Избранные задачи обработки изображений

Александр Вахитов
av38@yandex.ru

18 октября 2014 г.

1 Abstract

2 Введение

3 Детекция объектов на бинарных изображениях

3.1 Основные определения

Определение. Изображение $I(x, y)$ размера $N \times N$, каждый пиксель которого имеет значение 0 (является черным) либо 1 (белым), называется бинарным.

Определение. Два пикселя бинарного изображения называются соседними, если их координаты отличаются на единицу либо по горизонтали, либо по вертикали.

Определение. Отношение связности - это транзитивное замыкание отношения соседства.

Видно, что связность является отношением эквивалентности на множестве всех пикселей изображения.

Определение. Компонентой связности называется класс эквивалентности по отношению связности.

В дальнейшем мы будем считать, что компоненты связности определены только для белых пикселей изображения.

3.2 Базовая постановка задачи

Детекция объектов по бинарному изображению на основании их размеров, характеристик положения и движения применяется в охранном видеонаблюдении, обработке данных с радаров, поиске текста на отсканированных изображениях или на фотографиях печатных страниц.

Будем говорить, что пиксель бинарного изображения $I(x, y)$ размера $N \times N$ является свободным, если он принимает значение 0, и является занятым, если его значение равно 1. Компоненты связности изображения $I(x, y)$ будем называть объектами.

Пусть объект всегда занимает ровно два соседних друг с другом пикселя. Будем считать, что объект всегда находится на расстоянии 1 или более от края изображения.

Изображение $I(x, y)$ подвергается зашумлению и результат записывается в $J(x, y)$. Зашумление производится путем замены значений пикселя на противоположное с некоторой вероятностью, что можно выразить прибавлением по модулю 2 бинарного изображения с шумом $V(x, y)$ к исходному изображению $J(x, y)$. Изображение $V(x, y)$ является бинарным и все его пиксели являются независимыми одинаково распределенными случайными величинами. Оно формируется по правилу:

$$V(x, y) = \begin{cases} 1, \text{с вероятностью а,} \\ 0, \text{с вероятностью 1-а.} \end{cases}$$

Процесс зашумления можно описать как:

$$J(x, y) = I(x, y) + V(x, y) \pmod 2.$$

Нам необходимо разработать алгоритм, который по изображению $J(x, y)$ построит гипотезу относительно $I(x, y)$. Обозначим как $\hat{I}(x, y)$ результат работы алгоритма. Для отдельного пикселя он может быть одним из четырех перечисленных ниже.

Истинно позитивным (true positive) срабатыванием называется событие

$$TP = \{I(x, y) = 1 \cap \hat{I}(x, y) = 1\}.$$

По аналогии, следующие события называются:

- $TN = \{I(x, y) = 0 \cap \hat{I}(x, y) = 0\}$ - **истинно негативным (true negative),**
- $FN = \{I(x, y) = 1 \cap \hat{I}(x, y) = 0\}$ - **ложно негативным (false negative),**
- $FP = \{I(x, y) = 0 \cap \hat{I}(x, y) = 1\}$ - **ложно позитивным (false positive)**.

Видно, что для пикселя объекта

$$P(TP) + P(FN) = 1$$

Для пикселя фона:

$$P(FP) + P(TN) = 1$$

Можно формулировать различные требования к алгоритму, например:

- максимизация $P(TP)$ при условии $P(FP) < \epsilon$
- алгоритме, максимизирующем $P(TN)$ при условии $P(TP) > \epsilon$.

При этом эквивалентны требования максимизации $P(TP)$ и минимизации $P(FN)$.

Исследуем алгоритм для решения поставленной задачи, основанный на фильтрации компонент связности бинарного изображения по размеру.

- выделить все компоненты связности на $J(x, y)$
- присвоить $\hat{I}(x, y) = J(x, y)$
- для всех пикселей (x, y), принадлежащих компонентам связности размером меньше 2, присвоить $\hat{I}(x, y) = 0$
- вернуть $\hat{I}(x, y)$

Для определения компонент связности можно использовать операцию $floodFill$, доступную в библиотеке OpenCV.

В рамках исследования алгоритма вычислим вероятности истинно и ложно позитивного, а также - истинно и ложно негативного срабатываний. При этом $I(x, y)$ мы случайным не считаем.

1. Истинно позитивное

   Осуществляется событие $A_1 = \{\hat{I}(x, y) = 1 | I(x, y) = 1\}$. Обозначим как $N(x, y)$ множество пикселей, соседних с (x, y). Из свойств алгоритма следует, что событие A_1 это

   $$A_1 = \{J(x, y) = 1 | I(x, y) = 1\} \cap$$
   $$\{\exists (x', y') \in N(x, y) : \ J(x', y') = 1 | I(x, y) = 1\}.$$

   По определению,

   $$P\{J(x, y) = 1 | I(x, y) = 1\} = 1 - a;$$

   $$P\{\exists (x', y') \in N(x, y) : \ J(x', y') = 1 | I(x, y) = 1\} = 1 - a + a(1 - (1 - a)^3).$$

   Таким образом,

   $$P\{A_1\} = (1 - a)(1 - a(1 - (1 - a)^3)).$$

2. Истинно негативное Рассмотрим только случай, когда пиксель на изображении I является черным вместе со всеми своими соседями. Это событие обозначим как $B_2 = \{I(x,y) = 0, \forall (x',y') \in N(x,y) : I(x',y') = 0\}$. Тогда $A_2 = \{\hat{I}(x,y) = 0 | B_2\}$.

$$A_2 = \{J(x,y) = 0 | B_2\} \cup$$
$$(\{J(x,y) = 1 | B_2\} \cap \{\forall (x',y') \in N(x,y)\ J(x',y') = 0 | B_2\}).$$

   Для событий, входящих в выражение, вычислим:

$$P\{J(x,y) = 0 | B_2\} = 1 - a,$$

$$P\{J(x,y) = 1 | B_2\} = a,\ P\{\forall (x',y') \in N(x,y)\ J(x',y') = 0 | B_2\} = (1-a)^4,$$

   и в итоге получим:

$$P\{A_2\} = (1-a)(a(1-a)^3 + 1).$$

3. Ложно позитивное

   Будем рассматривать только случай, когда все соседи черного пикселя - черные.

$$A_3 = \{\hat{I}(x,y) = 1 | I(x,y) = 0\}.$$
$$P\{A_3\} = a(1 - (1-a)^4)$$

4. Ложно негативное

$$A_4 = \{\hat{I}(x,y) = 0 | I(x,y) = 1\}.$$
$$P\{A_4\} = a + (1-a)a(1-a)^3 = a(1 + (1-a)^4).$$