

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ
РАДИОТЕХНИКИ, ЭЛЕКТРОНИКИ И АВТОМАТИКИ



МЕТОДЫ И АЛГОРИТМЫ РАСПОЗНАВАНИЯ ОБРАЗОВ
С ИСПОЛЬЗОВАНИЕМ ДРЕВОВИДНЫХ ПРЕДСТАВЛЕНИЙ
МНОГОКАНАЛЬНЫХ ИЗОБРАЖЕНИЙ

Диссертация на соискание ученой степени кандидата технических наук
05.13.01 - Системный анализ, управление и обработка информации

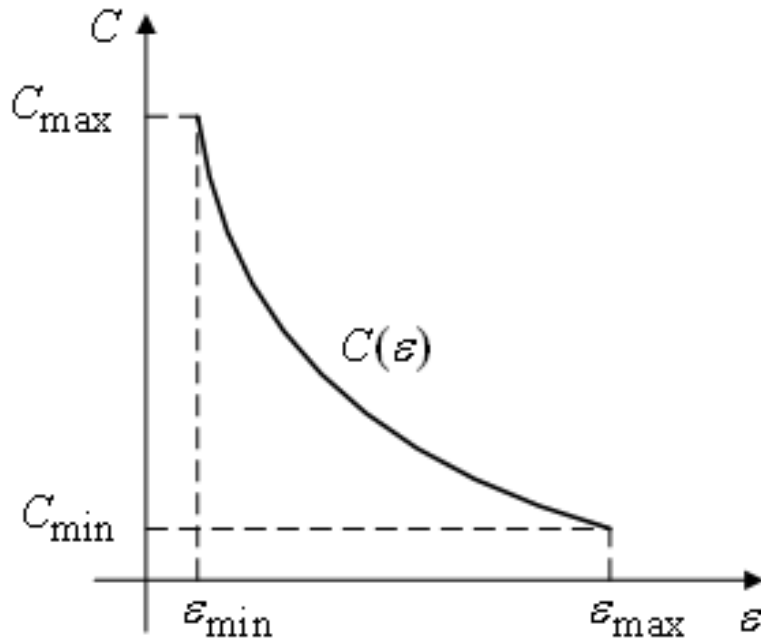
Автор:
Степанов Дмитрий Юрьевич

Научный руководитель:
к.т.н. Ланге Михаил Михайлович

Москва – 2013

1.1. Структурное описание геометрических форм

Схематическая зависимость достоверности и быстродействия алгоритмов распознавания



C – вычислительная сложность,
 ε – доля ошибок распознавания.

Подходы к структурному описанию образов

Квадродеревья и пирамиды
А. Rosenfeld (1980)

Иерархическая декомпозиция
Н. Row и G. Medioni (1993)

Скелетные представления
Л.М. Местецкий (2000, 2009)

Древовидные представления
S. Berretti и A. Del Bimbo (2004)
А. Torsello и E. Hancock (2004)

Деревья эллиптических примитивов
М.М. Ланге и С.Н. Ганебных (2004, 2007)

1.2. Цели и задачи диссертационной работы

- Цель работы заключается в разработке универсального метрического классификатора в пространстве древовидных представлений образов с многоуровневым разрешением для источников многоканальных изображений. Решаются следующие задачи:
- Разработка модели метрического классификатора объектов по многоканальным изображениям. Построение многослойных древовидных представлений образов и введение меры различия объектов на множестве представлений.
- Построение процедуры обучения классификатора на основе отбора эталонов и оптимизации параметров. Получение сравнительных оценок вычислительной сложности иерархического и переборного алгоритмов поиска решений.
- Подготовка базы изображений, разработка комплекса программ и проведение экспериментов по идентификации личности с использованием многоканальных данных, заданных полутонными изображениями подписей и цветными изображениями лиц.

