

АЛГОРИТМ ОПРЕДЕЛЕНИЯ АВТОРСТВА АРАБОГРАФИЧЕСКОГО ТЕКСТА ПО СТАТИСТИЧЕСКИМ ДАННЫМ

ШАЛЫМОВ Р.С. | МАЙ 2015

КАФЕДРА СИСТЕМНОГО ПРОГРАММИРОВАНИЯ

344 ГРУППА

Научный руководитель
проф. О.Н. Граничин

ПОСТАНОВКА ЗАДАЧИ

2

Наши цели и ожидаемые результаты

Реализация компьютерной системы для
автоматического распознавания автора
арабографического текста



Устойчивость алгоритма к различным почеркам

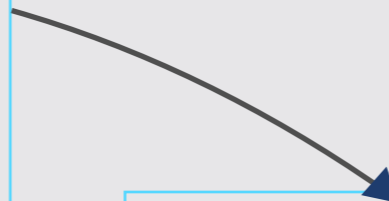
ОСНОВНЫЕ ЭТАПЫ

СЕГМЕНТАЦИЯ СТРАНИЦЫ
ДОКУМЕНТА НА СТРОКИ
(TEXT-LINE SEGMENTATION)

ИЗВЛЕЧЕНИЕ И ОБРАБОТКА
КОНТУРА СЛОВ В СТРОКЕ
(WORD CONTOUR
EXTRACTION)

КОДИРОВАНИЕ КОНТУРА
СПЕЦИАЛЬНОЙ
ПОСЛЕДОВАТЕЛЬНОСТЬЮ
СИМВОЛОВ
(CONTOUR HASHING)

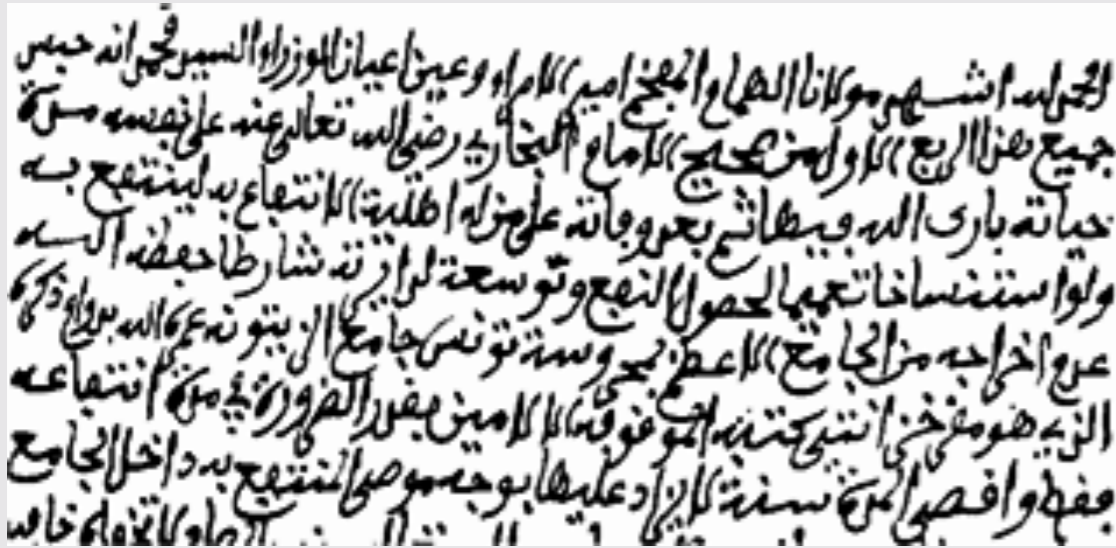
РАСЧЕТ КОРРЕЛЯЦИИ
МЕЖДУ ДВУМЯ
ГЕНЕРИРОВАННЫМИ
ПОСЛЕДОВАТЕЛЬНОСТЯМИ
(KOLMOGOROV-SMIRNOV
TEST)



TEXT-LINE SEGMENTATION

Этап 1: Классификация документа

Tightly Spaced Document (TSD)



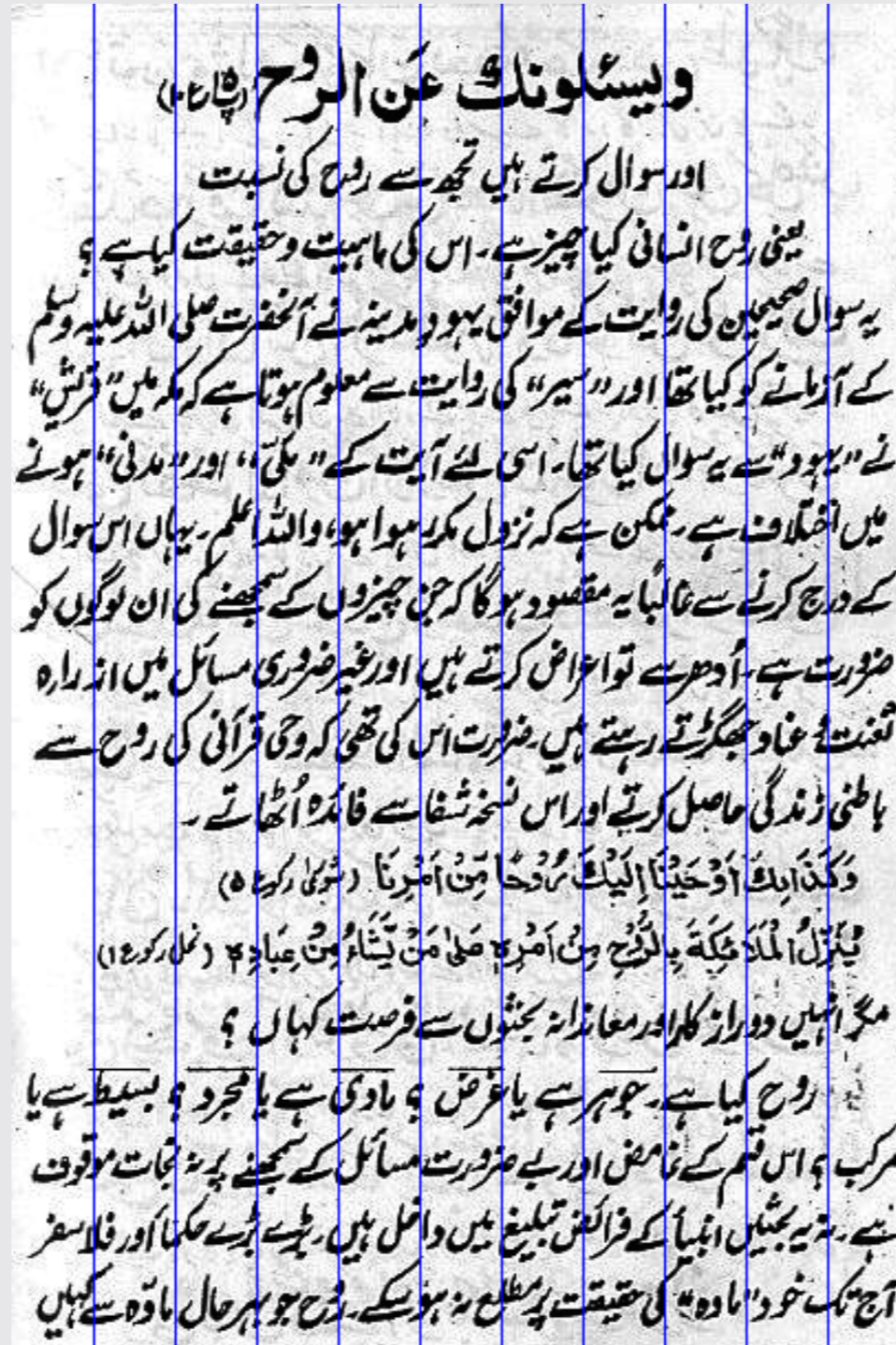
Widely Spaced Document (WSD)



