

На правах рукописи



КЛЕЩАРЬ Сергей Николаевич

МОДЕЛИ И МЕТОДЫ КОНТРОЛЯ ТЕХНИЧЕСКИХ СРЕДСТВ
В СИСТЕМАХ СТРАХОВОГО ХРАНЕНИЯ ИНФОРМАЦИИ

Специальность 05.13.01 – Системный анализ, управление и обработка
информации (технические системы)

Автореферат
диссертации на соискание учёной степени
кандидата технических наук

Рязань 2013

Работа выполнена в ФГБОУ ВПО «Тульский государственный университет».

Научный руководитель: Ларкин Евгений Васильевич – д.т.н. проф.,
ФГБОУ ВПО «Тульский государственный университет»

Официальные оппоненты Привалов Александр Николаевич – д.т.н. проф.,
ФГБОУ ВПО «Тульский государственный педагогический университет»

Балашов Олег Евгеньевич – к.т.н., доцент,
ФГБОУ ВПО «Рязанский государственный радиотехнический университет»

Ведущая организация: Казенное предприятие города Москвы «Московский научно-технический центр «Регион» при правительстве Москвы

Защита состоится *13 ноября* 2013 г. в 12.00 на заседании диссертационного совета Д 212.211.01 в ФГБОУ ВПО «Рязанский государственный радиотехнический университет» по адресу:

390005, г. Рязань, ул. Гагарина, 59/1, ауд. 235.

С диссертацией можно ознакомиться в библиотеке ФГБОУ ВПО «Рязанский государственный радиотехнический университет».

Автореферат разослан *10 октября* 2013 г.

Ученый секретарь
диссертационного совета
кандидат технических наук, доцент



Пржегорлинский В.Н

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы Документы - важнейшая часть информационных ресурсов любого предприятия и организации. Для их накопления, хранения, обработки и предоставления пользователям на предприятиях и в организациях существуют различные архивы, службы и системы хранения и обращения документов. В документах аккумулируется, хранится и передается потомкам научно-техническое, культурное и историческое наследие. Поэтому надежное и долговременное сохранение важнейших видов документов и обеспечение оперативного доступа к ним является важной задачей любого государства по обеспечению его экономической, военной и информационной безопасности, а также сохранению документированного научно-технического, культурного и исторического наследия его народов. Информационное страхование документации - необходимое условие стабильной деятельности предприятия в любых экстремальных ситуациях. Утрата документов может нанести серьезный экономический и технический урон предприятию или учреждению, имущественным и социальным интересам его сотрудников

Для надежного хранения и восполнения информационного содержания документированной информации создается специальный фонд. В России Постановлением Правительства Российской Федерации от 18 января 1995 г. № 65 "О создании единого российского страхового фонда документации", и от 26 декабря 1995 г. № 1253-68 "Об обеспечении создания единого российского страхового фонда документации" создается, хранится и используется специальный информационный ресурс – Единый российский страховой фонд документации.

Система страхового фонда документации предполагает перевод юридически значимых документов, представленных на бумажных носителях на микрофильм. Перевод осуществляется либо традиционным прямым микрофильмированием, либо с использованием гибридных технологий, заключающихся в сканировании документов и выводе результата сканирования на микрофильм.

Доступ к микрофильмированным копиям документов осуществляется через сканирование микрофильма и формирование цифровых копий. В обоих случаях в системе используются сканеры документов, представленных на бумажных носителях или микрофильмах. Собственно сканеры изображений с физических носителей представляют собой достаточно сложные системы, от параметров которых зависит качество электронных копий документов, размещаемых в страховом фонде документации, поэтому оценка параметров сканеров является важным фактором создания системы единого российского страхового фонда документации. Параметры определяются по тестовому сигналу, который в данном случае принимает вид тест-объекта, т.е. эталонного изображения, включающего элементы для измерения параметров отдельных узлов и сканера в целом.

Вопросы формирования тестовых изображений для оценки параметров сканеров изображений документов на бумажных носителях и изображения

микрофильмированных копий документов в настоящее время решены недостаточно, что определяет необходимость и **актуальность** исследований, проведенных в диссертации.

Объектом исследования диссертационной работы сканеры микрографических копий и документов на бумажных носителях в системе страхового фонда документации для долговременного хранения информации.

Предметом исследования диссертационной работы являются параметры сканеров и методы измерения параметров с использованием тест-объектов как эталонных изображений, ориентированных на оценку технических параметров сканеров микрографических и бумажных носителей, таких как точность цвето- и тонопередачи, точность измерения местоположения фрагментов изображений на сканируемом оригинале, и точность в передаче мелких деталей изображения.

Научные выводы диссертации основаны на трудах в области обработки и идентификации изображений Б.В.Анисимова, К.Блаттера, Р.Вудса, В.А.Виттиха, Р.Гонсалеса, У.Гренандера, Р.Дуды, В.В.Еремеева, В.К.Злобина, В.Д.Курганова, Дж.Купера, Л.В.Новикова; А.Папулиса, У.К.Прэтта, А.Розенфельда, Л.И.Розоноэра, В.В.Сергеева, В.А.Сойфера, П.Харта, В.М.Чернова, Л.П.Ярославского и др. При оценке потерь информативности изображений использовался подход, разработанный К.Шенноном.

Из всех существующих подходов к синтезу тестовых сигналов и оценки параметров сканеров по тест-объектам наиболее продуктивным представляется подход, основанный на **аналитических методах** математического моделирования, что позволяет целенаправленно планировать будущие свойства копий документов, закладываемых на длительное хранение в системе страхового фонда документации. Для этого в диссертации использованы: системный анализ, теория сигналов, теория функций комплексной переменной, теория обработки изображений, теория информации.

Цель диссертационной работы состоит в повышении эффективности функционирования системы страхового фонда документации за счет введения в микрографический процесс этапа контроля параметров сканера по специальным тестовым сигналам, нанесенным на тест-объект, и прогнозирования качественных параметров микрофильмов, закладываемых на хранение, по результатам сканирования тест-объектов.

В соответствии с поставленной целью, в диссертации поставлены и решены следующие **задачи**:

1) Проведен анализ существующих конструкций и параметров сканеров и сформированы типовые схемы сканеров изображений, нанесенных на бумажный носитель и /или микрофильм.

2) Проведен анализ подходов к оценке параметров сканеров и сделан вывод о целесообразности использования математического описания сканера, полученного с использованием законов геометрической оптики, с одной стороны, и пространственной динамики изображений, с другой стороны.

3) Построена аналитическая математическая модель сканера, как электронно-микрографической системы создания и использования долговременно

хранимых копий в страховом фонде документации.

4) Разработана типовая схема тестирования сканера, основанная на сравнении сигнала, полученного в результате сканирования тест-объекта реальным сканером, используемым для подготовки электронных форм и виртуальным сканером, полученным в результате имитационного моделирования процесса сканирования.

