

Алгоритм анализа и принятия решения в задаче селекции объектов на изображениях наземных сцен

Ю.Б. Блохинов, В.В. Гнилицкий,
В. В. Инсаров, А.С. Чернявский

ФГУП «ГосНИИАС», Москва

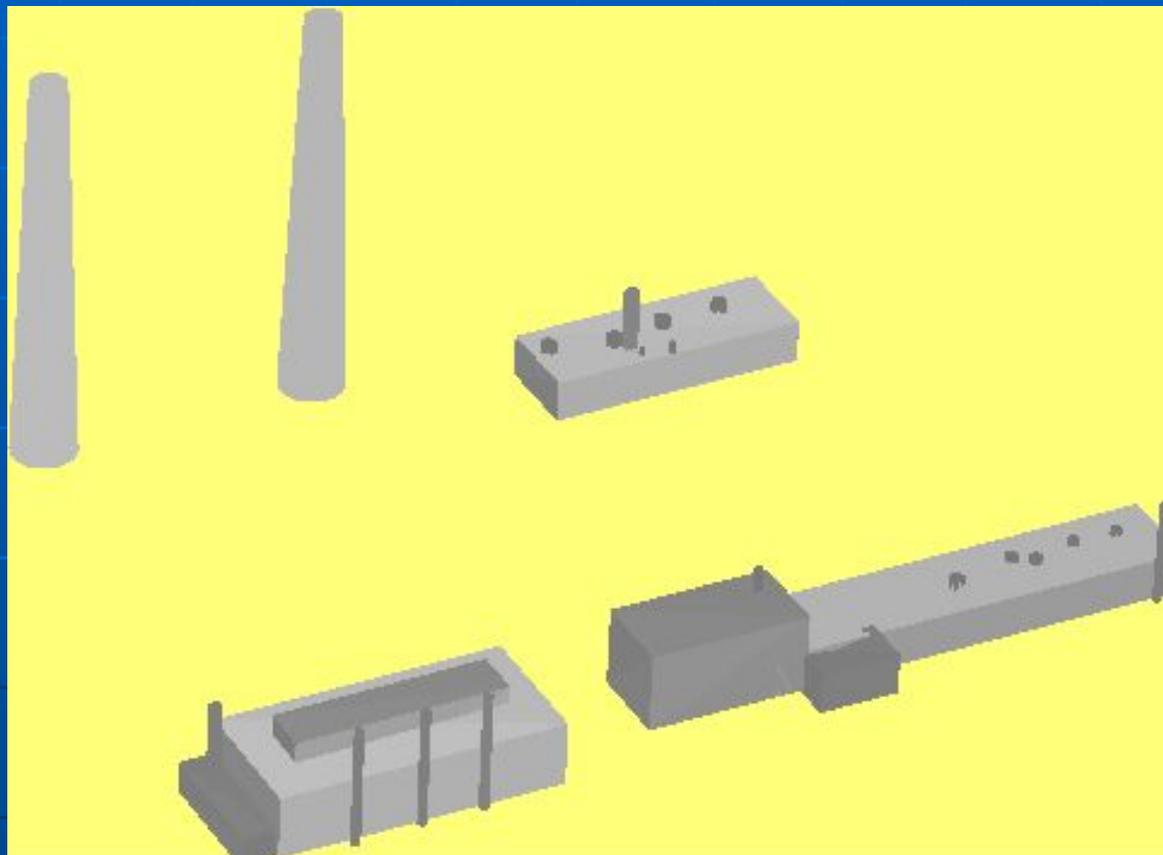


Постановка задачи

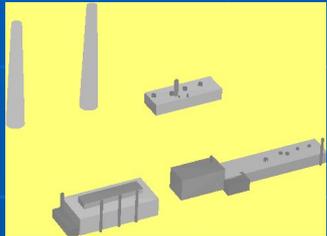
- Существуют 3D модели ряда наземных объектов, формирующих *сцену*
- Имеется цифровое *текущее изображение* (ТИ), полученное с летательного аппарата
- Заданы углы обзора и примерная область поиска объектов на изображении
- Предполагаем, что объект полностью находится в кадре

- Требуется определить местонахождение искомого объекта на ТИ
- При этом не использовать информацию о текстуре

