

РАЗРАБОТКА И ИССЛЕДОВАНИЕ МЕР БЛИЗОСТИ ДЛЯ ПРЕЦЕДЕНТНОЙ ИДЕНТИФИКАЦИИ ФРАГМЕНТОВ РУКОПИСНОГО ТЕКСТА

Н. О. Ефимов

Белгородский государственный национальный исследовательский университет

Поступила в редакцию 11.08.2018 г.

Аннотация. В работе предложены и исследованы меры близости для прецедентной идентификации фрагментов рукописного текста. Сформулирована задача, разработана методика проведения сравнительного анализа различных мер близости. Приведены результаты вычислительных экспериментов для мер близости на основе сравнения пикселей, мер близости на основе сопоставления проекций фрагментов и субполосных мер близости, на основе сопоставления значений энергии фрагментов в заданных частотных интервалах.

Ключевые слова: рукописный текст, обработка изображений, распознавание изображений, распознавание рукописного текста, прецедентная идентификация, субполосный анализ.

Annotation. In the article introduced and investigated proximity measures for precedential identification of words on handwritten text images. Main problem was formulated; methodology of various proximity measures comparative analysis was developed. Result of computational experiments on MatLab software for pixel comparison proximity measure, fragment projection mapping proximity measure and for subband proximity measure are presented. Subband proximity measure based on energy comparison and react on normalized Fourier transform differences in defined frequency slots. The efficiency of reviewed proximity measures is illustrated by calculated errors probability.

Keywords: handwritten text, image processing, image recognition, HWR, OCR, precedential identification, subband analysis.

ВВЕДЕНИЕ

Письменность является одной из устоявшихся и распространенных форм информационного обмена. Количество дошедших до наших времен текстов на различных языках не поддается счислению. К примеру, Российский государственный архив древних текстов содержит более 3 млн документов, за хронологический период XI – начало XX веков.

С развитием компьютерной техники для работы с текстовыми документами используется их представление в цифровой форме. Для этого бумажный носитель подвергают процедурам сканирования и сохранения полученных данных в память компьютера в формате цифрового изображения.

Такое представление существенно упрощает задачу хранения рукописных текстов и предоставления доступа к ним различных пользователей. Но есть и ряд недостатков: затруднены возможности ориентирования по тексту, практически невозможны операции редактирования, форматирования и поиск по тексту.

Существующие системы обработки изображений текста (A2iA, ParaScript, FineReader и др.) ориентированы на распознавание текста, то есть перевод изображений текста в текстовые данные. Причем, задача распознавания в настоящее время решена только для изображений печатного текста. В случае рукописного текста, существующие системы способны уверенно распознавать только машиночитаемые формы (анкеты, бланки, и т. д.), так как структурирование документов

и уменьшение диапазона вводимых в поле символов позволяет существенно упростить задачу. В основе большинства существующих систем лежат методы структурного описания символов в совокупности с моделью языка. Однако, до сих пор не решена задача сегментации слова на символы, ввиду их высокой вариативности в зависимости от языка, автора, положения в тексте и прочих факторов. Также, для работы с конкретным языком необходима своя языковая модель, либо долговременное обучение. Стоит отметить, что большинство существующих систем распознавания базируются на идее усреднения данных по ансамблю (вероятностных моделях), то есть на выявлении общего при начертании одних и тех же символов разными авторами. Такой подход эффективен при распознавании машинописного и рукопечатного текстов, но в случае рукописного ведет к потере индивидуальных характеристик почерка (тип, размер, наклон, разгон, связность и т. д.).

Поэтому, представляется целесообразным подход с позиций прецедентной идентификации фрагментов изображений рукописного текста. Тогда, **постановка задачи** имеет следующий вид:

1. Существует цифровое изображение рукописного текста, написанного одним автором и полученное в результате сканирования.
2. На изображении оператор интерактивно выделяет интересующий фрагмент в виде слова (далее – прецедент).
3. Необходимо автоматически найти и выделить в тексте все фрагменты, идентичные прецеденту в смысле некоторой меры близости.

В вышеобозначенной постановке задача в известной литературе подробно не рассматривалась. Вместе с тем, можно указать ряд применений, которые свидетельствуют о востребованности разработки:

1. Лингвистические исследования:
 - построение частотных словарей для одного автора;
 - построение частотных словарей для языка за указанный временной промежуток;
 - оценка изменений словарного состава языка.