TEXT-LINE SEGMENTATION

Этап 1: Классификация документа

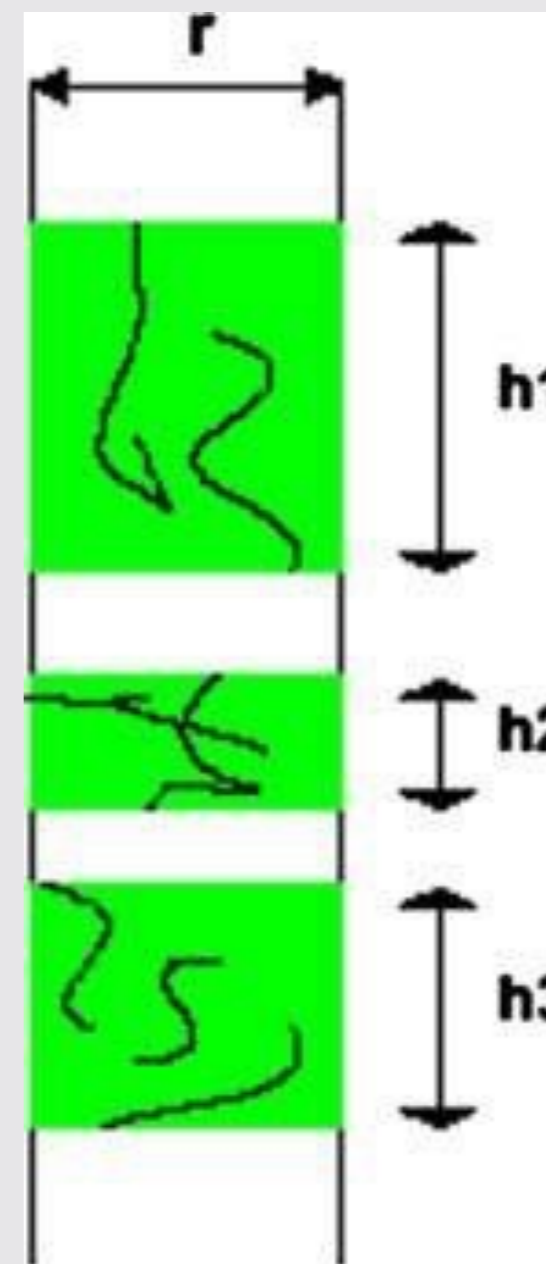
Разбиение документа на вертикальные полосы



TEXT-LINE SEGMENTATION

Этап 1: Классификация документа

Извлечение блоков текста из вертикальных полос



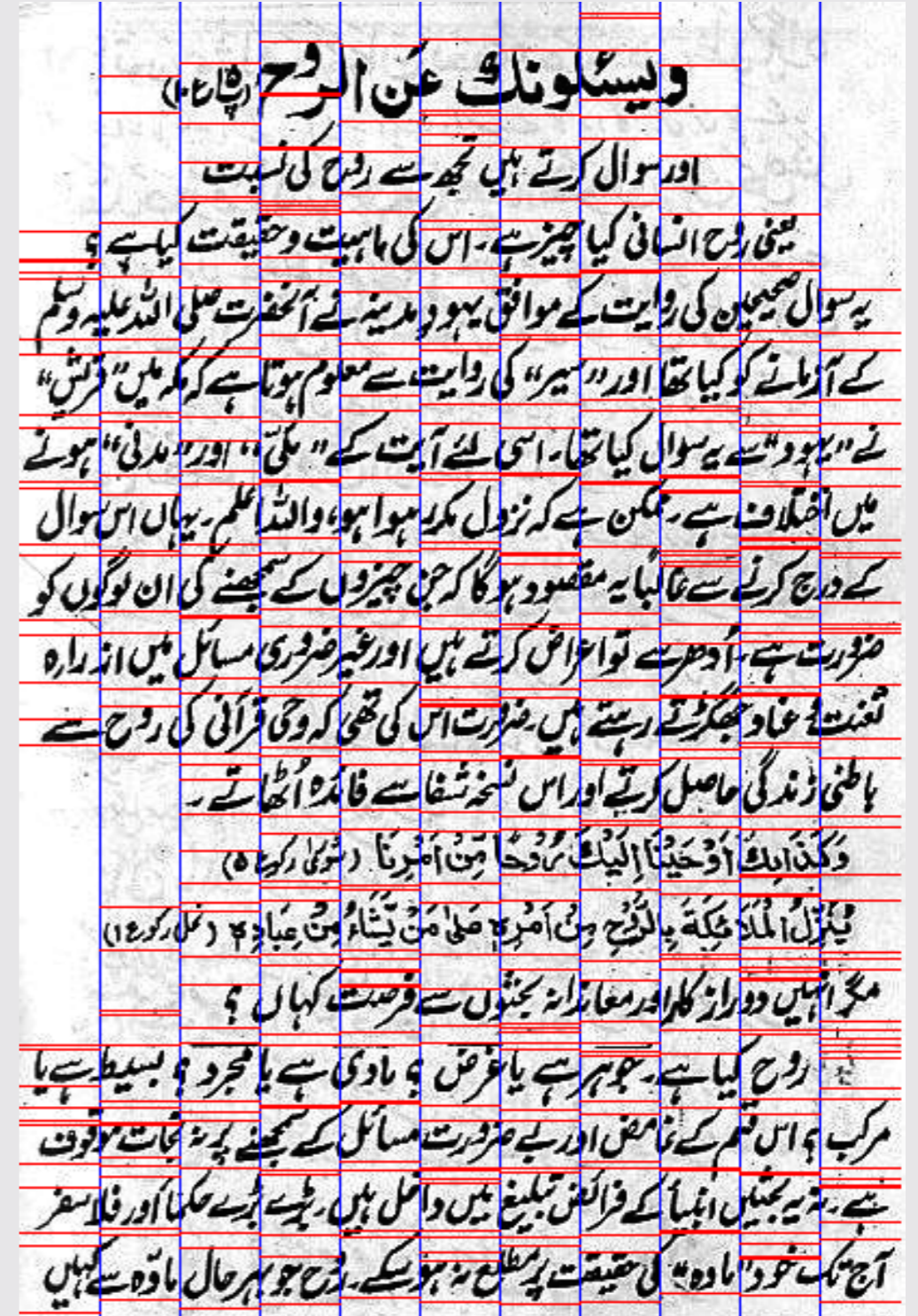
TEXT-LINE SEGMENTATION

Этап 1: Классификация документа

Хаусдорф определяет в d -мерном пространстве меру множества F как

$$H_d(F) = \lim_{\varepsilon \rightarrow 0} \inf_{r < \varepsilon} \left\{ \sum h(r) : C(r) \right\}$$

