На правах рукописи

Птицын Николай Вадимович

Алгоритмический метод защиты и идентификации маркированных печатных документов

Специальность 05.13.17
Теоретические основы информатики

Автореферат диссертации на соискание
ученой степени кандидата технических наук

Москва — 2006

Диссертационная работа выполнена в Московском Государственном Техническом Университете имени Н. Э. Баумана на кафедре Системы Обработки Информации и Управления (ИУ5).

Научный руководитель:	профессор, доктор технических наук Черненький В. М.
Официальные оппоненты:	профессор, доктор технических наук Фролов А. Б., кандидат технических наук Крайлюк А. Д.,
Ведущая организация:	ЗАО Концерн «БИТ»

Защита с на заседании Ученого совета Д212.141.10 ехническом Университете имени Н. Э. Бау , ул. 2-я Бауманская, дом 5. С Н. Э. Баумана.

ванов С. Р.

Общая характеристика работы

Актуальность проблемы Документооборот, бумажный и электронный, является неотъемлемой частью человеческой деятельности. Не смотря на то, что в последние десятилетия электронный документооборот существенно повысил эффективность экономической деятельности, «бумажные технологии» остаются на сегодняшний день важным средством обмена информацией (документы, книги, этикетки), взаиморасчётов (наличные деньги, чеки, ценные бумаги, билеты, акцизные марки) и идентификации личности (паспорта, пропуска, удостоверения). Однако, случаи незаконного копирования пока имеют место достаточно часто. Этому способствует то, что средства копирования становятся всё качественнее и доступнее по цене. В первую очередь, здесь следует выделить цифровые студии на базе профессионального сканера, компьютера с графическим пакетами и цифрового устройства вывода.

Технологии защиты и проверки печатной продукции развиваются в направлении удорожания печатного процесса, использования уникального оборудования и специальных расходных материалов. Это делает защиту менее доступной широкому кругу потребителей: малого и среднего бизнеса, а так же индивидуальных пользователям. Здесь по-прежнему используется традиционные средства информационной безопасности — подписи и чернильные печати. Для указанного круга пользователей являются актуальным следующие проблемы:

Защита от копирования ставит перед собой задачу эффективного *распознавания* оригинальных документов и уличения подделок, т. е. несанкционированных копий. Документу придаются такие свойства, которые утрачиваются при воспроизведении на копирующим оборудовании. Важно, чтобы эти свойства не менялись в процессе использования документа и не могли повторяться злоумышленниками.

Контроль целостности подразумевает установление факта несанкционированного изменения содержания (например, изменение номинальный стоимости ценной бумаги, фамилии, имени).

Машинная идентификация, как правило, реализуется с помощью специальной маркировки, содержащей код. Примером является широко распространённый штрихкод, используемый для маркировки розничных товаров.

Аутентификация есть подтверждение факта того, что отправитель (автор документа), действительно является тем лицом, за кого он себя выдает.

Эти задачи могут быть решены при помощи так называемых *алгоритмических* методов защиты документов. В основе алгоритмических методов лежат вычислительные алгоритмы обработки изображения и криптографии. Эти методы противопоставляются *традиционным* методам защиты, которые опираются на неповторимость процесса печати.

Объектом исследования являются алгоритмические методы защиты от копирования и методы распознавания защитной маркировки на печатных документах.

Целями диссертации являются:

1) создание теоретической базы алгоритмической защиты документов;
2) разработка технологии алгоритмической защиты документов.

Методы исследования В работе применены методы цифровой обработки сигналов, теория вероятности и математическая статистика, теория детерминированного хаоса, фрактальная геометрия, криптография и теория кодирования.