2. Поиск нечетких дубликатов:

- поиск документов, принадлежащих одному автору;
- поиск специальных символов, сокращений, характерных для автора (существует, рассматривается и т. д.) в документе;
- поиск по личным записям, конспектам;
- поиск документов, содержащих сходные фрагменты, подпись, печать;
- поиск текстового оригинала по имеющейся иллюстрации.

Предложенная гипотеза H_0 имеет вид: **сравниваемые фрагменты идентичны**. Для проверки гипотезы формируется решающая функция (РФ), представляющая собой меру близости сравниваемых фрагментов, попадание значений которой в критическую область воспринимается как основание отвергнуть гипотезу.

Эффективность процедур распознавания образов зависит от адекватности признаков, с помощью которых описываются объекты. Причем, пространство признаков должно удовлетворять следующим требованиям:

- Количество признаков должно быть минимально возможным, но достаточным для успешной (в смысле заданного значения вероятностей ошибок) идентификации фрагментов.
- Различие значений признаков для объектов, идентичных прецеденту, должно быть минимально.
- Значения признаков для объектов, не идентичных прецеденту, должны существенно отличаться от соответствующих значений для идентичных объектов.

До определенной степени постановку задачи можно рассмотреть с позиций обнаружения объектов на изображениях. В настоящее время широко распространены методы на основе вычисления ключевых точек, например, SIFT, SURF, детекторы SUM7, MSER, и др. Проведенные исследования показали (рис. 1), что применение подобного подхода затруднено большим количеством ошибок вида «ложная тревога». Это связано с тем, что метод SURF (к примеру) реагирует на пятна, углы и края линий, а рукописный текст состоит из небольшого количества постоянно повторяющихся элементов.

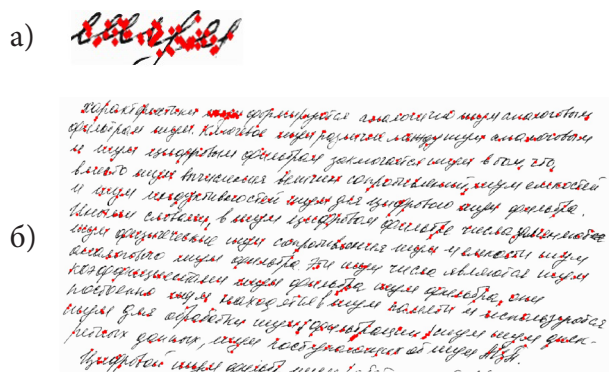


Рис. 1. Ключевые точки, вычисленные методом SURF: а) прецедентный фрагмент б) Фрагмент текста, размерностью 2546×1242

Таким образом, целесообразна разработка меры близости фрагментов рукописного текста, учитывающая свойства рукописного текста и позволяющая осуществлять прецедентную идентификацию заданных фрагментов.

МАТЕРИАЛЫ И МЕТОДЫ ИССЛЕДОВАНИЯ

Разработка меры близости осуществляется исходя из следующих **исходных условий**:

Φ_0 – цифровое изображение сканированного рукописного текста, $\Phi_0 = (f_{i,k})$, $i = 1, 2, \dots, N$, $k = 1, 2, \dots, M$; i, k – координаты на плоскости, а $f_{i,k}$ – интенсивность изображения в заданной точке.

На изображении текста (подразумевается сканированная страница текста) Φ_0 , в формате «оттенки серого», размерностью $M \times N$ выделяется прямоугольная область Φ_p , размерностью $M_p \times N_p$, которая в дальнейшем именуется прецедентным фрагментом (индекс p означает прецедент).

Далее из изображения Φ_0 формируется множество Z^{Φ_k} фрагментов Φ_0 , размерностью $M_p \times N_p$, где $k = 1, 2, \dots, K$ – количество полученных фрагментов, каждый из которых необходимо сопоставить с прецедентом используя меру близости.

В работе сформулирована **гипотеза H_0** : сравниваемые фрагменты идентичны. Для проверки справедливости гипотезы используется решающая функция, представляющая

собой меру близости и границу критической области. В соответствии с гипотезой естественно рассматривать два класса фрагментов: идентичные заданному и все остальные.

Для оценивания предлагаемых мер близости предлагается вычислять значения ошибок распознавания первого и второго рода:

$$P_1 = \frac{N_{o1}}{N_{пол}}, \quad P_2 = \frac{N_{o2}}{N_{пол2}}, \quad (1)$$

где N_{o1} – количество объектов, принадлежащих к классу распознавания, но не отнесенных к нему, $N_{пол}$ – количество объектов на изображении, которые принадлежат к классу распознавания;

N_{o2} – количество объектов, ошибочно отнесенных к интересующему классу распознавания, $N_{пол2}$ – количество не принадлежащих к классу распознавания объектов на изображении.