Модель

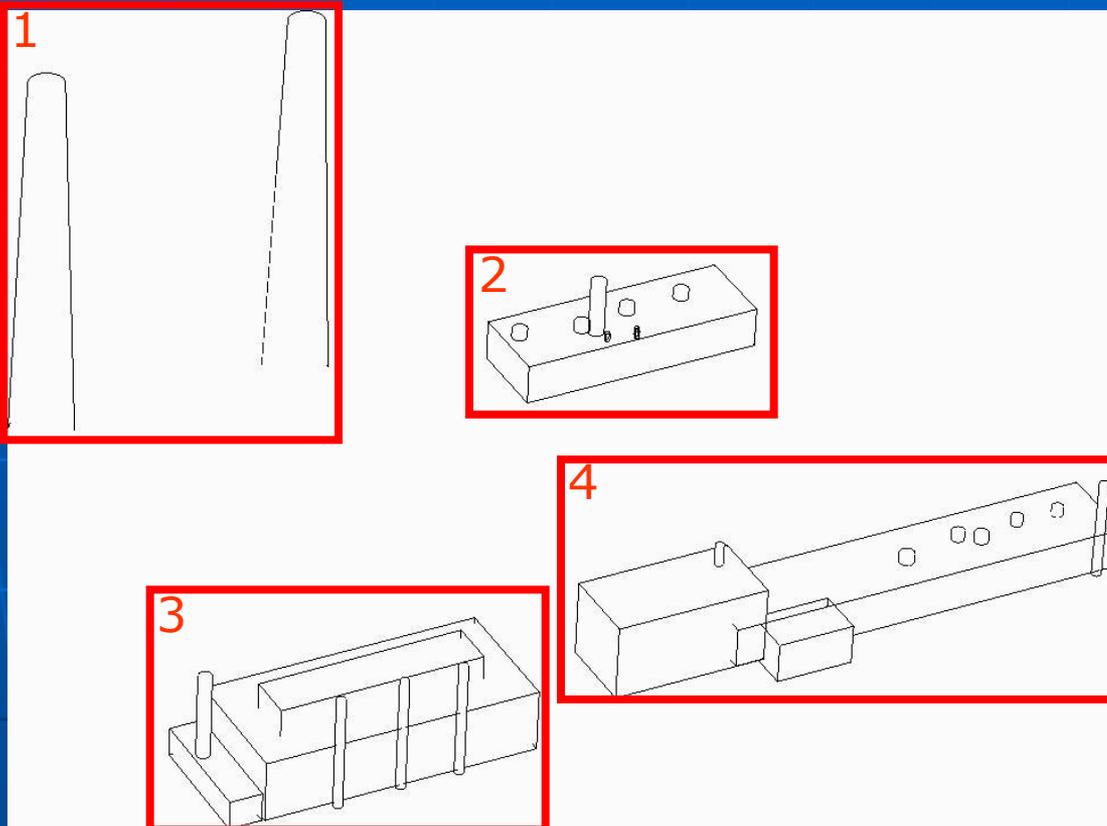
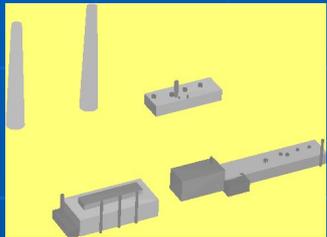


ТИ



По имеющейся информации о точке съемки генерируется конкретный вид эталона объекта

Контурные эталоны объектов



Для каждого контурного эталона (КЭ) строится функция сходства с ТИ

Предобработка ТИ

Предобработка I

- Нормально распределенный шум

$$\tilde{I}(x, y) = \max(0, \min(I(x, y) + N(0, \sigma), 255))$$

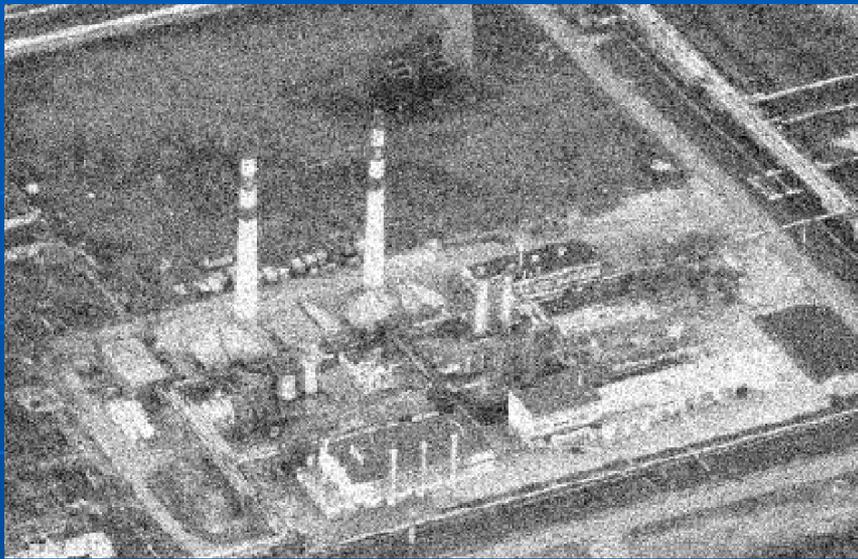
- Устраняется двусторонним фильтром

$$I_f(x_0, y_0) = \frac{1}{W} \sum_{(x, y) \in S} G_{\sigma_S} \left(\sqrt{(x - x_0)^2 + (y - y_0)^2} \right) G_{\sigma_I} (|I(x_0, y_0) - I(x, y)|) I(x, y)$$

$$G_{\sigma_S} = \frac{1}{\sqrt{2\pi\sigma_S}} e^{-\frac{(x-x_0)^2 + (y-y_0)^2}{2\sigma_S^2}} \quad G_{\sigma_I} = \frac{1}{\sqrt{2\pi\sigma_I}} e^{-\frac{|I(x_0, y_0) - I(x, y)|^2}{2\sigma_I^2}}$$