Задачи Для достижения указанных целей в диссертации выполнены:

- анализ существующих методов защиты;
- выбор математического аппарата для описания преобразования изображения в процессе печати, сканирования и несанкционированного копирования злоумышленником;
- выявление статистических различий между цифровыми изображениями оригинального и скопированного документов;
- моделирование «прохождения» изображения через систему принтер-сканер;
- разработка генератора псевдо-случайного фрактального изображения на базе итерационных уравнений;
- разработка и тестирование алгоритмов:
  - генерации фракталов;
  - кодирования и декодирования данных (битового массива) в изображении;
  - вычисления статистических параметров изображения, чувствительных к копированию;
  - принятия решения о достоверности документа на основе рассчитанных параметров; оценка разброса значений параметров;
- разработка макета программного обеспечения, позволяющего произвести адаптацию метода;
- экспериментальная проверка разработанного метода.

Научную новизну составляют:

1) математическая модель преобразования изображения в системе «принтер-сканер»;
2) теоретическое обоснование метода защиты;
3) фрактал на базе кластера плоской кристаллической решётки;

4) схема кодирования информации в виде кругового штрихкода;
5) метод дискретного интегрального преобразования изображения;
6) метод непараметрической нечёткой классификации, имеющий приложение в задачах распознавания образов.

Практическую ценность представляют

1) технология защиты печатных документов путём нанесения и сканирования цветной маркировки;
2) метод установления факта подделки для проведения экспертиз;
3) алгоритм поиска ЦП на документе и прецизионного вычисления её координат на отсканированном изображении.
4) алгоритм обнаружения точного местонахождения цветной маркировки на документе;
5) алгоритм быстрого интегрального преобразования изображения;
6) алгоритм непараметрической нечёткой классификации.

Реализация результатов работы Разработанные методы, модели, алгоритмы, программное обеспечение использованы ООО «Диахром» (РФ) и «ДТЛ» (Великобритания) при разработке комплекса защиты печатных документов. По теме диссертации были сделаны доклады и опубликованы тезисы на конференциях (в том числе международной) [6, 1], подана заявка на патент [4], опубликованы статьи [3, 2]. В работе использованы некоторые идеи и результаты, опубликованные в английской диссертации [5].

Структура диссертации Диссертация состоит из введения, семи глав, описания алгоритмов, выводов и списка литературы. Объем диссертации составляет 175 страниц и 50 рисунков.

Краткое содержание

Введение

Во введении обоснована актуальность диссертационной работы; сформулированы цель и задачи, исследовательский метод, научная новизна, практическая ценность; представлено краткое содержание диссертации.

Глава 1

В первой главе предлагается классификация, краткое описание и анализ существующих методов защиты печатной продукции. Рассмотрены два основных подхода к защите печатных документов:

1. Традиционные методы имеют в основе сложность (неповторимость) технологического процесса;
2. Алгоритмические методы широко используют принципы компьютерной обработки изображения, криптографии и стеганографии.

Основным недостатком традиционных методов защиты является использование специального и, следовательно, дорогого оборудования и специальных расходных материалов (бумаги, чернил). Другим недостатком является то, что традиционные методы ориентированы, главным образом, на зрительное восприятие человеком.

В настоящее время становится приоритетным автоматическое распознавание с помощью различного рода устройств, таких как сканеры и детекторы. Автоматические системы на базе алгоритмического метода позволяет избежать субъективности, свойственную человеческому восприятию, повысить надёжность проверки, а главное сократить время подготовки и верификации документов. Однако существующие алгоритмические методы не лишены недостатков: пока еще низкая степень защиты и невозможность адаптации к оборудованию. Адаптивные методы защиты должны обеспечить: (а) расширенную совместимость с устройствами вывода; и (б) привязку к устройству вывода пользователя, так чтобы несанкционированная копия, изготовленная на другом печатном устройстве, идентифицировалась бы детектором в процессе верификации документа.

Условия применения разрабатываемой технологии состоят в следующим:

1) малый тираж (защите/проверке подлежат единичные экземпляры документов или небольшая серия);
2) типовое оборудование (эмитент документов и проверяющая сторона имеют стандартизированное оборудование, по которому настраивается программное обеспечение; при изменении моделей оборудования на той или другой стороне выполняется процедура переобучения);
3) доступность ключей (проверяющая сторона должна иметь доступ к секретным ключам и обучающей информации);

Глава 2

Глава 2 содержит математическое описание предметной области. Кратко изложены элементы обработки изображения. Устанавливается взаимосвязь между физическими принципами и предлагаемой моделью.