«Гипотеза отвергается, когда значение решающей функции больше порога, величина которого выбирается исходя из требования обеспечения заданного уровня значений ошибок первого рода. Решающая функция должна отвечать принципу максимизации вероятности правильного принятия противоположной гипотезы, когда она верна (минимизации значений ошибок второго рода).»

Ошибка **первого рода** возникает в том случае, если гипотеза H_0 неверно отвергнута, то есть фрагмент близок к прецеденту, но отнесен в противоположный класс объектов. Ошибка **второго рода** возникает в том случае, если гипотеза H_0 неверно принята, то есть фрагмент не идентичен к прецеденту, но отнесен к идентичным.

Для проведения исследования была сформирована **база рукописных слов**: Количество авторов – 10; Количество различных слов – 10; Количество вариантов одного слова – около 100, в зависимости от автора и слова. Множество фрагментов текста, написанных одним автором обозначим как $Z_{author}^{\Phi_k}$, множество фрагментов текста, соответствующих одному слову одного автора обозначим как $Z_{word}^{\Phi_k}$. Тогда

$$Z_{author_1}^{\Phi_k} = \{Z_{word_1}^{\Phi_k}, Z_{word_2}^{\Phi_k}, \dots, Z_{word_W}^{\Phi_k}\}, \quad (2)$$

где W – количество слов в выборке для одного автора, $Z_{word_w}^{\Phi_k} = \{\Phi_1, \Phi_2, \dots, \Phi_{sl}\}$, sl – количество объектов, принадлежащих одному слову.

Таким образом, для автора 1 сформирована следующая выборка:

$Z_{author_1}^{\Phi_k} = \{\Phi_1, \Phi_2, \dots, \Phi_{1040}\}$, общее количество объектов в выборке $K = 1040$;
 $Z_{author_1}^{\Phi_k} = \{Z_{word_1}^{\Phi_k}, Z_{word_2}^{\Phi_k}, \dots, Z_{word_{10}}^{\Phi_k}\}$, количество слов в выборке $W = 10$, примеры исследуемых слов представлены на рис. 2;

$Z_{word_1}^{\Phi_k} = \{\Phi_1, \Phi_2, \dots, \Phi_{101}\}$,

$Z_{word_2}^{\Phi_k} = \{\Phi_{102}, \Phi_{103}, \dots, \Phi_{206}\}$, и т. д.

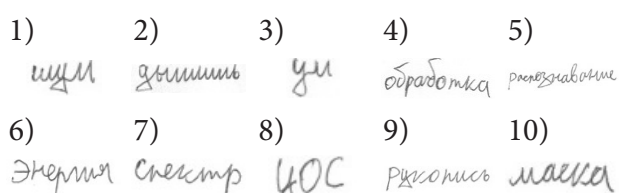


Рис. 2. Примеры исследуемых слов для автора 1

Методика исследования меры близости

1) Задается выборка слов для одного автора $Z_{author}^{\Phi_k}$. Количество фрагментов $k = 1, 2, \dots, K$; $K = 1040$ для автора 1.

2) Задается интересующее слово w , к примеру, для слова «шум» $w = 1$.

3) Из множества фрагментов (sl – их количество), соответствующих заданному слову $Z_{word_w}^{\Phi_k}$ выбирается один, который впоследствии считается прецедентом Φ_p .

4) Рассчитывается значение меры близости для прецедента Φ_p и каждого из фрагментов $\Phi_k \in Z_{author}^{\Phi_k}$.

5) Задается пороговое значение h , как максимум меры близости S в области идентичных фрагментов. Это позволяет фиксировать значение ошибки первого рода равной нулю, и рассчитать значения ошибки второго рода.

$$h = \max_{i=1, \dots, sl} S_i. \quad (3)$$

6) Шаги 3–5 выполняются для каждого фрагмента из области идентичных.

7) На основании полученных данных для заданного слова w определяются:

1. Минимальное значение P_2 :
 $P_2^{\min} = \min_{i=1, \dots, sl} P_{2_i}$;

2. Максимальное значение P_2 :
 $P_2^{\max} = \max P_{2_i}$;

3. Размах P_2 : $R_{P_2} = P_2^{\max} - P_2^{\min}$;

4. Среднее значение P_2 : $P_2^{\text{mean}} = \text{mean}(P_{2_i})$, где $i = 1, 2, \dots, sl$; mean – среднее арифметическое.