- В отличие от Гауссиана, сохраняет края

Нормально-распределенный шум - двусторонний фильтр -



$\sigma = 25$



$\sigma = 50$

до фильтрации

Нормально-распределенный шум - двусторонний фильтр -



$\sigma = 25$



$\sigma = 50$

после фильтрации

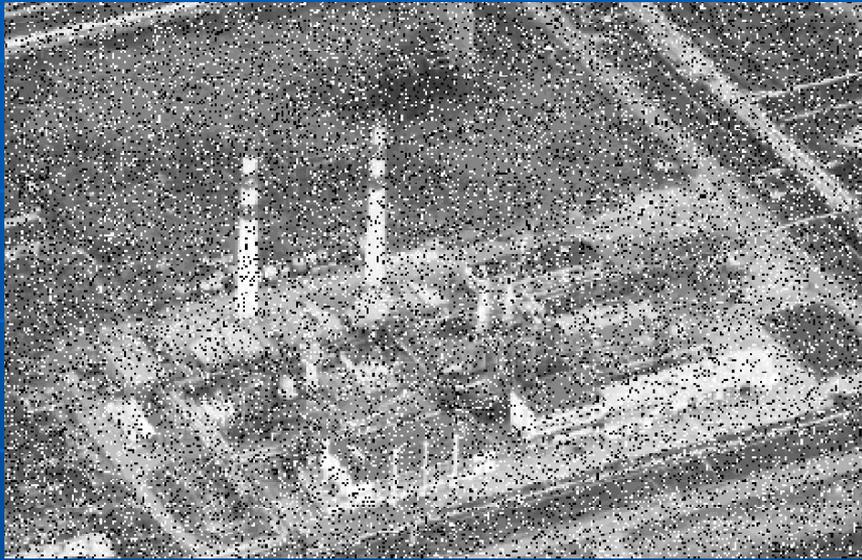
Предобработка II

- Импульсный шум («соль-перец»)

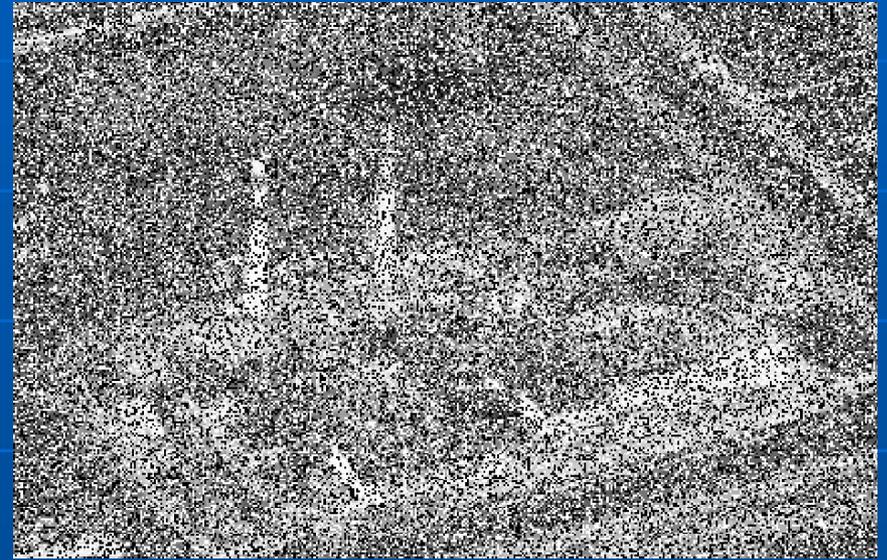
$$\tilde{I}(x, y) = \begin{cases} 0, & p = P_a \\ 255, & p = P_b \\ I(x, y), & p = (1 - P_s) \end{cases} \quad P_s = P_a + P_b \leq 1$$

- Устраняется медианной фильтрацией с адаптивным окном
- Лучше, чем обычная медиана, справляется с большим уровнем шума и меньше искажает форму краев

Импульсный шум - адаптивная медиана -



$P_s = 0.20$



$P_s = 0.50$

до фильтрации

Импульсный шум - адаптивная медиана -



$P_s = 0.20$

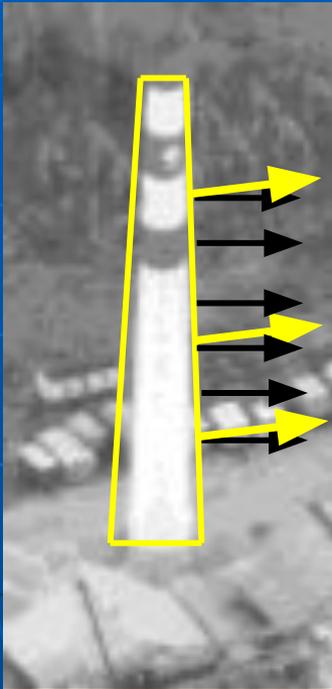


$P_s = 0.50$

после фильтрации

Поиск объекта

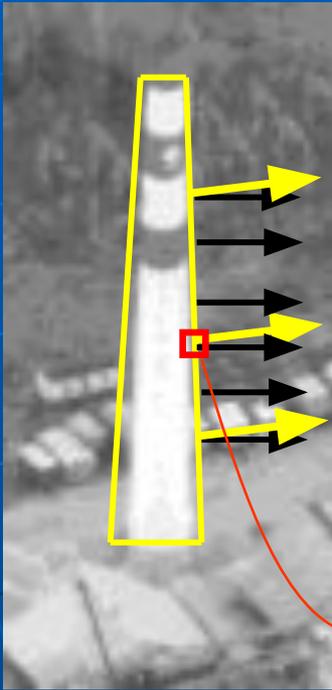
Функция сходства



- Контурный эталон $E(i, j)$ размывается Гауссианом
- Находится поле направлений градиентов (нормалей) эталона $R_I(i, j) = \arctg(I_y / I_x)$
- То же проводится для ТИ:

$$R_E(i, j) = \arctg(E_y / E_x)$$

Функция сходства



- Значения углов квантуются в M ячеек (например, $M=6$)

