Розділ 2: Алгоритм дифузії
Домашнє завдання

1 Основне (2 бали)

1.1 Супермодулярність і репараметризація

В алгоритмах пошуку оптимальної розмітки, що використовують перехід до двоїстих змінних, використовується такий факт, що для двох супермодулярних функцій $g_1 : K^2 \to \mathbb{R}$ і $g_2 : K^2 \to \mathbb{R}$ функція

$$g(k, k') = \max_{\varkappa \in K} \{g_1(k, \varkappa) + g_2(\varkappa, k')\}$$

також супермодулярна. Доведіть цю тотожність.

1.2 Супермодулярність і увігнутість

У деяких випадках зручно використовувати увігнуті функції в якості вагів. Доведіть, що для будь-якої увігнутої $f : \mathbb{R} \to \mathbb{R}$ та скінченної множини міток $K \subset \mathbb{R}$ (де впорядкованість та різниця визначаються природньо) функція

$$g(k, k') = f(k - k')$$

супермодулярна.

2 Додаткове (3 бали)

2.1 Навчання на задачах розмітки

Нехай $T = \{1, \dots, n\}^2$ — поле зору, $\tau \subset T^2$ — структура другого порядку, визначена на полі зору, $X \ni x : T \to \{0, 1\}$ — зображення, задане на полі зору T, $q : T \times X \times \{0, 1\} \to \mathbb{R}$ — відома унарна штрафна функція, що залежить від зображення, $g : \tau \times \{0, 1\}^2 \to \mathbb{R}$ — **не**відома бінарна штрафна функція, що **не** залежить від зображення. Маємо експоненційний розподіл виду

$$p(k; x) = \frac{1}{Z(x)} \cdot \exp\left\{\sum_{t \in T} q_t(k_t; x) + \sum_{tt' \in \tau} g_{tt'}(k_t, k_{t'})\right\}, \quad k : T \to \{0, 1\}, \quad x : T \to \{0, 1\}.$$

Побудувати ефективний алгоритм, що на вхід приймає таку множину пар $\left\{\langle x^i, k^i \rangle : i = \overline{1, m}\right\}$, що

$$p\left(k^i; x^i\right) > p\left(k; x^i\right), \quad \forall k \neq k^i, \quad \forall i = \overline{1, m},$$

а на виході дає бінарну штрафну функцію $g : \tau \times \{0, 1\}^2 \to \mathbb{R}$, для якої ці нерівності виконуться для всієї навчальної множини.

Вважати, що ми знаємо ефективний алгоритм пошуку найкращої розмітки, а також інший алгоритм, який для скінченної множини $X \subset \mathbb{R}^m$ довільної скінченної розмірності m, такої що існує невідоме $\alpha \in \mathbb{R}^m$, для якого

$$\alpha \cdot x > 0, \quad x \in X,$$

знайде одне з можливих α за скінченну кількість кроків (наприклад, алгоритм навчання персептрону або алгоритм Козинця).

3 Комп'ютерне (4 бали)

3.1 Основи текстурної сегментації

Задача

Для зашумленого кольорового зображення $x : T \to \{0, \dots, 255\}^3$ побудувати ітеративне наближення до оптимальної бінарної сегментації за допомогою алгоритму дифузії

$$k^* \in \operatorname*{argmax}_{k : T \to \{0,1\}} \frac{1}{Z(x)} \cdot \exp\left\{-\sum_{t \in T} \left[\!\left[\|x_t - c(k_t)\| > \left\|x_t - c\left(\tilde{k}_{t'}\right)\right\| \right]\!\right] - \sum_{tt' \in \tau} \alpha \cdot [\![k_t \neq k_{t'}]\!]\right\},$$

де $\tilde{k}$ означає заперечення до k, відображення $c : K \to \{0, \ldots, 255\}^3$ дає колір сегменту (студент самостійно обирає два кольори, на які хоче розбити зображення), константу $\alpha > 0$ студент обирає під конкретні зображення, $[\![\cdot]\!]$ — дужки Айверсона, $0 < Z$ — нормуючий множник розподілу, а τ — така структура сусідства, що сусідами пікселя t з координатами $\langle t_x, t_y \rangle$ є

$$N(t) = \{\langle t_x - 1, t_y \rangle, \langle t_x + 1, t_y \rangle, \langle t_x, t_y - 1 \rangle, \langle t_x, t_y + 1 \rangle\}.$$

Мета

Закріпити навички пошуку найбільш ймовірного стану випадкового поля.

Завдання

На вхід програмі подається

- шлях до зашумленого зображення;
- значення α;
- кольори двох сегментів (або інтенсивності у випадку чорно-білих зображень).

Програма малює результати кожної ітерації пошуку сегментації.