

LILIANE CHAVES DE RESENDE

**ABORDAGENS DA PRESERVAÇÃO DIGITAL E
ALGORÍTMOS DE REALCE PARA DOCUMENTOS HISTÓRICOS**

Trabalho de conclusão de curso apresentado ao Curso de Ciência da Computação.

UNIVERSIDADE PRESIDENTE ANTÔNIO CARLOS

Orientador: Prof. Marcelo Miranda Coelho

**BARBACENA
2003**

LILIANE CHAVES DE RESENDE

**ABORDAGENS DA PRESERVAÇÃO DIGITAL E
ALGORÍTMOS DE REALCE PARA DOCUMENTOS HISTÓRICOS**

Este trabalho de conclusão de curso foi julgado adequado à obtenção do grau de Bacharel em Ciência da Computação e aprovado em sua forma final pelo Curso de Ciência da Computação da Universidade Presidente Antônio Carlos.

Barbacena – MG, 03 de dezembro de 2003.

Prof. Marcelo Miranda Coelho - Orientador

Prof. Eduardo Macedo Bhering - Membro da Banca Examinadora

Prof^a. Lorena Sophia C. de Oliveira - Membro da Banca Examinadora

AGRADECIMENTOS

Agradeço a Deus a oportunidade de sempre recomeçar. O recomeço é sempre doloroso, mas traz infinitas mudanças na vida de cada um de nós.

Agradeço a Deus também pelo meu maior tesouro: minha família, sem ela hoje estaria numa situação que não consigo nem mesmo visualizar.

RESUMO

A motivação para realização deste trabalho veio a partir da necessidade de divulgar as preciosidades dos documentos históricos pertencentes ao Arquivo Eclesiástico da Diocese de São João del-Rei. Abordagens da Preservação digital de documentos históricos propõem as possibilidades oferecidas pelos recursos da tecnologia digital para preservação. Em seguida, apresenta os algoritmos de realce como uma das técnicas de processamento digital de imagens para atuar como ferramenta utilizada para melhorar a aparência dos documentos históricos digitalizados. Os testes desenvolvidos pelos algoritmos de realce demonstram que sua aplicação proporcionam uma melhora significativa na visualização dos documentos, possibilitando que seu conteúdo possa ser facilmente visualizado em relação ao documento original.

Palavras-chave: Preservação digital, processamento digital de imagens, algoritmos de realce

SUMÁRIO

<u>LISTAS.....</u>	<u>6</u>
<u>1 INTRODUÇÃO.....</u>	<u>8</u>
<u>2 ACERVOS DIGITAIS: CARACTERÍSTICAS E TENDÊNCIAS.....</u>	<u>10</u>
<u>3 DIGITALIZAÇÃO DE IMAGENS E ARMAZENAMENTO.....</u>	<u>15</u>
<u>4 INTRODUÇÃO AO PROCESSAMENTO DIGITAL DE IMAGENS.....</u>	<u>39</u>
<u>5 ALGORITMOS DE REALCE.....</u>	<u>49</u>
<u>6 CONCLUSÃO E TRABALHOS FUTUROS.....</u>	<u>75</u>
<u>REFERÊNCIAS BIBLIOGRÁFICAS.....</u>	<u>77</u>
<u>ANEXO A</u>	<u>79</u>
<u>PASSOS NECESSÁRIOS AOS ALGORITMOS DE REALCE.....</u>	<u>79</u>

LISTAS

FIG. 3.1 – IMAGEM DE VARREDURA. ADAPTADO DE [SCURI, 1999].....	21
FIG. 3.2 – RELAÇÃO ENTRE RESOLUÇÃO E IQ EM DOCUMENTOS TEXTUAIS. ADAPTADO DE [VALLE JR., 2003].....	23
FIG. 3.3 – MÚLTIPLOS FORMATOS DE CD'S. ADAPTADO [VALLE JR.; 2003].....	35
FIG. 3.4 – MÚLTIPLOS FORMATOS DE DVD'S. ADAPTADO [VALLE JR.; 2003].....	36
FIG. 4.1 – ELEMENTOS DE UM SISTEMA DE PROCESSAMENTOS DE IMAGEM, ADAPTADO DE [MARQUES FILHOS; VIEIRA NETO, 1989].....	41
FIG. 4.2 – O SISTEMA VISUAL HUMANO DETALHANDO [HTTP://WWW.CAA.UFF.BR/~ACONCI/SISVISI.HTM].....	44
FIG. 5.2 – FILTRO DA MÉDIA 3X3.....	53
FIG. 5.3 – FILTRO DA MÉDIA 3X3 COM K = 4 VIZINHOS MAIS PRÓXIMOS.....	54
FIG. 5.4 – FILTRO DA MODA 3X3.....	54
FIG. 5.5 – FILTRO DA ORDEM 3X3.....	55
FIG. 5.6 – FILTRO DA MEDIANA 3X3.....	55
FIG. 5.7 – FILTRO DA TAEX 3X3.....	56
FIG. 5.8 – FILTRO PASSA-ALTA.....	56
FIG. 5.9 – FILTRO PASSA-BAIXA.....	56
FIG. 5.10 – FILTRO DA ESCURECER 3X3.....	57

FIG. 5.11 – FILTRO DA CLAREAR 3X3.....	57
FIG. 5.12 – FILTRO DA TRATACLARA 3X3.....	58
FIG. 5.13 – FILTRO DA MÉDIACLARA 3X3.....	58
FIG. 5.14 – FILTRO REALÇARBORDA 3X3.....	58
FIG. 5.16 – IMAGEM COM CONTEÚDO APAGADO.....	62
FIG. 5.17– IMAGEM ENVELHECIDA COM AMARELADO CARACTERÍSTICO.....	63
FIG. 5.18– IMAEM ENVELHECIDA COM ESCURECIMENTO CARACTERÍSTICO.....	64
FIG. 5.19– IMAGEM ENVELHECIDA COM ESCURECIMENTO CARACTERÍSTICO.....	65
FIG. 5.20 – IMAGEM ENVELHECIDA E COM MUITOS BURACOS E RASGOS.....	66
FIG. 5.21– IMAGEM MUITO BORRADA.....	67
FIG. 5.22– IMAGEM CLARA.....	69
FIG. 5.23– IMAGEM CLARA.....	70
FIG. 5.24– IMAGEM CLARA.....	71
FIG. 5.25– IMAGEM CLARA.....	72
FIG. 5.26– IAGEM HOMOGÊNEA.....	73
FIG. 5.27– IMAGEM HOMOGÊNEA.....	74

1 INTRODUÇÃO

Vive-se em uma era influenciada pelo universo digital. Muitas são as influências impostas para se aderir à tecnologia digital. Quanto se vai a um banco, fica-se preso ou atrelado a sistemas computacionais que fazem a maior parte das tarefas, que eram realizadas por seres humanos.

Já se consegue hoje, realizar uma cirurgia onde médico e paciente se encontram em locais geograficamente diferentes, pelo simples uso dos recursos oferecidos pela tecnologia digital. Muitas instituições de ensino já estão aderindo às vantagens da tecnologia digital para ampliar as condições de aprendizado, colocando à disposição, para toda a comunidade que se interessar em aumentar seus conhecimentos culturais cursos e eventos on-line.

Da mesma forma, a preservação de documentos permanentes, objetivo de acervos documentais tradicionais, já está sendo avaliada e aplicada em termos do universo digital.

É neste contexto que se insere este trabalho. Como ponto inicial, serão abordadas as possibilidades oferecidas pelos recursos da tecnologia digital para a preservação.

Posteriormente, será apresentado a utilização de uma das técnicas do processamento digital de imagens, como ferramenta para melhorar a aparência de imagens digitalizadas.

O objetivo desse trabalho foi direcionado para a preservação de imagens de documentos históricos, com propostas para o armazenamento digital e de um tratamento inicial para os documentos permanentes do acervo documental do Arquivo Eclesiástico da Diocese de São João del-Rei. Não apresenta-se como um manual, mas como diretriz inicial na reformatação de documentos históricos existentes no formato tradicional para o universo digital.

A estrutura deste trabalho é composta por seis capítulos, que são descritos a seguir.

O Capítulo 2 aborda as características e tendências dos acervos digitais, bem como, a finalidade da preservação digital. Procurou-se, através de pesquisas selecionadas por autores da área, apresentar conceitos e justificativas para transformar as informações preservadas de acervos tradicionais para o formato digital.

No Capítulo 3 são apresentadas algumas opções para utilização dos recursos computacionais, cujo objetivo é a preservação das informações por acervos digitais de acordo com o objeto de preservação específico neste trabalho.

No Capítulo 4, são apresentados, resumidamente, os conceitos básicos das técnicas do processamento digital de imagens.

No Capítulo 5, é realizada uma abordagem do uso de algoritmos de realce como técnica de melhoramento da imagem.

Finalmente, o Capítulo 6 apresenta algumas conclusões e propostas de trabalhos futuros, que podem ser realizadas com documentos históricos digitalizados a partir do presente trabalho.