8) Повторить пункты 2–7 для каждого слова.

Предлагаемая методика позволяет сравнивать различные меры близости, величины P_2^{\min} и P_2^{\max} характеризуют полученные значения ошибки второго рода при фиксированном значении ошибки первого рода (п. 5 методики). Размах R_{P_2} характеризует чувствительность меры близости к изменчивости идентичных фрагментов. Некоторые фрагменты могут быть идентичны прецеденту, но существенно отличаться от него (в смысле характеристик написания – наклон, связность, размах и т. д.), для таких фрагментов значение P_2 максимально, поэтому необходима характеристика P_2^{mean} для оценки основного количества идентичных фрагментов.

Исследуемые меры близости

1) Мера близости на основе сравнения значений пикселей [1, 2].

1. Задана выборка слов для одного автора. Здесь и далее приведены результаты исследований для автора 1.

2. Задается прецедентный фрагмент Φ_p , размерностью M_p, N_p , остальные фрагменты выборки обозначим как Φ_k . Общее количество фрагментов равно K и соответствует длине выборки.

3. Так как размеры слов колеблются в широких пределах, для сравнения по пикселям приведем фрагмент Φ_k к размерности $M_p \times N_p$.

4. Для каждого $k = 1, 2, \dots, K$ осуществляется

$$\partial_k(i, j) = \begin{cases} 1, & \Phi_k(i, j) = \Phi_p(i, j) \\ 0, & \Phi_k(i, j) \neq \Phi_p(i, j) \end{cases}, \quad (4)$$

где $i = 1, 2, \dots, M_p, j = 1, 2, \dots, N_p$.

5. Вычислить

$$S_{pix}^k = \frac{100}{MN} \sum_{i=1}^M \sum_{j=1}^N \partial_k(i, j). \quad (5)$$

Выражение характеризует то, насколько близок фрагмент k к прецедентному в процентном соотношении.

6. Тогда, мера близости имеет вид:

$$S_{pix}^{pk} = 100 - s_{pix}^k, \quad (6)$$

и принимает значение 0 при полной идентичности фрагмента и 100 при полной неидентичности.

7. Вычислить значения ошибки второго рода, согласно изложенной выше методике.

2) Мера близости на основе Евклидовой нормы [1, 2]:

$$S_{Evkl}^{pk} = \sqrt{\frac{\sum_{i=1}^{M_p} \sum_{j=1}^{N_p} (\Phi_k(i, j) - a \cdot \Phi_p(i, j))^2}{\sum_{i=1}^{M_p} \sum_{j=1}^{N_p} \Phi_p^2(i, j)}}, \quad (7)$$

где a – коэффициент, позволяющий минимизировать различия в интенсивности написания слов, который вычисляется на основании соотношения:

$$a = \frac{\sum_{i=1}^{M_p} \sum_{j=1}^{N_p} \Phi_k(i, j) \Phi_p(i, j)}{\sum_{i=1}^{M_p} \sum_{j=1}^{N_p} \Phi_p^2(i, j)}. \quad (8)$$

3) Мера близости на основе сопоставления проекций фрагмента:

Поле фрагмента разбивается на четыре квадранта, рис. 3.

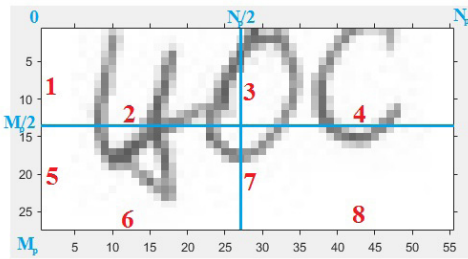


Рис. 3. Схема разбиения поля объекта

На основании преобразования Радона для цифровых изображений (9) определяются вертикальные и горизонтальные проекции контура символа в каждом квадранте на линии, разделяющие квадранты [3].

$$g(\rho, \theta) =$$

$$= \sum_{j=1}^N \sum_{i=1}^M \Phi(i, j) \delta(j \cos(\theta_d) + i \sin(\theta_d) - \rho_t), \quad (9)$$

где $\Phi(i, j)$ – цифровое изображение размерностью $M \times N$,

$j \cos(\theta_d) + i \sin(\theta_d) - \rho_t = 0 \sqrt{a^2 + b^2}$ – уравнение прямой, в форме представления нормалью, правая часть (9) обращается в нуль всюду, кроме точек, в которых δ – функция принимает нулевое значение, т. е. суммирование производится вдоль линии $j \cos(\theta_d) + i \sin(\theta_d) - \rho_t$, θ_d – угол наклона прямой, ρ_t характеризует положение прямой в пространстве, $t = 1, 2, \dots, T$ – количество прямых, для которых вычисляется значение (9), а значит – количество точек проекции на заданную ось.