1.3. Положения, выносимые на защиту

- Модель метрического классификатора по критериям взвешенного голосования и ближайшего эталона в пространстве многослойных древовидных представлений образов, заданных многоканальными изображениями.

- Процедура обучения классификатора на основе построения древовидно-структурированного покрытия обучающего множества сферами в пространстве представлений с заданной мерой.

- Способ комплексирования решений по обобщенной мере различия объектов на множестве многослойных древовидных представлений.

- Программный комплекс для реализации процедур обучения и тестирования классификаторов. Экспериментальные оценки качества идентификации личности по полутоновым изображениям подписей, цветным изображениям лиц и ансамблю.

2.1. Многослойное представление и мера различия

Многослойное представление

Структура представления объекта

$$A \rightarrow A^L = \{A_q^L = (a_q^0, \dots, a_q^l, \dots, a_q^L)\}_{q=1}^N,$$

N - число слоев (каналов),

A_q^L - полное бинарное дерево

q -го слоя глубины L ,

a_q^l - l -й уровень в дереве A_q^L .

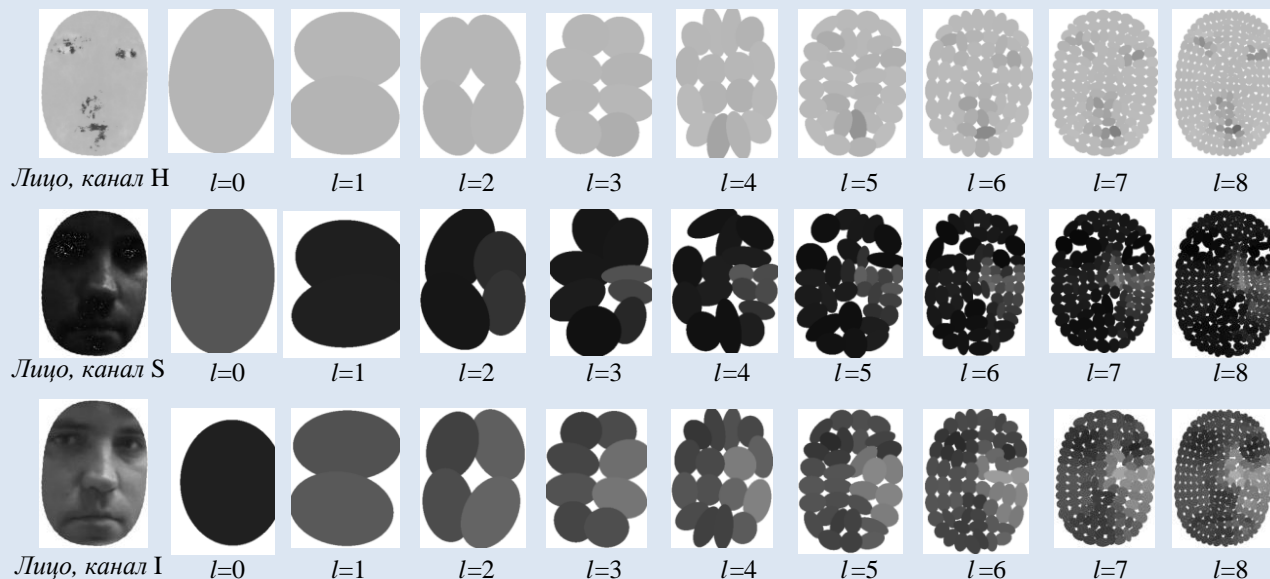
Мера различия по N представлениям

$$d_l(A, \hat{A}) = \sum_{q=1}^N \gamma^{(q)} d_l^{(q)}(A, \hat{A}),$$

$d_l^{(q)}(A, \hat{A})$ - мера различия по q -му слою представления,

$\{\gamma^{(q)} \geq 0: \sum_{q=1}^N \gamma^{(q)} = 1\}_{q=1}^N$ - нормализованные коэффициенты.