Изображение рассматривается как двумерный сигнал $f(x, y)$ или сообщение, передаваемое через канал связи. Функция $f(x, y)$ возвращает интенсивность изображения в точке (x, y). Интенсивность есть действительное значение в интервале $[0, 1]$. Канал связи представляет собой сложную систему, включающую устройство печати, среду обращения документа и устройство

сканирования, с помощью которого реализуется проверка:

$$f(x,y) \to \boxed{\text{Принтер}} \xrightarrow{\text{бумага}} \boxed{\text{Сканер}} \to s(x,y),$$

где $f(x,y)$ — исходное изображение (оригинал-макет документа), $s(x,y)$ — преобразованное изображение (сосканированный документ), такого же вида, что и $f(x,y)$. В случае несанкционированного копирования или изменения документа, изображение $s(x,y)$ содержит дополнительные искажения, которые должны идентифицироваться проверяющим устройством.

В качестве первого приближения математической модели взята линейная оптическая система

$$s(x,y) = f(x,y) \otimes\otimes k(x,y) + n(x,y), \tag{1}$$

где $\otimes\otimes$ — двумерная свёртка, $k(x,y)$ — ядро свёртки, характеризующее канал связи, и $n(x,y)$ — аддитивный шум. Модель (1) и дополняется нелинейным преобразованием интенсивности $q(\cdot)$:

$$s(x,y) = q[f(x,y)] \cdot f(x,y) \otimes\otimes k(x,y) + n(x,y),$$

где $q : [0,1] \mapsto [0,1]$. Приведены графики функции $q(\cdot)$. Предлагаемый метод защиты опирается на следующие положения:

Невозможность обратной свёртки при наличии шума Невозможность деконволюции является теоретическим и практическим препятствием, которое не позволяет злоумышленнику получить исходное цифровое изображение f из сканированного отпечатка s и напечатать идентичную копию документа. Основное уравнение (1), описывающее преобразование изображения в канале передаче, в общем случае (при $n \neq 0$) *не имеет стабильного решения* $\hat{f} \approx f$. То есть, небольшие возмущения в полученном изображении s приводят к существенным искажениям восстановленного изображения $\hat{f}$. Это перекликается с криптографическими методами защиты, в которых работает принцип *необратимости преобразования* из-за высокой чувствительности к начальным условиям. Другой проблемой злоумышленника является нахождение неизвестной функции k, характеризующей оборудование, на котором был изготовлен подлинный документ.

Конечная разрешающая способность Печать и сканирование изображения можно рассматривать как передачу сигнала через канал с ограниченным спектром пропускания. Теорема отсчётов Котельникова показывает, каким образом сигнал может быть точно представлен в виде последовательности значений. Применительно к исследуемой области, теорема определяет разрешающую способность системы принтер-сканер, которая обеспечила бы передачу изображения без потерь. В реальных системах, происходит ослабление или «обрезание» высоких частот.

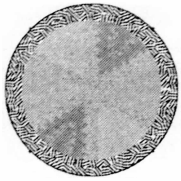

Рис. 1: Пример цифровой печати (черно-белая версия): (1) фрактальная основа (в центре) и (2) круговой штрихкод (вокруг)

Глава 3

Глава 3 описывает технологию алгоритмической защиты печатных документов. Технология обеспечивает защиту путём нанесения и считывания цветной маркировки. Эта маркировка называется *цифровой печатью* (ЦП) в смысле «цифрового» аналога гербовой печати. ЦП представляет собой сложное цветное изображение круглой формы. ЦП предназначена для защиты от копирования, идентификации, аутентификации и проверки целостности документа.

Пример ЦП представлен на рис. 1. В основе печати лежит симметричный *фрактал* круглой формы. Фрактал обладает достаточно сложной текстурой, которую не может идеально повторить злоумышленник. На фрактальную основу может быть наложен логотип или другая символика пользователя системы. Вокруг основы печати закодирована полезная информация в виде круглого штрихкода. ЦП внедряется в оригинал-макет документа, подлежащий защите.