Так как предлагается разбить исходное изображение на 4 области, для каждой из которых вычислить значение проекции на указанные оси (рис. 3), значение проекций p_{pr}^k для фрагмента Φ_k вычисляется как:

$$p_{pr}^k = \{g_{proj_1}, g_{proj_2}, \dots, g_{proj_8}\}, \quad (10)$$

где g_{proj_b} – значение проекции (9) в T точках, при $\theta_d = 90$ для нечетных b и $\theta_d = 0$ для четных; b – количество отрезков, на которые осуществляется проецирование (рис. 3).

Тогда, мера близости имеет вид:

$$S_{pr}^{pk} = \frac{1}{\sum_{i=1}^L p_{pr_i}^k} \sum_{i=1}^L |p_{pr_i}^k - p_{pr_i}^p|, \quad (11)$$

где p_{pr}^k, p_{pr}^p – значения проекций для фрагмента и прецедента соответственно, $L = 2M_p + 2N_p$ – количество элементов проекции.

4) Меры близости, разработанные на основе использования субполосных признаков прецедента.

Изображения рукописного текста обладают свойством квазипериодичности – для одного автора сохраняются такие характеристики почерка как наклон, разгон, связность, примерно одинаковый размер символов. Для анализа квазипериодических и периодических деталей на изображении, адекватным является представление изображений в области пространственных частот (ПЧ).

Под частотным представлением функции $f_{i,k}$ понимается выражение [4]:

$$f_{i,k} = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} F^{\Phi}(u, v) e^{ju(i-1)} e^{jv(k-1)} dudv, \quad (12)$$

где j – мнимая единица, $F^\Phi(u, v)$ – частотная характеристика, под которой обычно понимают трансформанту Фурье изображения Φ ,

$$F^\Phi(u, v) = \sum_{i=1}^M \sum_{k=1}^N f_{ik} e^{-ju(i-1)} e^{-jv(k-1)}, \quad (13)$$

Аргументы (u, v) называют пространственными частотами (ППЧ),
 $(u, v) \in V_{2\pi}$,

$$V_{2\pi} = \{(u, v) | -\pi \leq u < \pi, -\pi \leq v < \pi\}. \quad (14)$$

Соотношение (15) представляет собой аналог равенства Парсеваля [5] для двумерного случая, которое связывает энергии (нормы) функций, заданных в дискретных отсчетах, и значения их трансформант Фурье:»

$$\sum_{i=1}^N \sum_{k=1}^M f_{ik}^2 = \sum_{s=1}^S \sum_{r=1}^R \frac{1}{4\pi^2} \iint_{(u,v) \in V_{sr}} |F^\Phi(u, v)|^2 dudv. \quad (15)$$

В контексте субполосного анализа [5, 6, 7] равенство (15) целесообразно представить в виде:

$$\|\Phi\|^2 = \sum_{s=1}^S \sum_{r=1}^R E_{sr}(\Phi), \quad (16)$$

где

$$E_{sr}(\Phi) = \frac{1}{4\pi^2} \iint_{(u,v) \in V_{sr}} |F^\Phi(u, v)|^2 dudv. \quad (17)$$

$V_{sr} = V((u_{s1}, u_{s2}), (v_{r1}, v_{r2}))$ – подобласть пространственных частот (ППЧ), $s = 1, 2, \dots, S$, $r = 1, 2, \dots, R$, S, R – количество частотных интервалов для каждой оси.

Интегралы вида (17) будем называть значениями энергии изображения, попадающей в соответствующие ППЧ. Тогда величина $P_{sr}(\Phi)$:

$$P_{sr}(\Phi) = \frac{E_{sr}(\Phi)}{\|\Phi\|^2} \quad (18)$$

представляет собой долю энергии изображения, которая позволяет судить о ее сосредоточенности в выбранном интервале ППЧ V_{sr} .

