

Многоклассовое распознавание в пространстве представлений с многоуровневым разрешением

Ганебных С.Н., Ланге М.М.

**Федеральный исследовательский центр "Информатика и управление" РАН
Вычислительный центр им. А.А. Дородницына РАН**

**17 - я Всероссийская конференция "Математические методы распознавания образов"
(ММРО - 17, Россия, г. Светлогорск, 19 - 25 сентября 2015 г.)**

2. Содержание доклада

- Многоуровневое представление и мера различия образов
- Многоклассовая модель распознавания и задача исследования
- Решающий алгоритм на основе стратегии иерархического поиска
- Оценки вероятности ошибок и вычислительной сложности алгоритма
- Экспериментальные результаты
(ROC-кривые, характеристики AUC и зависимости вероятности ошибок от вычислительной сложности решающего алгоритма)
- Результаты и выводы

3. Представление образов с многоуровневым разрешением

Множество объектов источника (образов): X_S , где $x \in X_S$ – набор пикселей, образующих двумерное тело (плотность – функция яркости) с однозначно идентифицируемой системой собственных координат U, V .

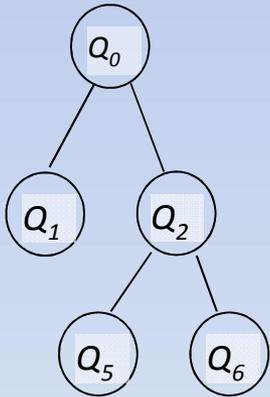
Представление образа $x \in X_S$ деревом эллиптических примитивов глубины L :

$x^L = (x_0, \dots, x_l, \dots, x_L)$, где $x_l = \{Q_n\}$ – набор концевых вершин (эллиптических примитивов) в поддереве $x^l = (x_0, \dots, x_l)$ глубины l .

Примеры:

дерево примитивов $x^2 = (x_0, x_1, x_2)$

эллиптический примитив $Q_n = (r_n, u_n, v_n, z_n)$



$$l=0: x_0 = Q_0$$

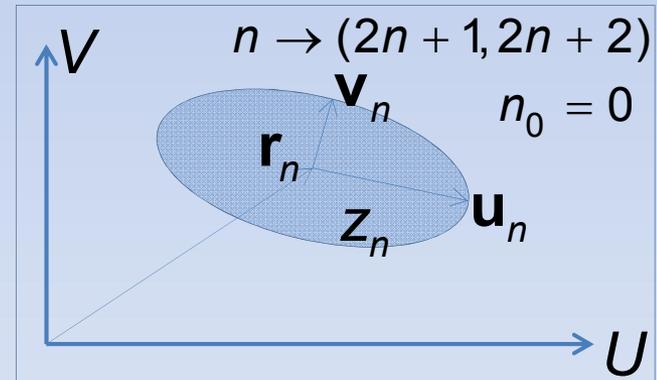
$$x^0 = x_0$$

$$l=1: x_1 = (Q_1, Q_2)$$

$$x^1 = (x_0, x_1)$$

$$l=2: x_2 = (Q_1, Q_5, Q_6)$$

$$x^2 = (x_0, x_1, x_2)$$

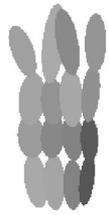


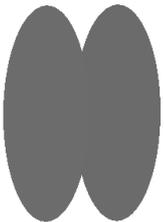
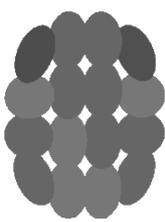
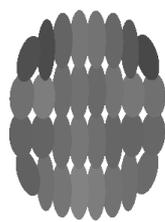
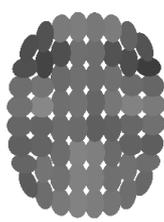
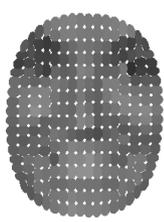
Инвариантность:

При достаточно малом размере пикселей и большом числе уровней квантования яркостей бинарное дерево эллиптических примитивов с точностью до размера пикселя и уровня квантования инвариантно к преобразованиям поворота, смещения, изменения масштаба и уровня яркости представляемого объекта.