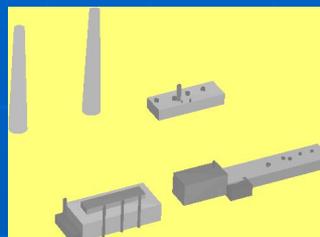
- Сходство между эталоном и ТИ в точке (i, j) :

$$p(i, j) = \begin{cases} 1, & \text{если } |\text{bin}(R_I(i, j)) - \text{bin}(R_E(i, j))| < T \\ 0, & \text{иначе} \end{cases}$$

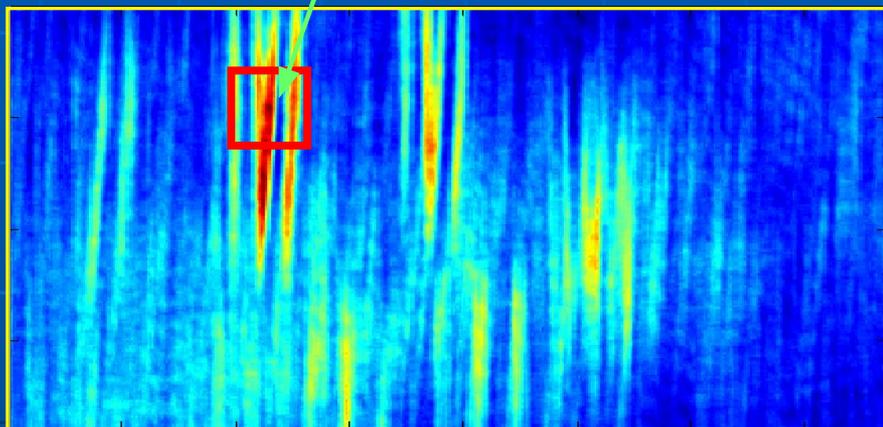
- Итоговая функция сходства

$$D(m, n) = \sum_{\substack{i \in [m, m+H_E] \\ j \in [n, n+W_E]}} p(i, j)$$

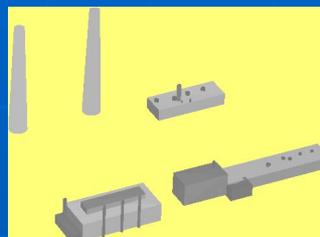
Пик функции
сходства
объекта №1



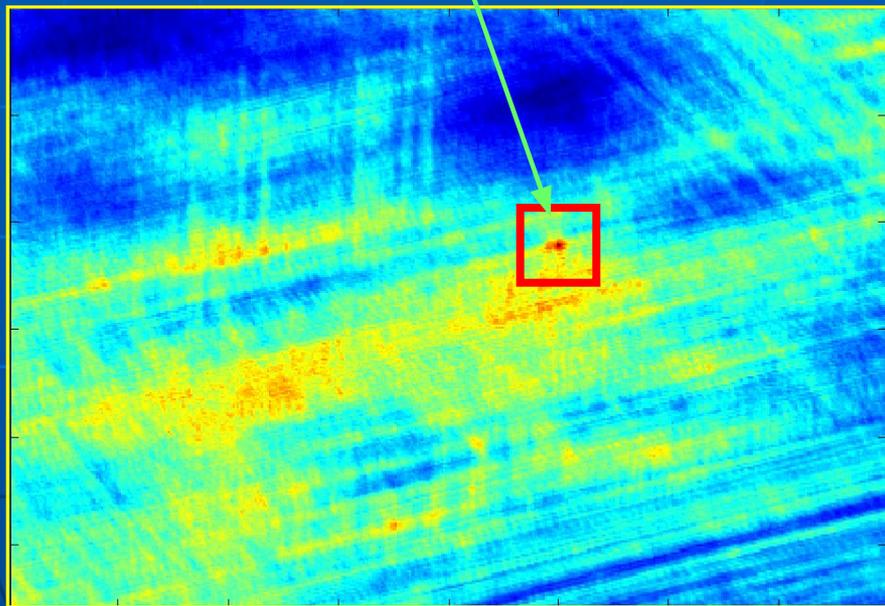
Наиболее
вероятное
расположение



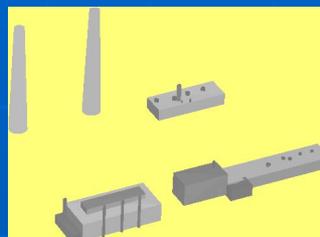
Пик функции
сходства
объекта №2



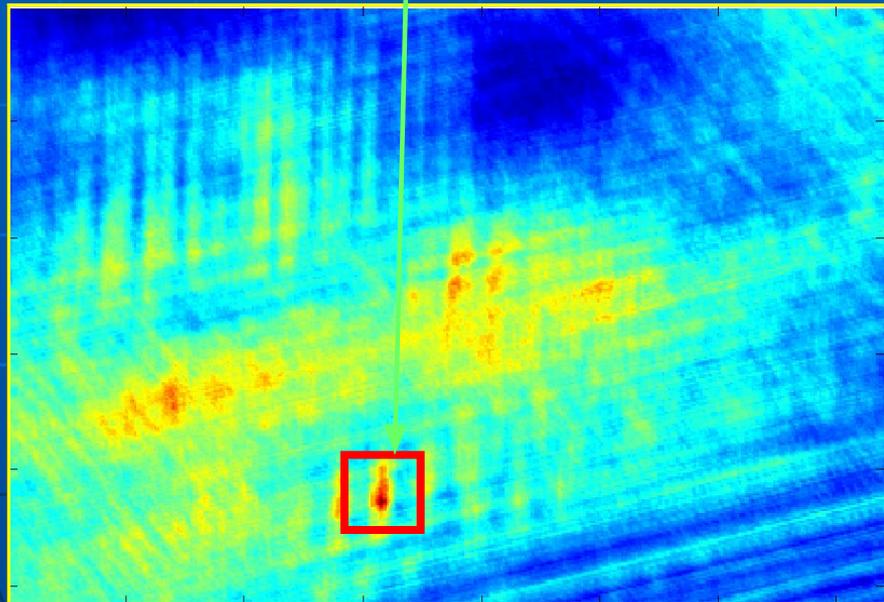
Наиболее
вероятное
расположение



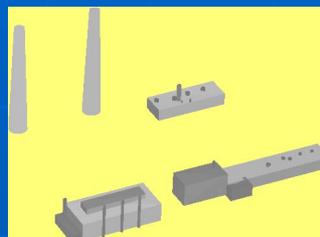
Пик функции
сходства
объекта №3



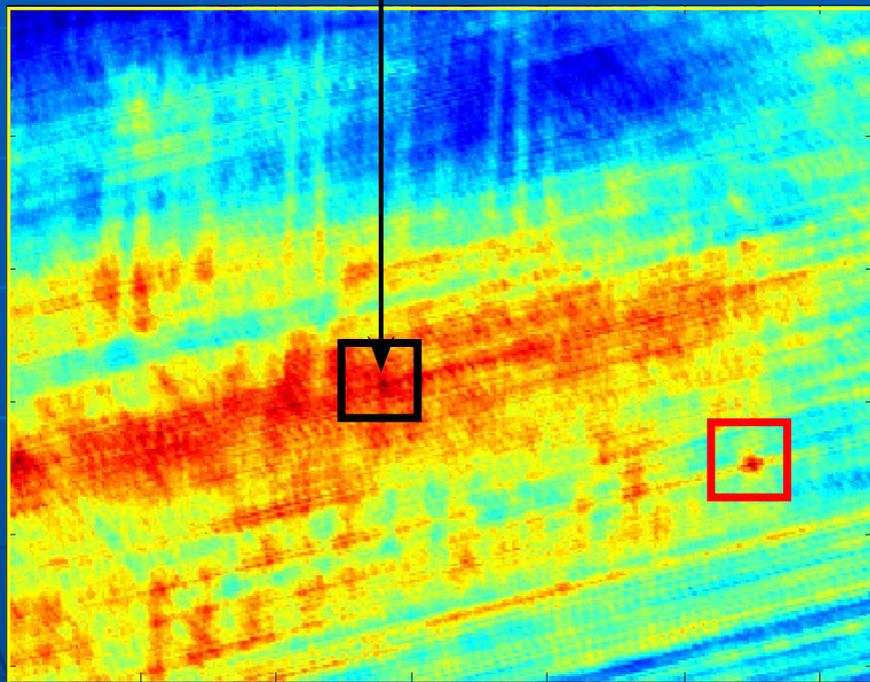
Наиболее
вероятное
расположение



Пик функции
сходства
объекта №4



Наиболее