Причем, (17) и (18) можно выразить через субполосные матрицы [6]:

$$E_{sr} = \text{tr}(A_s \Phi B_r \Phi^T), \quad (19)$$

$$P_{sr}(\Phi) = \frac{\text{tr}(A_s \Phi B_r \Phi^T)}{\text{tr}(\Phi \Phi^T)}, \quad (20)$$

где «tr» означает след матрицы, «T» – операция транспонирования матрицы, Φ – матри-

ца исходного изображения, A_s, B_r – субполосные матрицы $A_s = \{a_{in}^s\}$, $i, n = 1, 2, \dots, N$; $B_r = \{b_{km}^r\}$, $k, m = 1, 2, \dots, M$, элементы которых вычисляются на основании следующих соотношений:

$$a_{in}^s = \begin{cases} \frac{\sin(u_{s2}(i-n)) - \sin(u_{s1}(i-n))}{\pi(i-n)}, & i \neq n, \\ \frac{u_{s2} - u_{s1}}{\pi}, & i = n; \end{cases} \quad (21)$$

$$b_{km}^r = \begin{cases} \frac{\sin(v_{r2}(k-m)) - \sin(v_{r1}(k-m))}{\pi(k-m)}, & k \neq m, \\ \frac{v_{r2} - v_{r1}}{\pi}, & k = m. \end{cases} \quad (22)$$

В работах [7, 8] показано, что большая часть энергии (19) изображения рукописного текста содержится в малом количестве частотных интервалов, используя различные способы выделения информационных интервалов можно выделить фон, текст, и высокочастотные/шумовые компоненты. Так как ППЧ неравнозначны в смысле восстанавливаемых компонент изображения (рис. 4, 5) представляется целесообразной разработка меры близости, позволяющей учитывать различия энергии в заданных частотных интервалах.

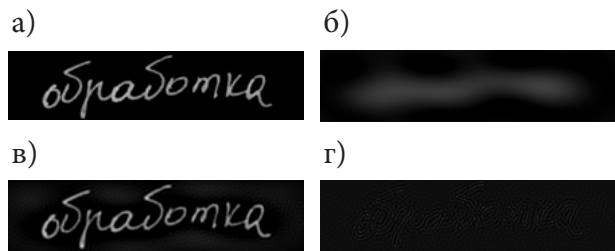


Рис. 4. Пример восстановления изображения из различных ППЧ

а) все ППЧ; б) изображение, восстановленное из частотных интервалов, содержащих 20 % энергии (соответствуют белым ППЧ рис. 5); в) изображение, восстановленное из остальных частотных интервалов, содержащих значение энергии выше среднего (соответствуют серым ППЧ рис. 5); г) изображение, восстановленное из оставшихся частотных интервалов (соответствуют черным ППЧ рис. 5)

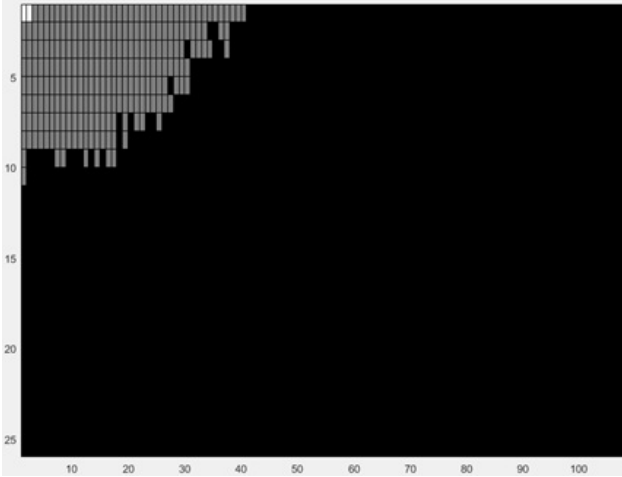


Рис. 5. Карта информационных интервалов для рис. 4а

Количество интервалов, на которое осуществляется разбиение по каждой оси, вычисляется согласно выражениям:

$$R_a = \frac{M - 2}{4}, \quad R_b = \frac{N - 2}{4}. \quad (23)$$

4.1. Мера близости на основе вычисления субполосного коэффициента корреляции в заданном частотном интервале:

$$S_{skk}^{pk} = 2 \left(\frac{1}{R_1 \cdot R_2} \sum_{r=1}^{R_b} \sum_{s=1}^{R_a} (1 - p_{sr}^{pk}) \right), \quad (24)$$

где R_a, R_b – количество частотных интервалов (23), p_{rm}^{pk} – субполосный коэффициент корреляции:

$$p_{rm}^{pk} = \frac{\text{tr}(A_s \Phi_p B_r \Phi_k^T)}{\sqrt{E_{sr}^{\Phi_p} \cdot E_{sr}^{\Phi_k}}}; \quad (25)$$

A, B – субполосные матрицы (21, 22); $E_{sr}^{\Phi_p}, E_{sr}^{\Phi_k}$ – значения энергии прецедента (19) и фрагмента в заданных частотных интервалах V_{sr} ; s, r – частотный интервал.