1. **Создание фрактала и обучение**
   Пользователь создаёт уникальный фрактал, делает тестовый отпечаток ЦП и «обучает» программное обеспечение на сканере, который будет использоваться в дальнейшим при верификации. Производится несколько тестовых отпечатков, которые многократно сканируются. Система «запоминает» *контрольные характеристики*, которые жёстко привязаны к конкретной основе ЦП и оборудованию (принтер-сканер).
2. **Создание ЦП**
   В момент «подписания» документа создается ЦП. При этом берется готовая основа ЦП (уникальная для данного пользователя), и формируются данные о документе (автор, дата, контрольная сумма, серийный номер и т.д.). Эти данные шифруются и кодируются в виде штрихкода вокруг основы (рис. 1). Затем ЦП печатается вместе с документом.
3. **Верификация ЦП**
   Для проверки документа, производиться сканирование документа и обработка полученного изображения: декодируются и дешифруются данные из штрихкода, статистически сравнивается характеристики распо-

знаваемого и эталонных документов. Пользователю выдаётся сообщение, содержащее: (a) вероятность подделки (b) декодированную информацию. Если степень близости считается удовлетворительной, и документ соответствует декодированной информации, то экземпляр признается оригиналом.

Контрольные характеристики Канал связи (принтер-сканер) описывается так называемыми контрольными характеристиками или признаками, представляющими собой последовательность действительных значений. В диссертации обоснован выбор характеристик и их адекватность модели канала связи. Показано, что подлинная и скопированная (поддельная) ЦП обладают достаточно различимыми характеристиками. С другой стороны, серия подлинных ЦП обладает близкими характеристиками.

Характеристики должны быть инвариантными к таким факторам как погрешность позиционирования (вычисления координат ЦП), износ оборудования, старения расходных материалов (бумаги, чернил), воздействие внешней среды. В процессе обучения определяются те допуски (коридор допустимых значений), в которых может измениться характеристика. Расширение коридора приводит к понижению степени защиты.

Пусть $f_{m,n}$ — цифровое изображение размера $M \times N$ и $S \in \{(m,n) \mid 0 \leq m < M, \quad 0 \leq n < N\}$ — множество точек ЦП. В диссертации предложены три типа характеристик:

1. Гистограммы интенсивностей (отдельно по каждому каналу):

$$\begin{aligned} H_i &= \frac{1}{|S|} \sum_{(m,n)\in S} \sum h_i(f_{m,n}), \quad i = 0, 1, \ldots, 255, \\ h_i(x) &= \begin{cases} 1, & x = i, \\ 0, & x \neq i; \end{cases} \end{aligned}$$

2. Спектр Фурье, усреднённый по радиусу (отдельное по каждому каналу):

$$\begin{aligned} \langle F_r \rangle &= \frac{1}{2\pi r} \sum_{u^2+v^2=r^2} |F_{u,v}|, \quad r = 1, 2, ..., \min(M, N), \\ F_{u,v} &= \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f_{m,n} \exp(-i2\pi(um/M + vn/N). \end{aligned} \tag{2}$$

3. Характеристика паттернов на базе локальной лакунарности Λ_k степени

$k = 2$ (объединяет три цветовых канала r, g и b):

$$\Lambda_k(m,n) = \left\langle \left(\frac{|\mathbf{f}_{m,n} - \langle \mathbf{f}_{m,n} \rangle_{S'}|}{|\langle \mathbf{f}_{m,n} \rangle_{S'}|} \right)^k \right\rangle_S^{\frac{1}{k}},$$

где

$$\begin{aligned} \mathbf{f}_{m,n} &= (r_{m,n}, g_{m,n}, b_{m,n}), \\ |\mathbf{f}_{m,n}| &= \sqrt{R^2 r_{m,n}^2 + G^2 g_{m,n}^2 + B^2 b_{m,n}^2}, \\ |\mathbf{f}_{m,n} - \mathbf{f}'_{m,n}| &= \sqrt{R^2 (r_{m,n} - r'_{m,n})^2 + G^2 (g_{m,n} - g'_{m,n})^2 + B^2 (b_{m,n} - b'_{m,n})^2}, \end{aligned}$$