4. Примеры древовидных представлений

									
signature	$l=0$	$l=1$	$l=2$	$l=3$	$l=4$	$l=5$	$l=6$	$l=7$	$l=8$

									
gesture	$l=0$	$l=1$	$l=2$	$l=3$	$l=4$	$l=5$	$l=6$	$l=7$	$l=8$

									
face	$l=0$	$l=1$	$l=2$	$l=3$	$l=4$	$l=5$	$l=6$	$l=7$	$l=8$

5. Многоуровневая мера различия образов

Компоненты меры различия образов $\mathbf{x} \in \mathbf{X}_s$ и $\hat{\mathbf{x}} \in \mathbf{X}_s$ l -го уровня:

$$\left\{ \begin{aligned} \rho_r(\mathbf{x}, \hat{\mathbf{x}}) &= \rho_r(\mathbf{x}', \hat{\mathbf{x}}') = \sum_{n=1}^{2^{l+1}-2} w_n \|\mathbf{r}_n - \hat{\mathbf{r}}_n\| \\ \rho_{uv}(\mathbf{x}', \hat{\mathbf{x}}') &= \sum_{n=1}^{2^{l+1}-2} w_n (\|\mathbf{u}_n - \hat{\mathbf{u}}_n\| + \|\mathbf{v}_n - \hat{\mathbf{v}}_n\|) / 2 \\ \rho_z(\mathbf{x}', \hat{\mathbf{x}}') &= \sum_{n=1}^{2^{l+1}-2} w_n \|\mathbf{z}_n - \hat{\mathbf{z}}_n\| \end{aligned} \right.$$

где $w_n \geq 0$: $\sum w_n = 1$ – нормированные веса примитивов; $\|\mathbf{y} - \hat{\mathbf{y}}\|$ – метрика L1

Обучающее множество:

$$\mathbf{X} = \{\mathbf{X}_i = \{\mathbf{x}_{ij}\}_{j=1}^{m_i}\}_{i=1}^c \subset \mathbf{X}_s$$

Среднее различие l -го уровня образов из \mathbf{X} относительно $\hat{\mathbf{x}} \in \mathbf{X}$:

$$\left\{ \begin{aligned} \sigma_r(\hat{\mathbf{x}}') &= \frac{1}{\|\mathbf{X}\| - 1} \sum_{\mathbf{x}': \mathbf{x} \in \mathbf{X}} \rho_r(\mathbf{x}', \hat{\mathbf{x}}'), & \sigma_{uv}(\hat{\mathbf{x}}') &= \frac{1}{\|\mathbf{X}\| - 1} \sum_{\mathbf{x}': \mathbf{x} \in \mathbf{X}} \rho_{uv}(\mathbf{x}', \hat{\mathbf{x}}') \\ \sigma_z(\hat{\mathbf{x}}') &= \frac{1}{\|\mathbf{X}\| - 1} \sum_{\mathbf{x}': \mathbf{x} \in \mathbf{X}} \rho_z(\mathbf{x}', \hat{\mathbf{x}}') \end{aligned} \right.$$

Мера различия образов \mathbf{x} и $\hat{\mathbf{x}}$ уровня l :

$$d^l(\mathbf{x}, \hat{\mathbf{x}}) = \frac{\rho_r(\mathbf{x}', \hat{\mathbf{x}}')}{\sigma_r(\hat{\mathbf{x}}')} + \frac{\rho_{uv}(\mathbf{x}', \hat{\mathbf{x}}')}{\sigma_{uv}(\hat{\mathbf{x}}')} + \frac{\rho_z(\mathbf{x}', \hat{\mathbf{x}}')}{\sigma_z(\hat{\mathbf{x}}')}$$

6. ТЕСТОВЫЙ ИСТОЧНИК

Источники изображений:

число изображений: 2800

число классов: 90

подписи: 40 объектов по 20 реализаций

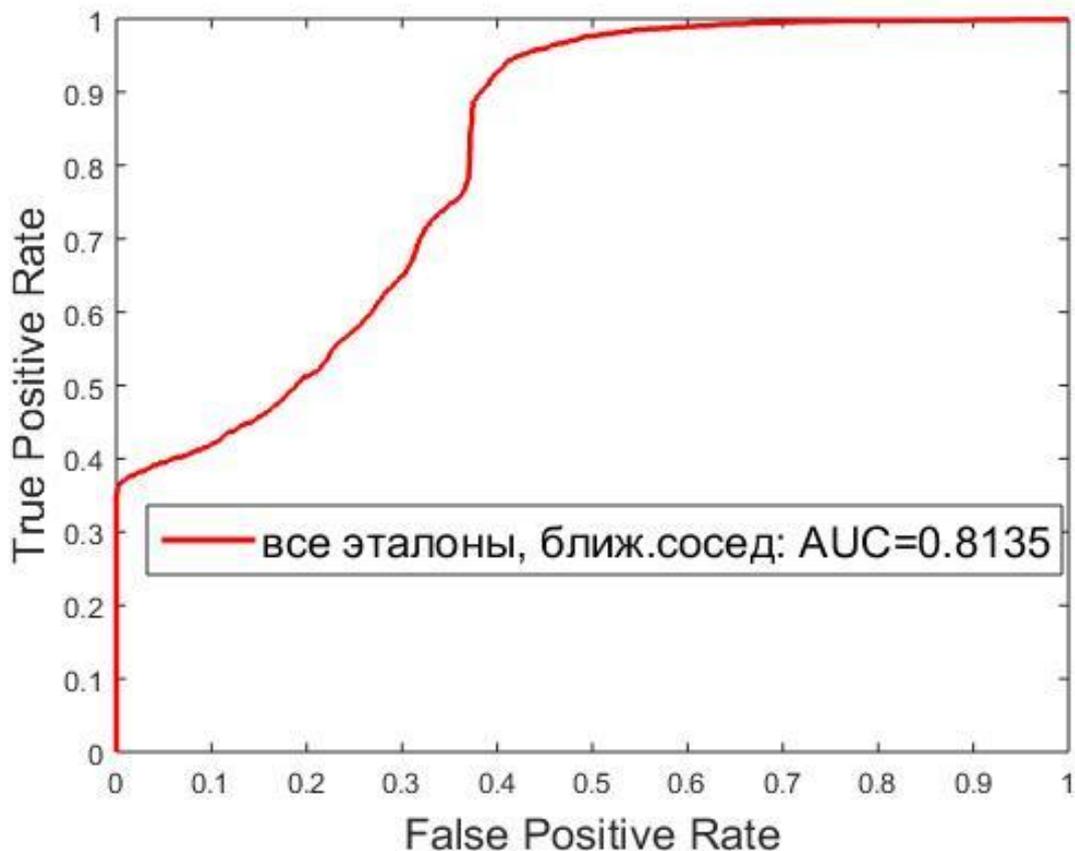
жесты руки: 25 объектов по 40 реализаций

лица: 25 объектов по 40 реализаций

TPR (True Positive Rate) – доля правильно распознанных объектов *positive* классов, LOO

FPR (False Positive Rate) – доля ошибок для объектов *zero* класса, Leave-One Class-Out

По $L=12$ -му уровню меры



7. Модель классификатора

Множество объектов источника: \mathbf{X}_s , $\mathbf{x} \in \mathbf{X}_s$ – образ на изображении

Множество классов: $\Omega = \Omega_c \cup \omega_0$, $\Omega_c = \{\omega_1, \dots, \omega_c\}$ – *positive*, ω_0 – *zero*, классы

Обучающее множество: $\mathbf{X} = \{\mathbf{X}_i = \{\mathbf{x}_{ij}\}_{j=1}^{m_i}\}_{i=1}^c \subset \mathbf{X}_s$