5) В сканерах выделены два канала измерения: канал измерения светового потока, отраженного от поверхности бумажного носителя или прошедшего через микрофильм, и канал измерения координат фрагментов изображения документа на плоскости носителя, проведен анализ статических передаточных характеристик каналов, и их влияние на точность передачи в факсимильной цифровой модели изображения указанных параметров.

6) Построена математическая модель пространственной динамики сканера с использованием понятия логарифмической амплитудной пространственно-частотной характеристики.

7) Произведена оценка информационных потерь каналов вследствие неидеальности статических передаточных характеристик измерительных каналов сканера и реальной логарифмической амплитудно-частотной характеристики сканера в целом.

8) Разработан научно обоснованный подход к формированию пространственных тестовых сигналов для оценки параметров измерительных каналов сканера и сканера в целом.

9) Разработаны и внедрены на уровне проекта ГОСТ тест-объекты для оценки качества основных параметров работы сканеров создания цифровых копий документов с микрографических и бумажных носителей в страховом фонде документации.

Научная новизна диссертации заключается в следующем.

1) Аналитическая математическая модель сканера впервые представлена как виртуальная система, состоящая из двух каналов: измерения фотометрических параметров сканируемой сцены и измерения пространственного местоположения предметов сцены, что позволило сформулировать для каждого из выделенных каналов понятие статической передаточной характеристики, а для сканера в целом - понятие пространственной логарифмической амплитудно-частотной характеристики.

2) Проведена оценка потерь информации в каналах сканера, как результата отклонения статических передаточных характеристик каналов и пространственной логарифмической амплитудно-частотной характеристики от идеальных значений, что позволяет сформулировать требования к элементам тест-объектов для оценки параметров каналов сканера, и сканера в целом.

3) Научно обоснованы и разработаны требования к элементам тест-объектов, ориентированным на построение статических передаточных характеристик каналов измерения фотометрических параметров и пространственных координат сканера, а также на оценку логарифмической амплитудной пространственно-частотной характеристики сканера в целом.

4) Предложен тест-объект для оценки параметров сканера и разработаны методы обработки изображений, получаемых в результате сканирования тест-объекта, для прогнозирования качественных параметров электронных форм документов, закладываемых на хранение в системе страхового фонда документации.

Практическая ценность работы заключается в том, что разработанные в диссертации тест-объект и методы обработки изображений ориентированы на использование их в системе страхового фонда для доступа к документам, представленным на микрофильмах и бумажных носителях, что позволяет повысить надежность передачи информации во времени при ее длительном хранении.

Достоверность полученных теоретических результатов подтверждается результатами натуральных испытаний сканирования оригиналов документов на бумажных носителях и их микрофильмированных копий, а также внедрением результатов в народное хозяйство на уровне проекта ГОСТ.

Положения, выносимые на защиту.

1) Аналитическая математическая модель сканера, как виртуальной системы, состоящей из канала измерения фотометрических параметров и канала измерения пространственного местоположения предметов сцены.

2) Оценка потерь информации в каналах сканера, как результата отклонения статических передаточных характеристик каналов и пространственной логарифмической амплитудно-частотной характеристики от идеальных значений.

3) Требования к элементам тест-объектов, ориентированным на оценку статических передаточных характеристик каналов и логарифмической амплитудной пространственно-частотной характеристики сканера.

4) Тест-объект для измерения параметров сканеров и методы обработки изображений тест-объекта, для прогнозирования качественных параметров электронных форм документов, закладываемых на хранение в системе страхового фонда документации.

Реализация и внедрение результатов. Предложенные в диссертации методы реализованы автором в ГОСТ 33007 "ЕР СФД. Микрофильмы страхового фонда, изготовленные с электронных документов. Общие технические требования и методы контроля", проекте ГОСТ 33..... "ЕР СФД. Электронные копии документов, полученные сканированием бумажных оригиналов. Показатели качества и методы контроля"; Методических рекомендациях по выбору сканирующего оборудования для нужд Российского архивного агентства.

Ряд теоретических положений внедрен в учебный процесс Тульского государственного университета на кафедре «Робототехника и автоматизация производства» в лекционных курсах по дисциплинам: «Информатика», «Информационные устройства и системы в робототехнике», «Системы технического зрения роботов».

Апробация работы. Основные положения диссертации докладывались на следующих конференциях и семинарах.

1. XXXI Научная сессия, посвященная Дню радио. - Тула, Тульский госу-

дарственный университет, 2012.

2. Системы управления электротехническими объектами СУЭТО-2012 - Тула, Тульский государственный университет, 2012.

По теме диссертации опубликовано 20 работ, включенных в список литературы, в том числе: 3 статьи, представляющих собой материалы межрегиональных научно-технических конференций, 4 статьи в сборнике, рекомендуемом ВАК РФ для публикаций материалов диссертаций.

Структура и объем работы. Диссертационная работа состоит из введения, четырех разделов, изложенных на 154 страницах машинописного текста и включающих 58 рисунков и 7 таблицы, заключения, списка использованной литературы из 117 наименований.

СОДЕРЖАНИЕ РАБОТЫ

Первый раздел содержит определение системы страхового фонда документации, обеспечивающей сохранность документов и передачу информации во времени, приведено понятие гибридных технологий, даны различные схемы сохранения и передачи во времени документов, способы записи информационного содержания документов на микрофильмы, используемые в системе страхового фонда документации, показано, какое место занимает в процессе изготовления микрофильмов страхового фонда сканирование готовых документов с проверкой качества полученных электронных копий.

На основании проведенного обзора технических средств получены обобщенные функциональные схемы сканеров бумажного носителя и микрофильма. Показано, что важным этапом создания страховых микрофильмов длительного хранения является этап сканирования бумажных оригиналов, определяющих качество страховых копий.

Приведена общая схема контроля электронной копии документа предназначенной для изготовления микрофильма страхового фонда (рисунок 1), где эталонный тестовый сигнал $S_T(Y, Z, \lambda)$ подается на входы реальной $f_P(S_T)$ и эталонной $f_Э(S_T)$ систем, в результате чего формируются сигналы S_P и $S_Э$, нормы разности сходных характеристик которых $\varepsilon = \|\mathbf{h}(S_Э) - \mathbf{h}(S_P)\|$ сравнивается с порогом ε_{Π} . по результатам сравнения информационная система признается настроенной правильно, или признается, что она подлежит настройке.

