\right)^{2}.$$

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

$$q^{*}\left(t\right)=\arg\min_{x'}\left\{\int_{X}\sum_{i=1}^{n}\beta_{i}^{2}\cdot\left(x'_{i}-x_{i}\right)^{2}dF\left(t,x\right)\right\}.$$

Щоб мінімізувати неперервну функцію від параметрів x'_{i}, можна взяти по них похідну

$$\frac{\partial\int\limits_{X}\sum\limits_{i=1}^{n}\beta_{i}^{2}\cdot\left(x'_{i}-x_{i}\right)^{2}dF\left(t,x\right)}{\partial x'_{i}}=2\cdot\int_{X}\beta_{i}^{2}\cdot\left(x'_{i}-x_{i}\right)dF\left(t,x\right),\qquad i=1..n$$

та прирівняти до нуля

$$\sum_{X}\left(x'_{i}-x_{i}\right)dF\left(t,x\right)=0,\qquad i=1..n.$$

Значення компоненти

$$x'_{i}=\frac{\int\limits_{X}x_{i}dF\left(t,x\right)}{\int\limits_{X}dF\left(t,x\right)}=\frac{\int\limits_{X}x_{i}dF\left(t,x\right)}{\mathbb{P}\left(t\right)}$$

Результуюча стратегія

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

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

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

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

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

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

$$\frac{\partial \int\limits_X \sum\limits_{i=1}^{n} (x'_i - x_i)^2 \, dF(t,x)}{\partial x'_i} = 2 \cdot \int\limits_{x \in X} dF(t,x) \cdot (x'_i - x_i), \qquad i = 1..n$$

та прирівняти до нуля

$$\int\limits_X (x'_i - x_i) \, dF(t,x) = 0, \qquad i = 1..n.$$

Значення компоненти

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

Результуюча стратегія

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

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 Інтервальна функція витрат

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

Дискретизуємо простір параметрів X. Введемо невелике δx 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

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

де

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

Далі треба розбити простір на гіперкуби зі стороною δx 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

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

Введемо зручне позначення для шуму в пікселі i

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

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

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

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

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

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

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

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

Помножимо на -2 та винесемо логарифм дисперсії за знак суми

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

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

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

Розв'яжемо цю нерівність відносно дисперсії шуму

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

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

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

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

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

Візьмемо експоненту від обох частин

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

Перепишемо суму квадратів параметрів моделі як їх середнє помножене на кількість

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

Для зображень, що використовуються на практиці, дріб буде приймати настільки мале значення, що їм можна знехтувати

$$\sigma^2 < e^{-1}.$$

Це означає, що внеском ймовірності параметрів x теж можна знехтувати на не зовсім малих зображеннях (наприклад, 500×500).

Повернемося до (1.1) та замінимо дисперсію на її оцінку

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

Позбавимося константних доданків

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

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

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

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

$$\widetilde{y}^2 \to \min_{x \in X}.$$

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

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

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.

1.4.2 Різниця параметрів

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

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

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

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

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

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

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

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

Стратегія розпізнавання

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

У неперервному випадку сума змінюється на інтеграл та зникають об’єми гіперкубів

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

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

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

1.5 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

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

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 для полегшення машинних обчислень

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

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 $z_\gamma > 0$

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

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

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

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

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

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

$$N = \left\lceil \frac{z_\gamma^2 \cdot D\left(Q\right)}{\varepsilon^2 \cdot \overline{Q}_N^2} \right\rceil, \qquad z_\gamma = \Phi^{-1}\left(\frac{\gamma+1}{2}\right). \tag{1.2}$$

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 (1.2) 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. 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. 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.

10. Айвазян, С.А. Прикладная статистика: Классификация и снижение размерности : Справочное издание / С.А. Айвазян. — М.: Финансы и статистика, 1989. — 606 с.

11. Berger, J.O. Statistical Decision Theory: Foundations, Concepts, and Methods / J.O. Berger // Springer Series in Statistics. — Springer New York, 1980. — P. 428.

12. 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.

13. Wald, A. Selected Papers in Statistics and Probability / A. Wald, I.M. Statistics // No. 10. — McGraw-Hill, 1955. — P. 702.