где $C(r)$ - любое конечное покрытие F объектами радиуса $r < \varepsilon$



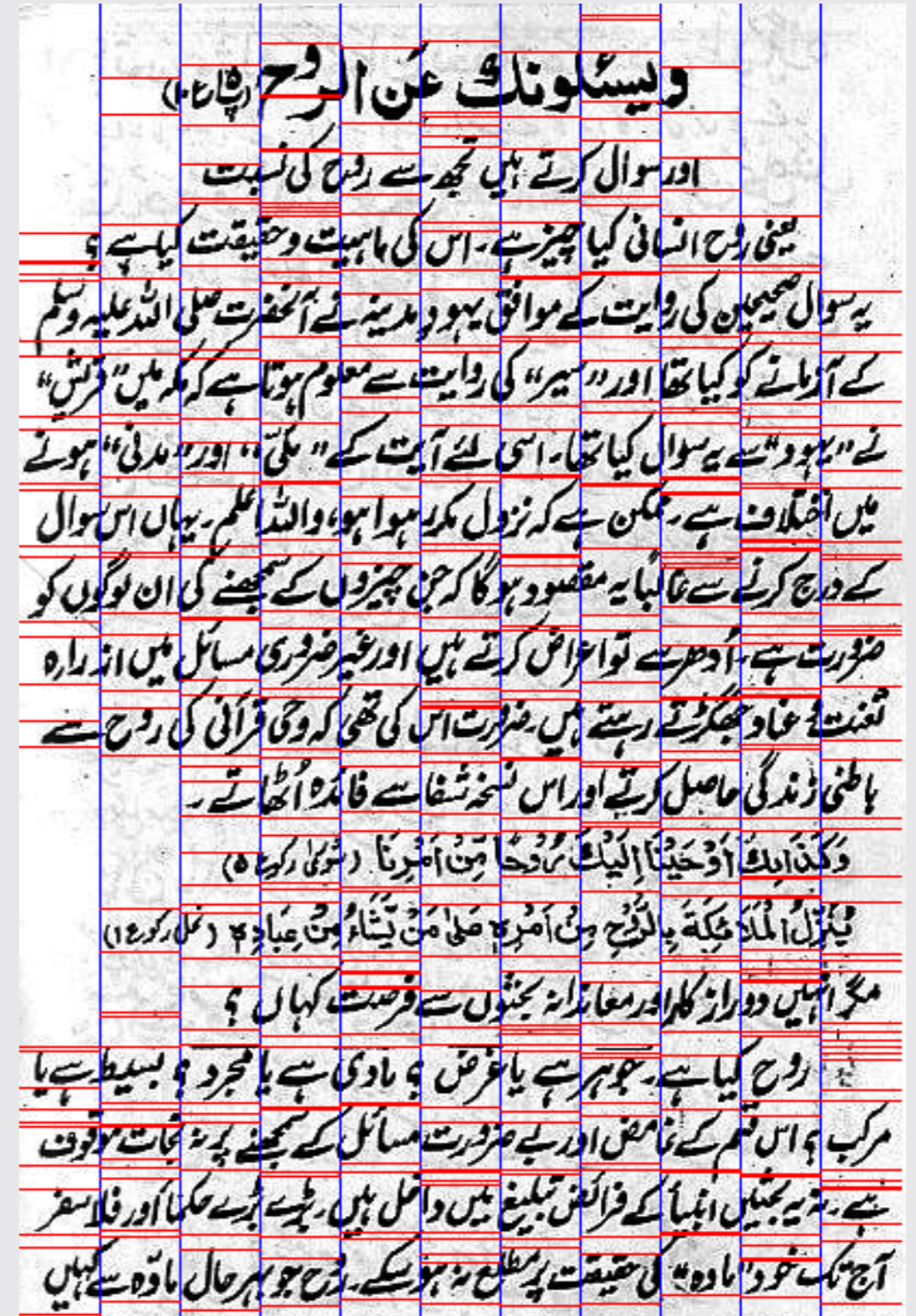
TEXT-LINE SEGMENTATION

Этап 1: Классификация документа

Хаусдорф показывает, что для любого множества F в d -мерном пространстве существует величина D_H (Хаусдорфова размерность) такая, что:

$$H_d(F) = \begin{cases} \infty & \text{if } d < D_H \\ 0 & \text{if } d > D_H \end{cases}$$

