

Відновлення просторової конфігурації людського обличчя за його фотознімком на основі генеративної моделі

Тищенко М.А.

Міжнародний науково-навчальний центр інформаційних технологій та систем
irtc.org.ua/image

Анотація

В даній роботі запропоновано метод відновлення тривимірної моделі обличчя людини за його фотознімком. Метод базується на припущенні, що множина всіх можливих форм людського обличчя є в певному сенсі опуклою. Форма конкретного обличчя апроксимується опуклою комбінацією деяких опорних облич. Набір опорних облич називатимемо генеративною моделлю.

Якщо генеративна модель побудована, на основі вхідного зображення запропонований алгоритм визначає такі параметри форми обличчя, а також ракурс зйомки та освітлення, які мінімізують відмінність між вхідним та згенерованим зображеннями.

1. Вступ

Відновлення просторової конфігурації людського обличчя займає окрему, досить велику нішу серед завдань комп'ютерного зору. Роботи, присвячені цій тематиці можна розділити на два великі класи: геометрично-орієнтовані та орієнтовані на генеративну модель.

Підходи, засновані на генеративній моделі, експлуатують методологію аналізу через синтез. Генеративну модель обличчя можна уявляти у вигляді ящика з ручками, які можна підкручувати. Залежно від положення цих ручок ящик генерує те або інше зображення того або іншого людського обличчя. Завдання полягає в то-

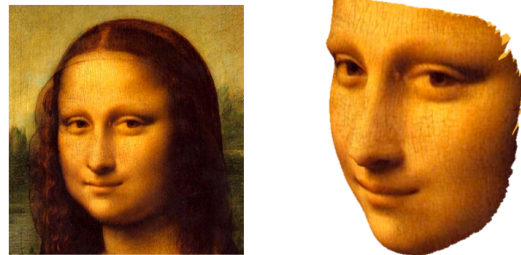


Рис. 1: Приклад відновлення тривимірної моделі обличчя за одним зображенням.

му, щоб встановити ручки цього ящика в таке положення, щоб він згенерував саме вхідне зображення, або, принаймі, якомога більш близьке до нього. При цьому тривимірною моделлю обличчя однозначно визначається положенням ручок цього ящика, тобто параметрами генеративної моделі. Істотною перевагою такого підходу є можливість якісного відновлення поверхні обличчя навіть за одним зображенням.

На рис. 1 наведено приклад відновлення просторової конфігурації обличчя за одним його зображенням за допомогою запропонованого методу.

2. Генеративна модель людського обличчя

Як ми вже казали, генеративну модель людського обличчя найпростіше уявляти собі у вигляді ящика з ручками, які можна підкручувати. Залежно від положення ру-

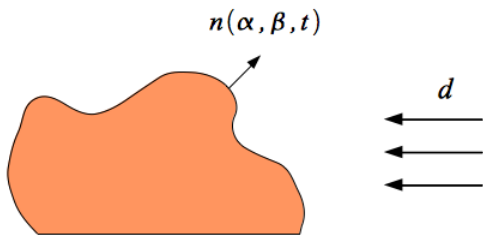


Рис. 2: Модель спрямованого освітлення.

чок, цей ящик генерує те чи інше зображення того чи іншого людського обличчя. У цьому ящику є ручки трьох типів. Деякі з них відповідають за форму поверхні (α), інші за положення камери в просторі (β), треті за освітлення (γ). Слід зауважити, що коли ми уявляємо поверхню обличчя, положення камери і освітлення у вигляді ручок цього ящика, ми фактично стверджуємо, що всі ці величини можуть бути описані скінченною кількістю параметрів. З положенням камери все просто. Воно дійсно задається скінченною, і, більше того, малою кількістю параметрів. А от з освітленням і поверхнею обличчя все дещо складніше. Адже зовсім не очевидно, що взагалі кажучи нескінченну множину різних варіантів освітлення і не менш нескінченну множину всіх можливих поверхонь людських облич можна описати скінченною кількістю параметрів. Про те як параметризувати множину поверхонь людських облич мова піде в наступних підрозділах. А з параметризацією освітлення розберемося вже зараз.