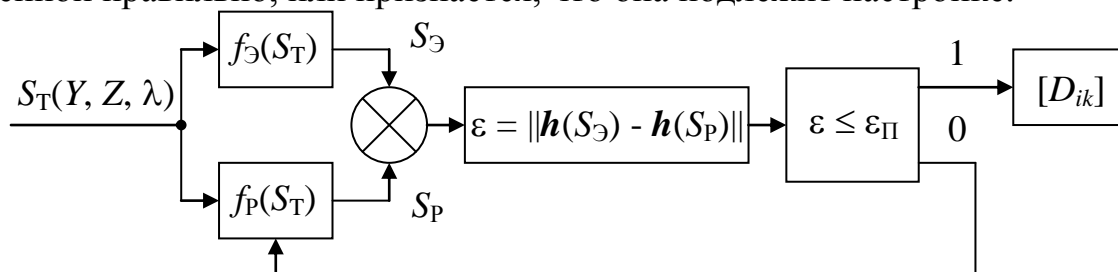


Рисунок 1 - Схема контроля СЭД

Проведен анализ тестовых сигналов, используемых в известных схемах настройки, в том числе метод, разработанный в Корнуэльском (Cornwell) университете США, метод QI (Quality Index), тест-объекты ISO 12653-1 и ISO

12653-2, универсальный тест-объект (УТТ), разработанный в компании Image Engineering, тест-объекты в соответствии с ГОСТ 13.1.703-91 и ГОСТ 13.1.701-95. Показано, что существенным недостатком известных тестовых сигналов является их ориентация на последующее визуальное восприятие сигналов S_p и S_Σ и субъективную оценку нормы ε .

Проведен анализ изображений на бумажных документах, сформированных с помощью технических средств. Показано, что независимо от способа нанесения изображения на бумажную основу возможны следующие виды модуляции светового потока: изменение отражательной способности непрерывных участков поверхности бумаги; текстурированная модуляция; амплитудная первого и второго рода; частотная (1.14); стохастическая или квазистохастическая; комбинированная

Разработан подход к формированию имитационной модели эталонного виртуального сканера, по которой может быть произведена оценка основных параметров сканирования. Определено, что модель должна включать две составляющие: модель статических геометрических преобразований эталонного сигнала и модель, описывающая пространственную динамику преобразований, приводящую к потерям разрешения.

Для исследования статических характеристик сканирующих устройств предложено использовать геометрические соотношения, вытекающие из особенностей конструкции элементов сканера, модели которых строятся в системе координат $xOyz$ предметной плоскости и системе координат $XO'YZ$, связанной с фотоэлектронным преобразователем.

Показано, что исследование пространственной динамики преобразований сигнала в сканирующих системах связано с решением дифференциальных уравнений вида $F(\Phi(y, z), E(Y, Z), \nabla E(Y, Z), \nabla^2 E(Y, Z), \dots) = 0$, где $\Phi(y, z)$ - оптический сигнал в предметной плоскости; $E(Y, Z)$ - оптический сигнал в плоскости расположения фотоэлектронного преобразователя; ; $\nabla^2 = \left(\frac{\partial^2}{\partial Y^2}, \frac{\partial^2}{\partial Z^2} \right), \dots$

Во **втором разделе** разработана модель эталонного сканера, в которой учтены особенности носителя изображений, как источника сигнала; способ формирования изображений на носителе, как фактор, определяющий общие особенности сигнала, и погрешности, вносимые в сигнал, узлами и блоками сканера на оптическом, фотоэлектронном и электронном уровнях.

Источник света, используемый в сканере, представлен в виде плоской излучающей поверхности, которая описывается уравнением $\Omega(x', y', z') = 0$, где x', y', z' - координаты системы координат, связанной с источником; Ω - некоторая скалярная функция. Излучающая поверхность ограничена линией пересечения с поверхностью $\zeta_n(x', y', z') \leq 0$. Излучение из каждой точки поверхности обладает осевой симметрией относительно нормали к излучающей поверхности, причем интенсивность излучения из точки K с координатами (x'_K, y'_K, z'_K) в направлении угла α_n к нормали, определяется по зависимости $I_n(x'_K, y'_K, z'_K, \alpha_n, \lambda)$.

Поверхность документа, описана зависимостью $x=0$; $|y| \leq \Delta_y$, $|z| \leq \Delta_z$ где $2\Delta_y \times 2\Delta_z$ - размеры области, которая в текущий момент сканируется. N источников света создают в точке k сканируемого документа освещенность

$$E_1(0, y_k, z_k, \lambda) = \sum_{n=1}^N \kappa_{1n} \int_{\substack{\hat{\Omega}_n(x, y, z)=0, \\ \hat{\zeta}_n(x, y, z) < 0}} \iota_n[x, y, z, \alpha_n(x, 0, y, y_k, z, z_k), \lambda] dx dy dz, \quad (1)$$

где $\iota_n[x, y, z, \alpha_n(x, 0, y, y_k, z, z_k), \lambda]$ - индикатриса интенсивности излучения n -го источника из точки с координатами x, y, z в направлении угла $\alpha_n(x, 0, y, y_k, z, z_k)$ между нормалью к поверхности в указанной точки и лучом, направленным в точку k ; κ - коэффициент пропорциональности.

Отмечается, что освещенность $E(0, y_k, z_k, \lambda) = \text{const}_{y_k, z_k}$ в пределах всей предметной плоскости для сканеров планшетного типа и в пределах линии, проецируемой объективом на линейку фотоэлектронных преобразователей, для сканеров рулонного типа. Неравномерность освещенности может появляться за счет особенностей конструкции многоэлементных осветителей, появления боковых засветок, затенения части предметной плоскости элементами конструкции, образования непрозрачного налета на поверхности осветителя, и т.п.

Бумажная основа вместе с нанесенным на нее изображением документа представляет собой отражающую поверхность, которая осуществляет пространственную модуляцию падающего на нее светового потока. Отражение света от каждой точки поверхности бумажного носителя в данном случае зависит от угла падения, под которым понимается угол между нормалью к отражающей поверхности и падающим лучом. Отражение характеризуется индикатрисой, которая обладает осевой симметрией относительно прямой, лежащей в одной плоскости с падающим лучом, и расположенной под углом относительно нормали, равным углу падения.

Показано, что точкой k создается следующий световой поток, попадающий во входной зрачок объектива:

$$\Phi_o(y_k, z_k) = \sum_{n=0}^N \kappa_{2n} \cdot \iint_{y''^2 + z''^2 \leq o^2} \iota_R(\alpha_{Rn}(y_k, z_k, y'', z''), \beta_F, \lambda) dy'' dz'', \quad (2)$$

где κ_{2n} - коэффициент пропорциональности; y'', z'' - координаты системы, связанной со входным зрачком объектива; β_F - угол между падающим лучом и нормалью к поверхности; $\alpha_{Rn}(y_k, z_k, y'', z'')$ - угол между выбранным направлением и отраженным лучом; ι_R - индикатриса отражения.

Вследствие того, что индикатриса отражения $\iota_R(\alpha_{Rn}, \beta_F, \lambda)$ зависит от состояния отражательной поверхности носителя и длины волны, что определяется нанесенным на носитель красителем, доля освещенности входного зрачка, обусловленная отражением света от участка носителя в точке k , модулируется по интенсивности, что и образует оптический сигнал на входных окнах фотоэлектронных ячеек фотоэлектронного преобразователя.