вероятное
расположение

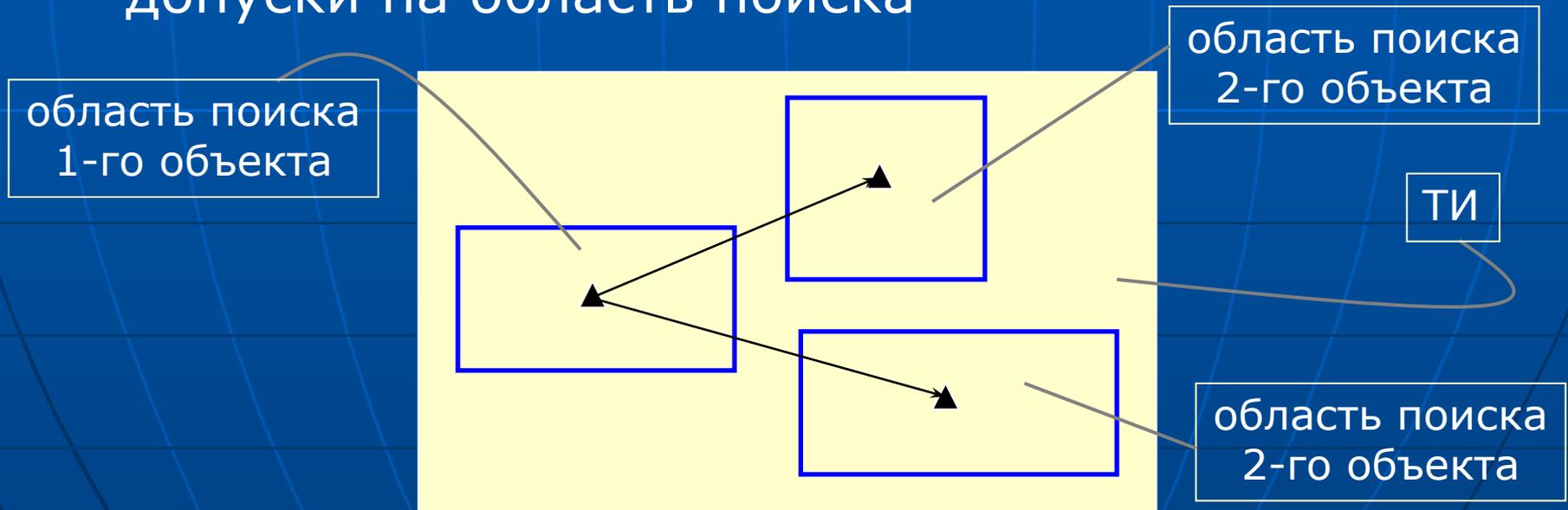


Ложное
обнаружение

Сборка сцены

Сборка сцены

- По ряду причин (шум, загораживание) функция сходства может иметь **ложные максимумы**
- Необходимо учитывать информацию о взаимном расположении объектов, формирующих сцену – это расстояния и допуски на область поиска



Сборка сцены

- Пусть контурный эталон перенесен в положение k в системе координат ТИ
- Пусть эталон состоит из M ненулевых пикселей
- Пусть в данном положении в N^+ случаях из M направление градиента ТИ и нормали к КЭ совпадает; назовем множество таких точек

$$\Omega = \{(i, j) \mid КЭ_k(i, j) \uparrow\uparrow ТИ(i, j)\}, \quad |\Omega| = N^+, \quad N^+ \in [0, M]$$

- Пусть $P_{об} = P\{(i, j) \in (\Omega \boxtimes \text{Объект})\}$ N^+ - обычная функция сходимости
- $P_{фон} = P\{(i, j) \in (\Omega \boxtimes \neg\text{Объект})\}$