2 ACERVOS DIGITAIS: CARACTERÍSTICAS E TENDÊNCIAS

A importância de avaliar as possibilidades oferecidas pela tecnologia digital é fundamental, pois no “universo digital, tecnologias de formação de imagem digital deixam instituições em uma situação similar àquela oferecida pelo rádio, na comunidade mundial, na década de 60”. [CONWAY, 2001]

Acervos digitais atuam como uma das importantes possibilidades da tecnologia digital para guardar e analisar a história e a cultura do ser humano. Sua tarefa principal é atuar com responsabilidade na preservação da memória documental na história de uma sociedade. Entretanto, essa tarefa não se faz tão simples. É necessário que se obtenha uma forma sistemática e um severo compromisso de conservar e garantir a otimização ao acesso das informações, sem que estas se percam pela fragilidade dos artefatos computacionais. [VALLE JR.; 2003]

É importante ressaltar que a tecnologia digital ainda não possui uma total e segura afirmação de que a informação digitalizada é, e continuará sendo por séculos, a tecnologia que melhor se aplica no sentido da preservação.

A respeito do que existe e é utilizado pelos acervos tradicionais para a preservação de documentos em papel degradados, ou em processo de degradação, não é o último passo

para a preservação. É apenas um passo intermediário no processo para a utilização da tecnologia digital. [WATERS, 2001]

A tecnologia digital ainda carece de soluções no sentido de armazenamento e acesso à informação digitalizada. Seus suportes são bem mais frágeis e perecíveis do que sua contraparte analógica. Além disso, os componentes de hardware e software se tornam rapidamente obsoletos, tanto para visualização quanto para manipulação das informações. Por causa disso, transformam-se numa ameaça constante de perda de acesso aos acervos digitais.

Devido a essas questões ligadas a tecnologia digital, cientistas e profissionais da área de preservação acreditam que a combinação das tecnologias tradicional e digital é uma alternativa viável que pode resolver tanto questões de armazenamento para preservação a longo prazo, quanto questões de otimização de acesso e manipulação dos dados digitalizados.

2.1 ACERVOS CORRENTES X ACERVOS PERMANENTES

É importante entender o contexto vital de um documento dentro da instituição que possui a sua custódia. Esses conceitos servem como base na diretiva a ser dada ao se colocar um acervo documental no formato digital.

Eduardo Valle Jr. [VALLE JR.; 2003] apresenta o conceito de arquivo corrente e arquivo permanente, de acordo com as normas da arquivística convencional. Estes conceitos são considerados a partir de duas fases importantes e distintas da vida dos documentos:

- *fase corrente*: corresponde à etapa de criação, modificação e tramitação do documento. Nesta fase registram-se todas as informações do desenvolvimento das tarefas desempenhadas numa instituição;
- *fase permanente*: corresponde à etapa de recolhimento e arquivamento do documento. Nesta fase, o documento julgado importante passa a estar disponível apenas para consulta. Essa fase é a mais importante e que concentra os conceitos tradicionais de preservação.

2.2 ACERVOS DIGITALIZADOS

As vantagens oferecidas pela tecnologia digital (como, por exemplo, melhoria de serviços de acesso e manipulação de informações digitalizadas, redução de custos, etc.), juntamente com a tendência ao seu uso, têm sido uma motivação para museus, bibliotecas, arquivos, a disponibilizarem seus acervos em formato digital [WATERS, 2001].

Eduardo Valle Jr. [VALLE JR.; 2003] apresenta uma lista de algumas instituições internacionais e nacionais que mais se destacaram ao aderir a tecnologia digital:

- ***The Making of América*** – coleção que inclui mais de 1,5 milhões de páginas de texto digitalizadas e disponibilizadas na WWW¹. Este projeto registra a história norte-americana do século XIX, levando a um segundo projeto - *The Making of América II* - que participaram não só as Universidade de Michigan e Cornell como no primeiro, mas também a Universidade da Califórnia, Stanford, Penn State e a Biblioteca Pública de Nova Iorque. Este segundo tinha como foco coleções acerca de transportes da época da Corrida do Ouro. Baseou-se num complexo conjunto de metadados para permitir o acesso direto a versões digitalizadas das fontes primárias.
- ***The Minnesota Historical Society*** – Uma instituição privada, cultural e sem fins lucrativos, com um acervo de caráter misto de biblioteca, coleções museológicas e fundos arquivísticos. Baseou-se numa iniciativa de proteção de registros eletrônicos e uma preocupação acerca da normalização do uso de metadados para identificação, preservação e acesso aos registros governamentais.
- ***Open Archives Initiative e Eprints*** – Repositórios de trabalhos acadêmicos e artigos científicos. Baseou-se na padronização dos metadados de descrição. Têm sido considerados como projetos pilotos ideais para a implantação de bibliotecas digitais.

¹ Word WideWeb, rede mundial de computadores.

- ***Biblioteca do Congresso Norte-americano*** – Importante acervo documental dos Estados Unidos com gigantescas coleções bibliográficas. Oferece apenas uma pequena fração em formato digital: 7,5 milhões de itens digitais, amostrados de 100 coleções históricas. Baseou-se na diversidade de pesquisas para diferentes tipos de usuários e compartilhamento de metadados com outras importantes bibliotecas digitais.
- ***Projeto Nabuco*** – Projeto promovido pela cooperação da Fundação Joaquim Nabuco e o Departamento de Informática da Universidade Federal do Paraná, o qual disponibiliza documentos compostos de fotografias, cartas e escritos do importante político recifense. Baseou-se na preocupação com o aspecto de compressão de imagens e reconhecimento de texto, dando ênfase à interface de busca por palavra-chave.
- ***Biblioteca Nacional Brasileira*** - Oferece parte de seu acervo em formato digital. Baseou-se na diversidade de pesquisas, disponibilizando um acesso contextualizado a objetos digitais.

2.3 A PRESERVAÇÃO DIGITAL

De acordo com Hedstrom [LEONE, 2002], entende-se por preservação tradicional “...assegurar proteção à informação de valor permanente para acesso pelas gerações presentes e futuras.” E ainda citando Webb “... parte do trabalho central de uma instituição arquivística que busca manter suas coleções para uso permanente, protegendo-as de ameaças, ou salvando-as e reparando-as para compensar seus impactos”.

A proposta quanto à preservação digital é a mesma, diferenciando-se apenas pela metodologia empregada. De acordo com a definição anterior, pode-se ampliar o entendimento à preservação digital como sendo “... o planejamento, alocação de recursos e aplicação de

métodos e tecnologias para assegurar que a informação digital de valor contínuo permaneça acessível e utilizável.” [LEONE, 2002]

Assim como conservar um documento na forma de papel exige inúmeras considerações quanto aos métodos e diretivas de preservação, a preservação digital passa por dificuldades nas questões do desenvolvimento de padrões e soluções quanto à escolha dos formatos digitais, as técnicas de migração e emulação das bases de documentos digitais e o desuso dos componentes físicos e lógicos utilizados para preservação digital.

O que se considera mais importante no tocante à preservação digital é avaliar as possibilidades oferecidas pela tecnologia digital, para então escolher aquela que melhor se encaixe no contexto da preservação tradicional.

No universo digital, torna-se possível a preservação. Mas é importante, que se entenda, que no contexto digital, preservação deve ser relacionada “a um sistema como um todo e não a um determinado componente, tal como um filme ou a armazenagem de um disco”. Ela é uma interligação de software, hardware e subsistemas de serviço [CONWAY, 2001].

A opção digital deve estar inserida num contexto onde seja possível o reaproveitamento do que já se tem como tecnologia de preservação, sem que com isso haja perda de trabalho ou insegurança quanto ao objeto de preservação das instituições.

3 DIGITALIZAÇÃO DE IMAGENS E ARMAZENAMENTO

A digitalização de acervos históricos pode ser realizada pelo uso de sistemas de informação voltados especificamente para a preservação digital. Muitos desses sistemas de gestão documental já estão disponíveis, porém mais voltados para os documentos em fase corrente do que para os documentos em fase permanente. Isso pode ocorrer devido às normas e exigências da qualidade arquivística segundo a definição tradicional [VALLE JR.; 2003].

Para sistemas de preservação digital, a produtividade será refletida de acordo com o avanço tecnológico. Mesmo assim, um sistema de informação para gerenciar imagens digitais tende a ser a solução mais promissora para documentos impressos em papel [WILLIS, 2001].

Os sistemas de informação para preservação de documentos históricos permanentes podem se separar em duas categorias. A primeira se refere à preservação de artefatos de origem digital, que mesmo apresentando grandes desafios e complexas soluções, é de fundamental importância pela crescente gama de produções técnicas, intelectuais e artísticas que vêm sendo criados digitalmente. A segunda é o uso de digitalização como técnica de reformatação aplicada à preservação do conteúdo legal de documentos históricos,

motivada por facilidades de acesso, recuperação dos dados digitais e cópias fiéis e de alta qualidade dos documentos. Porém precisam atender às exigências da qualidade arquivística ou então atuar em parceria com as mesmas.