Пример представления лица в модели HSI с уровнями разрешения $l = 0, \dots, 8$



2.2. Модель метрического классификатора

Априорное распределение объектов

Множество объектов $\mathbf{A} = \{\mathbf{A}_i\}_{i=0}^c$, $c+1$ - классов с вероятностями $P(\mathbf{A}_i)$, $i = 0, \dots, c$.

Априорные вероятности «своих» и «чужих» объектов $P_{own} = \sum_{i=1}^c P(\mathbf{A}_i)$, $P_{alien} = P(\mathbf{A}_0) = 1 - P_{own}$.

Обучающее множество и множество эталонов

Обучающее множество

$$\mathbf{B} = \{\mathbf{B}_i = \{B_{ij}\}_{j=1}^{m_i}\}_{i=1}^c \subset \mathbf{A}, m_i = \|\mathbf{B}_i\|.$$

Множество эталонов

$$\hat{\mathbf{B}} = \{\hat{\mathbf{B}}_i = \{\hat{B}_{ij}\}_{j=1}^{\hat{m}_i}\}_{i=1}^c \subset \mathbf{B}, \hat{m}_i \leq m_i.$$

Сфера влияния эталона \hat{B}_{ij}
по мере $d_L(A, \hat{B}_{ij})$

$$S_L(\hat{B}_{ij}, D_L(\hat{B}_{ij})) = \{A \in \mathbf{A} : d_L(A, \hat{B}_{ij}) \leq D_L(\hat{B}_{ij})\},$$

\hat{B}_{ij} - центр, $D_L(\hat{B}_{ij}) \geq 0$ - радиус.

Решающее правило для объекта A

Мера сходства L -го порядка между объектом A и подмножеством эталонов $\hat{\mathbf{B}}_i$

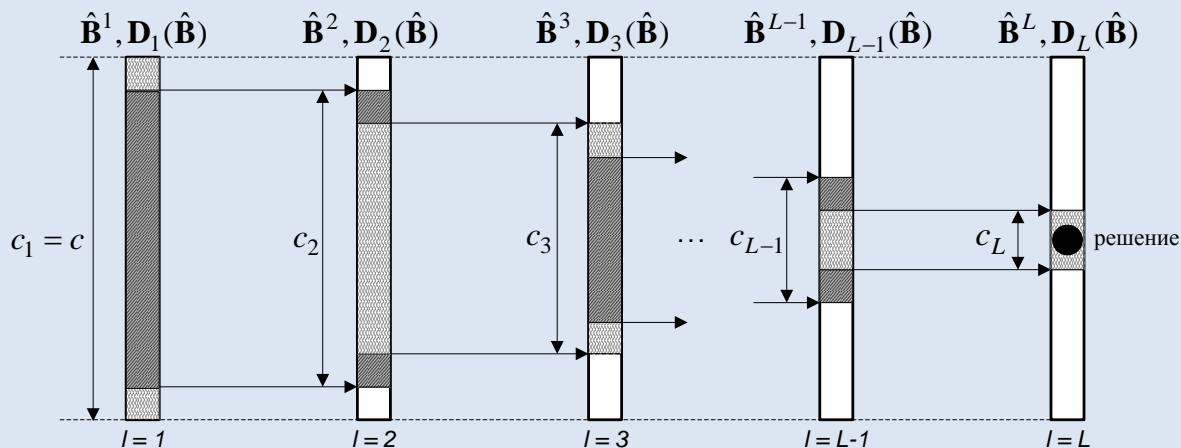
$$\mu_L(A, \hat{\mathbf{B}}_i) = \sum_{j=1}^{\hat{m}_i} q(\hat{B}_{ij}) 2^{\frac{-d_L(A, \hat{B}_{ij})}{D_L(\hat{B}_{ij})}} [d_L(A, \hat{B}_{ij}) \leq D_L(\hat{B}_{ij})].$$