Информация с микрофильма считывается в проходящем свете. Показано, что модуляция светового сигнала, падающего на микрофильм, зависит от концентрации микрокристаллов серебра в информационном слое. Функция поглощения единственного микрокристалла серебра может быть представлена в виде гауссиана с параметром, определяющим ширину, определяемым по размерам микрокристалла серебра. Если вероятность попадания некоторого количества микрокристаллов серебра в информационный слой определяется распределением Пуассона, то отношение освещенности площади перед слоем диспергента к освещенности той же площади за слоем диспергента определяется законом Бугера-Ламберта-Бера, имеющим вид

$$d = \lg \left[\frac{E(dS)}{E'(dS)} \right] = \kappa_{\mu} \mu,$$

где κ_{μ} - коэффициент пропорциональности; μ - концентрация микрокристаллов серебра (диспергент) в информационном слое; $E(dS)$ и $E'(dS)$ - освещенности площади dS перед и за слоем диспергента, соответственно; d - оптическая плотность микрофильма.

Показано, что на потери информации влияет виньетирование объектива, вследствие чего неравномерность освещенности плоскости расположения фотоэлектронного преобразователя даже при равномерном освещении предметной плоскости может достигать 10÷15 %.

Определено, что пространственная динамика реального оптического и оптоэлектронного тракта преобразования сигнала в сканере может быть описа-

на в виде функции Гаусса, $w_G(Y, Z) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{Y^2 + Z^2}{2\sigma^2}\right)$, показатель σ кото-

рой определяется диаметром кружка рассеяния объектива, размерами фоточувствительных ячеек фотоэлектронного преобразователя, динамикой транспортировки заряда в фоточувствительном приборе с зарядовой связью и работы усилительно-согласующих схем при аналого-цифровом преобразовании.

В третьем разделе предложено разделить функциональную схему сканера на канал измерения уровня сигнала, получаемого от предметов сцены и канал измерения местоположения предметов сцены, что позволило произвести оценку потерь информации в процессе создания микрофильмов страхового фонда с электронных документов.

Показано, что на каждом этапе формирования электронной копии документа сканером в эталонную факсимильную цифровую модель изображения документа вносятся изменения, которые названы потерями информации. Часть потерь, вызванная конструктивными особенностями используемого сканера и характеристиками его узлов и блоков, исследованными в разделе 2, закладывается на этапе проектирования, изготовления и настройки прибора. Другая часть потерь возникает в процессе сканирования, за счет отклонения климатических факторов (температура, влажность) от идеальных значений, существования внешних электрических и магнитных полей, внешнего электромагнитного из-

лучения, радиационного фона и т.п. Третья часть потерь возникает за счет естественной деградации деталей и узлов оптической, механической и электронной компонент сканера в процессе его длительной эксплуатации.

Все потери информации предложено разделять на два вида: потери, связанные с неидеальными статическими передаточными характеристиками сканера; потери, связанные с пространственной динамикой элементов, осуществляющих оптико-электронное преобразование сигнала. В свою очередь, статические передаточные характеристики сканеров разделяются на статическую передаточную характеристику канала измерения модулирующего фактора фрагментов изображения и статическую передаточную характеристику канала измерения пространственных координат фрагментов изображения по координатам y, z . Обе статические характеристики в общем случае являются нелинейными. На потери информации в сканере влияют многие факторы, каждый из которых формирует свои, первичные, или частные потери сканера ξ_i .

Суммарные потери определяется как взвешенная сумма первичных потерь $\xi = \sqrt{\sum_i a_i \xi_i^2}$, где

$a_i > 0$ - коэффициент, учитывающий степень влияния i го фактора на потери информации. Суммарные потери ξ выражаются в изменении местоположения отдельных элементов изображения в электронной форме относительно их местоположения в изображении на сканируемом документе и в изменении цвета или тона элементов изображения относительно цвета или того же элемента на сканируемом документе. Погрешности каждого вида электронной копии документа должны идентифицироваться своим тестовым сигналом, представленным в тест-объекте.

Статическая передаточная характеристика канала измерения модулирующего фактора фрагментов изображения приведена на рисунке 2, где по оси абсцисс откладывается модулирующий фактор Q (отражательная способность поверхности бумажного носителя или концентрация микрокристаллов серебра в информационном слое микрофильма), а по оси ординат - результаты D' измерения и цифрового представления величины Q . В идеальном случае статическая характеристика имеет вид ступенчатой функции 1, в которую изначально заложен шум квантования, оцениваемый как половина расстояния между двумя соседними порогами квантования. В том случае, если передаточная характеристика является нелинейной (кривые 2 и 3 на рис. 3.2), это приводит к информационным потерям. Действительно, оценим количество информации, содержащееся в одном пикселе ФЦМИ, как энтропию сообщения, состоящего из одного слова, которое может принимать P равновероятных значений: $H = \log_2 P$ [бит], где P - количество дискретных значений, на которые разбивается и модулирующий фактор Q , и код D , в соответствии с рис. 2. Если кривая, представляющая передаточную характеристику канала измерения модулирующего фактора, имеет наклон меньший, чем угол ν , то это означает, что на этом участке два дискретных значения фактора Q соответствуют одному значению кода D . И наоборот, если на некотором участке кривая, представляющая передаточную ха-

рактическую характеристику канала измерения модулирующего фактора, имеет наклон больший, чем угол υ , то это означает, что два разных кода D соответствуют одному дискретному значению фактора Q . В любом случае, количество разных значений кода D , либо дискретных значений модулирующего фактора Q при изменении угла υ уменьшается. Если уменьшение составляет $m < P$ единиц, то максимальное количество информации, содержащейся в одном пикселе, оцененное как информационная энтропия, становится равным $H = \log_2(P - m) < \log_2 P$, что и означает потерю информации.

Соответственно, тест-объект должен иметь компоненты, позволяющие строить кривую типа 2 и проверять среднюю статическую и статическую чувствительность сканера.

Статическая передаточная характеристика канала измерения пространственных координат фрагментов изображения приведена на рис. 3, где по оси абсцисс откладываются значения координат на носителе изображения, нормированные в интервале $0 \leq y/2, z/2 \leq 1$, а по оси ординат откладываются значения координат на плоскости расположения фоточувствительных ячеек фотоэлектронного преобразователя, нормированные в интервале $0 \leq Y/2, Z/2 \leq 1$.

В идеальном случае статическая передаточная характеристика канала измерения координат представляет собой ступенчатую функцию, аналогичную ступенчатой функции, приведенной на рисунке 2, в которую изначально заложен шум дискретизации. В том случае, если передаточная характеристика является нелинейной (кривая 2), это приводит к информационным потерям.