При $r \rightarrow 0$ $D \rightarrow D_H$



TEXT-LINE SEGMENTATION

Этап 1: Классификация документа

$$\log N_B = \log H_B + D * \log(1/r)$$

Выделяем вектор свойств

$$(\Delta \log H, D)$$

Пример

1

2

D

1.05

1.76

$\Delta \log H$

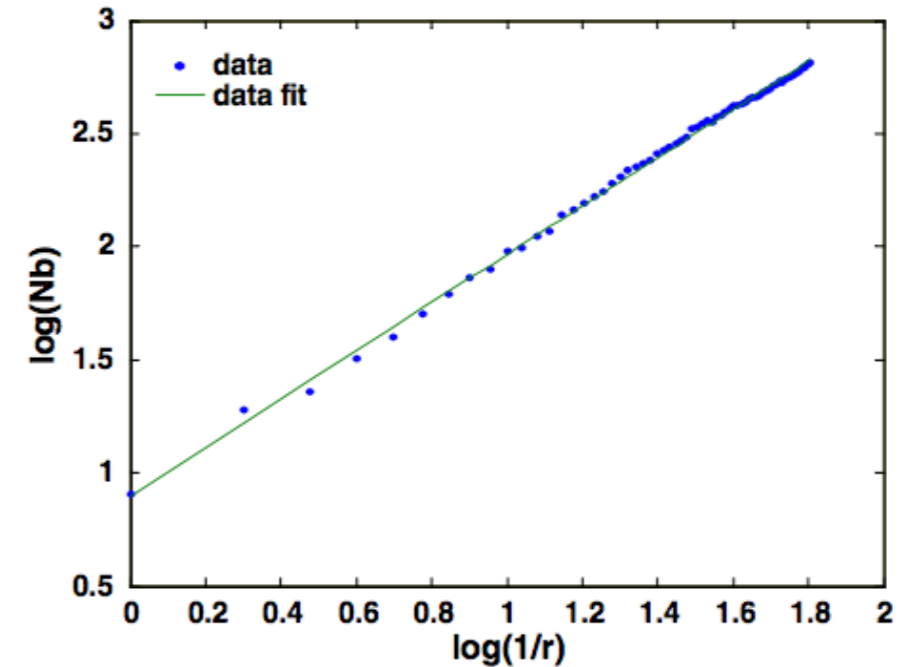
0.06

0.76

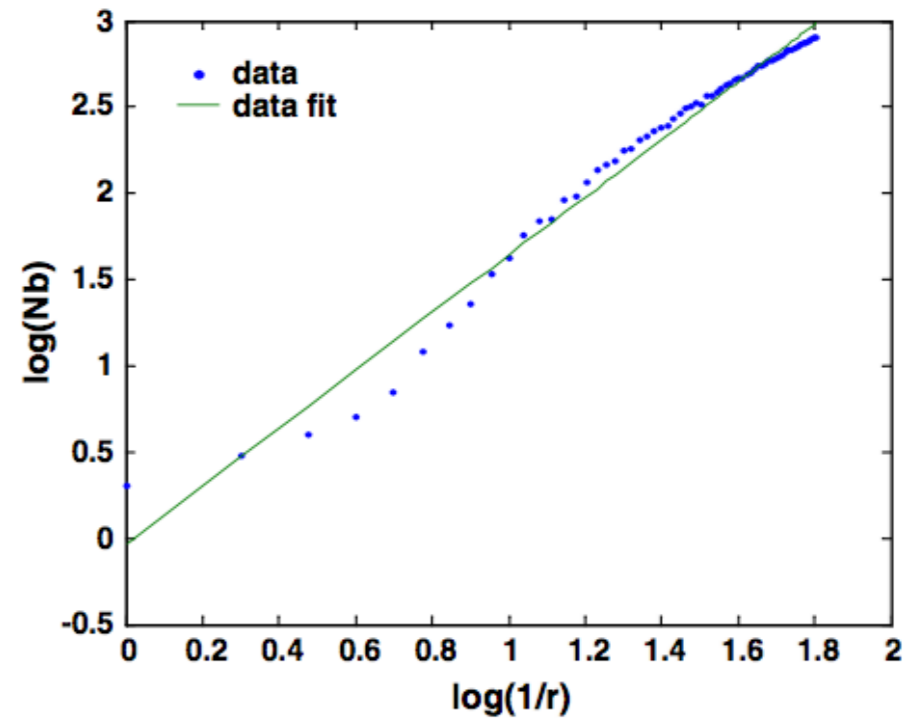
مورة شهيرة بالالعاب التي كانت تشتهر بها كل اربع سنين وتسمى الاولمبية
وهي قريبة من مدينة بيزوقدا خطأ بعض الجغرافيين حيث قال انهم ما بيننا
واحدة وهي مروية بنهر الالفه وبها جبل اولبيا ويحمل هذه المدينة الآن قريه
تسمى حراكا

اولثنة

مدينة موضوعة بالبحر بين جزيرة التي بين جوفى ترمابيق واسترمونيق وهو
جزء من مملكة مقدونيا وكانت فوق حرا تقع من الارض في اسفل خليج تريفنا ورو
وفي الجنوب الشرقي من ترمالونيق



الجملة انشدهم وكان الشراع المفلج امير الامراء وعين اعيان الوزراء السبعين انهم حين
جميع هذا الراج الكاوي من جميع الاماكن المتخارجه رضوا له تعاديه على نفسه مسرة
خدائته باري الدهر وبطانتهم بعروفاة على منزه اطلدته بالانتفاع به ليستمتع به
ولو استغنى ساخانتهما الحصول النفع وتوسعة لرايته شارب حافته السه
عنه اخرى من الجاهل اعطى نعيمه ونسمة تومر جامع ان يتونه عوالة بوليا ذكره
الذي هو من خشي ان يتركه اتمو غوفه الكا امين بفور الضرر والنجمة من انتفاعه
بفقه وافصحا من نسبة الكا اذ عليه ما يوجد هو صواب النفع به داخل الجامع



TEXT-LINE SEGMENTATION

11

Этап 2: Кластеризация (классификация) блоков

Fuzzy C Means



Tightly Spaced
Document (TSD)

Widely Spaced
Document (WSD)

- Маленькие блоки
- Средние блоки
- Крупные блоки

- Маленькие блоки
- Средние блоки

TEXT-LINE SEGMENTATION

12

Этап 3: Определение наилучшего количества вертикальных полос изображения

Вводим параметр оптимальной плотности между и внутри кластеров блоков ComposingDensity:

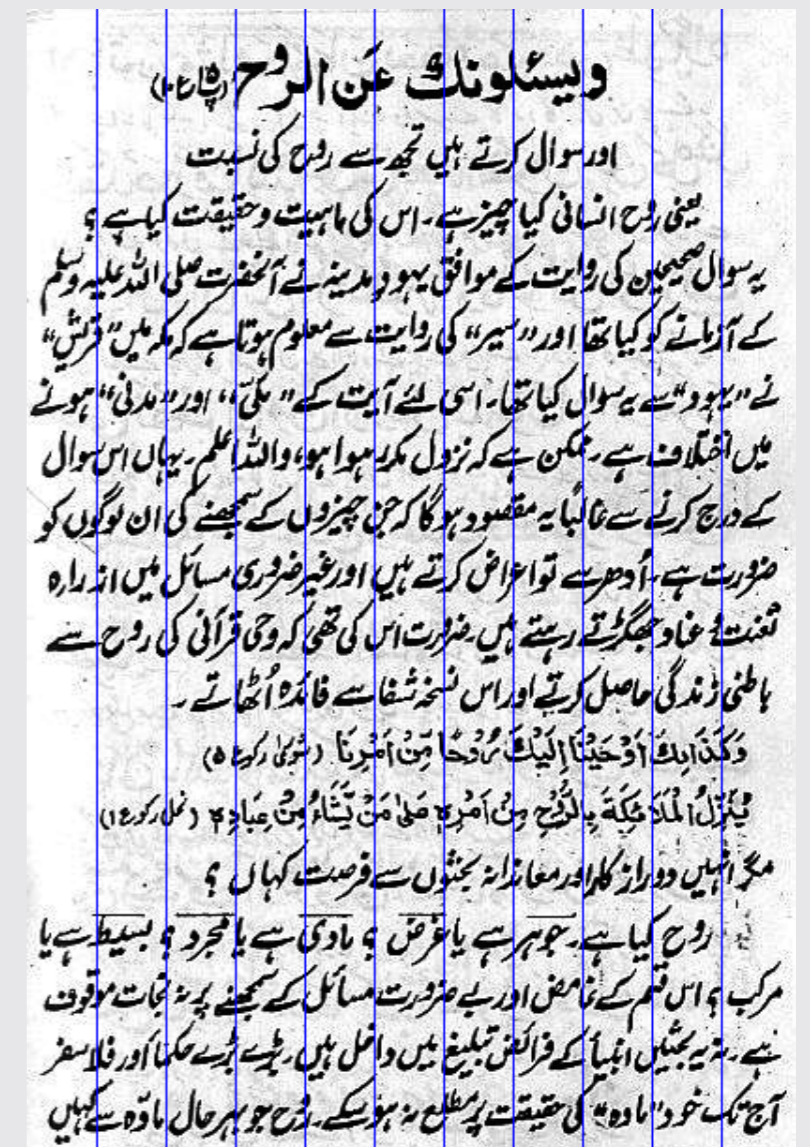
$$CDbw(vsn) = intra_den(vsn) * sep(vsn)$$