Критерий классификации объекта A

$$k = \arg \max_{i=1}^c \mu_L(A, \hat{\mathbf{B}}_i), \quad i^* = k[\mu_L(A, \hat{\mathbf{B}}_k) > 0].$$

2.3. Алгоритмы поиска и вычислительная сложность

Многоуровневая структура сети эталонов



Алгоритмы поиска

Полный перебор
(Exhaustive Search)

$$c_L = c.$$

Направленный перебор
(Guided Search)

$$c_l = \lfloor c 2^{-\beta(l-1)} \rfloor,$$

$$l = 1, \dots, L, \beta \geq 0.$$

Утверждение 1

При $\beta = \frac{1}{L} \log_2 c \geq 1$ вычислительная сложность алгоритмов Exhaustive Search и Guided Search удовлетворяет оценкам

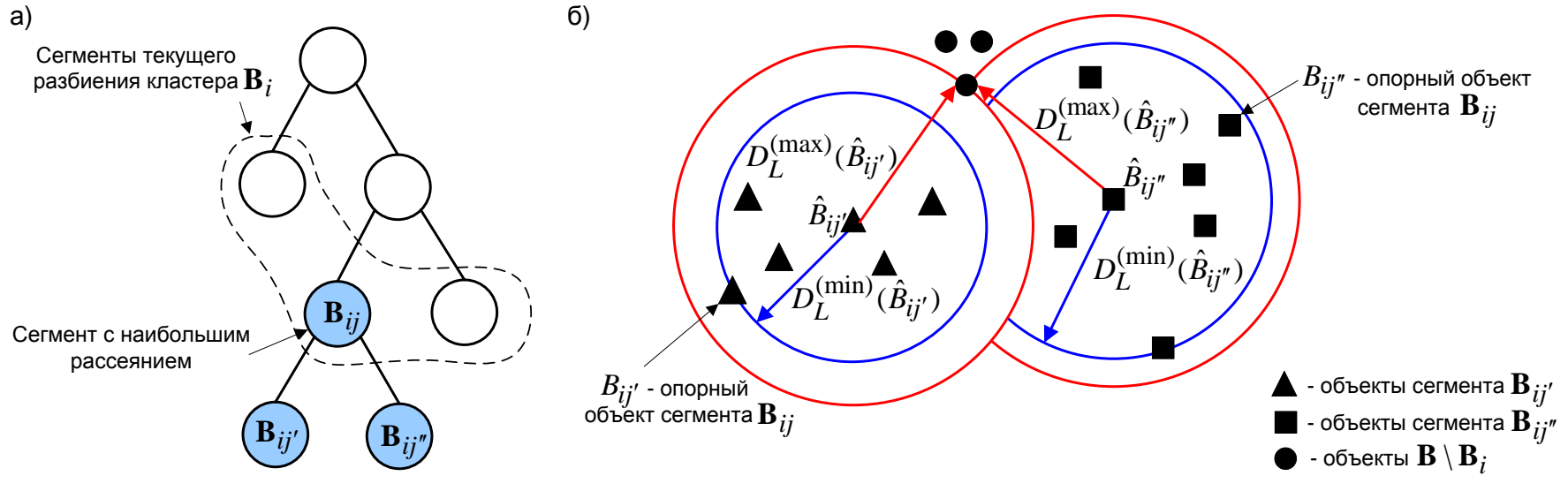
$$C_{ES} = 2N(2^L - 1) \sum_{i=1}^c \hat{m}_i = O(c^2),$$

$$C_{GS} \leq 2NcL \max_{i=1} \hat{m}_i = O(c \log_2 c).$$

С ростом числа классов c алгоритм Guided Search дает выигрыш $C = \frac{C_{ES}}{C_{GS}}$ в быстродействии по сравнению с алгоритмом Exhaustive Search порядка $\frac{c}{\log_2 c}$.