Сборка сцены

- Процесс вычисления функции сходства между КЭ и ТИ можно рассматривать проведение N независимых испытаний по схеме Бернулли
- Основная гипотеза - в k -м положении действительно находится искомый объект
- Альтернативная гипотеза – совпадение направлений вызвано шумом (фоном)
- Отношение правдоподобия двух гипотез:

$$\log l_k = N \left[\frac{N_k^+}{N} \log \frac{P_{об}}{P_{фон}} + \left(1 - \frac{N_k^+}{N} \right) \log \left(\frac{1 - P_{об}}{1 - P_{фон}} \right) \right]$$

Сборка сцены

- На практике используем оценки $P_{об} = \frac{N_k^+}{N}$ и $P_{фон} = \frac{N_k^+}{N}$

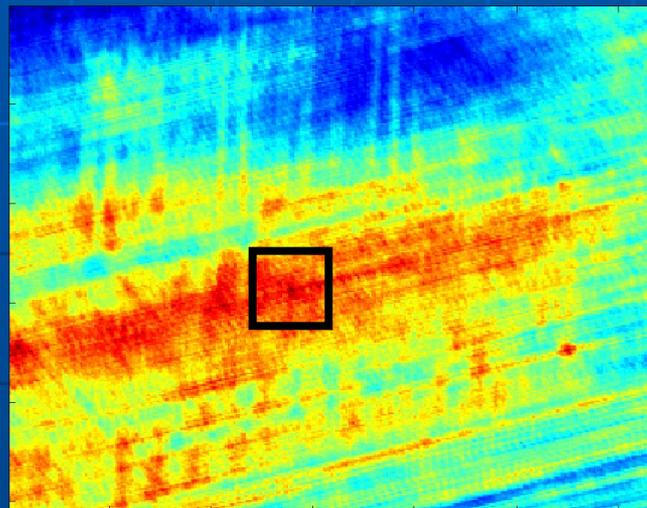
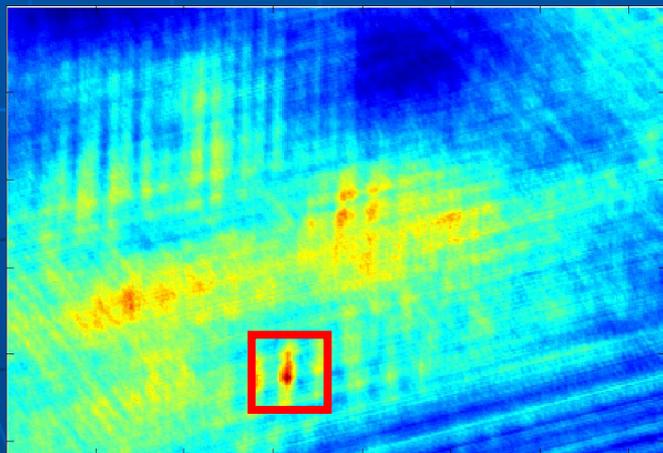
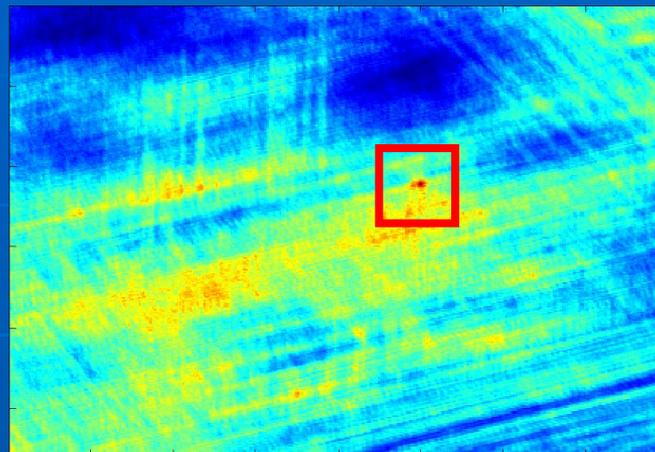
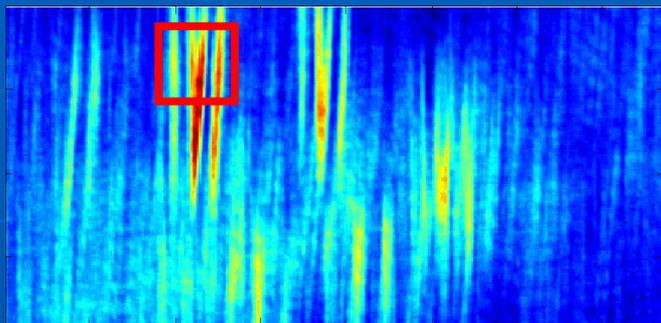
- Если объектов несколько, то для каждого объекта в каждом пикселе его зоны поиска определяется N_k^+ и N_k^+