где *intra_den* - мера плотности блоков внутри кластеров

sep - мера разрозненности кластеров

vsn - количество вертикальных полос при текущем измерении

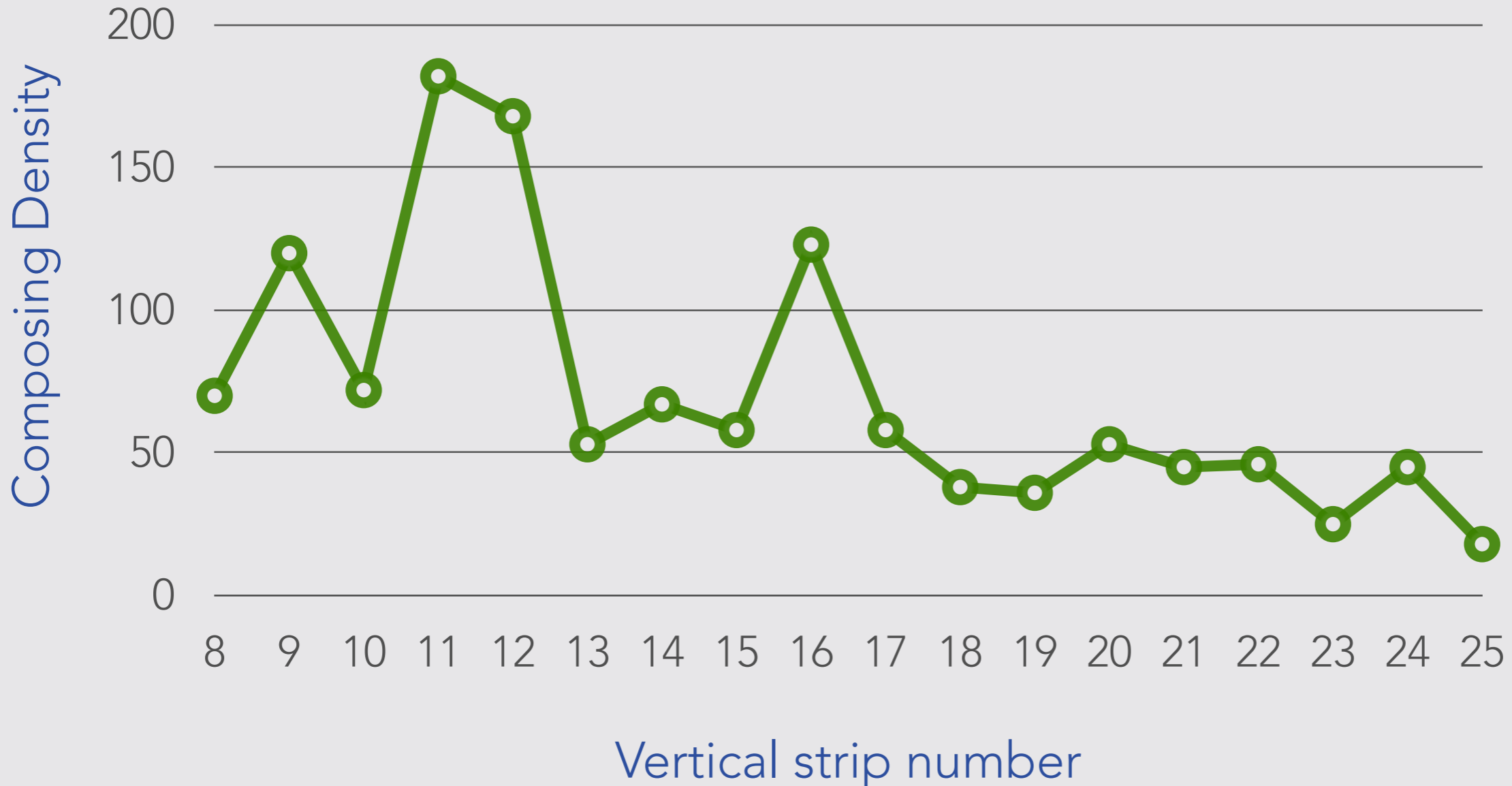
Строим зависимость *CDbw* от *vsn*



TEXT-LINE SEGMENTATION

Этап 3: Определение наилучшего количества вертикальных полос изображения

COMPOSING DENSITY PLOT



Этап 6: Построение текстовой кривой

رسائل ابن عربي

ولا نظرت عيني إلى حسن وجهها
ولا سمعت أذناي قط لها ذكرا
إلى أن تراءى البرق من جانب الحمى
فنعمني يوماً وعدبني دهرًا
وقلنا فيه أيضاً:

ولم أدر من هذا الذي قال: لا أدري
وقد حارت الحيرت في وفي أمري
أترجم عن حب يعانقه سبي
ولم أدر من هذا الذي ضمّه صدري
كمثل سحاب الليل أسفر عن بدر
بنيّة عين القلب بنت أخي الصدر
فليلي بها أربى على ليلة القدر
علقتُ بمن أهواه من حيث لا أدري
قد حلت في حالي وحالت خواطري
فبيننا أنا من بعد عشرين حجة
فلم أدر من أهوى ولا أعرف اسمه
إلى أن بدا لي وجهها من نقابها
فقلت لهم: من هذه؟ قيل: هذه
فكبرتُ إجلالاً لها ولأصلها

واختلف الناس في حدّه، فما رأيت أحداً حدّه بالحدّ الذاتي، بل لا يتصور ذلك، فما حدّه من حدّه إلا بنتائه وآثاره ولوازمه، ولا سيما وقد اتصف به الجذاب العزيز، وهو الله عزّ وجل، وأحسن ما سمعتُ فيه ما حدثنا غير واحد عن أبي العباس بن الصنهاجي رحمه الله تعالى، قالوا سمعناه يقول وقد شغل عن المحبة فقال: الغيرة من صفات المحبة، والغيرة تأتي إلا الستر، فلا تحدّ.

واللطف ما في الحب وجدته، وهو أن تجد عشقاً مفرطاً، وهوى وشوقاً مقلقاً وغراماً ونحولاً، وامتناع نوم، ولذّة طعام، ولا تدري فيمن، ولا بمن؟ ولا يتعين لك محبوبك، وهذا اللطف ما وجدته ذوقاً، ثم بعد ذلك بالاتفاق، أما يبدو لك تجلّي في كشف فيتعلق الحبّ به، أو ترى شخصاً فيتعلق ذلك الوجد تجده به عند رؤيته، فتعلم أن ذلك كان محبوبك [١ظ] وأنت لا تشعر، أو يُذكر الشخص فتجد الميل إليه بذلك الهوى، فتعلم أنه صاحبك، وهذا من أخفى دقائق استشراف النفوس على الأشياء من خلف حجاب الغير، فيجهل حالها، ولا تدري بمن هامت، ولا فيمن هامت وما هيّمتها؟ وتجد الناس في ذلك القبض والبسط الذي لا يعرف له سبب، فعند ذلك إما يأتيه ما يحزنه، فيعرف أن ذلك القبض كان لذلك الأمر، أو يأتيه ما يسره فيعرف أن ذلك البسط كان لهذا الأمر، وذلك لاستشراف النفوس على الأمور من قبل تكوينها في تعلق الحواس الظاهرة، وهي مقدمات التكوين، وينسب ذلك أخذ الميثاق على

رسائل ابن عربي

ولا نظرت عيني إلى حسن وجهها
ولا سمعت أذناي قط لها ذكرا
إلى أن تراءى البرق من جانب الحمى
فنعمني يوماً وعدبني دهرًا
وقلنا فيه أيضاً:

ولم أدر من هذا الذي قال: لا أدري
وقد حارت الحيرت في وفي أمري
أترجم عن حب يعانقه سبي
ولم أدر من هذا الذي ضمّه صدري
كمثل سحاب الليل أسفر عن بدر
بنيّة عين القلب بنت أخي الصدر
فليلي بها أربى على ليلة القدر
علقتُ بمن أهواه من حيث لا أدري
قد حلت في حالي وحالت خواطري
فبيننا أنا من بعد عشرين حجة
فلم أدر من أهوى ولا أعرف اسمه
إلى أن بدا لي وجهها من نقابها
فقلت لهم: من هذه؟ قيل: هذه
فكبرتُ إجلالاً لها ولأصلها

واختلف الناس في حدّه، فما رأيت أحداً حدّه بالحدّ الذاتي، بل لا يتصور ذلك، فما حدّه من حدّه إلا بنتائه وآثاره ولوازمه، ولا سيما وقد اتصف به الجذاب العزيز، وهو الله عزّ وجل، وأحسن ما سمعتُ فيه ما حدثنا غير واحد عن أبي العباس بن الصنهاجي رحمه الله تعالى، قالوا سمعناه يقول وقد شغل عن المحبة فقال: الغيرة من صفات المحبة، والغيرة تأتي إلا الستر، فلا تحدّ.