и весовые коэффициенты $R + G + B = 1$. Пусть S — множество точек ЦП. Среднее значение $\langle \mathbf{f}_{m,n} \rangle_{S'}$ рассчитывается в некоторой небольшой окрестности точки (m,n) одним из следующих способов:

- Окно «диск» радиуса R (быстрый, но менее точный метод)

$$\begin{aligned} \langle \mathbf{f}_{m,n} \rangle &= \frac{1}{|S'|} \sum_{(m',n') \in S'} \mathbf{f}_{m,n}, \\ S'(m,n) &= \left\{ (m',n') \in S \mid (m'-m)^2 + (n'-n)^2 \le R \right\}; \end{aligned}$$

- Гауссовское окно (более точный, но ресурсоёмкий метод)

$$\mathbf{f}_{m,n} = \frac{1}{\sigma\sqrt{2\pi}} \sum_{(m',n') \in S} \mathbf{f}_{m,n} \exp\left(-\frac{(m'-m)^2 + (n'-n)^2}{2\sigma^2} \right).$$

Лакунарность (англ. lacunarity, лат. lacuna — впадина, яма) — понятие, введённое Б. Мандельбротом во фрактальной геометрии. Вместе с фрактальной размерностью, лакунарность определяет шероховатость, неровность сигнала. В качестве контрольной характеристики преложена *гистограмма* лакунарностей $\Lambda_k(m,n)$, рассчитанная по всей площади ЦП:

$$\begin{aligned} H(i) &= \sum_{(m,n) \in S} h_i\big(\Lambda_k(m,n)\big), \quad i = 0, 1, 2 \ldots, I-1; \\ h_i(\Lambda) &= \begin{cases} 1, & i \le I \frac{\Lambda - \Lambda_{\min}}{\Lambda_{\max} - \Lambda_{\min}} < (i+1), \\ 0, & \text{в противном случае;} \end{cases} \\ \Lambda_{\min} &= \min_{(m,n) \in S} \Lambda_k(m,n); \\ \Lambda_{\max} &= \max_{(m,n) \in S} \Lambda_k(m,n); \end{aligned}$$

Контрольная характеристика на базе локальной лакунарности имеет преимущества при распознавании защитной маркировки:

- Суммируются *относительные* отклонения $|\mathbf{f}_{m,n} - \langle \mathbf{f}_{m,n} \rangle| \,/\, |\langle \mathbf{f}_{m,n} \rangle|$, поэтому характеристика не зависит от общего уровня яркости освещения. Если изображение предварительно инвертировать, так чтобы минимальное значение f соответствовало максимальной интенсивности, то характеристика будет инвариантна к выцветанию чернил.
- Настройка параметров преобразования (степень k, размер окна R, σ) позволяют получить максимальную чувствительность к интересующим свойствам изображения.

Принятие решение о подлинности документа Пусть $\mathbf{x} = (x_1, \ldots, x_K)$ — характеристика распознаваемого документа и

$$\mathbf{y}_n = (y_{n,1}, \ldots, y_{n,K}), \quad 1 \le n \le N,$$

— контрольные характеристики, где N — число уроков. Рассмотрим значения

$$y_{1,k}, y_{2,k}, \ldots y_{N,k}$$

для некоторого фиксированного $k = 1, 2, \ldots, K$. Будем считать эти значения выборкой случайной величины, распределённой по некоторому вероятностному закону. Случайность этих значений обуславливается следующими свойствами сканера:

- неидеальность механической части;
- ошибки при дискретизации уровней сигнала;
- шум светочувствительных элементов.

