

Поиск зрачка на изображении глаза методом проекций яркости.

В. В. Панкратов

Изучается проблема быстрого распознавания зрачка на изображении. Мы работаем с изображением как с функцией яркости от пары аргументов. Чтобы улучшить скорость ее обработки, предполагается применение метода проекции, который использует функцию проекции интенсивности на горизонтальные и вертикальные оси и используя ее строит положение центра глаза. Предполагается модифицировать и исследовать как можно оптимизировать этот алгоритм, используя нейросети.

1 Введение

Одна из задач, которую ставит перед собой обработка компьютерных изображений - определение лица человека на фотографии. Такая формулировка вносит неточность, поэтому принято определять положения лица путем его внешне заметных черт, одной из которых является центр зрачка. Эту задачу решает, например, статья [2], однако нас интересует более простой и быстрый алгоритм. В статье [1] предложен быстрый алгоритм нахождения центра, который определяет верхнюю нижнюю правую и левую "границы" глаза путем построения функции проекции, определенной в [3]. В данной работе наша цель модифицировать этот алгоритм, изменив положение осей, построив проекции на них, и наблюдая за изменением результатов. Также мы применим нейросеть для этой задачи и определим, какой вклад она окажет на результат работы предложенного алгоритма.

2 Постановка задачи

Дано цифровое изображение лица. Предполагается, что оно черно-белое, то есть каждой пиксель определяется функцией яркости $I(i, j)$. Проводится сокращение размерности данных до одномерной картины, используя функции проекции (для простоты укажем их определение для горизонтальной и вертикальной проекций).

$$P_H(j) = \frac{1}{i_M - i_m + 1} \sum_{i=i_m}^{i=i_M} I(i, j)$$

$$P_V(i) = \frac{1}{j_N - j_n + 1} \sum_{j=j_n}^{j=j_N} I(i, j)$$

Требуется решить задачу нахождения центра глаза по этим проекциям.

Пусть дана тренировочная выборка \mathcal{D} объектов и ответов: (\bar{x}_i, y_i) . Будем рассматривать отдельно задачу для левого и правого глаза. Необходимо определить оптимальную функцию f , дающую оптимальный ответ на \mathcal{D} , то есть решить задачу минимизации функционала качества:

$$Q(f, \mathcal{D}) = \frac{1}{|\mathcal{D}|} \sum \mathcal{L}(\bar{x}, y)$$

Функционал качества отсюда определим как сумму квадратов расстояний до действительных их координат:

$$Q(f, \mathcal{D}) = \sum_{\mathcal{D}} (y_i - f(\bar{x}_i))^2 = \sum_{\mathcal{D}} \varepsilon^2$$

Получаем задачу оптимизации

$$f^* = \arg \min_f Q(f, \mathcal{D})$$

Ее можно решить при помощи градиентного спуска

$$-dQ = \sum (y_i - f(\bar{x}_i)) df$$

На тестовой выборке качество полученных результатов оцениваем иначе:

$$\varepsilon = \frac{\max(\varepsilon_L, \varepsilon_R)}{D_{LR}}$$

Здесь D_{LR} – расстояние между центрами глаз, а $\varepsilon_{L,R}$ – ε левого или правого глаза соответственно.

Литература

- [1] G. C. Feng and P. C. Yuen. Variance projection function and its application to eye detection for human face recognition. *Pattern Recognition Letters*, 19(9):899–906, July 1998.
- [2] L. Florea, C. Florea, and C. Vertan. Robust eye centers localization with zero-crossing encoded image projections. *CoRR*, 2015.
- [3] J. Song, Z. Chi, and J. Liu. A robust eye detection method using combined binary edge and intensity information. *Pattern Recognition*, 39(6):1110–1125, 2006.
- [4] Z.-H. Zhou and X. Geng. Projection functions for eye detection. *Pattern Recognition*, 2004.

Поступила в редакцию