3.1. Построение покрытий и отбор эталонов

TSC-покрытия кластеров сферами (Tree-Structured Covers)



Покрывание кластера сферами: а) бинарное дерево сегментов;
б) сферы покрытия сегментов

Центры и радиусы сфер покрытия

Центр сферы покрытия сегмента \mathbf{B}_{ij}

$$\hat{B}_{ij} = \arg \max_{\hat{B}_{ik} \in \mathbf{B}_{ij}} \min_{B_{ik'} \in \mathbf{B}_{ij}} d_L(B_{ik'}, \hat{B}_{ik})$$

Радиус сферы покрытия сегмента \mathbf{B}_{ij}

$$D_L(\hat{B}_{ij}) = (1 - \alpha_i) D_L^{(\min)}(\hat{B}_{ij}) + \alpha_i D_L^{(\max)}(\hat{B}_{ij}).$$

Параметры покрытия кластера \mathbf{B}_i

$$(\hat{m}_i, \alpha_i).$$

3.2. Функция ошибки и оптимизация покрытий

Утверждение 2

Множество сфер покрытий с центрами $\hat{\mathbf{B}} = \{\{\hat{B}_{ij}\}_{j=1}^{\hat{m}_i}\}_{i=1}^c$ и радиусами $\mathbf{D}_L(\hat{\mathbf{B}}) = \{\{D_L(\hat{B}_{ij})\}_{j=1}^{\hat{m}_i}\}_{i=1}^c$ образует TSC-классификатор в пространстве древовидно-структурированных представлений.

Утверждение 3

Вероятность ошибки обучения TSC-классификатора с параметрами (\hat{m}_i, α_i) , $i = 1, \dots, c$, оценивается средним значением

$$\varepsilon_{\mathbf{B}}(\hat{\mathbf{B}}, \mathbf{D}_L(\hat{\mathbf{B}})) = \frac{1}{c} \sum_{i=1}^c \varepsilon_{\mathbf{B}}^{cv}(\hat{m}_i, \alpha_i),$$

где
$$\varepsilon_{\mathbf{B}}^{cv}(\hat{m}_i, \alpha_i) = \left(\frac{(c-1)}{c} \varepsilon_{\mathbf{B} \setminus \mathbf{B}_i}^{cv}(\hat{m}_i, \alpha_i) + \varepsilon_{\mathbf{B}_i}^{cv}(\hat{m}_i, \alpha_i) \right) P_{own} + \varepsilon_{\mathbf{B} \setminus \mathbf{B}_i}^{cv}(\hat{m}_i, \alpha_i) P_{alien}$$

- функция ошибки скользящего контроля при предъявлении покрытию с параметрами (\hat{m}_i, α_i) «своих» \mathbf{B}_i и «чужих» $\mathbf{B} \setminus \mathbf{B}_i$ объектов обучающего множества.

Оптимизация покрытий кластеров

$$(\hat{m}_i, \alpha_i) = \arg \min_{1 \leq \hat{m}'_i \leq m; 0 \leq \alpha'_i \leq 1} \varepsilon_{\mathbf{B}}^{cv}(\hat{m}'_i, \alpha'_i),$$

где m - наибольшее заданное число эталонов.

3.3. Оценки параметров меры и способ комплексирования

Параметры меры различия объектов по представлениям q -го слоя

$$\omega_k^{(q)} = \frac{\log_2 \varepsilon_{\mathbf{B}}^{(q)}(\hat{\mathbf{B}}, \mathbf{D}_{Lk}(\hat{\mathbf{B}}))}{\sum_{k=1}^3 \log_2 \varepsilon_{\mathbf{B}}^{(q)}(\hat{\mathbf{B}}, \mathbf{D}_{Lk}(\hat{\mathbf{B}}))}, \quad k = 1, 2, 3,$$

$\varepsilon_{\mathbf{B}}^{(q)}(\hat{\mathbf{B}}, \mathbf{D}_{Lk}(\hat{\mathbf{B}}))$ - доля ошибок обучения, полученная на покрытиях с параметрами (\hat{m}_i, α_i) по мере $d_{Lk}^{(q)}$.