Метод 1: Отклонение от коридора Определим «коридор» допустимых значений:

$$\begin{aligned} y_k^{\min} &= \min_{1 \le n \le N}, \quad 1 \le k \le K, \\ y_k^{\max} &= \max_{1 \le n \le N} y_{n,k}, \quad 1 \le k \le K, \end{aligned}$$

Рассмотрим среднеквадратическое отклонение характеристики от контрольного коридора:

$$D = \frac{1}{K} \sum_{k=1}^{K} d^2(x_k),$$

где

$$d(x_k) = \begin{cases} y_k^{\min} - x_k, & x_k < y_k^{\min}, \\ 0, & y_k^{\min} \le x_k \le y_k^{\max} \\ x_k - y_k^{\max}, & x_k > y_k^{\max}. \end{cases}$$

Тогда, решение принимается согласно правилу:

$$\text{решение} = \begin{cases} \text{подлинный}, & D < D_{\text{кр}}, \\ \text{поддельный}, & \text{в противном случае}, \end{cases}$$

где $D_{\text{кр}}$ — максимально допустимое отклонение.

Метод 2: Проверка гипотез по статистике χ^2 . Будем считать, что источников отклонений много, и они независимы. Тогда последовательность $y_{n,k}$ для фиксированного k есть выборка случайной величины, распределённой по нормальному закону. Оценим параметры распределения (соответственно, математического ожидания и дисперсии):

$$\begin{aligned} \hat{\mu}_k &= \frac{1}{N}\sum_{n=1}^{N} y_{n,k}, \quad 1 \le k \le K, \\ \hat{\sigma}_k &= \frac{1}{N-1}\sum_{n=1}^{N} \left(y_{n,k} - \hat{\mu}_k\right)^2, \quad 1 \le k \le K. \end{aligned}$$

Рассмотрим гипотезы

- H_0: документ подлинный (характеристики близки);
- H_a: документ поддельный (характеристики сильно различаются);

Статистика

$$\hat{\chi}^2 = \sum_{k=1}^{K} \frac{\left(x_k - \hat{\mu}_k\right)^2}{\hat{\sigma}_k^2}, \tag{3}$$

имеет распределение χ^2 с $(K-1)$ степенями свободы. Значения математического ожидания и дисперсии заменены соответственно оценками $\hat{\mu}_k$ и $\hat{\sigma}_k$. Величина $P(\hat{\chi}^2 \le \chi^2)$ есть вероятность того, что $\hat{\chi}^2$ — сумма квадратов случайных величин с распределением $N(0,1)$ — будет меньше χ^2. Тогда

$$\text{решение} = \begin{cases} \text{подлинный } (H_0), & P(\hat{\chi}^2 \le \chi^2) < \alpha, \\ \text{поддельный } (H_a), & \text{в противном случае}. \end{cases}$$

Вероятность гипотезы H_0 может быть рассчитана как

$$P(\hat{\chi}^2 \le \chi^2) = \int_{\hat{\chi}^2}^{\infty} f(\chi^2)\, d\chi^2 = \Gamma\left(\frac{K-1}{2}, \frac{1}{2}\hat{\chi}^2\right),$$

где $\Gamma(a,x)$ — нормированая неполная гамма-функция:

$$\Gamma(a,x) \equiv \frac{1}{\Gamma(a)}\int_0^x e^{-t}t^{a-1}\, dt, \quad \Gamma(a) \equiv \int_0^{\infty} t^{a-1}e^{-t}\, dt.$$

При большом числе степеней свободы ($K > 30$) распределение χ^2 близко к нормальному. Уровень значимости α выбирается экспериментально, исходя из оптимального соотношения ошибок первого и второго рода.

Определение положения ЦП на документе Для вычисления контрольных характеристик ЦП необходимо иметь точные координаты ЦП. В случае неправильного позиционирования возникает существенное отклонение, приводящая к ошибочному распознаванию подлинного документа. Координаты ЦП варьируются в небольшом диапазоне из-за погрешности позиционирования документа на сканере. Разработан алгоритм двухэтапного поиска ЦП, включающий грубое (предварительное) и прецизионное позиционирование:

Грубое позиционирование использует метод наименьших квадратов для поиска центра ЦП. В каждой точке изображения вычисляется локальная гистограмма с помощью разностного алгоритма [2]. Минимальное значение среднеквадратической разности локальной гистограммы и контрольной характеристики соответствует искомым координатам ЦП.