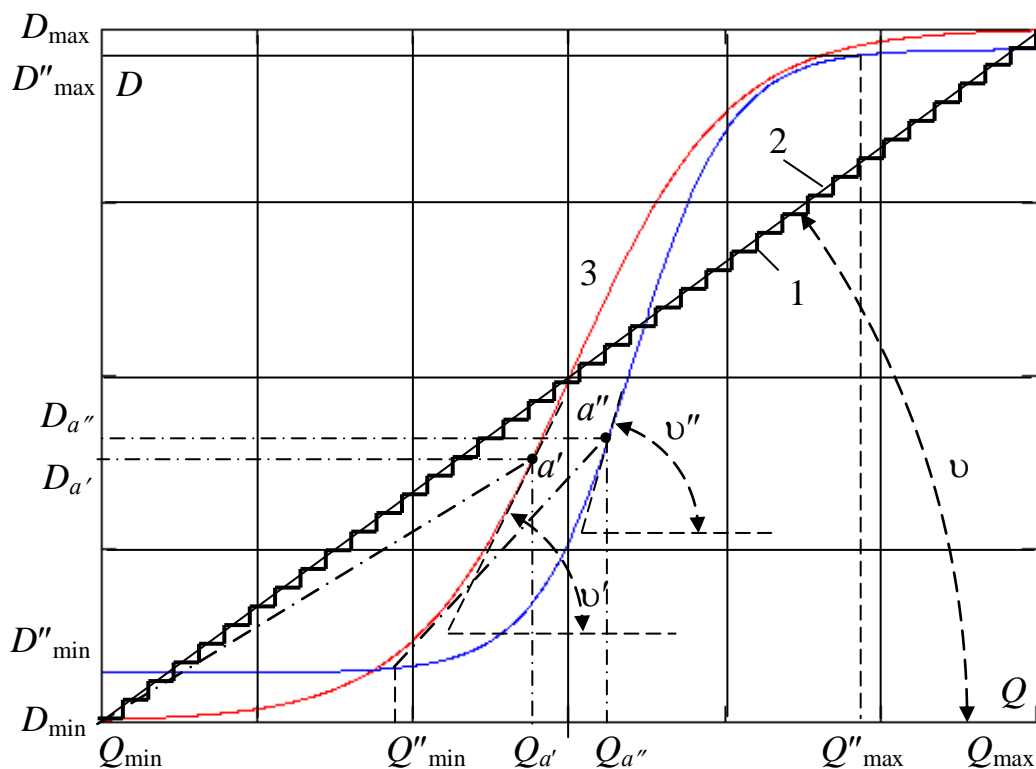


Рисунок 2 - Погрешности цвето- и тонопередачи сканера

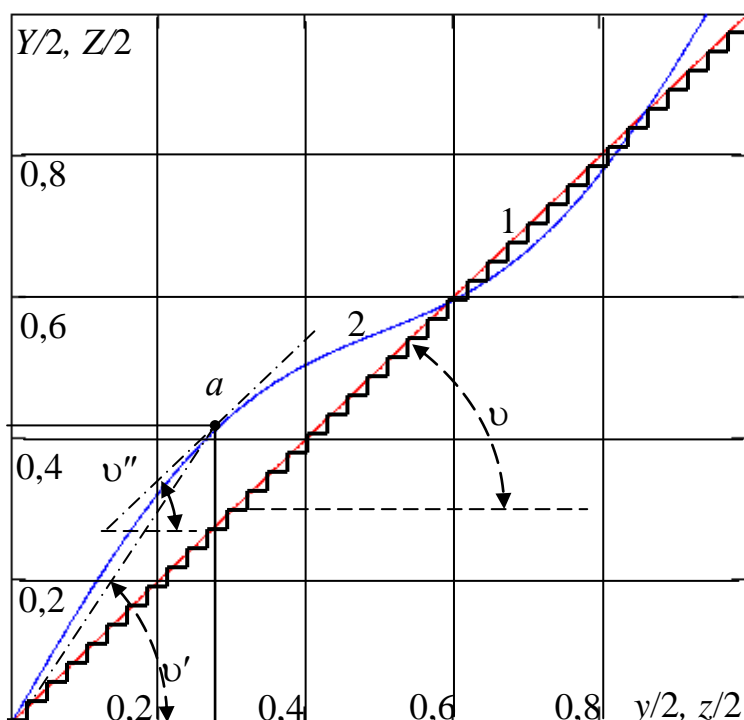


Рисунок 3- Погрешности измерения координат элементов ФЦМИ

Действительно количество информации, содержащееся в одном пикселе ФЦМИ, может быть оценено по зависимости (3.5) В паре соседних пикселей информации содержится вдвое больше. Если кривая, представляющая передаточную характеристику канала измерения координат, имеет наклон меньший, чем угол υ , то это означает, что два соседних пикселя имеют одинаковые значения, т.е. в них содержится то же количество информации, что и в одном пикселе канала с линейной характеристикой. И наоборот, если на некотором участке кривая, представляющая передаточную характеристику канала измерения координат, имеет наклон больший, чем угол υ , то это означает, что один из пикселей пропускается при выполнении операции дискретизации, что также приводит к потерям.

Генезис информационных потерь, связанных с пространственной динамикой сканера, поясняется рисунок 4, где изображен результат сканирования изображения документа, содержащего две рядом стоящие одинаковые геометрические точки, образ которых представляет собой сумму функций Гаусса, нормированных по объему. В диссертации показано, что с ростом соотношения b/a , где a - расстояние между геометрическими точками, а b - фактор, определяющий ширину гауссиана, динамический диапазон, понимаемый как $|E'_1 - E'_2|$, сокращается, что приводит уменьшению количества градаций при оцифровке сигнала в промежутке между центрами точек, а следовательно к потерям информации.

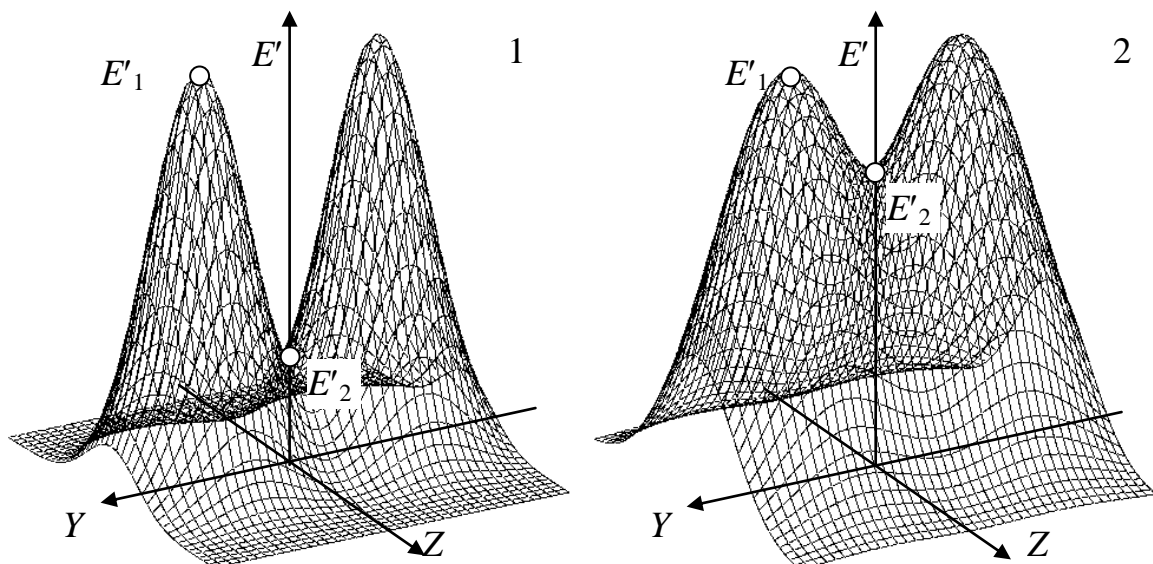


Рисунок 4 - Информационные потери, связанные с пространственной динамикой сканера

Показано, что требования к элементам тест-объекта для канала измерения соответствующих факторов напрямую зависят от решаемой задачи и характера сигнала на выходе соответствующего канала измерения.