Não se pretende com este trabalho aprofundar em conceitos voltados para sistemas de informação para preservação, mas apresentar os pontos mais relevantes na fase inicial da digitalização de documentos, desde os aparatos computacionais físicos para aquisição e armazenamento da imagem digital até o uso de técnicas aplicadas ao seu melhoramento visual.

3.1 PRESERVAÇÃO DE DOCUMENTOS DE ORIGEM DIGITAL

O computador digital tem origem recente, sendo seu uso popularizado a partir de 1960. Contudo, dados estatísticos realizados por estudiosos apontam que, nos dias atuais, a maior partes dos trabalhos de impressão e escrita de textos são feitos por computador. As gravações profissionais de som são quase todas digitais. As câmeras fotográficas com filmes analógicos estão sendo substituídas por câmeras fotográficas digitais. Enfim, com o fato de a tecnologia ser muito recente e estar evoluindo com muita rapidez, traz consequências drásticas para a preservação dos documentos criados sob sua égide. [VALLE JR.; 2003]

Neste sentido, a microfilmagem e, mais recentemente, a digitalização, são técnicas adotadas para atuarem em conjunto, podendo tornar-se a melhor solução frente ao objetivo de preservar e conservar documentos de origem históricas em processo de degradação.[LEONE; 2002]

Profissionais e estudiosos, frente às possibilidades de aderir à tecnologia digital, sugerem que os conceitos da preservação vêm assumindo, pelo menos, três diferentes significados [CONWAY, 2001]:

- *Possibilitar o uso:* documentos deteriorados com o tempo podem ser digitalizados, melhorados e viabilizados para consulta, através de recursos oferecidos pela tecnologia digital;

- *Proteger o item documental:* documentos digitalizados podem ser copiados, em alta qualidade, através dos recursos da tecnologia digital, tal qual o documento original e deixados à disposição dos consulentes, como recursos de preservação dos originais;
- *Manter os objetos digitais:* o desafio de proteger cópias da deterioração, ou mesmo da destruição, torna-se o foco principal do trabalho de preservação.

Focado na tecnologia digital e na alta velocidade em que esta se insere, deve-se dar muito mais atenção aos problemas causados pela obsolescência dos componentes digitais. Entende-se por obsolescência, no contexto deste trabalho, o desuso dos componentes de hardware ou de software, necessários para resguardar os documentos digitais de sua degradação e para evitar perda de acesso.

De acordo com [VALLE JR.; 2003], uma forma encontrada para contornar, parcialmente, estes problemas foi proposta pela *Commission on Preservation and Access* e o *Research Libraries Group*, os quais consideram que a preservação dos dados digitais fica sujeita aos seguintes fatores:

- *Conteúdo:* considera complexa e com diferentes níveis de abstração a preservação do conteúdo digital.;
- *Imutabilidade:* considera fundamental a segurança dos documentos digitais, para que se possa salvaguardar a autenticidade do registro digital;
- *Referência:* considera importante a criação de mecanismos consistentes de referência aos dados, e que estes possam estar ligados a identificadores e recuperados através deles;
- *Proveniência:* considera importante a integridade da informação digital, estabelecendo o registro de sua origem e cadeia de custódia;
- *Contexto:* considera importante a preservação do contexto da informação digital como principal requisito para manter sua integridade. No que se

refere à informação digital, é necessário também um contexto técnico que possa fazer referência ao hardware e software.

Já para a obsolescência dos artefatos da tecnologia digital, as políticas das instituições e da indústria de informática, bem como o comportamento dos usuários, não facilitam a tarefa de preservação, levando a problemas severos da fragilidade de suporte ao acesso e manipulação dos dados digitais. Dessa forma, esses problemas foram caracterizados da seguinte maneira [VALLE JR.; 2003]:

- *Visualização*: os dados digitais são visualizados em software específico, e um hardware é necessário para que se possa executar este software. Não basta apenas preservar os dados, é preciso também preservar todo os componentes de hardware e software necessários à visualização e manipulação desses dados;
- *Embaralhamento*: a necessidade de resolver problemas a curto prazo, no uso da informação digital, acabam resultando em problemas de preservação. Este fato pode ser exemplificado pelo uso de técnicas de compressão e criptografia. Usando-se a compressão é necessário guardar não só o software de compressão mas também o de descompressão. Na criptografia, além da proteção dos dados, tornando-os incomprensíveis, é preciso guardar também a chaves de decodificação fornecidas aos usuários, que facilmente podem ser perdidas;
- *Inter-relação*: dados digitais cada vez mais são inter-relacionados, através de recurso de incorporação e hiperligação. Como exemplo, a disponibilização dos dados digitais pelas páginas da web, compostas por diversos arquivos, que muitas vezes residem em diversos computadores distantes entre si. Isso pode ocasionar na dificuldade de delimitar os itens de informação e de seus contextos;
- *Custódia*: a maior parte dos dados digitais ainda não possui atribuições legais dos responsáveis pelas custódias das informações. Isso porque as instituições responsáveis focalizam mais manutenção e preservação dos vários tipos de materiais analógicos do que os digitais;

- *Tradução*: a migração dos dados para o formato digital pode ocasionar uma perda de contexto real da informação, fazendo com que erros e imprecisões de tradução sejam constantes e inoportunos.

De acordo com que tem apresentado a bibliografia, fica comprovada a necessidade de uma estratégia de proteção aos mecanismos de manipulação, acesso e visualização dos dados digitais, sujeitos a se tornarem indisponíveis pela obsolescência dos artefatos computacionais.

Mas apesar da enorme dificuldade de se manter a longevidade digital, sistemas de informação proporcionam certos benefícios que se tornam atraentes e chegam a compensar os riscos de sua utilização como meios, ou mesmo um papel auxiliador, na preservação dos acervos históricos. Estes benefícios são sugeridos por [VALLE JR.; 2003], da seguinte maneira:

- *Na perfeita replicabilidade dos dados digitais*: as cópias dos dados digitais são perfeitamente fiéis ao original;
- *Na facilidade de manipulação e análise dos dados digitais por causa da sua natureza numérica*: os dados digitais podem ser aprimorados por técnicas utilizadas no processamento digital de imagens, tais como: operações de realce e restauração. Disponíveis para imagens analógicas somente através de morosas e delicadas operações de laboratório;
- *Na possibilidade de resguardar os documentos originais da manipulação desnecessária*, postergando sua deterioração;
- *No oferecimento de uma forma mais conveniente de representar seu conteúdo intelectual*: uma vez que os documentos a serem preservados podem ser de natureza bastante complexa, como por exemplo, preservação de sítios arqueológicos;

Na utilização de sistemas digitais para fins de preservação, deve-se considerar três áreas de análise primária: captura do documento, armazenamento e acesso. Na captura, considera-se o mecanismo de captura, a qualidade da imagem, a resolução, a velocidade de aquisição, o custo do sistema, o custo operacional e a necessidade de indexação. No

armazenamento considera-se o meio físico a ser utilizado, bem como, sua necessidade de renovação e atualização do mesmo, sua obsolescência, seu custo, sua confiabilidade, sua performance e suas referências ao acesso da informação. No acesso deve-se considerar a capacidade, qualidade e velocidade de recuperação da informação, a capacidade de transmissão e a distribuição. Enfim, com o que se tem de sistemas micrográficos e um bom planejamento de um sistema digital, voltado para preservação, podem, juntos, contemplar e suprir as necessidades das áreas de análise primárias da melhor forma. [WILLIS, 2001]

No que concerne à tecnologia digital, uma abordagem ampla, no que se refere à digitalização de imagens, vem sendo difundida pelo grande número de pesquisadores e estudiosos da área. Assim, vêm sendo apresentadas propostas de mecanismos para resolver questões quanto a formatos de armazenamento, captura, qualidade para visualização, acesso e recuperação das informações digitalizadas, seleção de componentes de hardware para preparação dos documentos, objetivando a preservação digital. [VALLE JR.; 2003]

3.2 A NATUREZA DAS IMAGENS DIGITAIS

A imagem é uma mídia que possui um conteúdo semântico que precisa ser descrito. Para isso, são importantes e necessários, o estudo e o entendimento do que venha ser uma imagem digital. [SANTOS, 1995]

3.2.1 IMAGENS DIGITAIS

A imagem digital possui uma maneira convencional de ser representada e armazenada, denominada imagem de varredura. Ela é formada por uma grade bidimensional, onde o encontro de uma linha com uma coluna é denominado pixel, ou mais especificamente, como um elemento da imagem (Figura 3.1). O pixel possui valores de tonalidade no caso de imagens coloridas, e de luminosidade no caso de imagens monocromáticas.