- Формируются оценки $P_{фон} = \frac{\sum_{i=1}^m N_i^+}{\sum_{i=1}^m N_i}$ и $P_{об} = \left(N_1^+ + \sum_{i=2}^m \max N_i^+ \right) / \sum_{i=1}^m N_i$

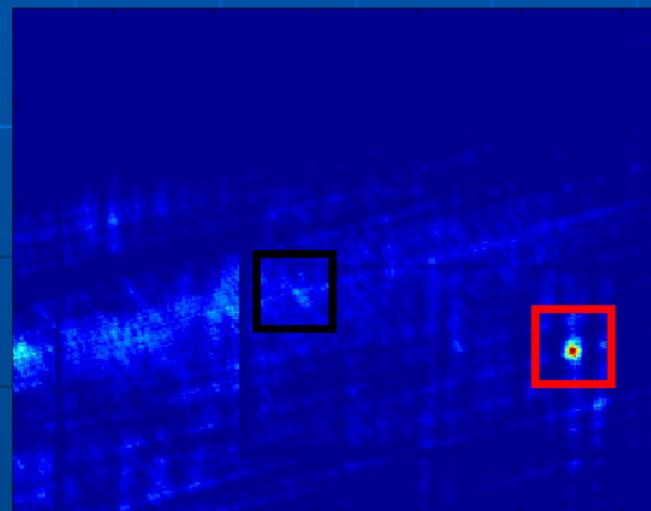
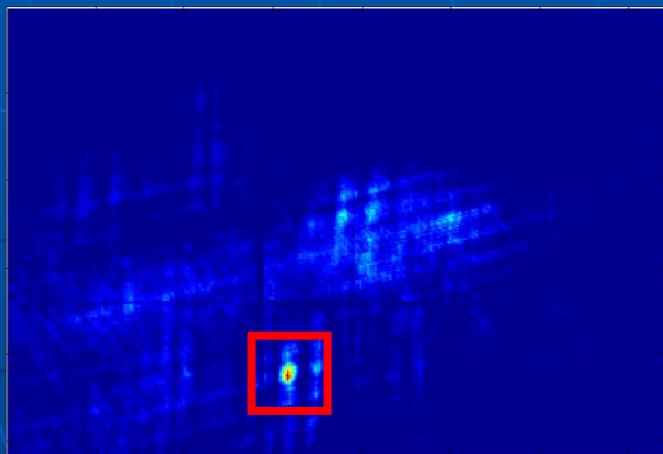
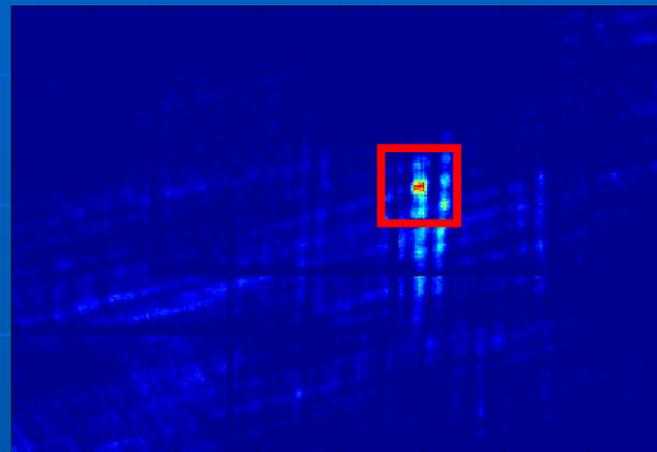
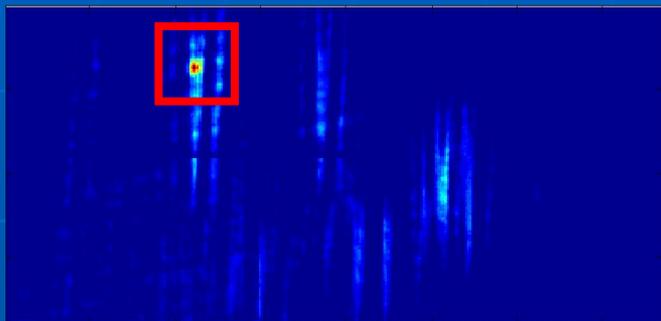
- Итоговая агрегированная функция сходства:

$$\log l_k = \sum_{i=1}^m N_i \cdot \left[P_{об} \log \frac{P_{об}}{P_{фон}} + (1 - P_{об}) \log \left(\frac{1 - P_{об}}{1 - P_{фон}} \right) \right]$$