При построении экспериментальной кривой измерения цвето- и тонопередачи должен быть сформирован тест-объект, в котором уровень градаций цветовых составляющих в сигнале RGB, и/или оттенков серого в документах, не имеющих цветового кодирования, нарастал бы по определенному закону, например, по линейному (оптический клин). При этом с целью повышения точности измерений тестовый сигнал Q должен меняться ступенчато и иметь значения $\{Q^0, Q^1, \dots, Q^{n(Q)}, \dots, Q^{N(Q)-1}\}$. Многократные измерения в рассматриваемом канале организуются естественным образом, за счет сканирования предметной плоскости по строкам и столбцам при формировании ФЦМИ, поэтому для многократных измерений величины $D(Y, Z)$ элемент для измерения конкретных значений должен иметь конечные размеры, позволяющие проводить множество измерений одного и того же значения модулирующего фактора. Результат измерения должен быть представлен в виде гистограммы значений величин $\{D^0, D^1, \dots, D^{n(Q)}, \dots, D^{N(Q)-1}\}$, по которым могут быть найдены

их математические ожидания $\bar{D}^{n(Q)} = \frac{\sum_{m[n(Q)]=1}^{M[n(Q)]} D_m[n(Q)]}{M[n(Q)]}$ и среднеквадратичные

отклонения $\bar{\Delta}^{n(Q)} = \sqrt{\frac{1}{M[n(Q)]} \cdot \sum_{m[n(Q)]=1}^{M[n(Q)]} [D_m[n(Q)] - \bar{D}^{n(Q)}]^2}$, которые, в свою

очередь, должны укладываться в заданные пределы. Вследствие возможной неравномерности освещенности предметной плоскости и эффекта виньетирования оптический клин должен быть размещен в центре и по краям тест-объекта.

Показано, что при автоматическом определении координат элементов тест-объекта, предназначенных для тестирования канала измерения координат Y, Z предметов сцены, указанные предметы прежде всего должны быть найдены на фоне шума, а затем уже должна решаться задача вычисления их координат. В соответствии с принципом оптимальной фильтрации таковыми предметами могут являться функции Гаусса, распределенные некоторым упорядоченным образом по площади кадра (например, находящиеся на пересечении диагоналей сцены и концентрических окружностей, с линейно возрастающим радиусом и центром, совпадающим с центром сцены), что показано на рисунке 5.

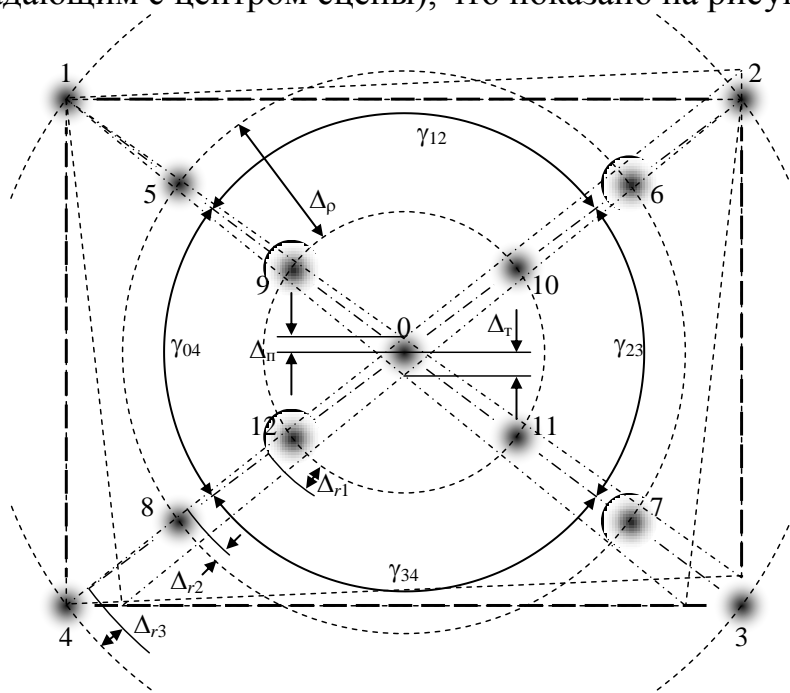


Рисунок 5 - Структура тест-объекта для определения статических передаточных характеристик канала измерения координат

В диссертации предложено по данному тест-объекту оценивать величину статических aberrаций сканера типа «подушка», «бочка», «трапеция», «параллелограмм» и др.

Пространственную динамику сканера предложено оценивать с помощью прямоугольной четырехшпальной миры, описываемой зависимостью

$$Q(Y, Z) = [\eta(Y + U) - \eta(Y - U)] \cdot \sum_{n=-2}^1 [\eta(Z - T - 4nT) - \eta(Z - 3T - 4nT)], \text{ где } \eta(\dots) -$$

единичная функция Хевисайда; $2U$ - длина штриха; $2T$ - ширина штриха.

Четвертый раздел посвящен разработке тест-объекта и программы автоматизированного контроля качества электронных документов, поставляемых на страховое микрофильмирование. Используемые в тест-объекте тестовые элементы позволяют количественно (инструментальными средствами) или визуально оценить качество формируемого изображения и изготавливать микрофильмы предназначенные для длительного страхового хранения.

К основным параметрам сканеров отнесены следующие: равномерность передачи градаций серого по площади сканируемого документа; линейность

оптического клина при передаче градаций серого; частотно-контрастная характеристика; геометрические искажения изображения. В соответствии с этим предложена структура тест-объекта, приведенная на рисунке 6.