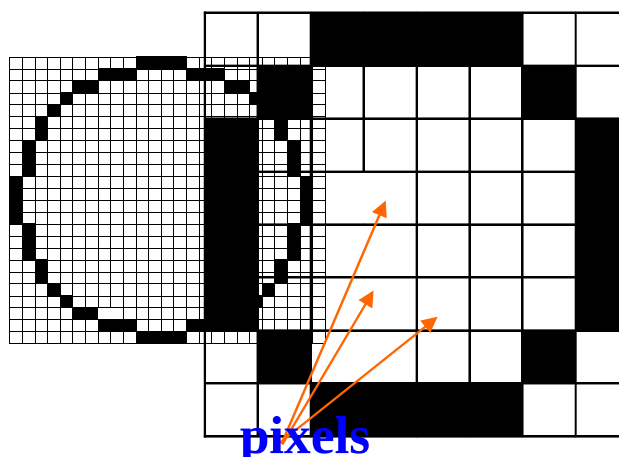


Fig. 3.1 – Imagem de Varredura. Adaptado de [SCURI, 1999]

Imagens monocromáticas têm suas características voltadas para o valor numérico de cada pixel, referente à luminosidade que vai desde o preto até o branco. Nos monitores de vídeo, utilizados como padrão de visualização de imagens digitais, a ausência de luz se refere à cor preta e a presença de luz total se refere à cor branca. Os estágios intermediários, que vão do preto ao branco, representam o percentual limítrofe de luminosidade, apresentando tons de cinza na prática.

Para imagens coloridas, a incidência de luz para gerar as cores é feita através da combinação de três faixas de comprimento de onda das cores primárias: vermelho, azul e verde. É necessário a informação numérica não apenas da luminosidade, mas também da tonalidade de cada amostra. Isto é, a informação dos diferentes comprimentos de onda combinados num raio luminoso de cada amostra.

A tecnologia de imagem digital propicia a codificação da imagem, digitalmente capturada, de documentos para armazenamento, acesso e recuperação por sistemas computadorizados. [WATERS, 2001]

3.2.2 RESOLUÇÃO

A imagem ao ser digitalizada assume um tamanho adimensional em pixels e pode-se conhecer sua medida de qualidade conhecendo-se a razão entre o número de pixels obtido e o tamanho da imagem real. Denominamos a esse procedimento de resolução. Normalmente, a resolução é medida em DPI, ou seja, *dots per inch* – pontos por polegada [SCURI, 1999].

A escolha da resolução é a primeira decisão a ser tomada em um projeto de digitalização de documentos em acervos. Obviamente, quanto maior, a resolução melhor a qualidade da imagem digitalizada e, conseqüentemente, maior será o tamanho do arquivo gerado pela digitalização, requerendo um sofisticado componente computacional de hardware que armazene este arquivo. [VALLE JR.; 2003]

Para documentos onde o valor visual é importante, é preciso que sua captura seja feita guardando o máximo possível de detalhes da imagem original. E nestes casos, utiliza-se uma resolução igual ou superior à resolução do original.

Sugere-se uma resolução de no mínimo 300 dpi (pixels por polegada) e cor de 24 bits para documentos manuscritos de “qualquer tipo”. Contudo, é apenas uma sugestão, pois poucas foram as definições precisas acerca da resolução específica na digitalização de documentos. Alguns fatores devem ser considerados ao se determinar resolução de imagens. [KENNEY, CHAPMAN; 2001]

- *Quanto ao acesso para visualização:* Se a visualização for monitor de vídeo, uma resolução de 450dpi ultrapassa a necessária, que é de 150dpi;
- *Quanto ao acesso para reprodução:* uma possibilidade é aumentar a qualidade da imagem, por exemplo, 600dpi para reprodução, se não for a imagem matriz de digitalização. Não há necessidade de maior resolução, pois a imagem ficará limitada à resolução máxima do componente de reprodução.

Sugere-se ainda uma adaptação através de um método de medição de qualidade em microfilmes, denominado *índice de qualidade* - IQ, para estabelecer requisitos de resolução para imagens de texto. [KENNEY, CHAPMAN; 2001]

Tanto na microfilmagem quanto na tecnologia digital, o IQ é baseado na legibilidade do texto e na resolução do sistema, ou seja, a habilidade de capturar detalhes finos dos documentos. O IQ é utilizado para predizer níveis de qualidade da imagem, que podem ser: marginal (ou baixo), médio e alto. Este método relaciona a altura do menor caractere de interesse no texto, normalmente a letra “e” minúscula, e a resolução do documento. Segundo eles, se o menor caractere for formado por um número razoável de pontos por polegada, então o texto, como um todo, será legível. A Figura 3.2, apresenta algumas resoluções ideais, de acordo com os diferentes tamanhos dos tipos (caracteres), para atingir o IQ desejado:

Altura do Menor Tipo H (em mm)	Resolução (em dpi) necessária ao IQ		
	<i>Marginal (3,6)</i>	<i>Médio (5,0)</i>	<i>Alto (8,0)</i>
0,5	366	508	813
1,0	183	254	406
1,5	122	169	271
2,0	91	127	203
2,5	73	102	163

Fig. 3.2 – Relação entre Resolução e IQ em Documentos Textuais. Adaptado de [VALLE JR., 2003]

É importante lembrar que, a resolução aplicada ao cálculo de índice de qualidade é efetivamente obtida pela resolução de saída da captura da imagem. Ou seja, a resolução é menor do que a oferecida pelo componente de captura devido aos chamados erros de registros, ocasionados durante a captura, tais como: problemas de foco, trepidação mecânica, e muitos outros, resultando numa resolução inferior à nominal.

3.3 CONSIDERAÇÕES SOBRE O DE USO DAS IMAGENS

A primeira consideração a ser feita, referente ao uso das imagens digitais, é seu propósito. Esta consideração é a mais importante porque ela pode definir os componentes de captura, resolução, armazenamento e o formato de arquivo. [VALLE JR.; 2003]

Alguns propósitos de imagens digitais são:

- *Índice*: imagens com pouca resolução e de pequenos tamanhos, utilizadas como miniaturas em resultados de pesquisa ou em listagens impressas;
- *Acesso*: imagens com uma resolução um pouco melhor, apresentando detalhes mais ricos e próximos à imagem original. Utilizadas para visualização, mas não substituem a aparência da original;
- *Reprodução*: imagens da alta qualidade e resolução, utilizadas como recurso digital para duplicação fiel da imagem original. Proporcionam uma aparência satisfatória em relação à imagem original;
- *Preservação*: imagens da mais alta qualidade e resolução possível, utilizadas para guardar a aparência idêntica da imagem original. Sua captura é feita com a maior riqueza de detalhes, preservadas para possíveis aplicações futuras.

Outras considerações devem ser feitas, tais como: a análise que se espera fazer com a imagem digital; o custo da preparação e remoção do material a ser digitalizado, que acaba sendo incerto, devido à natureza da digitalização; como também, um esquema que proporcione uma captura de imagens com a maior qualidade possível, visando A criação de uma imagem matriz, chamada de *master* ou matriz digital de preservação.

A *master*, por ser um arquivo grande demais para ser acessado, será armazenada separadamente do sistema de informação. Este irá manipular essa imagem em dispositivos terciários, como uma fita magnética ou um disco óptico, por exemplo, e só será acessada para dar origem a novas imagens de manipulação e visualização, caso seja necessário.

3.4 DIGITALIZAÇÃO

A digitalização de documentos, no contexto de preservação de documentos históricos, tem uma importância ímpar no que diz respeito a evitar o manuseio. Consequentemente a degradação, sem impedir o acesso à cópia fiel dos mesmos.

A digitalização de documentos é uma tecnologia inovadora, e vem crescendo velozmente, a partir da necessidade das instituições de preservar as informações originais em papel. O processo de digitalizar imagens é um dos recursos oferecidos pela tecnologia digital.

A qualidade da imagem digitalizada é determinada pela qualidade dos periféricos específicos de captura, tais como scanner's ou câmeras digitais. Após efetuação de sua "leitura", os documentos capturados são convertidos para a forma digital, para serem, posteriormente, manipulados e acessados pela tecnologia digital específica.

Normalmente, as imagens digitalizadas são armazenadas em aparatos computacionais convencionais através de componentes de hardwares específicos, como por exemplo, fitas magnéticas, discos rígidos, discos ópticos, etc. São acessadas e recuperadas através de sistema de gerenciamento de documentos ou um sistema gerenciador de base de dados.