Параметры меры различия объектов по N -слойным представлениям

$$\gamma^{(q)} = \frac{\log_2 \varepsilon_{\mathbf{B}}^{(q)}(\hat{\mathbf{B}}, \mathbf{D}_L(\hat{\mathbf{B}}))}{\sum_{q=1}^N \log_2 \varepsilon_{\mathbf{B}}^{(q)}(\hat{\mathbf{B}}, \mathbf{D}_L(\hat{\mathbf{B}}))}, \quad q = 1, \dots, N,$$

$\varepsilon_{\mathbf{B}}^{(q)}(\hat{\mathbf{B}}, \mathbf{D}_L(\hat{\mathbf{B}}))$ - доля ошибок обучения, полученная на покрытиях с параметрами (\hat{m}_i, α_i) по мере $d_L^{(q)}$.

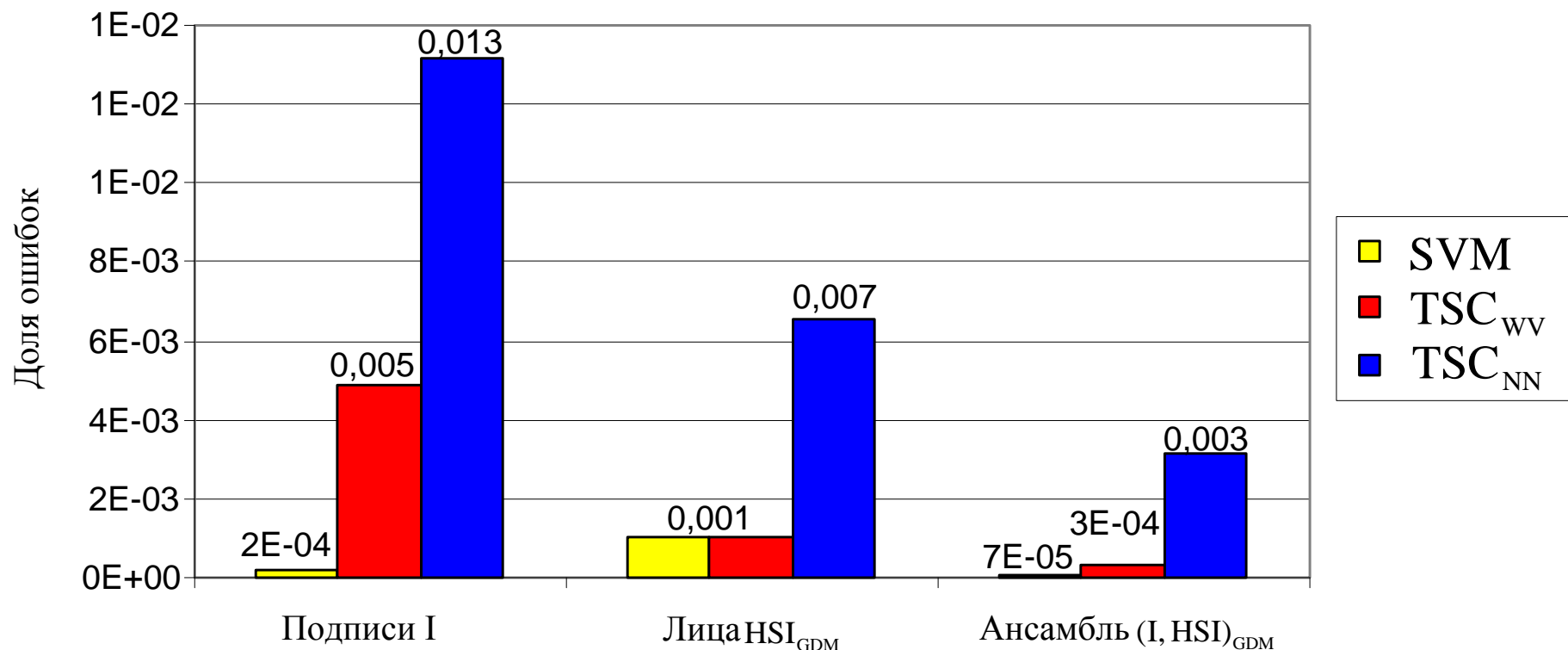
Способ комплексирования по обобщенной мере