4.2 Мера близости на основе вычисления разности энергии фрагментов в заданном частотном интервале:

$$S_{Energ}^{pk} = \frac{1}{\sum_{r=1}^{R_b} \sum_{s=1}^{R_a} E_{sr}^{\Phi_p}} \sum_{r=1}^{R_b} \sum_{s=1}^{R_a} |E_{sr}^{\Phi_k} - E_{sr}^{\Phi_p}|. \quad (26)$$

4.3 Мера близости на основе вычисления разности долей энергии фрагментов в заданном частотном интервале.

$$S_{Doly}^{pk} = \sum_{r=1}^{R_b} \sum_{s=1}^{R_a} |P_{sr}^{\Phi_k} - P_{sr}^{\Phi_p}|. \quad (27)$$

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЙ И ИХ ОБСУЖДЕНИЕ

В соответствии с вышеизложенной методикой были проведены вычислительные эксперименты в программном комплексе MATLAB для каждой меры близости. Основные результаты на примере четырех прецедентов (слов) приведены в табл. 1.

Методы сравнения пикселей (S_{pix}^{pk} (6), S_{Evkl}^{pk} (7)) показывают лучший результат в случае бинаризации исходного изображения, однако среднее значение ошибки для некоторых слов («маска», «дышишь») достигает 97 %.

Метод проекций S_{pr}^{pk} обладает меньшими значениями минимума, размаха и среднего P_2 . Существенным преимуществом данного метода является возможность вычисления значений проекций в заданном количестве точек. Наиболее эффективен метод проекций для коротких слов, содержащих до 4х символов, что связано со схемой разбиения (рис. 3). Возможно использование неравномерного разбиения в смысле увеличения количества подобластей в соответствии с размерами фрагмента, при этом важно учитывать различия размерностей идентичных фрагментов (длина и ширина одного слова колеблется в пределах 50 % от максимальной для одного автора).

Так как изображения рукописного текста обладают свойством концентрации энергии в малом количестве частотных интервалов целесообразна мера близости, основанная на использовании субполосного коэффициента корреляции (СКК). Наименьшие значения ошибок получены при вычислении значений меры близости в информационных интервалах, содержащих 80 % энергии. Большой размах меры близости, объясняется необходимостью приведения фрагмента и прецедента к одной размерности. В случае использования окна анализа, размером с прецедент существенно возрастает количество вычислений, возрастают ошибки как первого, так и второго рода. В статье представлены значения меры близости на основе СКК для фрагментов, приведенных к размерности прецедента путем бикубической интерполяции.

Таблица 1

Сравнение предложенных мер близости

	P_2^{\max}	P_2^{\min}	R_{P_2}	P_2^{mean}		P_2^{\max}	P_2^{\min}	R_{P_2}	P_2^{mean}
	«Шум»					«Дышишь»			
S_{pix}^{pk}	0.94	0.31	0.64	0.52		0.99	0.80	0.19	0.86
S_{Evkl}^{pk}	0.87	0.61	0.26	0.74		0.98	0.97	0.01	0.97
S_{pr}^{pk}	0.22	0.00	0.22	0.04		0.68	0.22	0.47	0.48
S_{skk}^{pk}	0.90	0.02	0.88	0.37		0.92	0.18	0.74	0.48
$S_{\text{Enerг}}^{pk}$	0.36	0.01	0.35	0.18		0.77	0.04	0.74	0.46
S_{Doly}^{pk}	0.71	0.01	0.70	0.24		0.67	0.01	0.66	0.31
	«Ум»					«Обработка»			
S_{pix}^{pk}	0.58	0.01	0.58	0.23		0.96	0.13	0.83	0.32
S_{Evkl}^{pk}	0.55	0.21	0.33	0.28		0.43	0.29	0.14	0.35
S_{pr}^{pk}	0.04	0.00	0.04	0.00		1.00	0.56	0.44	0.92
S_{skk}^{pk}	0.79	0.06	0.73	0.43		0.99	0.19	0.80	0.58
$S_{\text{Enerг}}^{pk}$	0.11	0.00	0.11	0.01		0.56	0.12	0.44	0.36
S_{Doly}^{pk}	0.67	0.09	0.58	0.21		0.62	0.07	0.55	0.37

Таблица 2

Значения ошибок 2-го рода субполосных мер близости $S_{\text{Enerг}}^{pk}$, S_{Doly}^{pk} и их комбинации