O fato da longevidade ser um fator preocupante na preservação de acervos documentais proporcionou à digitalização não ser um método de reformatação tão confiável, devido a obsolescência dos aparatos computacionais. A associação do uso da microfilmagem junto com a digitalização, pela tecnologia digital, vem resolvendo a questão problemática da preservação e acesso aos dados digitais. [VALLE JR.; 2003]

O microfilme é mais seguro para a preservação, chegando a durabilidade de 500 anos, e a digitalização possibilita maior facilidade ao acesso e manipulação das informações. Utilizando essa metodologia, [WILLIS, 2001] apresenta três formas de realizar a combinação da microfilmagem e a digitalização:

- *Digitalizar o material e imprimir a imagem digital em microfilme:* esta abordagem possibilita que a imagem digital seja processada antes que seja gerado o filme. Vem se tornando uma alternativa aceitável, porém a resolução da imagem fica presa à do componente utilizado na captura. A vantagem é que a digitalização sendo feita primeiro pode-se guardar as referências de cores e realçar os dados digitais;
- *Digitalizar e microfilmar os materiais simultaneamente:* esta pode resolver o problema da abordagem anterior: capturar a imagem com altíssima qualidade através da microfilmagem e guardar informação de cor aplicando técnicas de realce para melhorar o visual da imagem original, através da digitalização. O que ainda é um empecilho para esta abordagem é que a tecnologia ainda não é desenvolvida para propósitos de acervos permanentes, uma vez que os equipamentos atuais têm qualidade apenas para microfilmagem tradicional e comercial;
- *Microfilmar o material e digitalizar o microfilme:* Esta é ainda a abordagem mais utilizada porque permite a geração de um filme com altíssima resolução e, posteriormente, é feita a digitalização semi-automatizada. A maior desvantagem é o descarte da informação de cor, uma vez que a captura de filmes é monocromática.

A microfilmagem para posterior digitalização deve preparar a imagem, capturando-a da maneira o mais fiel possível e sem fazer nenhuma alteração de melhoria da legibilidade. A legibilidade é uma das soluções implementadas pela tecnologia digital. Mas o passo mais importante na microfilmagem deve ser a indexação do material num banco de dados com a correspondência entre as páginas dos documentos e os fotogramas do filme, para que sejam associados automaticamente, durante digitalização. Esse passo, também facilita a recuperação da imagem quando for necessária a pesquisa do documento original armazenado no microfilme. [VALLE JR.; 2003]

3.5 CAPTURA DE IMAGEM DIGITAL

A captura da imagem digital é feita por dispositivos de leitura de dados analógicos que a transforma em dados digitais. Os dispositivos mais conhecidos e utilizados para a digitalização são os scanner's e a câmera digital.

3.5.1 SCANNER

O funcionamento do scanner é semelhante ao da fotocopadora de mesa. O documento a ser digitalizado fica sobre uma superfície de vidro, virado para baixo, enquanto é varrido por uma luz, cuja lente focaliza a imagem em uma superfície fotossensível, composta por centenas de fotodetectores denominados CCD – *Charge-coupled Device*. A quantidade de CCD por polegada determina a resolução da imagem escaneada. A resolução determinada pelo scanner é estabelecida por pontos por polegada, dpi – *dots per inch*, onde cada ponto se correlaciona com um CCD. Quanto maior a quantidade de pontos maior a exatidão da imagem escaneada. Um scanner padrão possui capacidade de captura de 300 a 400 dpi. Existem vários tipos de scanner's: [OGDEN e OUTROS; 2001]

- ♦ *Scanner de Mesa:* são os mais conhecidos, pois são amplamente utilizados no mercado, cujos modelos são direcionados ao uso doméstico ou comerciais. A qualidade oferecida por estes modelos é de pouca fidelidade;
- ♦ *Scanner de Filmes:* desses modelos são oferecidos duas variações: de uso amador e de uso profissional, que podem atingir altíssima qualidade na captura de negativos. Podem detectar, durante a captura, alguns defeitos de captura tais como marcas de dedos e poeira;
- ♦ *Scanner de Microfilmes:* são modalidades de scanner's preparados para lidar com as inúmeras variações de microfilmes (rolos de filme não perfurado, filmes de alto contrastes e outros). Estes modelos simplesmente fazem a captura das informações sem nenhuma trabalho adicional, do tipo remoção de marcas de dedos e poeira.

- ♦ *Scanner de Tambor*: é o mais sofisticado das modalidades de scanner's. Utilizado para capturar informações com altíssima riqueza de detalhes, precisão e fidelidade. Podem atingir uma resolução de até 11.000 dpi. Muito utilizado na captura de documentos degradados para a preservação.

De acordo com a bibliografia, nem sempre os scanner's são indicados para captura de documentos históricos em processo de deterioração. Devido à incidência de luz sobre os documentos na captura para a informação digital, aumentam o nível de degradação dos documentos, aumentando também a deterioração do papel, há muito oxidado.

3.5.2 CÂMERA DIGITAL

As câmeras digitais, como nos scanner's, possuem sensores CCD que atua como um filme fotográfico, porém com uma resolução bastante acima das máquinas fotográficas tradicionais, cerca de 5 MP (milhões de pixels). A captura da informação é rápida e de imediato acesso [SCURI, 1999].

À medida que a tecnologia digital vai evoluindo, através de recursos de transmissão de dados, realce de imagens e os computadores mais rápidos e potentes, as câmeras digitais vêm surgindo como um dos meios mais utilizados para a captura do objeto arquivístico. Isso é devido à sua alta capacidade e qualidade de captura das informações e de um pré-processamento embutido nesses dispositivos, como controle de contraste de luz e definição do formato de arquivo. [KENNEY, CHAPMAN; 2001]

3.6 TIPOS DE FORMATOS DE IMAGENS

Após definir o propósito e a captura, é necessário definir o formato de armazenamento do arquivo da imagem digital. Para fins de preservação, é necessário escolher um formato que atenda, no mínimo, às seguintes características:

- *um formato compatível*, bem documentado e de caráter aberto, ou seja, um formato cujas especificações técnicas são amplamente conhecidas, e que preserve fielmente as informações coletadas na captura;
- *um formato que permita o acréscimo de informações à imagem digital*, no mínimo com os parâmetros utilizados na digitalização e uma identificação do documento original;

Entretanto, de acordo com [VALLE JR.; 2003], não existe um formato que atenda plenamente a essas especificações.

Um formato que se aproxima do exposto acima é o TIFF – Tagged image / Interchange File Format, ou seja, formato de arquivo etiquetado / formato de intercâmbio de arquivo de imagem, de propriedade da Adobe Systems. A empresa divulga amplamente sua documentação e especificação técnica. O formato TIFF permite armazenar informações da imagem através do conceito de etiquetas, que são informações que podem ser acrescentadas junto a um cabeçalho de arquivo. Além de preservar a integridade dos dados da imagem, como a tonalidade de cor, por exemplo, tem uma boa aceitação na sua utilização por uma vasta comunidade de pesquisadores e usuários.

Por ser tão flexível, torna-se um problema no que se refere à compatibilidade. Nem todos os sistemas conseguem manipular arquivos do formato TIFF. Na tentativa de solucionar este problema, é utilizado um formato da família TIFF, o TIFF/IT. Mas mesmo assim, o TIFF na sua forma mais avançada é normalmente o mais escolhido para a preservação digital.

Para a manipulação e acesso rápido da imagem digital, o formato escolhido normalmente é o PNG – *Photographic Network Graphics*, formato de gráficos portáteis de rede. O PNG, variação do formato GIF, é de propriedade da W3C – *World Web Consortium*, Consórcio da Cadeia de Escala Mundial, que publicou suas especificações técnicas. Possui um formato comprimido e menor variação de implementação e menor compatibilidade em relação ao TIFF. Entretanto, permite que informações textuais possam ser embutidas dentro do arquivo. É um formato comprimido sem perdas para diminuir o tamanho da imagem, por

isso pode ser reversível e não altera a sua apresentação visual. Porém afeta a longevidade digital pela incompatibilidade que o formato pode oferecer.

Um outro formato para manipulação é o JFIF – formato de arquivo para imagem padrão JPEG – Joint Photographic Experts Groups (grupo de desenvolvimento de arquivo comprimido). É um formato comprimido e com perdas, chegando a ser de 10 a 50 vezes menor que o arquivo original. Este formato descarta parte da informação do arquivo original por causa da sua compressão, mas ainda assim é uma boa escolha para manipulação e apresentação ao sistema visual humano. [MARÇAL, 2001]

O fato do formato JFIF ser do tipo com perdas torna irreversível o retorno às características originais, fazendo com que este seja utilizado mais para visualização e acesso. O que não acontece com o PNG, que mesmo comprimido, é reversível ao retorno das características originais da imagem. Por isso este formato é utilizado tanto para acesso quanto para reprodução.

3.7 ACESSO

Na arquivística tradicional, as idéias de preservação e acesso são distintas entre si. Uma vez que, a microfilmagem não resolve as questões de acesso, mesmo resolvendo as questões de preservação e da longevidade. O acesso é sequencial e é necessário um treinamento de um operador para manipular os componentes físicos necessários à visualização dos microfimes. Já na digitalização ocorre, em teoria, o oposto. Possibilita uma facilidade no que tange o acesso, a recuperação da informação e visualização, apesar de possuir uma longevidade bem pequena devido à rápida obsolescência dos aparatos computacionais.