Прецизионное позиционирование уточняет значения, полученные на предыдущем шаге, а так же определяет угол наклона ЦП. Реализовано с помощью корреляции паттерном кругого штрихкода.

Кодирование данных реализовано при помощи кругового штрихкода. Штрикод состоит из паттернов. Каждый паттерн кодирует $q = 4$ бита информации. Изображение одного паттерна задано функцией

$$f_{m,n} = \sin\left[\frac{\pi}{l}\Big(m\sin(\phi) + n\cos(\phi)\Big)\right],$$

$$\phi = \phi_0 + k\beta, \quad \beta = \frac{\pi}{K}, \quad k = \{0, 1, \ldots, K-1\}$$

где (m, n) — индексы пикселов, $K = 2^q$ — число состояний паттерна, ϕ_0 — угол, определяющий положение центра паттерна на кольце, l — расстояние между штрихами. Штрихкод содержит две копии блока данных. При обнаружении ошибки (соответствующие элементы двух копий не совпадают), выбирается элемент с максимальной энергией, вычисляемой при корреляции паттернов.

Глава 4

В главе 4 существующие методы непараметрической классификации — k-NN и окно Парзена — расширены для приложений с нечётким обучающим набором. Введение нечёткости в обучающий набор контрольных характеристик позволяет не только задавать классификацию в терминах теории вероятности

или мягких множеств, но и учитывать такие факторы как точность измерений.

Получен гибридный метод, объединяющий подходы k-NN и Парзена. Гибридный метод повышает точность распознавания образа за счёт адаптивной подстройки оконного параметра Парзена в зависимости от плотности и веса элементов обучающего набора в окрестности анализируемой точки.

Глава 5

В главе 5 предложен метод интегрального преобразования, основанный на вычислении разности между соседними суммами. Разностный метод преобразования изображения эффективен, когда необходимо выполнить точное суммирование внутри ограниченной области, заданной маской или другим способом. То есть, приложение этого метода оправдано для бинарных ядер простой формы (круг, кольцо, лепестки и пр.), и не целесообразно для неоднородных ядер (вейвлеты, гауссовская функция).

Разностный метод интегрального преобразования использован для определения координат ЦП (позиционирования) и вычисления характеристики паттернов на базе лакунарности.

Глава 6

В глава 6 обсуждается приложение фрактальной геометрии для создания защитной маркировки. Разработан генератор фракталов, строящий уникальные (неповторяющиеся) фракталы с достаточной сложной текстурой. Использовано семейство фракталов на базе кластера плоской кристаллической решётки[1]. Фракталы получены путём применения итерационного метода Ньютона

$$z \mapsto z - \frac{f(z)}{f'(z)}, \quad z = x + iy,$$

к точкам решётки на комплексной плоскости. Например, для многочлена седьмой степени

$$f(z) = z(z^6 - 1)$$

имеем итерации

$$z \mapsto \frac{6z^7}{7z^6 - 1},$$

которые порождают фрактал с элементами бассейна Ньютона.

Рассмотрены способы раскраски фрактала:

[1]Семейство итерационных уравнений предложено д. ф. н. П. Н. Антонюком (МГТУ им. Баумана, факультет ФН).

1. Тон (цвет) точка определяется номером аттрактора (корня комплексного многочлена), к которому притягивается эта точка.
2. Интенсивность точки определяется минимальным числом итераций, необходимым для достижения заданной окрестности аттрактора.
3. Интенсивность точки определяется плотностью траекторий динамической (итерационной) системы в этой точке.
4. Цвет/интенсивность резко меняются в зависимости от чётности/нечётности счетчика итераций («зебра»).

В итерационное выражение включается возмущающая функция:

$$z \mapsto z - g\left(\frac{f(z)}{f'(z)}, \gamma\right),$$

такая что

$$\lim_{z \to 0} \frac{g(z,\gamma)}{z} = 1,$$

что обеспечивает сходимость итераций к корням многочлена. Возмущающая функция имеет параметр γ, изменяемый псевдо-случайным образом в процессе построения фрактала. Это позволяет получать стахостический фрактал с уникальной текстурой, которая использована в качестве защитной маркировки. Примерами возмущающих функций такого рода являются $g(z,\gamma) = \sin(\gamma z)$ и $g(z,\gamma) = z/(1+\gamma|z|)$.

