

Оптимална бинаризация на цветни изображения, чрез линеен дискриминант на Фишер в цветовото (RGB) пространство

Иван Арабаджийски, ФН: 5МІ3400052

Мотивация за избор на темата

- Харесах темите за бинаризация по време на курса
- Искам да науча повече методи за бинаризация
- Не сме правили бинаризация на цветно изображение
- Звучеше интересно

Каква е основната цел на проекта?

Бинаризацията е:

- процес на отделяне на обекта от фона, в дадено изображение чрез намиране на 'праг' T , който го разделя. Всички пиксели с по-малка стойност от прага стават обект, а всички с по-голяма - фон.
- Основната задача на един бинаризиращ алгоритъм е да намери оптималния праг.
- Върху това ще се съсредоточим и ние.

Ще:

- имплементира линеен дискриминант на Фишър върху полутонови изображения и ще го сравнява с други известни алгоритми.
- имплементира линеен дискриминант на Фишър върху изображения в цветовото (RGB) пространство.

Какво е линеен дискриминант на Фишър?

Това е метод, използван в статистиката за намиране на линейна комбинация от качества (features), които характеризират или отделят два или повече класа от обекти.

Фишър дефинира разделението между два класа като отношението между междукласовата дисперсия и вътрешнокласовата дисперсия. (the variance between the classes to the variance within the classes). А ние ще искаме да максимизираме това отношение.

В нашата задача класовете са два - обект и фон. Нека имаме полутоново изображение с L нива на сивото (256 в случая) - $[0, 1, \dots, L-1]$. Да отбележим броят на пикселите с интензитет на сивото i с n_i , като броят на всички пиксели $N = n_0 + n_1 + n_2 + \dots + n_{L-1}$. Вероятността за това ниво на сивото i да се появи в изображението е

$$p_i = \frac{n_i}{N}, \quad p_i \geq 0, \quad \sum_{i=0}^{L-1} p_i = 1.$$

Малко математика

Сега си представяме, че границата, която търсим, е k . И разделяме пикселите на изображението в двата класа C_1 и C_2 (C_1 е множеството от пиксели с нива $[0, 1, \dots, k]$, а останалите принадлежат на C_2). Тогава вероятностите за тези класове

$$w_1(k) = \sum_{i=0}^k p_i, \quad w_2(k) = 1 - w_1(k)$$

Медианите на нивата на сиво:

$$m_1(k) = \sum_{i=0}^k \frac{ip_i}{w_1}, \quad m_2(k) = \sum_{i=k+1}^{L-1} \frac{ip_i}{w_2},$$

Съответните дисперсии:

$$\sigma_1^2 = \sum_{i=0}^k \frac{(i - m_1)^2 p_i}{w_1}, \quad \sigma_2^2 = \sum_{i=k+1}^{L-1} \frac{(i - m_2)^2 p_i}{w_2}.$$

Вътрешно класовата дисперсия:

$$\sigma_w^2 = w_1 \sigma_1^2 + w_2 \sigma_2^2,$$

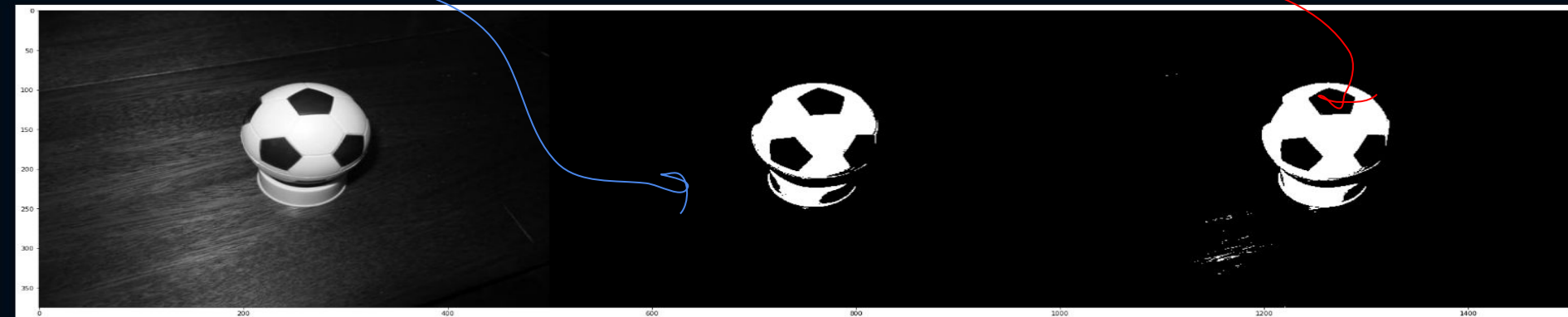
И накрая целевата функция, която определя колко добре разделя границата t : $\lambda(k) = \sigma_w^{-2} * (m_2(k) - m_1(k))^2$,

Алгоритъмът?

Пробваме за всички k от 1 до $L - 1$ и намираме, в коя стойност на k λ се максимизира. Това ще е оптималната ни граница t .

Резултати до момента върху полутонови изображения

Otsu: 118.0
Fisher's linear discriminant: 144



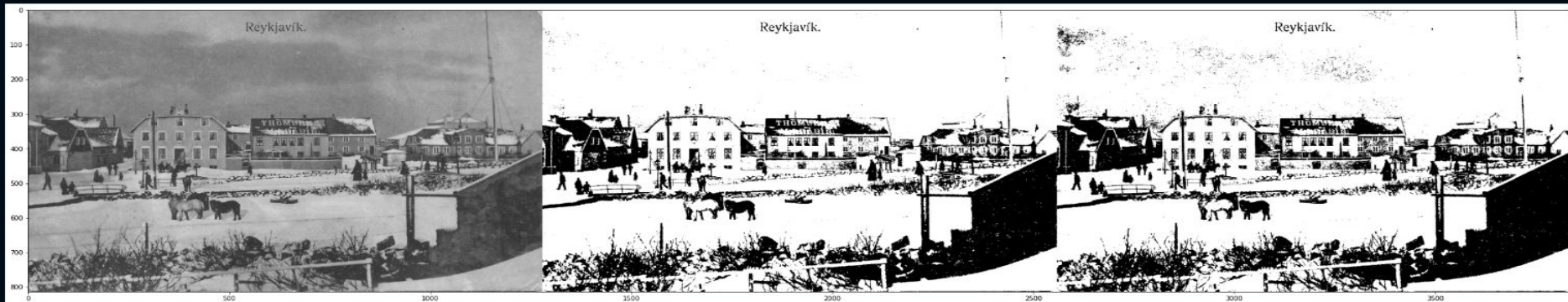
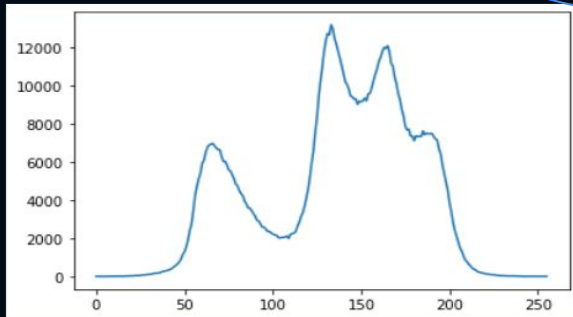
Otsu: 137.0
Fisher's linear discriminant: 131



И върху някои по-трудни изображения

Otsu: 117.0

Fisher's linear discriminant: 109



Сега за цветни изображения?

To be continued...

Идеи:

- Определяне на граница по всеки един от трите канала (red, green, blue) и обединяването им с логическо **&&**. Това не би било ефикасно, защото начинът, по който хората възприемат дадено изображение и как то се представя в RGB пространството се различават. По тази причина HSV и HSL моделите са по-подходящи.