Neste contexto, a aplicação da tecnologia digital transforma radicalmente o compromisso entre preservação e acesso, pois faz com essas dimensões se tornem cooperantes e atuem em conjunto.

Paul Conway [CONWAY; 2001] contradiz a arquivística tradicional, argumentando sobre questões acerca da tecnologia digital. Ele afirma que “a preservação no universo digital descarta toda e qualquer noção dúbia que entenda a preservação e acesso como atividades

distintas”. Sendo assim, uma eficaz atividade estratégica de recuperação da informação é a mais importante para que a digitalização possa ser considerada uma atividade de preservação. Esta estratégia deve ser bem estruturada, classificando e indexando bem as informações digitais. Uma base de dados que não esteja bem estruturada pode resultar numa grande quantidade de informações amorfas que contribui para o empobrecimento da manipulação dos dados digitais.

Vale ressaltar, que a idéia de preservação de acervos, há muito pregadas por pesquisadores e estudiosos da área, diz que “o objetivo da política de arquivos em um país democrático não pode ser a simples guarda de documentos. Deve ser nada menos que promover o enriquecimento da consciência histórica dos povos como um todo”. [VALLE JR.; 2003]

Enfim, um acervo de dados digitais deve contar com mecanismos de arranjo e indexação de dados e, até mesmo, com a criação de instrumentos de pesquisa que garantam a recuperação da informação de forma mais coerente.

3.8 PRESERVAÇÃO E ARMAZENAMENTO

A maior vantagem da preservação, advinda da natureza da tecnologia digital, é a replicabilidade das informações digitais pela sua natureza numérica. Obviamente, alguma imperfeição de um documento original também seria vista no documento digital. Em contrapartida, como já foi exposto anteriormente, o maior problema da tecnologia digital para a preservação é a obsolescência dos aparatos computacionais que interpretam os dados digitais. [VALLE JR.; 2003]

O propósito do armazenamento, no contexto da preservação digital, é propiciar a utilização de um aparato físico, ou mais de um, de alta capacidade e confiável para os dados digitais. Isso deve ocorrer tanto para acesso e recuperação da informação quanto para a reprodução. [WATERS; 2001]

Nesta seção, serão explanados os três suportes ao armazenamento da informação digital mais utilizados atualmente pelos acervos digitais.

3.8.1 SUPORTES DA INFORMAÇÃO DIGITAL

Ao contrário do que se tem como informação no ambiente mercadológico, estudos apontam que a maioria das mídias digitais possuem uma vida útil tão baixa que pode não chegar a 5 anos de durabilidade.

As mídias digitais, como componentes de armazenamento, são consideradas, por natureza, muito frágeis e seus delicados recursos de leitura e gravação podem condenar a vida útil de dados digitais. Esses suportes digitais são indicados para poucas décadas de uso, o que para as técnicas de preservação de acervos tradicionais é muito pouco, levando em consideração que estão acostumados a lidar com suportes que chegam a durabilidade de séculos.[VALLE JR.; 2003]

Em [CONWAY, 2001], este problema é apresentado através do que ele chama de “dilema dos suportes modernos. As variedades das máquinas e meios do universo digital intensifica a urgência, nas instituições, de uma atividade de preservação responsável. Para ajudar a descobrir como satisfazer esta necessidade intensificada, este relato primeiramente apresenta uma estrutura de conceitos fundamentais de preservação, conforme seus desdobramentos em décadas passadas. A partir daí, redefine e transforma-os em uma nova estrutura para ação no universo digital”.

3.8.2 ARMAZENAMENTO SECUNDÁRIO X TERCIÁRIO

Os suportes de armazenamento digital são classificados de duas formas [VALLE JR.; 2003]:

- *secundários*: disponibilidade mais ou menos imediata dos dados sem intervenção do operador;
- *terciários*: disponibilidade dos dados necessita da intervenção do operador antes que se tornem disponíveis ao usuário.

Ao armazenamento terciário são necessários os chamados procedimentos de montagem, isto é, procedimentos de preparação da mídia no dispositivo de leitura/gravação

para serem disponibilizados para o usuário. Alguns desses dispositivos, mais comuns aos usuários, como o CD-ROM, possuem procedimentos triviais, podendo ser até mesmo realizado pelo próprio usuário. No caso de fitas magnéticas em rolo, a montagem é mais complexa, necessitando assim de um operador já treinado.

Normalmente, utiliza-se o critério de frequência de acesso para determinar quais informações serão armazenadas em dispositivos secundários e terciários. Para os dispositivos secundários ficam informações que são acessadas com maior frequência e para dispositivos terciários as informações de acesso casual e de uso menos freqüente.

3.8.3 DISCOS RÍGIDOS

Discos rígidos possibilitam boas opções para o armazenamento, pois [VALLE JR., 2003]:

- acoplam mídia dos dados ao dispositivo de leitura/gravação como componente único;
- permitem grandes velocidades de transferências de dados;
- são confiáveis para operações contínuas;
- dispensam montagem.

Discos rígidos constam da formação de um ou mais discos de material rígido e magnetizado que, unidos a um eixo central, rodam em altíssimas velocidades (da ordem de 5000rpm – *rotações por minuto*). O aparato de leitura e gravação, denominado cabeça de leitura e gravação, flutua sobre os discos a distâncias micrométricas e movimentam-se, radialmente, medindo e/ou alterando seu campo magnético. Estatisticamente, chegam a possuir uma durabilidade de 4 a 5 anos, motivo pelo qual aconselha-se minimizar seu transporte e mantê-los em ambientes livres de vibrações e de altas temperaturas.

Duas opções podem ser feitas na escolha do uso de discos rígidos como suporte ao armazenamento secundário:

- pode-se utilizar um único disco rígido, de alta performance e com alta capacidade de armazenamento, para guardar os dados. Este esquema conhecido como SLED – *Single Large Expensive Disk*, somente um disco grande e caro. Se houver muitos dados a serem armazenados, pode-se utilizar um esquema denominado JBOD – *Just a Bunch of Drivers*, somente um monte de dispositivos, ou seja, adicionam-se outros discos formando um arranjo simples de discos. O problema principal em escolher entre o esquema SLED e JBOD consiste na perda de todos os dados, se houver falha de algum disco;
- diversos dispositivos podem ser utilizados, esquematizados para que a tarefa de armazenamento dos dados seja dividida e, ao mesmo tempo replicada, entre os discos. Esta solução resolve o problema do esquema anterior, fazendo com que, no caso de perda, a informação possa ser recuperada por uma cópia armazenada num outro disco. Este esquema é denominado RAID – *Redundant Array of Inexpensive Disk*, arranjo redundante de discos baratos.

A configuração do esquema RAID pode ser feita de diversas maneiras. Um dos mais populares é o RAID-5, que pode ser feito utilizando-se de 3 ou mais discos, sendo que o espaço para armazenamento dos dados será de $n-1$ discos. Isso aumenta a segurança no que se refere à perda, mas não é tolerante à falha de dois ou mais discos simultaneamente.

3.8.4 FITAS MAGNÉTICAS

Tendo seu uso iniciado na década de 50, o armazenamento terciário de informações por meio de fitas magnéticas é o mais antigo utilizado por sistemas digitais. Possuem vantagens como baixo custo, alta capacidade de armazenamento, confiabilidade e uma longa expectativa de vida útil. Mas, têm como desvantagens o acesso seqüencial, alto custo dos dispositivos de leitura e gravação, a fragilidade e a necessidade de treinamento para sua utilização.

Fitas magnéticas são formadas por uma superfície de gravação. São disponibilizadas em diferentes formatos: em rolo, cassetes ou cartuchos. Possuem tecnologia

de gravação longitudinal, que utiliza uma cabeça estática para gravar as trilhas no sentido de deslocamento da fita, ou helicoidal que utiliza cabeças rotativas, acopladas a um tambor que gira em altas velocidades, gravando as trilhas no sentido diagonal da fita. Possuem capacidade de armazenamento que vai de 20GB a 80GB no modo comprimido, ficando limitadas à quantidade de passagens pela cabeça de leitura e/ou gravação.[VALLE JR.; 2003]

A correta manipulação e armazenamento das fitas magnéticas, em locais de baixa temperatura e umidade relativa do ar, longe de poluição, poeira, tabaco, campos magnéticos fortes e gases corrosivos, garantem sua longevidade, podendo ultrapassar três décadas.

3.8.5 DISCOS ÓPTICOS E ÓPTICOS-MAGNÉTICOS

Introduzidos a partir da década de 80 e desenvolvidos pela Philips Electronics e pela Sony Inc., disco óptico e óptico-magnético são tecnologias recentes no suporte de armazenamento terciário de informações digitais. [VALLE JR, 2003]

A primeira mídia óptica lançada foi o compact disc ou, mais popularmente conhecido como CD – compact disc (disco compacto). Disco de policarbonato de 12 cm revestido por uma fina superfície refletora de alumínio e por uma camada de verniz para proteção. No início, foi concebido para registro de áudio digital e, posteriormente, um de seus variantes, o CD-ROM – compact disc read only memory (disco compacto somente para leitura), para uso específico em computadores, aderindo a outros tipos de dados digitais. A Figura 3.3 apresenta as variações encontradas na bibliografia a respeito dos discos CD's.