Глава 7

Глава 7 посвящена апробации разработанной технологии. Описаны процесс испытания (в том числе методы изготовления поддельной копии), использованное оборудования и статистические результаты. В следующих таблицах приведены количества ошибок после обучения системы на 5 эталонных документах, отсканированных по 4 раз каждый.

Уровень значимости $\alpha = 0.1$

Документ	**Всего** M	**Распознано** M_r	**Пропущено** $M_e = M - M_r$	**Ошибка** M_r/M
Подлинный (хорошее качество)	40	40	0	0%
Поддельный	20	20	0	0%
Подлинный (низкое качество)	10	7	3	30%

Уровень значимости $\alpha = 0.05$

Документ	Всего M	Распознано M_r	Пропущено $M_e = M - M_r$	Ошибка M_r/M
Подлинный (хорошее качество)	40	40	0	0%
Поддельный	20	20	0	0%
Подлинный (низкое качество)	10	3	7	70%

Приложение А

В приложении А представлены алгоритмы, реализующие технологию алгоритмической защиты: (1) генерация фрактала; (2) создание ЦП; (3) обучение и (4) верификация.

Выводы

Разработана технология алгоритмической защиты печатных документов. Адекватность технологи обоснована при помощи математической модели и проведённых экспериментов. Перспективным приложением разработанной технологией является область проведения экспертиз ценных бумаг, деловых документов, страховых полисов, контрактов гарантийного обслуживания, сертификатов и разрешений. Ограничением на массовое применение технологии является невысокая скорость сканирования, доступная в настоящее время, и необходимость в обучении системы.

1. Проведен анализ традиционных и алгоритмических методов защиты печатной продукции.
2. Обоснован подход, в котором изображение рассматривается как сигнал, передаваемый по каналу связи с шумом. Это позволяет моделировать системы принтер-сканер и принтер-сканер-принтер-сканер.
3. Предложена математическая модель преобразования изображения, которая показывает как видоизменяется изображение печатного документа при изготовлении поддельной копии.
4. Определены свойства канала связи и оценён «шум», вносимый средой обращения документа. Разработаны статистические методы сравнения двух каналов связи. Создан алгоритм идентификации каналов связи, соответствующих жизненным циклам подлинного или поддельным документам.
5. Реализована технология защиты печатных документов в виде программного обеспечения. С помощью программного обеспечения проведены испытания, в ходе которых разработанная технология позволила

безошибочно идентифицировать 40 подлинных и 20 скопированных (поддельных) документов.

Список литературы

1. *Птицын Н. В.* Метод защиты маркированной печатной продукции и компакт дисков // Технологии Microsoft в теории и практике программирования. Всероссийская конференция студентов, аспирантов и молодых ученых. — М: 2005. — С. 102–103.
2. *Птицын Н. В.* Разностный метод интегрального преобразования изображения // *Вестник МГТУ им. Н. Э. Баумана. Приборостроение.* — 2005. — № 3. — С. 58–64.
3. *Черненький В. М., Птицын Н. В.* Гибридной метод непараметрической нечёткой классификации в распознавании образов // *Вестник МГТУ им. Н. Э. Баумана. Приборостроение.* — 2005. — № 3. — С. 49–58.
4. Patent application 0503629.8, GB. Method and apparatus for automated analysis of biological specimen. /*N. V. Ptitsyn.* — 2005.
5. *Ptitsyn N. V.* Deterministic chaos in digital cryptography: Ph.D. thesis / Institute of Simulation Sciences, De Montfort University. — Leicester LE1 9BH, UK, 2003. — 152 p.
6. *Ptitsyn N. V., Blackledge J. M., Chernenkiy V. M.* Deterministic chaos in digital cryptography // Proceedings of IMA Conference on Fractal Geometry. — West Sussex, UK: Horwood Publishing, 2002. — P. 189–222.