Используется GDM-стратегия на основе критерия принятия решения по обобщенной мере различия (General Dissimilarity Measure) в пространстве многослойных представлений объектов.

4.1. Состав экспериментов и показатели эффективности

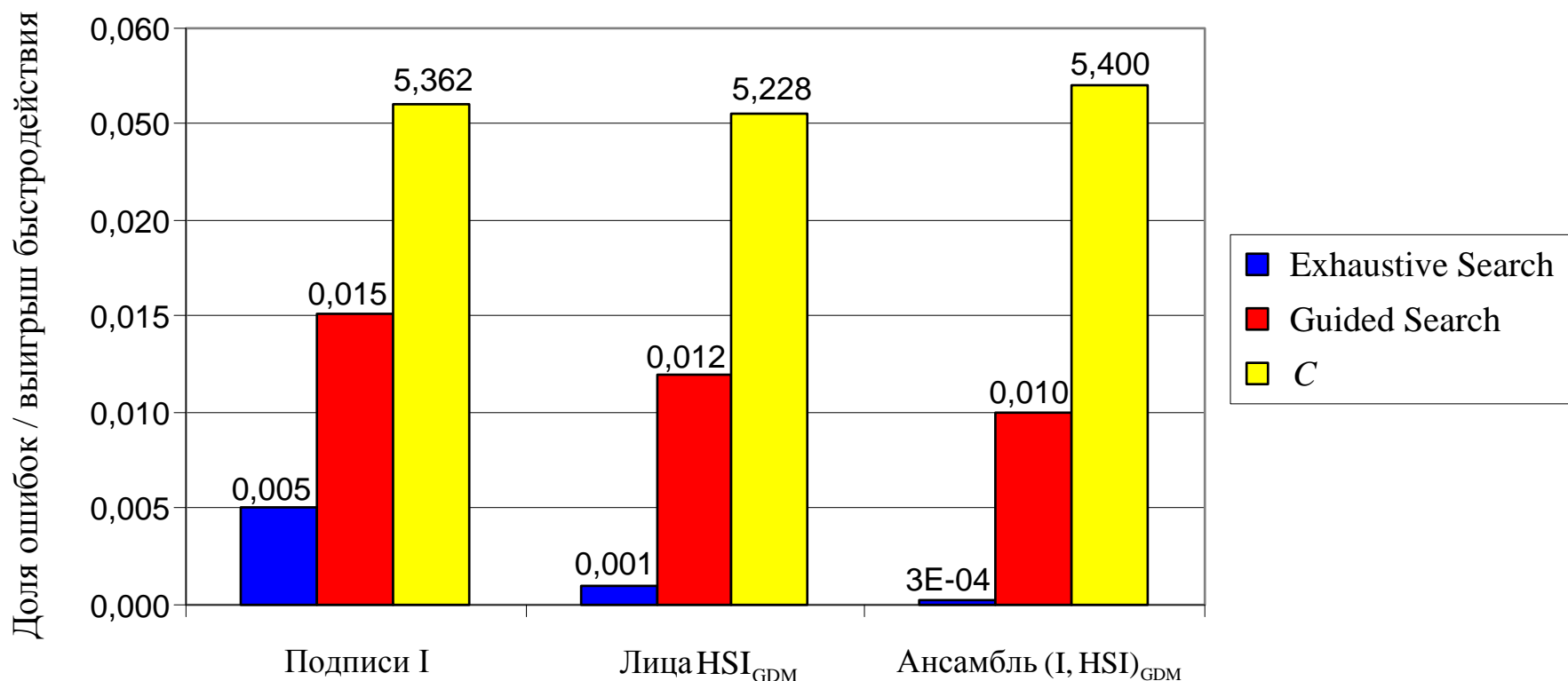
Характеристики классификации	Значение
Источники	Подписи (канал I)
	Лица (каналы H, S и I)
	Ансамбль (подписи + лица)
Классификаторы	Support Vector Machine (SVM)
	Tree Structured Cover, Nearest Neighbor (TSC _{NN})
	Tree Structured Cover, Weighted Vote (TSC _{WV})
Стратегии комплексирования	Weighted Majority Vote (WMV)
	General Dissimilarity Measure (GDM)
Алгоритмы поиска	Полный перебор эталонов (Exhaustive Search)
	Направленный перебор эталонов (Guided Search)
Вероятность "своих" объектов, P_{own}	1, 0.75, 0.5, 0.25
Схема эксперимента	30 times 2 fold Cross-Validation
Характеристики эффективности	Значение
Средняя доля ошибок	$\varepsilon = \frac{1}{30} \sum_{i=1}^{30} \varepsilon_i$
Среднеквадратическое отклонение	$\sigma = \sqrt{\frac{1}{30} \sum_{i=1}^{30} (\varepsilon_i - \varepsilon)^2}$
Вычислительный выигрыш	$C = C_{ES} / C_{GS}$