	P_2^{\min}	R_{P_2}	P_2^{mean}		P_2^{\min}	R_{P_2}	P_2^{mean}
	«Шум»				«Дышишь»		
$S_{\text{Enerг}}^{pk}$	0.011	0.347	0.184		0.037	0.737	0.462
S_{Doly}^{pk}	0.008	0.704	0.245		0.009	0.663	0.307
$S_{\text{Enerг}}^{pk} \& S_{\text{Doly}}^{pk}$	0.003	0.313	0.112		0.005	0.536	0.237
	«Ум»				«Обработка»		
$S_{\text{Enerг}}^{pk}$	0.001	0.113	0.012		0.118	0.444	0.361
S_{Doly}^{pk}	0.092	0.583	0.215		0.068	0.554	0.365
$S_{\text{Enerг}}^{pk} \& S_{\text{Doly}}^{pk}$	0.001	0.112	0.011		0.030	0.532	0.255

Однако, для вычисления значений энергии и долей энергии достаточно изображения фрагмента и количества частотных интервалов, что делает возможным адекватное сравнение фрагмента с прецедентом. Также, стоит отметить, что вычисление значений энергии

и вычисление значений долей энергии отличаются на одну операцию деления (26, 27). Это делает возможным одновременное использование мер близости в виде разности энергий и разности долей энергии без существенного увеличения вычислительной слож-

ности. Такой подход позволяет уменьшить среднее значение P_2 до 0.01, а размах до 0.11 для слова шум. Разность энергий/долей энергии показывает меньшие значения ошибок при вычислении значений во всех частотных интервалах.

ЗАКЛЮЧЕНИЕ

Таким образом, были разработаны, реализованы и исследованы различные меры для прецедентной идентификации фрагментов рукописного текста. Наибольшей инвариантностью к изменениям фрагментов рукописного текста обладают меры близости, основанные на субполосном подходе. Причем, характеристики точных значений энергии в заданном частотном интервале и доли общей энергии фрагмента в заданном частотном интервале дополняют друг друга с позиций качества идентификации – для коротких слов (менее 5 символов) меньшие значения ошибок показала S_{Energy}^{pk} , для длинных (5 и более символов) – S_{Doly}^{pk} , что иллюстрируется данными табл. 1, 2.

СПИСОК ЛИТЕРАТУРЫ

1. Мерков, А. Б. Основные методы, применяемые для распознавания рукописного текста / А. Б. Мерков. – Москва : Лаборатория распознавания образов МЦНМО, 2004.

Ефимов Н. О. – ассистент кафедры информационно-телекоммуникационных систем и технологий, институт инженерных и цифровых технологий, Белгородский государственный национальный исследовательский университет.
E-mail: n.o.efimov@gmail.com

2. Фу, К. Структурные методы в распознавании образов / К. Фу. – Москва : Мир, 1977. – 320 с.

3. Гонсалес, Р. Цифровая обработка изображений / Р. Гонсалес, Р. Вудс. – Москва : Техносфера, 2006. – 1072 с.

4. Семенов, Ю. А. Алгоритмы телекоммуникационных сетей / Ю. А. Семенов; в 3 частях. Часть 1. Алгоритмы и протоколы каналов и сетей передачи данных. – Москва : БИНОМ. Лаборатория знаний, 2007. – 640 с.

5. Черноморец, А. А. Субполосные модели и вариационные методы обработки изображений земной поверхности: дис... д-ра техн. наук. – Белгород : НИУ «БелГУ», 2017. – 402 с.

6. Жилияков, Е. Г. Вариационные алгоритмы анализа и обработки изображений на основе частотных представлений: монография / Е. Г. Жилияков, А. А. Черноморец. – Белгород: Изд-во ГИК, 2009. – 146 с.

7. Жилияков, Е. Г. О субполосном анализе изображений / Е. Г. Жилияков, Н. О. Ефимов // Научные ведомости БелГУ. Сер. Экономика. Информатика. – 2015. – 33/1. – С.118–124.

8. Жилияков, Е. Г. Распознавание фрагментов рукописного текста. Характерные частотные интервалы / Е. Г. Жилияков, Н. О. Ефимов // Информационные технологии. – 2018. – Т. 24, № 7. – С. 481–486.

Efimov N. O. – assistant, department of Information And Telecommunication Systems And Technologies, Institute of Engineering and Digital Technologies, Belgorod National Research University.
E-mail: n.o.efimov@gmail.com