Mídia	Formato de Dados	Descrição	Lançamento
CD	CD-DA	Áudio Digital	1980
	CD-ROM	Dados de computador	1984
	Vídeo CD	Vídeo digital	1993
CD-MO		Óptico-magnético / Regravável	1990
CD-R		Óptico / Gravável	1993
CD-RW		Óptico / Regravável	1996

Fig. 3.3 – Múltiplos formatos de CD's. Adaptado [VALLE JR.; 2003]

Os primeiros discos CD's a serem desenvolvidos foram:

- *CD-MO*: permite múltiplas gravações e sua leitura é feita a laser. A gravação é uma combinação de aplicação de laser com um campo magnético que vai alterando suas propriedades refletoras da superfície de registro;
- *CD-R* e *CD-RW*: apresenta uma camada adicional entre a superfície refletora e a base de polycarbonato, que durante gravação tem suas propriedades ópticas alteradas. *CD-R* permite apenas uma gravação e múltiplas leituras e o *CD-RW* permite a regravação.

Na utilização do CD-ROM pode-se atingir a uma capacidade de armazenamento de aproximadamente 700MB (milhões de bytes), considerado insuficiente nos dias de hoje como suporte às necessidades de armazenamento. Como resposta a essa necessidade, foram desenvolvidos discos com inovações em relação aos primeiros CD-ROM's. Discos com maior capacidade de armazenamento com uma codificação de dados mais eficiente, com recursos duplicados da capacidade leitura e gravação em múltiplas camadas, conhecidos por DVD – digital versatile disk (disco digital versátil). A figura 3.4 apresenta as variações encontradas na bibliografia acerca dos múltiplos formatos de DVD's.

Tecnologia	Aplicação	Lançamento
DVD-ROM	Apenas leitura	1996
DVD-R	Gravável	1997
DVD-RAM	Regravável	1998
DVD-RW	Regravável	2000
DVD+RW	Regravável	2001

Fig. 3.4 – Múltiplos formatos de DVD's. Adaptado [VALLE JR.; 2003]

Discos ópticos tornaram mídias como suportes de armazenamento mais populares devido a uma série de fatores [VALLE JR., 2003]:

- Simplicidade e rapidez do procedimento de montagem, sem necessidade de maiores treinamentos para manuseio;
- Possibilitam acesso direto das informações gravadas;

- Resistência maior em relação aos discos rígidos e fitas magnéticas, limitada principalmente ao manuseio do usuário e ao ambiente externo;

Apesar de possibilitar uma boa solução para o suporte ao armazenamento, mídias ópticas têm sido alvos de controvérsias. Conservadas em um ambiente de temperatura e umidade razoáveis em torno de 30° C e menos de 60% da umidade relativa do ar, podem durar de 10 a 20 anos. Mas devido aos componentes químicos desses discos ópticos gerarem oxidações, esta estimativa faz cair a sua vida útil, podendo chegar a poucos anos de utilização. Estudos mostram que existe uma correlação entre qualidade imediata dos discos e sua durabilidade ao longo do prazo. É necessário que sejam utilizados hardware e softwares especiais que acompanhem a qualidade dos discos após a gravação e que os que apresentarem qualidade excepcionalmente boa sejam reservados para o armazenamento a longo prazo.

3.9 MÉTODOS DE MIGRAÇÃO DE DADOS

Um dos maiores problemas que a tecnologia digital enfrenta, em relação à preservação dos dados digitais, é a obsolescência dos componentes de hardware e software.

Quando as mídias digitais se “degradam”, ou seja, começam a ficar obsoletas, são sugeridos pela bibliografia alguns métodos de preservação e conservação, na tentativa de diminuir a perda de informação ocasionada pela mudança de um sistema digital para outro. Estes métodos vão desde o acondicionamento físico das mídias até a avaliação da conversão lógica dos dados digitais.

Eduardo Valle Jr. [VALLE JR.; 2003] propõe os principais métodos preventivos e conservativos conhecidos na bibliografia e que podem ser aplicados na preservação digital. São eles: refrescamento, migração e emulação.

O refrescamento faz uma simples cópia dos dados digitais de uma mídia para outra, sem que haja perda ou alteração dos dados. A desvantagem desse método é que se a densidade das mídias são diferentes, o total aproveitamento da nova mídia pode se tornar mais trabalhoso.

A migração é um processo, quase impossível de se prever com antecedência. É custoso e complexo, e com grandes riscos de perdas de dados durante a conversão. Também, torna-se difícil prever como será a tecnologia digital daqui a alguns anos. A migração é uma tarefa urgente porque, numa mudança de paradigma computacional, pode impossibilitar a conversão dos dados digitais. Nesta tarefa, pode-se utilizar um procedimento de levantamento de riscos para a análise da migração, como um recurso de previsão de perdas das informações durante a reformatação de mídias no contexto lógico. Por exemplo, conversão de um arquivo no formato TIFF 5.0 para TIFF 6.0.

A emulação praticamente tenta atuar em conjunto ou substituir a migração. Propõe manter as mídias em seus formatos originais simulando ambientes operacionais utilizados pelas mídias que se tornaram obsoletas.

“A preservação da imagem digital é uma tarefa delicada, devido à própria natureza dessa informação e aos rápidos ciclos de obsolescência da tecnologia que a suporta” [VALLE JR.; 2003]. O que se objetiva na preservação digital é tentar aumentar a vida útil dos aparatos computacionais para extrair o que de melhor eles possam oferecer.

4 INTRODUÇÃO AO PROCESSAMENTO DIGITAL DE IMAGENS

A área de processamento digital de imagens - PDI vem crescendo bastante com a evolução da tecnologia digital. O crescente interesse viabiliza aplicações de duas categorias distintas: o aprimoramento de informações pictóricas para interpretação humana e a análise automática por computador de informações extraídas de uma cena. [MARQUES FILHOS; VIEIRA NETO, 1989]

As aplicações iniciais das técnicas de processamento digital de imagens foram ilustrações de jornais enviadas por cabo submarino entre Londres e Nova York, por volta de 1920. O primeiro sistema Bartlane de transmissão de imagens por cabo submarino entre Londres e Nova York era utilizado para codificar imagens com o objetivo de diminuir o tempo de transporte de imagens através do oceano Atlântico, de mais ou menos uma semana para menos de três horas. Nesta época já se conseguia com esse sistema codificar uma imagem com 5 níveis de intensidade distintos e ampliados para 15 a partir de 1929. [GONZALEZ; WOODS, 2000]

Mas as técnicas de processamento digital de imagens realmente tiveram seu marco evolutivo na década de 60 com o advento de computadores digitais e com o programa espacial norte-americano. As imagens da Lua eram transmitidas pela sonda Ranger 7. Eram aprimoradas de distorções para serem mostradas por uma câmara de TV levada com a tripulação do foguete espacial. As técnicas utilizadas para o processamento das imagens serviram de base para o realce e restauração de imagens de outros programas espaciais posteriores como, por exemplo, as expedições tripuladas da série Apollo. [MARQUES FILHOS; VIEIRA NETO, 1989]

De 1964 até hoje, a área de processamento digital de imagens com suas técnicas e métodos ainda apresentam crescimentos expressivos devido ao fato de que com suas aplicações pode-se resolver uma variedade enorme de problemas que requerem melhorias da informação visual para a análise e interpretação ou mesmo a percepção automática por máquinas computacionais.

Como exemplo da análise e interpretação humana de imagens melhoradas por tratamento digital, têm-se as aplicações na área médica com imagens de raios-X, na geologia, no estudo de imagens de satélites, na restauração de figuras em arqueologia, na melhora de imagens físicas, na medicina nuclear, na melhoria de imagens transmitidas por ondas de rádio. Como exemplo da percepção por máquinas, têm-se o reconhecimento automático de caracteres, o processamento automático de impressões digitais, o reconhecimento automático de rostos, a visão computacional, a guiagem de robôs e o processamento automático de imagens de satélites para reconhecimento de queimadas. [BUSQUIM, 2002]

O processamento digital de imagens inclui diversos conceitos e técnicas tais como: técnicas utilizando a transformada de Fourier, técnicas de realce, como filtragem espacial, com o propósito de melhorar o resultado da imagem original, restauração de imagens, com foco na recuperação ou reconstrução de uma imagem degradada, compressão de imagens, cujo objetivo é comprimir imagens apresentando ou não erros, técnicas de segmentação, que tratam da divisão de uma imagem em para posterior análise, técnicas de reconhecimento e interpretação de imagens. Após a aplicações das técnicas anteriores de tratamento da imagem, deve-se concentrar na análise automática de padrões e características que a imagem traz como contexto semântico. [BUSQUIM, 2002]

4.1 ELEMENTOS DE UM SISTEMA PDI

O processamento digital de imagens constitui-se de um conjunto de elementos de hardware e software que, amparados por fundamentos teóricos, executam as tarefas necessárias ao processamento de imagens. [GONZALEZ;WOODS, 1999]

A Figura 4.1 permite ilustrar os elementos de sistemas de baixo custo até sofisticadas estações de trabalho, utilizadas em aplicações específicas de processamento de imagens. Ela apresenta as principais operações que se podem realizar sobre uma imagem, tais como: aquisição, armazenamento, processamento e exibição.