2.1. Модель спрямованого освітлення

Вважатимемо, що поверхня освітлена одним джерелом спрямованого освітлення. Це таке джерело, яке світить з нескінченності паралельними променями, і характеризується одним вектором d , напрям якого задає напрям освітлення, а довжина – його інтенсивність. Освітленість поверхні в деякій точці t визначається як скалярний добуток нормалі до поверхні в цій точці на

вектор d :

$$L(\alpha, \beta, d, t) = \sum_{i=1}^3 d_i \cdot n_i(\alpha, \beta, t).$$

Вважатимемо, що є ще так зване фонове освітлення – це просто константа, яка додається до освітленості всіх точок поверхні:

$$L(\alpha, \beta, d, c, t) = \sum_{i=1}^3 d_i \cdot n_i(\alpha, \beta, t) + c. \quad (1)$$

Подивимось, що станеться, якщо маємо не одне джерело спрямованого освітлення, а скажімо, два. Тоді освітленість деякої точки поверхні буде рівна

$$\sum_{i=1}^3 d_i^1 \cdot n_i(\alpha, \beta, t) + \sum_{i=1}^3 d_i^2 \cdot n_i(\alpha, \beta, t), \quad (2)$$

де d^1 – вектор, що характеризує перше джерело, а d^2 – друге. Очевидно, що (2) можна переписати у вигляді:

$$\sum_{i=1}^3 (d_i^1 + d_i^2) \cdot n_i(\alpha, \beta, t). \quad (3)$$

Таким чином, в цій моделі освітлення все одно маємо ми одне джерело освітлення, два, або тисячу. Всі вони еквівалентні деякому одному джерелу.

Таким чином, ми параметризували освітлення вектором d та скаляром c :

$$\gamma = \{d, c\}. \quad (4)$$

2.2. Опукла комбінація облич

Ми виходимо з припущення, що множина поверхонь людських облич є в певному сенсі опуклою. Іншими словами, якщо ми візьмемо дві поверхні, що є людськими обличчями, та усереднимо їх – то отримаємо також обличчя.

Припустимо, що є набір поверхонь характерних людських облич $\{s_1, s_2, \dots, s_n\}$, які називатимемо *базисними поверхнями*.

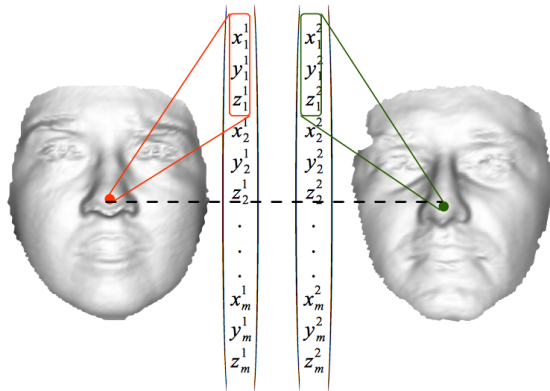


Рис. 3: Відповідні точки різних моделей мають знаходитись на однакових позиціях в своїх векторах.

Будемо апроксимувати поверхню довільного людського обличчя опуклою комбінацією базисних поверхонь:

$$s = \sum_{i=1}^n \alpha_i \cdot s_i, \quad \sum_{i=1}^n \alpha_i = 1. \quad (5)$$

Коефіцієнти опуклої комбінації $\{\alpha_i : i = 1, 2, \dots, n\}$ якраз і є параметрами поверхні обличчя α , про які ми говорили раніше.

Але що саме слід розуміти під усереднюванням двох поверхонь? Прояснимо цю обставину на максимально спрощеному прикладі.

2.3. Генеративна модель обличчя

Вважатимемо, що базисні поверхні $\{s_1, s_2, \dots, s_n\}$ задані у вигляді векторів, що містять тривимірні координати точок поверхні:

$$s_i = (x_1^i, y_1^i, z_1^i, x_2^i, y_2^i, z_2^i, \dots, x_{m_i}^i, y_{m_i}^i, z_{m_i}^i)^T, \quad (6)$$

де m_i – кількість точок в i -й поверхні.

Для того, щоб до таких поверхонь можна було застосовувати операцію (5) слід забезпечити виконання принаймні двох вимог:

1. Кількість точок в усіх поверхнях має бути однаковою:

$$m_i = m, i = 1, 2, \dots, n.$$

2. Компоненти векторів мають бути впорядковані так, щоб однакові за змістом точки стояли на однакових позиціях у своїх векторах. Скажімо, якщо координата кінчика носа першої моделі (базисної поверхні) розташована на першій позиції у векторі s_1 , то координати кінчика носа всіх інших поверхонь також повинні розташовуватися на першій позиції у своїх векторах s_2, s_3, \dots, s_n (див. мал. 3).

Перша з вказаних вимог гарантує, що всі вектори, що входять в суму (5) мають однакову розмірність, таким чином забезпечуючи коректність визначення операції (5).