واللطف ما في الحب وجدته، وهو أن تجد عشقاً مفرطاً، وهوى وشوقاً مقلقاً وغراماً ونحولاً، وامتناع نوم، ولذّة طعام، ولا تدري فيمن، ولا بمن؟ ولا يتعين لك محبوبك، وهذا اللطف ما وجدته ذوقاً، ثم بعد ذلك بالاتفاق، أما يبدو لك تجلّي في كشف فيتعلق الحبّ به، أو ترى شخصاً فيتعلق ذلك الوجد تجده به عند رؤيته، فتعلم أن ذلك كان محبوبك [١ظ] وأنت لا تشعر، أو يُذكر الشخص فتجد الميل إليه بذلك الهوى، فتعلم أنه صاحبك، وهذا من أخفى دقائق استشراف النفوس على الأشياء من خلف حجاب الغير، فيجهل حالها، ولا تدري بمن هامت، ولا فيمن هامت وما هيّمتها؟ وتجد الناس في ذلك القبض والبسط الذي لا يعرف له سبب، فعند ذلك إما يأتيه ما يحزنه، فيعرف أن ذلك القبض كان لذلك الأمر، أو يأتيه ما يسره فيعرف أن ذلك البسط كان لهذا الأمر، وذلك لاستشراف النفوس على الأمور من قبل تكوينها في تعلق الحواس الظاهرة، وهي مقدمات التكوين، وينسب ذلك أخذ الميثاق على

Этап 6: Построение текстовой кривой

بما طه طبائها لاس الامراض مختلطه وجون معها اعراضه
عناها اسمها المجران قال ابرقظ ان في الحرفين يون
احدا جون وامل في الامراض فاما الرسم فاصح الاوقات موصاه اقطا
سونا في التفسير الاول ان فهم من قوله احدا ايردا وصلات
في الحرف ايردا وامل وفي الرسم بلصند منه لوز الرسم على عايه
ولوز الحرف محلقا لانه ساجب على الهواء في اليوم الواحد مسر
ولانه سلوا الحصف وقد تحسرو مع الاخلاط وصعقت الهوى
الاخلاط التي تات محل من قبل ميلها الى ناحية الجلد مصدها ردي الحرف



بما طه طبائها لاس الامراض مختلطه وجون معها اعراضه
عناها اسمها المجران قال ابرقظ ان في الحرفين يون
احدا جون وامل في الامراض فاما الرسم فاصح الاوقات موصاه اقطا
سونا في التفسير الاول ان فهم من قوله احدا ايردا وصلات
في الحرف ايردا وامل وفي الرسم بلصند منه لوز الرسم على عايه
ولوز الحرف محلقا لانه ساجب على الهواء في اليوم الواحد مسر
ولانه سلوا الحصف وقد تحسرو مع الاخلاط وصعقت الهوى
الاخلاط التي تات محل من قبل ميلها الى ناحية الجلد مصدها ردي الحرف

Этап 6: Построение текстовой кривой

بكر ولا عمر ولا أحداً رضي الله عنهم فإن أبا بكر من جملة أحواله كونه صديقاً وقد شاركه في هذا المقام غيره من الصديقين ولذلك قال تعالى: ﴿أُولَئِكَ هُمُ الصَّادِقُونَ﴾ [الحديد: ١٩] وقد فضل الصديق بسر وقر في صدره أعطاه الله إياه وشهد له به رسول الله ﷺ فعندنا بين الصديقية والرسالة مقام وهذا هو المقام الذي ذكرناه والذي أقول به أنه ليس بين أبي بكر رضي الله عنه وبين النبي ﷺ رجل ولا نذكر الصديقية فأرفع الأولياء أبو بكر رضي الله عنه فاجتهدوا رضي الله عنكم في تحصيله وأنا أنبهكم على العلامات التي تستدلون بها عليه، وذلك أنكم إذا قمتم بشرائط الخلوة كما ذكرناها في كتاب الخلوة ورفعت لكم أعلام المشاهد وقطعتموها وشاهدتم وعابنتم وأطلعتم ونزهتم ووقفتم المواقف المقدسة وقلبتم العوارف العرفانية فأنتم من أهل الولاية العظمى والدائرة المحيطة الكبرى لا تسلطوا في التحكم في العالم بالهمم أو بالصورة الظاهرة إن كانت لكم قوة سلطان أصلاً لعلو المقام الذي أنتم عليه فإن الله مستدرجكم فيه من حيث لا تعلمون وقد قال: ﴿وَمَنْ خَلَقْنَا أُمَّةً يَهْدُونَ بِالْحَقِّ وَبِهِ يَعْدِلُونَ﴾ [١٨١] وَالَّذِينَ كَذَّبُوا بِآيَاتِنَا سَنَسْتَدْرِجُهُمْ مِنْ حَيْثُ لَا يَعْلَمُونَ ﴿١٨٢﴾ وَأُمْلِي لَهُمْ إِنَّ كَيْدِي مَتِينٌ ﴿١٨٣﴾ [الأعراف: ١٨١ - ١٨٣] ولم يقل من الدنيا فقد يملي لكم من هذه الصنف فإنه سبحانه يملي أنه يملي لكل طائفة من حيث ما تشتهي وتتشق به واستوى في ذلك أبناء الدنيا وأبناء الآخرة والاستدراج والمكر لهذه الطائفة أسرع وأنفذ من غيرهم من الطوائف فالله لا تنفذوا حكماً ولا تتعدوا حداً من الحدود المعلومة عند أهل الرسوم وإن اختلفوا في ذلك وحرّم الواحد عين ما حلله الآخر فلا تتقلد هذا الرسمي في شيء من ذلك ولا تخالفه واعمل ما توجه عليك في وقتك مما فيه سلامتك واشتغل بنفسك شغلاً كلياً واهرب إلى محل إجماعهم فإن لم تجد إجماعاً فكن مع أكثرهم فإن لم تجد كثرة فكن مع أصحاب الحديث في تلك المسألة المطلوبة، وقل أن يحتاج أهل الطريق إلى مثل هذا لأنهم قد زهدوا في الدنيا فقللت أفعالهم فقل الحكم عليهم فإذا بدت لكم وفقكم الله حضرة الأحكام وتنزلات الشرائع ورأيتم خازنها جبريل عليه السلام فذلك أول أعلام تحصيل هذا المقام فإن مد بين يديك هذا اللوح الذي يتضمن الأحكام فستعابن الأوضاع والشرائع الحكمية والنبوية وستعابن الأعصار والأماكن وستعابن الأحوال وستعابن توجه هذه الأحكام على الأحوال لقيامها بالأشخاص فينفذ الحكم في الشخص للحال لا لعينه فاحفظ ما تراه.