Функция сходства



Функция сходства после сборки, «голосуют» 4 объекта



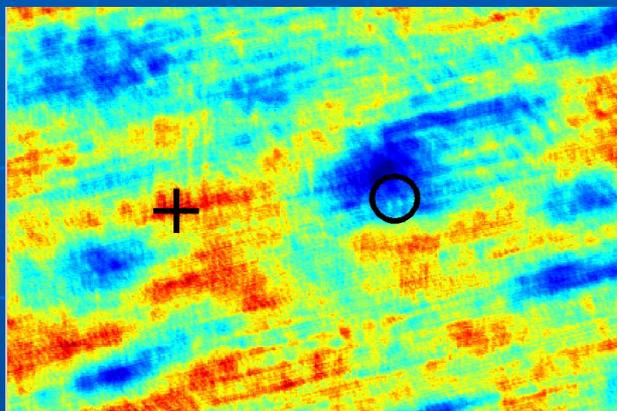
Ложный пик «погашен»



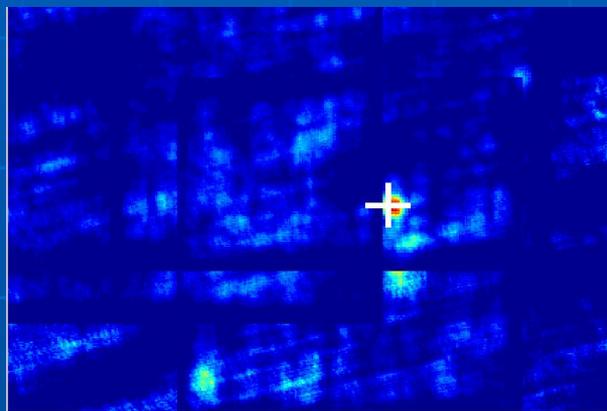
Учет информации о взаимном расстоянии между объектами
позволил точно локализовать искомые объекты

Два примера с загоразиванием искомого объекта

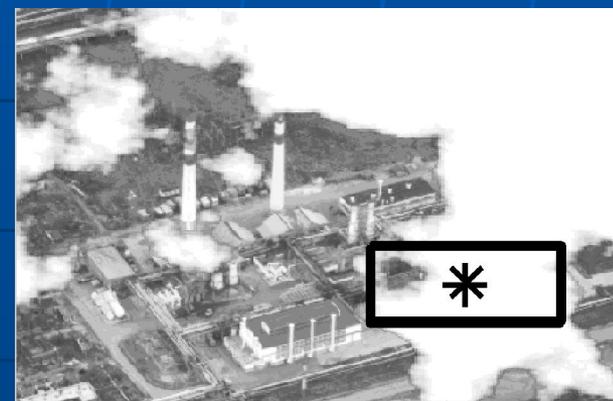
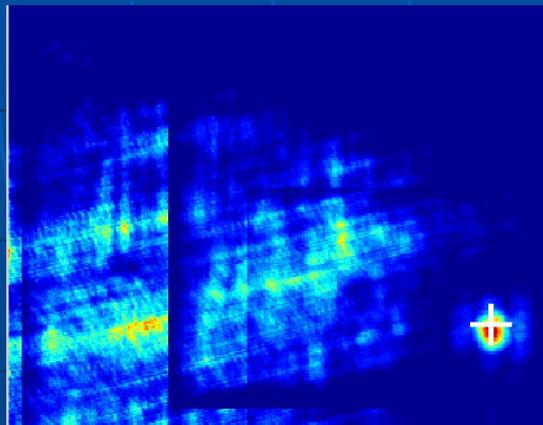
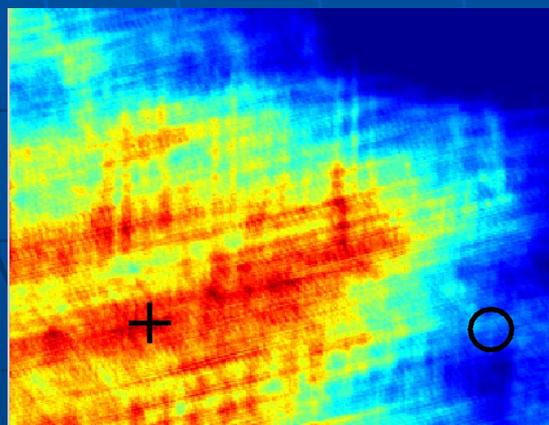
Функция сходства



Сборка



Локализация объекта



Выводы

- Предложенная функция сходства позволяет обнаружить объект на изображении по его контурному эталону
- Использование информации о взаимном расположении объекта повышает устойчивость алгоритма селекции объектов
- Методы предобработки позволяют существенно снизить влияние шума