Друга вимога забезпечує правильне усереднювання поверхонь, гарантуючи, що в результаті застосування операції (5) вийде саме людське обличчя, а не довільна поверхня.

Виконання цих двох вимог не викликає труднощів якщо вказані відповідності між точками різних базисних поверхонь. Саме відшукування таких відповідностей і є основною складністю у виконанні вказаних вимог. Спосіб встановлення таких відповідностей описано в [3]. Схожий підхід було використано в [4] для генерування міміки людського обличчя.

3. Відновлення поверхні обличчя за фотознімком

В попередньому розділі ми детально описали внутрішній устрій генеративної моделі людського обличчя. Тепер нам знову зручніше піднятися на такий рівень абстракції, на якому генеративна модель – це просто ящик з трьома групами параметрів: тими, які відповідають за форму поверхні (α), тими, які відповідають за положення камери (β), і тими, які відповідають за освітлення (γ). При конкретних значеннях цих параметрів генеративна модель синтезує зображення людського обличчя що має форму, задану параметрами α , знятого з ракурсу, заданого параметрами β , в умо-

вах освітлення, заданих параметрами γ .

Тепер самий час задатися питанням: як же виглядатиме це синтезоване зображення? Згадаємо, що наша генеративна модель синтезує нетекстуровану поверхню. Для визначеності припустимо, що всі точки поверхні зафарбовані в білий колір. Тоді на згенерованому зображенні буде щось схоже на фотографію гіпсової статуї. Опишемо спочатку спосіб відновлення тривимірної моделі для випадку, коли вхідним зображенням є саме фотографія гіпсової статуї людського обличчя. Після цього ми вкажемо як модифікувати запропонований підхід для роботи з фотографіями реальних людських облич.

3.1. Відновлення гіпсової статуї

Нехай маємо вхідне зображення

$$I : T \rightarrow RGB, \quad (7)$$

де T – множина пікселів зображення, $RGB = \{0, 1, \dots, 255\} \times \{0, 1, \dots, 255\} \times \{0, 1, \dots, 255\}$.

Нехай задано зображення, синтезоване генеративною моделлю за параметрів α, β і γ в моделі освітлення (1):

$$L_{\alpha, \beta, \gamma} : T \rightarrow RGB. \quad (8)$$

Визначимо відмінність вхідного зображення від згенерованого:

$$F(I, L_{\alpha, \beta, \gamma}) = \sum_{t \in \hat{T}} (I(t) - L_{\alpha, \beta, \gamma}(t))^2, \quad (9)$$

де \hat{T} – підмножина пікселів, в які проектується хоча б одна точка поверхні обличчя.

Задача відновлення тривимірної моделі обличчя за його зображенням полягає у відшукуванні таких параметрів α, β і γ , при яких відмінність вхідного зображення від згенерованого є мінімальною:

$$F(I, L_{\alpha, \beta, \gamma}) \rightarrow \min_{\alpha, \beta, \gamma}. \quad (10)$$

У моделі освітлення (1) функція (10) при фіксованих параметрах α і β мінімізується по γ методом найменших квадратів. Мінімізація по α і β виконується методом Нелдера-Мида [5].

3.2. Відновлення людського обличчя

Для того, щоб обробляти фотографії реальних людських облич, а не зображення одноколірних гіпсових статуй, вирізуватимемо з фотознімку всі елементи, відмінні від шкіри обличчя: брови, очі і губи. Оскільки однакові за змістом точки всіх базисних поверхонь розташовані на однакових позиціях у своїх векторах, то виконати видалення вказаних фрагментів досить тільки для однієї з базисних поверхонь. На всі інші моделі ця операція переноситься автоматично.

4. Література

- [1] Alex Rav-Acha, Pushmeet Kohli, Carsten Rother, Andrew Fitzgibbon. Unwrap mosaics: a new representation for video editing // In Proc. of SIGGRAPH 2008.
- [2] V. Blanz, T. Vetter. "Face identification across different poses and illumination with a 3D morphable model", IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 25, pp. 1063-1074, 2003.
- [3] Тищенко М.А. Відновлення тривимірної форми людського обличчя за одним зображенням. // Праці 8 міжнародної конференції "Шевченківська весна 2010", Київ, 22-26 Бер. 2010.
- [4] Недашківський О.В., Тищенко М.А. Генерування зображень різних виразів людських облич. // Праці 9 Всеукраїнської міжнародної конференції УкрОбраз-2008, Київ, 3-7 Лист. 2008.
- [5] J.A. Nelder and R. Mead, A simplex method for function minimization // Computer Journal, 1965, vol 7, pp 308–313.