واعلم أن جبريل لا ينزل على غير رسول يوحى أبداً ولا ينسخ شريعة فتعمل هناك في وسيلة ورقيقة تكون من ذلك اللوح إلى قلبك إن أردت تحصيل هذا المقام فتسجد صورة جبريل وما هي بجبريل وهي مختصة بالأولياء فانظر إليها فإن رأيته ناظرة إليك فاعلم أنك منهم وإن لم ترها ناظرة إليك فاعلم أنك غير مراد لذلك المقام فتأدب وانصرف وكن من

بكر ولا عمر ولا أحداً رضي الله عنهم فإن أبا بكر من جملة أحواله كونه صديقاً وقد شاركه في هذا المقام غيره من الصديقين ولذلك قال تعالى: ﴿أُولَئِكَ هُمُ الصَّادِقُونَ﴾ [الحديد: ١٩] وقد فضل الصديق بسر وقر في صدره أعطاه الله إياه وشهد له به رسول الله ﷺ فعندنا بين الصديقية والرسالة مقام وهذا هو المقام الذي ذكرناه والذي أقول به أنه ليس بين أبي بكر رضي الله عنه وبين النبي ﷺ رجل ولا نذكر الصديقية فأرفع الأولياء أبو بكر رضي الله عنه فاجتهدوا رضي الله عنكم في تحصيله وأنا أنبهكم على العلامات التي تستدلون بها عليه، وذلك أنكم إذا قمتم بشرائط الخلوة كما ذكرناها في كتاب الخلوة ورفعت لكم أعلام المشاهد وقطعتموها وشاهدتم وعابنتم وأطلعتم ونزهتم ووقفتم المواقف المقدسة وقلبتم العوارف العرفانية فأنتم من أهل الولاية العظمى والدائرة المحيطة الكبرى لا تسلطوا في التحكم في العالم بالهمم أو بالصورة الظاهرة إن كانت لكم قوة سلطان أصلاً لعلو المقام الذي أنتم عليه فإن الله مستدرجكم فيه من حيث لا تعلمون وقد قال: ﴿وَمَنْ خَلَقْنَا أُمَّةً يَهْدُونَ بِالْحَقِّ وَبِهِ يَعْدِلُونَ﴾ [١٨١] وَالَّذِينَ كَذَّبُوا بِآيَاتِنَا سَنَسْتَدْرِجُهُمْ مِنْ حَيْثُ لَا يَعْلَمُونَ ﴿١٨٢﴾ وَأُمْلِي لَهُمْ إِنَّ كَيْدِي مَتِينٌ ﴿١٨٣﴾ [الأعراف: ١٨١ - ١٨٣] ولم يقل من الدنيا فقد يملي لكم من هذه الصنف فإنه سبحانه يملي أنه يملي لكل طائفة من حيث ما تشتهي وتتشق به واستوى في ذلك أبناء الدنيا وأبناء الآخرة والاستدراج والمكر لهذه الطائفة أسرع وأنفذ من غيرهم من الطوائف فالله لا تنفذوا حكماً ولا تتعدوا حداً من الحدود المعلومة عند أهل الرسوم وإن اختلفوا في ذلك وحرّم الواحد عين ما حلله الآخر فلا تتقلد هذا الرسمي في شيء من ذلك ولا تخالفه واعمل ما توجه عليك في وقتك مما فيه سلامتك واشتغل بنفسك شغلاً كلياً واهرب إلى محل إجماعهم فإن لم تجد إجماعاً فكن مع أكثرهم فإن لم تجد كثرة فكن مع أصحاب الحديث في تلك المسألة المطلوبة، وقل أن يحتاج أهل الطريق إلى مثل هذا لأنهم قد زهدوا في الدنيا فقللت أفعالهم فقل الحكم عليهم فإذا بدت لكم وفقكم الله حضرة الأحكام وتنزلات الشرائع ورأيتم خازنها جبريل عليه السلام فذلك أول أعلام تحصيل هذا المقام فإن مد بين يديك هذا اللوح الذي يتضمن الأحكام فستعابن الأوضاع والشرائع الحكمية والنبوية وستعابن الأعصار والأماكن وستعابن الأحوال وستعابن توجه هذه الأحكام على الأحوال لقيامها بالأشخاص فينفذ الحكم في الشخص للحال لا لعينه فاحفظ ما تراه.

واعلم أن جبريل لا ينزل على غير رسول يوحى أبداً ولا ينسخ شريعة فتعمل هناك في وسيلة ورقيقة تكون من ذلك اللوح إلى قلبك إن أردت تحصيل هذا المقام فتسجد صورة جبريل وما هي بجبريل وهي مختصة بالأولياء فانظر إليها فإن رأيته ناظرة إليك فاعلم أنك منهم وإن لم ترها ناظرة إليك فاعلم أنك غير مراد لذلك المقام فتأدب وانصرف وكن من