4.2. Распознавание подписей, лиц и ансамбля источников



Доли ошибок классификации изображений подписей, лиц и ансамбля «подписи, лица» при $P_{own} = 1$ (алгоритм Exhaustive Search)

4.3. Распознавание подписей, лиц и ансамбля источников



Доли ошибок тестирования TSC_{wv} - классификатора с решающими алгоритмами Exhaustive Search и Guided Search на источниках подписей, лиц и ансамбля «подписи, лица» при $P_{own} = 1$

5.1. Основные результаты и выводы

- Разработан способ древовидного представления многоканальных изображений многослойными наборами эллиптических примитивов. Введена мера различия объектов, и построена модель метрического TSC-классификатора.
- Получены оценки вычислительной сложности иерархического и переборного алгоритмов поиска решений. Показано, что иерархический алгоритм обеспечивает большее быстроедействие по сравнению с алгоритмом полного перебора.
- Разработана процедура обучения TSC-классификатора, включающая отбор эталонов и оценивание параметров. Для отбора эталонов и оценки параметров используется метод скользящего контроля. Предложен способ объединения классификаторов.
- Реализован программный комплекс в среде Turbo C++ для проведения экспериментов с различными источниками изображений. Комплекс включает модули предобработки и представления данных, обучения и тестирования классификаторов.
- Выполнены эксперименты по распознаванию изображений подписей, лиц, ансамбля изображений «подписи, лица». Получены оценки качества классификации и вычислительного выигрыша иерархического алгоритма относительно переборного.
- Разработанная модель классификатора универсальна и применима для широкого класса источников многоканальных изображений. В перспективе планируется исследовать модификации схемы отбора эталонов и алгоритма иерархического поиска.

5.2. Апробация, публикации и внедрения

Апробация

- 2 международные конференции
- 3 всероссийские конференции
- 4 внутривузовские конференции
- 3 конкурса и 1 семинар
- дипломы I, II, III степени
- премия имени д.т.н. В.С. Корсакова-Богаткова
- грамоты конкурсов

12 публикаций

- 4 статьи в журналах из перечня ВАК
- 2 свидетельства регистрации программ
- 2 доклада в трудах международных конференций
- 2 доклада в трудах всероссийских конференций
- 2 доклада в трудах внутривузовских конференций

Внедрения

- МГТУ МИРЭА
- ОАО «Чувашнефтепродукт»
- Проект РФФИ №09-01-00573-а