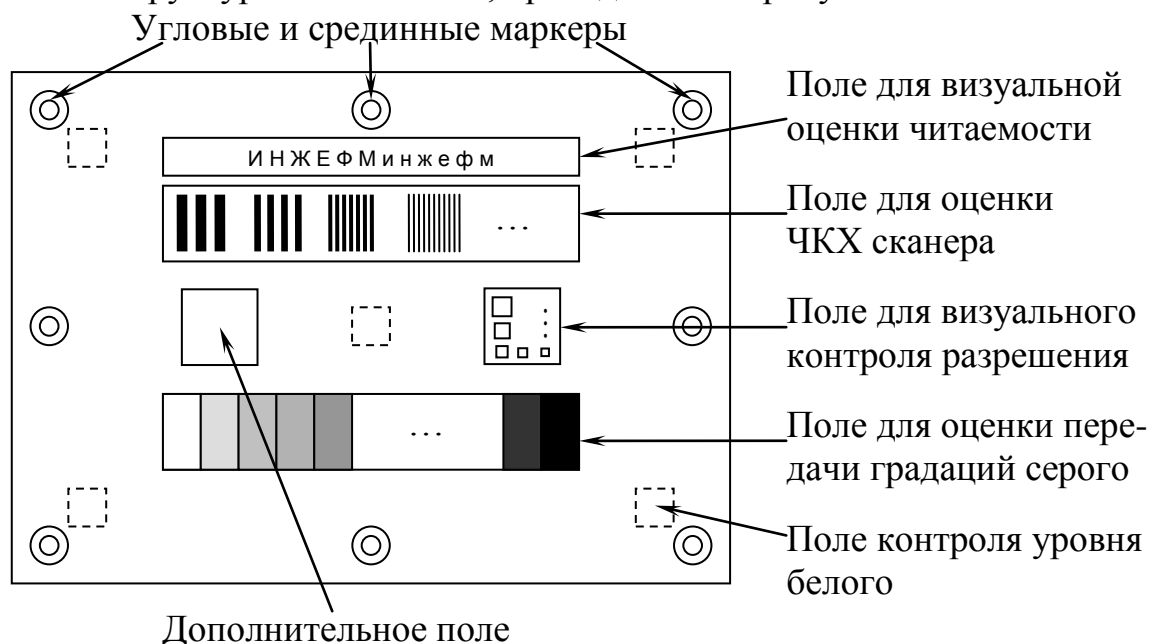


Рисунок 6 - Общая структура тест-объекта

Маркеры на изображении тест-объекта используются в качестве реперов при автоматическом поиске других фрагментов тест-объекта, а также для оценки статических aberrаций. В качестве маркеров использованы две концентрические окружности, что позволяет отличать их при согласованной фильтрации от дефектов изображения типа «темное пятно на светлом фоне». Усреднённый коэффициент трапецидальных искажений вычисляется по зависимости

$k_t = \left(\frac{L_{06} - L_{24}}{L_{06} + L_{24}} + \frac{L_{02} - L_{64}}{L_{02} + L_{64}} \right) \cdot 100\%$, где L_{ij} - расстояние между i -м и j -м маркерами.

Тангенциальные искажения («бочка», «подушка» и т.п.) оцениваются по величине и знаку отклонения срединных маркеров от прямых, соединяющих соответствующие угловые маркеры верхнего ряда $\delta_t = \frac{A_t n_7 + B_t m_7 - 1}{\sqrt{A_t^2 + B_t^2}}$, где

$$A_t = \frac{m_6 - m_0}{m_6 n_0 - m_0 n_6}; \quad B_t = \frac{n_0 - n_6}{m_6 n_0 - m_0 n_6}; \quad m_i, n_i - \text{оцененные значения абсциссы и ординаты } i\text{-го маркера.}$$

Аналогичные зависимости получены для маркеров, нанесенных на боковые и нижнюю границу сцены.

Метод оценки aberrаций был экспериментально проверен на сканере PS7000С МКII. Получен усредненный коэффициент трапецидальных искажений, равный $k_t = 0,14\%$. и коэффициент тангенциальных искажений равный $k_{tg} = 0,12\%$. Таким образом, тестируемый сканер показал высокую точность

передачи размеров и форм объектов при формировании изображения. Расчётные значения коэффициентов искажений не превышают 0,2%.

Поле для оценки частотно-контрастной характеристики сканера представляет собой набор групп пар чёрных и белых линий различной ширины, соответствующей различным пространственным частотам. Для формальной оценки качества передачи высоких пространственных частот предложен метод поперечного сканирования штрихов и определения усредненного максимального и минимального значений сигнала по зависимости $K = \frac{\bar{P}_{\max} - \bar{P}_{\min}}{P_{\max} - P_{\min}} \cdot 100\%$, где

где P_{\max} и P_{\min} - усредненное максимальное и минимальное значения яркости при выбранном способе кодирования яркостей в цифровом изображении.

Поля контроля уровня белого представляют собой неокрашенные участки бумажного листа, на котором исполнен тест-объект. Они позволяют контролировать равномерность освещённости по полю кадра, что на практике является одним из важнейших показателей качества сканирования. Для минимизации влияния шумов яркость усредняется по всем пикселям, входящим в данное поле. В результате формируется множество яркостей, соответствующих четырём угловым и одной центральной зонам $\{P_{w1}, P_{w2}, \dots, P_{w5}\}$. Неравномерность рассчитывается по среднеквадратичному отклонению, отнесённому к ширине диапазона воспринимаемых сканером яркостей.

Для оценки диапазона оптических плотностей, воспринимаемого и передаваемого сканером, а также точности передачи различных уровней яркостей в тест-объект включен оптический клин, состоящий из 15 полей, имеющих переменную оптическую плотность в диапазоне от 0,05 до 1,95

Для обеспечения возможности использования разработанного тест-объекта в задачах оценки качества сканирования документов с частотным методом модуляции предлагается включить в структуру дополнительное поле, представляющее собой решётку чередующихся чёрных и белых линий с постоянной шириной чёрных линий (1 пиксель) и переменной шириной белых (1 через 1; 1 через 2; 1 через 3 и т.п.). Каждый тестовый элемент содержит пять строк вертикальных линий толщиной в 1 пиксель, разделённых промежутками в 1, 2, 3, 4 и 5 белых пикселей соответственно. Длина каждого штриха составляет 48 пикселей. Первая строка содержит 180 штрихов, вторая — 120, третья — 90, четвёртая — 72 и пятая — 60.

Предложенная методика оценки качества сканирования была реализована в виде программ, работающих с изображениями в графическом формате BMP (BitmapPicture), имеющими цветовое разрешение 8 бит/пиксель (градации серого). Перед началом работы из заголовка графического файла считывается информация о пространственном разрешении, с которым осуществлялось сканирование. Это позволяет, зная структуру изображения тест-объекта и его физические размеры, произвести примерный переход от реальных размеров в миллиметрах к размерам в пикселях на изображении.

Проведенные исследования позволяют сделать следующие выводы.

1) Разработана типовая функциональная схема сканера, используемого в системе страхового фонда документации для формирования электронных форм документов для последующего микрофильмирования и оперативного доступа.

2) Сформулировано понятие контроля процесса сканирования и разработана типовая структура системы контроля, основанная на сравнении параметров электронных форм тест-объектов, формируемых на выходах реального и виртуального сканеров.

3) На основании анализа существующих тест-объектов и методов контроля качества процессов сканирования предложен принцип разработки тест-объекта, основанный на формировании полей с фиксированными координатами для оценки различных параметров электронных форм для долговременного хранения и последующего микрофильмирования в страховом фонде документации.

4) Разработана математическая модель виртуального сканера, включающая модель многокомпонентного источника света, модель пространственной модуляции света при работе сканера на отражение и на просвет, модель формирования изображения объективом и модель фотоэлектронного преобразования.

5) Показано, что каждый этап преобразования сигнала вносит статические и пространственно-динамические погрешности в первоначальную пространственную модуляцию светового потока, а следовательно приводит к информационным потерям в формируемой электронной форме документа.