WORD CONTOUR EXTRACTION

17

دوار اللواتة

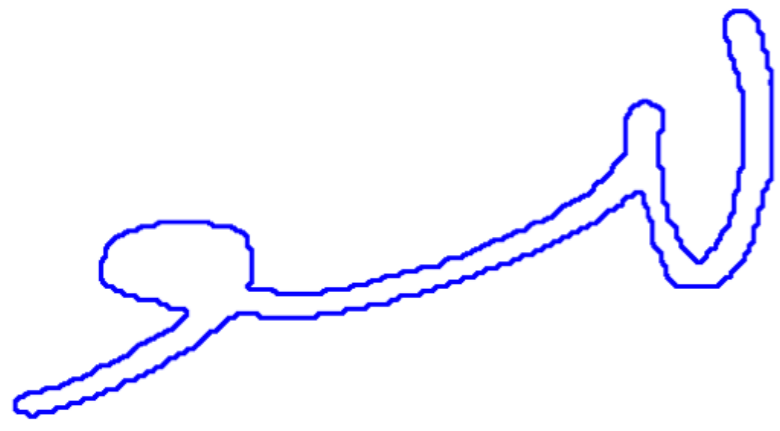
Увеличение
изображения

دوار اللواتة

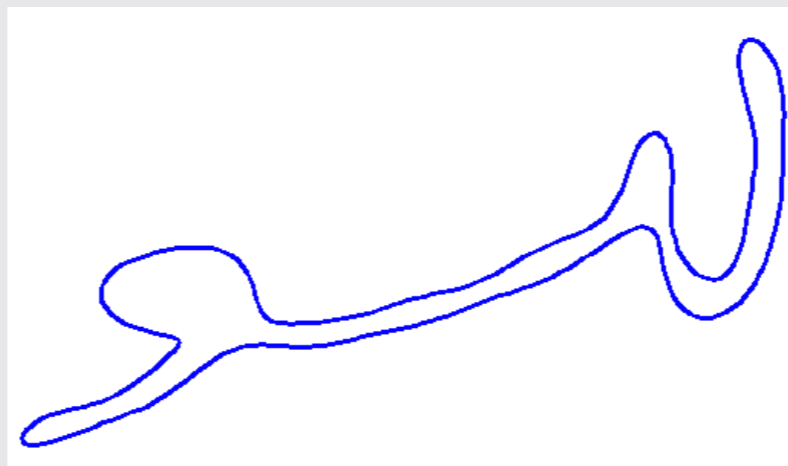
Image Dilation

دوار اللواتة

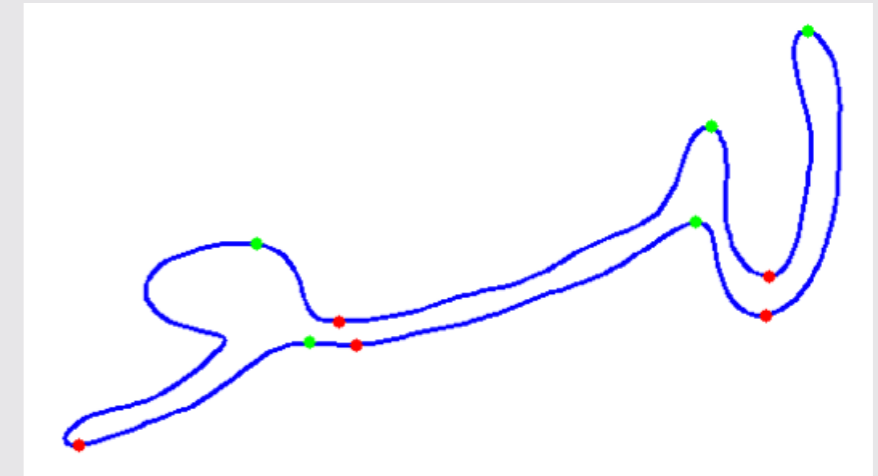
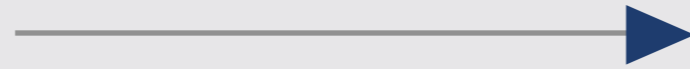
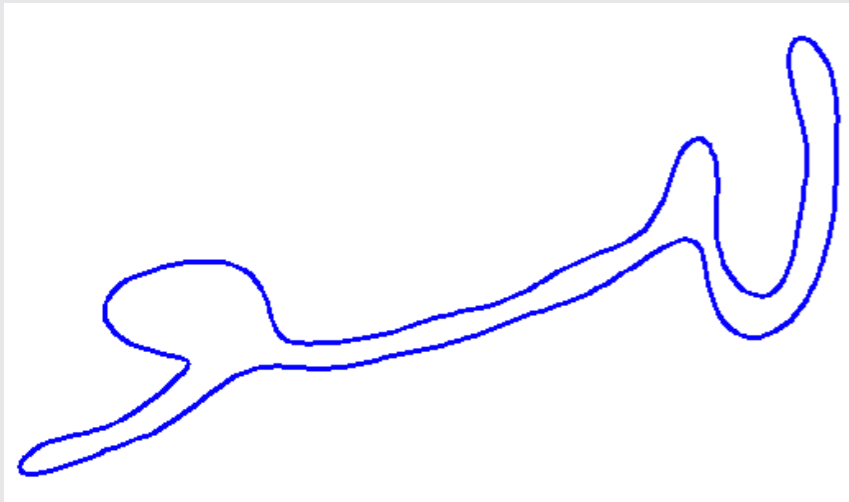
Contour
Extraction



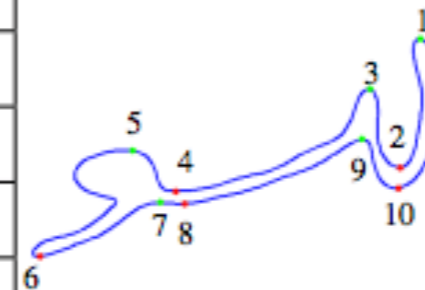
Contour
Smoothing



CONTOUR HASHING



№	type	x	x code	y	y code
1	'4'	0.916	'W'	-1	'a'
2	'3'	0.815	'U'	0.185	'D'
3	'4'	0.667	'Q'	-0.54	'k'
4	'3'	-0.293	'r'	0.403	'J'
5	'4'	-0.506	'l'	0.026	'A'
6	'1'	-0.964	'a'	1	'Z'
7	'2'	-0.369	'p'	0.502	'M'
8	'1'	-0.249	's'	0.517	'M'
9	'2'	0.627	'P'	-0.08	'w'
10	'1'	0.807	'T'	0.374	'I'



Word Hash: 4Wa3UD4Qk3rJ4lA1aZ2pM1sM2Pw1TI

DOCUMENT HASHING

Код строки равен:

$$H(s) = \prod_{k=1}^n H(w_k)$$

где

$$H(w_k)$$

- код k-ого слово в строке

وَيْسَلُونَكَ عَنِ الرُّوحِ قُرْآنًا
اور سوال کرتے ہیں تجھ سے روح کی نسبت
یعنی روح انسانی کیا چیز ہے۔ اس کی ماہیت و حقیقت کیا ہے؟
یہ سوال صحیحین کی روایت کے موافق یہود و مدینہ نے آنحضرت صلی اللہ علیہ وسلم
کے آزمائے کو کیا تھا اور "سیر" کی روایت سے معلوم ہوتا ہے کہ مکہ میں "قریش"
نے "یہود" سے یہ سوال کیا تھا۔ اسی لئے آیت کے "علی" اور "مدنی" ہونے
میں اختلاف ہے۔ ممکن ہے کہ نزول مکہ ہوا ہو، واللہ اعلم۔ یہاں اس سوال
کے درج کرنے سے غالباً یہ مقصود ہوگا کہ جن چیزوں کے سمجھنے کی ان لوگوں کو
ضرورت ہے۔ اُدھر سے تو اعراض کرتے ہیں اور غیر ضروری مسائل میں اذراہ
تغنت و عناء بھگوتے رہتے ہیں۔ ضرورت اس کی تھی کہ وحی قرآنی کی روح سے
باطنی زندگی حاصل کرتے اور اس نسخہ شفا سے فائدہ اٹھاتے۔
وَكَذَٰلِكَ أَوْحَيْنَا إِلَيْكَ رُوحَنَا مِن تَرْفَعِهَا (تو کیا رکھو؟)
يُنزِّلُ الْمَلَائِكَةَ بِالرُّوحِ مِنْ أَمْرِهِ صَلَّى مَنْ يَشَاءُ مِنْ عِبَادِهِ (نمل رکرو؟)
مگر ہمیں دو راز کلمہ اور معارفانہ بحثوں سے فرصت کہاں؟
روح کیا ہے۔ جو ہر ہے یا عرض؟ مادی ہے یا مجرد؟ بسیط ہے یا
مرکب؟ اس قسم کے نامض اور بے ضرورت مسائل کے سمجھنے پر نہ نجات موقوف
ہے۔ نہ یہ بحثیں انبیاء کے فرائض تبلیغ میں داخل ہیں۔ بڑے بڑے حکما اور فلاسفر
آج تک خود "مادہ" کی حقیقت پر مطلع نہ ہو سکے۔ روح جو بہر حال مادہ سے کہیں

وَيْسَلُونَكَ عَنِ الرُّوحِ قُرْآنًا

1sa2...1ef

+

اور سوال کرتے ہیں تجھ سے روح کی نسبت

3et4...2sR

+

یعنی روح انسانی کیا چیز ہے۔ اس کی ماہیت و حقیقت کیا ہے؟

4eY1...3ze

+

⋮

+

آج تک خود "مادہ" کی حقیقت پر مطلع نہ ہو سکے۔ روح جو بہر حال مادہ سے کہیں

3dN...2Fb

Гипотеза: Каждый автор является генератором псевдослучайный символов

Тест: KS-тест пытается определить являются ли два образца из одного источника с известным распределением

$$X = X_1, X_2, \dots, X_m$$

$$Y = Y_1, Y_2, \dots, Y_n$$

$$F(x) = \frac{1}{n} \sum_{i=1}^n I_{\{X_i \leq x\}}$$

$$D = \sup_t |F(t) - F_0(t)|$$

- Полностью реализован алгоритм сегментации рукописного документа на строки
- Полностью реализовано извлечение и обработка контуров слов в строке
- Получение хэш-кода документа

Реализация: C++, OpenCV