Множество эталонных объектов: $\hat{\mathbf{X}} = \{\hat{\mathbf{X}}_i = \{\hat{\mathbf{x}}_{ik}\}_{k=1}^{\hat{m}_i} \subseteq \mathbf{X}_i\}_{i=1}^c$, $1 \leq \hat{m}_i \leq m_i$, $i = 1, \dots, c$

Разделяющие функции по мере $d(\mathbf{x}, \hat{\mathbf{x}}) \geq 0$, $\mathbf{x} \in \mathbf{X}_s$, $\hat{\mathbf{x}} \in \hat{\mathbf{X}}$ для множества классов Ω_c :

$$g_i(\mathbf{x}) = e^{-\min_{j=1}^{\hat{m}_i} d(\mathbf{x}, \hat{\mathbf{x}}_{ij})}, \quad i = 1, \dots, c, \quad \hat{g}_i(\mathbf{x}) = \frac{g_i(\mathbf{x})}{g_i^*},$$

при $g_i^* \equiv 1$ – алгоритм ближайшего соседа

Правило классификации объекта $\mathbf{x} \in \mathbf{X}_s$:

$$i^*(\mathbf{x}) = \max_{i=1}^c [\hat{g}_i(\mathbf{x}) > \delta] \arg \max_{i=1}^c \hat{g}_i(\mathbf{x}), \quad \delta - \text{порог}$$

Предварительное обучение:

оптимизировать g_i^* , $i = 1, \dots, c$, по критерию максимизации AUC

$$g_i^* = \arg \max_{\mathbf{x} \in \mathbf{X} \setminus \mathbf{X}_i} \text{AUC}(g_i^k = \text{mean}(\text{sort } g_i(\mathbf{x}))(1:k))$$

8. Тестовый источник

Источники изображений:

число изображений: 2800

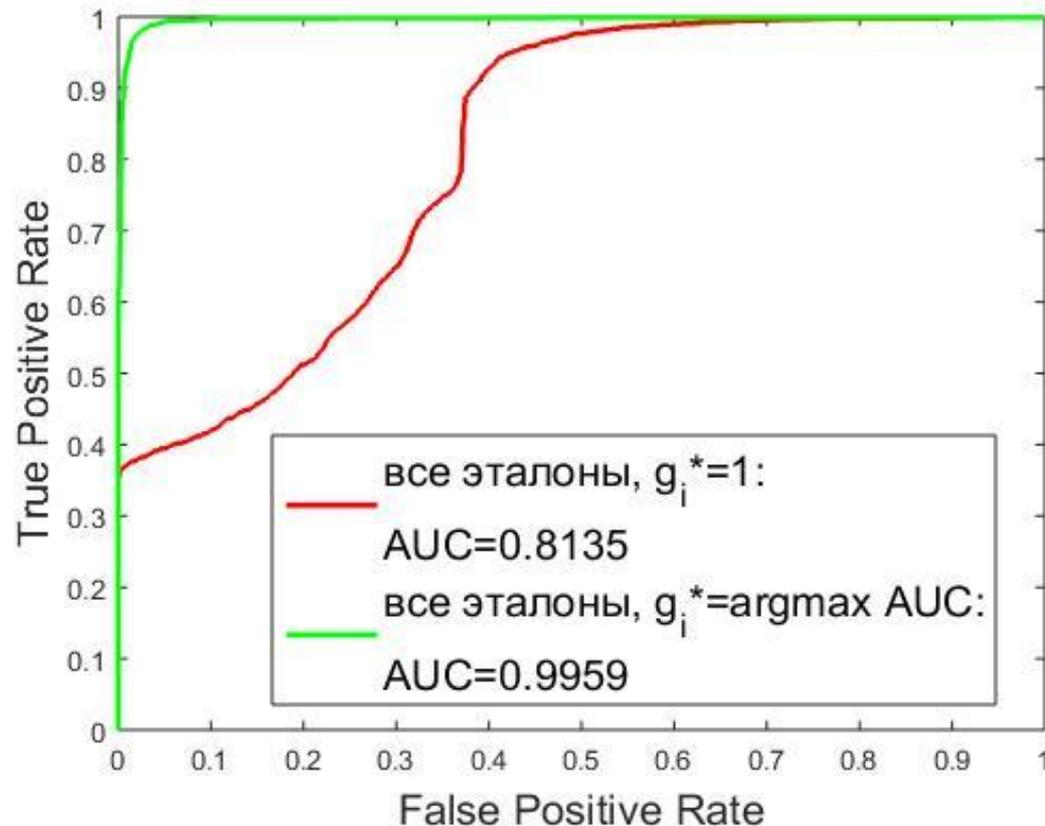
число классов: 90

подписи: 40 объектов по 20 реализаций

жесты руки: 25 объектов по 40 реализаций

лица: 25 объектов по 40 реализаций

TPR (True Positive Rate) — доля правильно распознанных объектов *positive* классов, LOO
FPR (False Positive Rate) — доля ошибок для объектов *zero* класса, leave-one Class-out



9. Оценки вероятности ошибок

Априорные вероятности: $\theta = \Pr(\Omega_c)$, $1 - \theta = \Pr(\omega_0)$, $0 \leq \theta \leq 1$,

Вероятность ошибок при условии, что

θ – неизвестный параметр источника, $p(\theta) = 1$ – равномерная плотность :

$$\varepsilon_{\text{err}}(\hat{\mathbf{X}}, \delta | \theta) = \theta \varepsilon_{\text{err}}(\hat{\mathbf{X}}, \delta | \Omega_c) + (1 - \theta) \varepsilon_{\text{err}}(\hat{\mathbf{X}}, \delta | \omega_0)$$

Функционал качества классификации:

$$\begin{aligned} \varepsilon_{\text{err}}(\hat{\mathbf{X}}, \delta) &= \int_0^1 \varepsilon_{\text{err}}(\hat{\mathbf{X}}, \delta | \theta) p(\theta) d\theta = \frac{1}{2} \left(\varepsilon_{\text{err}}(\hat{\mathbf{X}}, \delta | \Omega_c) + \varepsilon_{\text{err}}(\hat{\mathbf{X}}, \delta | \omega_0) \right) = \\ &= \frac{1}{2} \left((1 - \text{TPR}(\hat{\mathbf{X}}, \delta)) + \text{FPR}(\hat{\mathbf{X}}, \delta) \right) \end{aligned}$$

Обучение:

выбор эталонов $\hat{\mathbf{X}}$ и порога δ , минимизирующих оценку для $\varepsilon_{\text{err}}(\hat{\mathbf{X}}, \delta)$

10. Обучение классификатора

Оптимизация δ , минимизацией суммы ошибок для *positive* и *zero* объектов

$$\delta_1^* = \arg \min_{\delta} \frac{1}{2} ((1 - \text{TPR}(\mathbf{X}, \delta)) + \text{FPR}(\mathbf{X}, \delta))$$

Оптимизация числа эталонов $\{\hat{m}_i\}_{i=1}^c$ (алгоритм отбора $\hat{\mathbf{X}}_i(\tilde{m})$ задан), в каждом классе, путем минимизации взвешенной суммы ошибок для элементарного классификатора "класс против всех" с оптимальным параметром δ_1^* :

$$\hat{m}_i^1 = \arg \min_{1 \leq \tilde{m} \leq m_i} \left(\frac{1}{c} (1 - \text{TPR}_i(\hat{\mathbf{X}}_i(\tilde{m}), \delta_1^*)) + \text{FPR}_i(\hat{\mathbf{X}}_i(\tilde{m}), \delta_1^*) \right)$$

Оптимизация δ , минимизацией суммы ошибок для *positive* и *zero* объектов

$$\delta_2^* = \arg \min_{\delta} \frac{1}{2} ((1 - \text{TPR}(\hat{\mathbf{X}}(\{\hat{m}_i^1\}_{i=1}^c), \delta)) + \text{FPR}(\hat{\mathbf{X}}(\{\hat{m}_i^1\}_{i=1}^c), \delta))$$

и т.д. до сходимости или заданного числа итераций.