6) В соответствии со структурной схемой этапов формирования электронной формы документа в сканерах выделены два измерительных канала: канал измерения модулирующего фактора и канал измерения пространственных координат, а все потери информации разделены на потери, связанные со статическими характеристиками каналов сканеров и потери, связанные с динамическими характеристиками сканеров в соответствующих каналах.

7) Исследована статическая передаточная характеристика канала измерения модулирующего фактора, показана связь потерь информации в канале с параметрами статической передаточной характеристики.

8) Исследована статическая передаточная характеристика канала измерения пространственных координат предметов сцены, показана связь потерь информации с параметрами статической передаточной характеристики.

9) Исследована связь потерь информации с пространственной динамикой сканера и показано, что информационные потери в данном случае определяются глубиной модуляции сигнала от близко расположенных предметов на сканируемом изображении документа, т.е. величины динамического диапазона сканера.

10) Определены требования к тест-объекту для оценки потерь информации в сканерах и синтезирован тест-объект, позволяющий оценивать статические и пространственно-динамические характеристики сканеров.

11) На основании проведенных теоретических исследований разработан проект ГОСТ 33. 007 «ЕР СФД. Микрофильмы, изготовленные с электронных документов» Общие технические требования и методы контроля" ГОСТ 33 ...

«ЕР СФД. Электронные копии бумажных документов. Показатели качества и методы контроля», по контролю процесса подготовки электронных форм документов.

ОПУБЛИКОВАННЫЕ РАБОТЫ ПО ТЕМЕ ДИССЕРТАЦИИ

1. Клещарь С.Н. Формирование изображений в СОМ-системах // Интеллектуальные и информационные системы. Материалы Всероссийской научно-технической конференции «Интеллект-2011». - Тула: Изд-во ТулГУ, 2011. - С. 66 - 69.

2. Клещарь С.Н., Ларкин Е.В., Котов В.В. К вопросу параметрической точности микрографических систем // XXIX Научная сессия, посвященная Дню радио. - Тула: НТО РЭС им. А.С. Попова, 2011. - С. 172 – 176.

3. Клещарь С.Н. Характеристики глаза как оптического прибора в системе визуального контроля качества микрофильма // XXIX Научная сессия, посвященная Дню радио. - Тула: НТО РЭС им. А.С. Попова, 2011. - С. 140 – 144.

4. Клещарь С.Н., Ларкин Е.В. Восприятие миры в системе контроля качества микрофильма // XXIX Научная сессия, посвященная Дню радио. - Тула: НТО РЭС им. А.С. Попова, 2011. - С. 8 – 13.

5. Клещарь С.Н., Котов В.В. Совмещение цветowych плоскостей при восстановлении полноцветных изображений // Приборы и управление. Вып. 9. - Тула: Изд-во ТулГУ, 2011. - С. 41 - 47.

6. Клещарь С.Н. Ларкин Е.В. Проблемы цветопередачи при микрофильмировании документов с цветowym кодированием // Приборы и управление. Вып. 9. - Тула: Изд-во ТулГУ, 2011. - С. 47 - 52.

7. Клещарь С.Н. Контроль качества совмещения цветowych плоскостей при цветоделенном микрофильмировании // Приборы и управление. Вып. 9. - Тула: Изд-во ТулГУ, 2011. - С. 52 - 59.

8. Горшков А.А., Клещарь С.Н., Кузнецова Т.Р. Восприятие изображения на экране монитора // Приборы и управление. Вып. 10. - Тула: Изд-во ТулГУ, 2012. - С. 36 - 43.

9. Клещарь С.Н. Формирование изображения на микрофильме // Приборы и управление. - Вып. 10. Тула: Изд-во ТулГУ, 2012. - С. 81 - 86.

10. Клещарь С.Н., Котов В.В. Контроль параметров светового потока В СИМ-устройствах // Сборник научных трудов Шестой Всероссийской научно-практической конференции «Системы управления электротехническими объектами «СУЭТО-6» - Тула: Изд-во ТулГУ, 2012. - С. 235 - 237.

11. Клещарь С.Н., Ларкин Е.В. Выявление очагов диффузного рассеяния при длительном хранении микрофильмов устройствах // Сборник научных трудов Шестой Всероссийской научно-практической конференции «Системы управления электротехническими объектами «СУЭТО-6». - Тула: Изд-во ТулГУ, 2012. - С. 237 - 239.

12. Будков С.А., Клещарь С.Н. Статическая передаточная характеристика канала измерения пространственных координат // XXX Научная сессия, посвященная Дню радио. - Тула: НТО РЭС им. А.С. Попова, 2012. - С. 20 - 23.
13. Клещарь С.Н., Котов В.В. Микрофильм как носитель изображений документов // XXX Научная сессия, посвященная Дню радио. - Тула: НТО РЭС им. А.С. Попова, 2012. - С. 207 - 211.
14. Клещарь С.Н., Котов В.В. Гистограммный анализ изображений на микрофильме // XXX Научная сессия, посвященная Дню радио. - Тула: НТО РЭС им. А.С. Попова, 2012. - С. 200 - 207.
15. Клещарь С.Н., Котов В.В., Аршакян А.А. Методика оценки статической передаточной характеристики сканера Известия ТулГУ. Технические науки. - Тула: Изд-во ТулГУ. 2013. Вып. 1. - С. 268 – 274.
16. Котов В.В., Клещарь С.Н., Котова Н.А., Будков С.А. Методы оценки качества сканирующих систем // Известия ТулГУ. Технические науки. - Тула: Изд-во ТулГУ. 2013. Вып. 1. - С. 299 – 306.
17. Аршакян А.А., Будков А.И., Клещарь С.Н. Информационные потери, связанные с пространственной динамикой сканера // Известия ТулГУ, Технические науки. - Тула: Изд-во ТулГУ. 2013. Вып. 3. - С. 383 – 388.
18. Аршакян А.А., Клещарь С.Н., Ларкин Е.В. Оценка статических потерь информации в сканирующих устройства // Известия ТулГУ, Технические науки. - Тула: Изд-во ТулГУ. 2013. Вып. 3. - С. 388 – 391.
19. Аршакян А.А., Будков С.А., Клещарь С.Н. Динамические потери информации в сканирующих устройствах // Вестник ТулГУ, Системы управления. Тула: Изд-во ТулГУ. 2013. Вып. 1. - С. 89 – 94.
20. Аршакян А.А., Будков С.А., Клещарь С.Н. Статические потери информации в сканирующих устройствах // Вестник ТулГУ, Системы управления. Тула: Изд-во ТулГУ. 2013. Вып. 1. - С. 94 – 102.

Подписано в печать 14.01.2013
Бумага офисная. Формат бумаги 60x84 1/16. Гарнитура Таймс.
Усл. печ. л. 1,0. Тираж 100 экз.
Печать офсетная. Заказ №

Отпечатано в ООО «Интермета»,
390000, Рязань, ул. Семинарская, д.3.