11. Решающий алгоритм

Многоуровневая сеть представлений множества эталонов $\hat{\mathbf{X}} = \{\hat{\mathbf{X}}_i\}_{i=1}^c$:

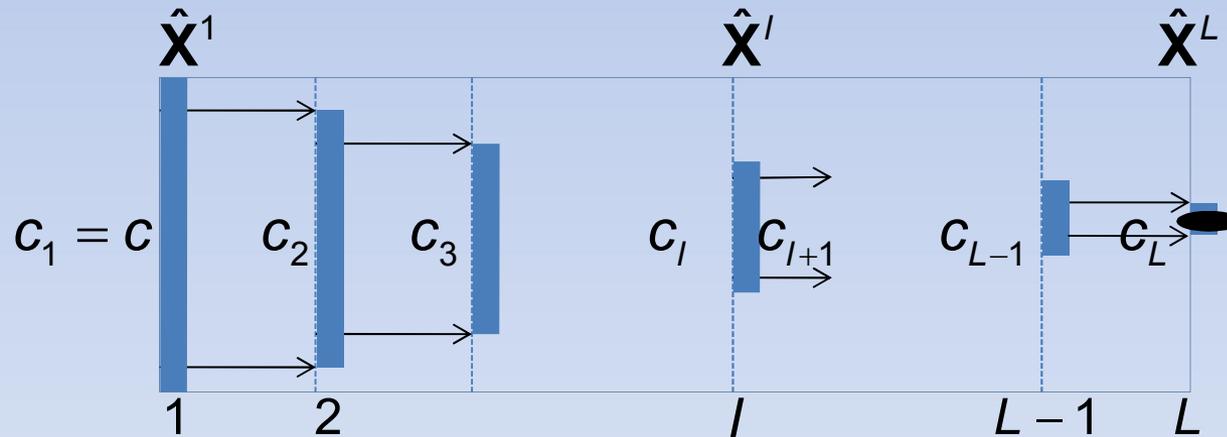
$\hat{\mathbf{X}}^1, \dots, \hat{\mathbf{X}}^l, \dots, \hat{\mathbf{X}}^L$, где $\hat{\mathbf{X}}^l = \{\hat{\mathbf{X}}_i^l\}_{i=1}^c$, $l = 1, \dots, L$

Стратегия поиска решения в сети эталонов :

$c_l = \lfloor c 2^{-\alpha(l-1)} \rfloor$, $l = 1, \dots, L$, - число классов, из которых отбираются c_{l+1} классов с наибольшими значениями разделяющих функций $\{g_i^l(x)\}$ по мере порядка l ;

$\alpha \geq 0$ – свободный параметр; принятие решения на уровне l $\begin{cases} < L, & \text{если } c_{l+1} < 2 \\ = L, & \text{если } c_L \geq 2 \end{cases}$

Схема алгоритма поиска :



Вычислительная сложность алгоритма

$$C_\alpha(\hat{\mathbf{X}}) \begin{cases} = 2(2^L - 1) c \operatorname{mean}_{i=1}^c \|\hat{\mathbf{X}}_i\|, & \alpha = 0 \text{ (переборный поиск)} \\ \leq 2 \frac{2^{(1-\alpha)L} - 1}{2^{(1-\alpha)} - 1} c \max_{i=1}^c \|\hat{\mathbf{X}}_i\|, & \alpha > 0 \text{ (иерархический поиск)} \end{cases} \Rightarrow \frac{C_{\alpha \geq 1}}{C_{\alpha = 0}} = O\left(\frac{L}{2^L - 1}\right)$$

12. Тестовый источник

Источники изображений:

число изображений: 2800

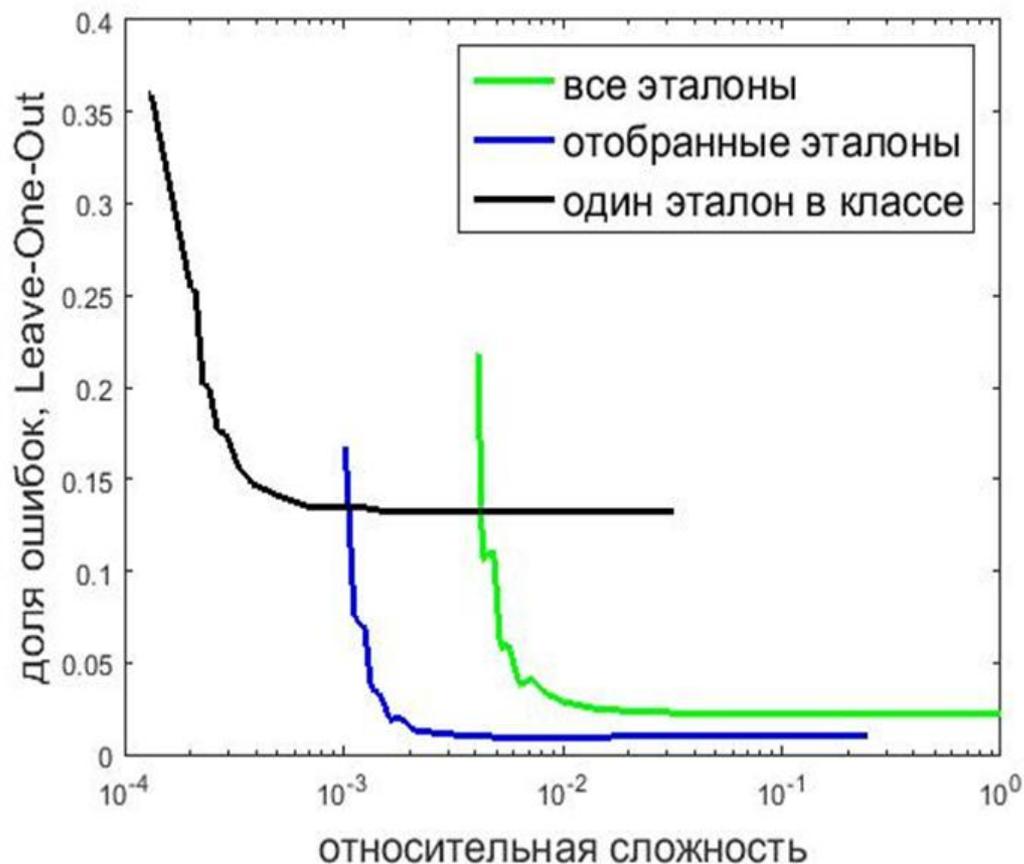
число классов: 90

подписи: 40 объектов по 20 реализаций

жесты руки: 25 объектов по 40 реализаций

лица: 25 объектов по 40 реализаций

Зависимость ошибки $((1-TPR)+FPR)/2$ от сложности (LOO + Leave-One Class-Out)



13. Результаты и выводы

- В пространстве представлений образов с многоуровневым разрешением разработан c -классовый ($c \geq 1$) метрический классификатор по критерию голосования наборов эталонов по классам, который допускает отказ от классификации.
- Для c -классового классификатора введена однопараметрическая ROC -кривая в терминах соотношения *True Positive Rate* и *False Positive Rate*.
- Процедура обучения сведена к отбору наборов эталонов по классам, которые минимизируют доли ошибок элементарных классификаторов типа "класс против всех", и нахождению оценки параметра, минимизирующей ошибку вида $((1 - TPR) + FPR)/2$.
- Используя сеть наборов эталонов по классам с многоуровневым разрешением, предложен решающий алгоритм на основе параметрической стратегии сужения зоны поиска решения на последовательных уровнях разрешения.
- Эффективность классификаторов с различными наборами эталонов продемонстрирована сравнительными ROC -кривыми и зависимостями "ошибка-сложность", полученными для составного источника изображений лиц, жестов и подписей.