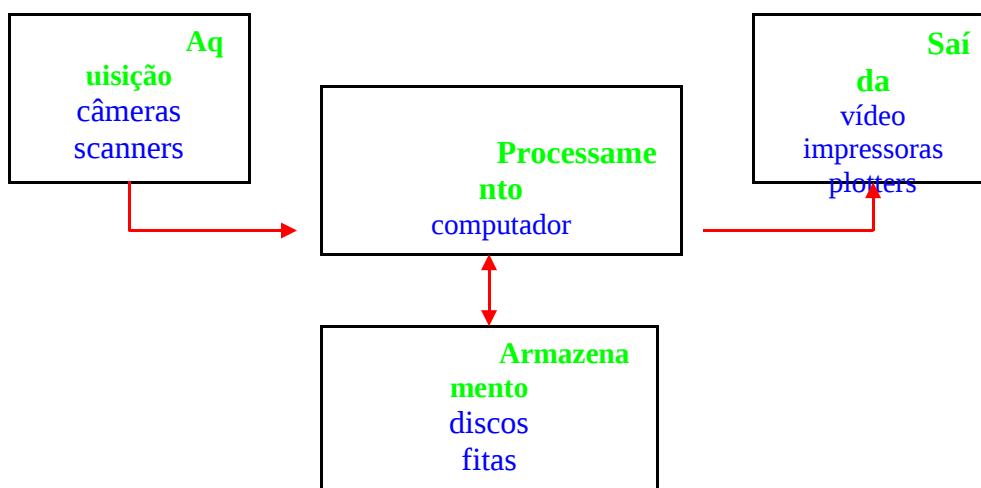


Fig. 4.1 – Elementos de um sistema de processamentos de imagem, adaptado de [MARQUES FILHOS; VIEIRA NETO, 1989]

Esses elementos podem ser resumidos em:

- *Aquisição da imagem*: envolve um dispositivo físico sensível a uma banda do espectro da energia eletromagnética que produz um sinal elétrico atuando em conjunto com um digitalizador que identifica esse sinal elétrico convertendo-o para a forma digital;

- *Armazenamento*: envolve dispositivos próprios de acordo com a categoria de manipulação das imagens. É um dos maiores desafios no projeto de sistemas de processamento de imagens por tratar de grandes quantidades de bytes;
- *Processamento*: envolve procedimentos específicos das técnicas de processamento de imagem expressos algoritmicamente e implementados em uma linguagem de programação;
- *Exibição*: envolve dispositivos físicos para demonstração dos resultados das técnicas implementadas de processamento digital de imagens.

4.2 ELEMENTOS DE PERCEPÇÃO VISUAL

Muitas técnicas desenvolvidas no processamento digital de imagens foram motivadas pelo processo de percepção visual humano.

4.2.1 A ESTRUTURA DO OLHO HUMANO

O olho humano é próximo de uma esfera de diâmetro médio de cerca de 20mm e envolvidos por três membranas:

- A córnea, membrana resistente e transparente que cobre a superfície anterior do olho;
- A coróide, membrana cuja função é nutrir o olho, se divide em corpo ciliar e íris;
- A retina, membrana mais interna que se estende por toda a posição posterior do olho;

Quando a visão está bem focalizada, a luz do objeto fora do olho se torna imagem na retina, que conta com o auxílio de receptores discretos de luz distribuídos por toda a superfície da retina. Esses receptores são denominados de: cones e bastonetes.

Os cones, que são cerca de 6 a 7 milhões em cada olho, são conectados individualmente na fóvea, motivo pelo qual são bastante sensíveis a detalhes e seu tipo de

visão é denominado de fotópica ou de luz clara. Já os bastonetes, que são cerca de 75 a 100 milhões em cada olho, são distribuídos por toda a retina. Vários bastonetes são ligados a uma única fibra nervosa, motivo pelo qual são usados para identificar baixos níveis de luminosidade e seu tipo de visão é denominado de escotópica ou visão de luz escura.

4.2.2 A FORMAÇÃO DA IMAGEM NO OLHO

O cristalino, formado por camadas concêntricas de células fibrosas, é sustentado por fibras que se conectam ao corpo ciliar e possuem pigmentação de cor levemente amarela, é uma lente óptica comum e bastante flexível.

Quando o olho focaliza um objeto mais de perto ou de longe, o cristalino exerce papel fundamental ao cálculo do tamanho da imagem.

Após a imagem refletida na retina, principalmente na área da fóvea, a percepção acontece pela relativa excitação dos receptores de luz (cones) que a transformam em impulsos elétricos que são posteriormente decodificados pelo cérebro, para onde são enviados através do nervo ótico, como apresenta a Figura 4.2 [GONZALEZ- WOODS, 1992].

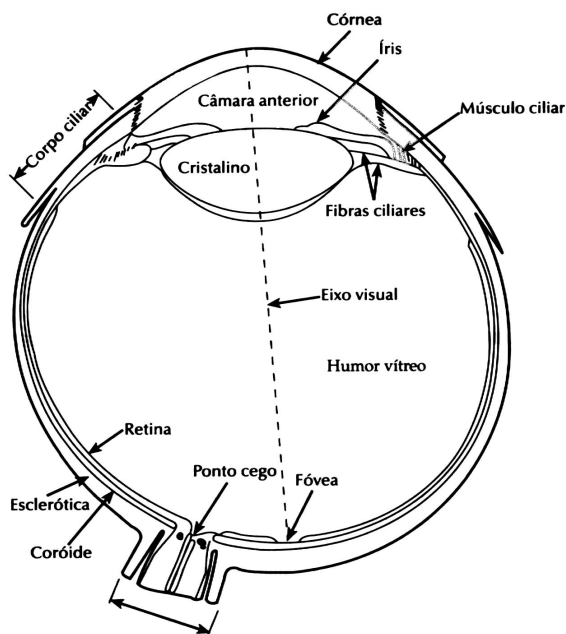


Fig. 4.2 – O sistema visual humano detalhando [<http://www.caa.uff.br/~aconci/sisvisi.htm>]

4.2.3 ADAPTAÇÃO AO BRILHO E A DISCRIMINAÇÃO

O processamento digital de imagens se baseia no modelo do sistema visual humano – SVH, devido à aguçada percepção do mesmo em relação aos níveis discretos de brilho e contraste. A adaptação da escala de níveis de luminosidade pelo SVH é enorme, podendo chegar até cerca de 10^{10} níveis de adaptação ao brilho.

O que se consegue com o processamento digital de imagens é utilizar um pouco dos níveis de escala de brilho em relação ao olho humano para reconhecer e processar imagens.

4.3 TÉCNICAS UTILIZADAS NO PROCESSAMENTO DE IMAGEM

Por Processamento Digital de Imagens entende-se a manipulação de imagens pelo computador de modo que a entrada e saída sejam imagens. O motivo de se processar imagens é melhorar o aspecto visual de uma imagem.

Na manipulação de imagens realizada pelo processamento podem se encontradas duas classes de transformações [MASCARENHAS; VELASCO, 1982]:

- Transformações radiométricas, onde a alteração da imagem será feita pixel a pixel sem alterar a sua geometria.;
- Transformações geométricas, onde a alteração da imagem será feita alterando-se a geometria da mesma tentando ao máximo manter os níveis de intensidade próprios a ela.

Essas transformações possuem a finalidade de eliminar distorções ou enfatizar as características da imagem introduzida ao sistema de imageamento.

A seguir serão descritas, de forma bastante resumida, cada uma das técnicas utilizadas pelo processamento digital de imagens encontradas em [GONZALEZ- WOODS, 1992]. Cada uma dessas técnicas, através das transformações de imageamento, realizará uma função específica.

Realce

O objetivo das técnicas de realce é processar uma imagem original cujo resultado dependerá da natureza de uma aplicação específica. Podem ser divididas em duas categorias:

- Métodos de domínio espacial, baseados na manipulação direta ao agregado de pixels que compõem uma imagem.

- Métodos de domínio por frequência, baseadas especificamente nas transformadas de Fourier como fator base utilizada na filtragem da imagem.

Restauração de Imagens

Esta técnica tenta reconstruir ou recuperar uma imagem que foi degradada, usando-se de algum conhecimento inicial do fenômeno de degradação.

Para isso, são utilizados procedimentos heurísticos projetados para manipular uma imagem. O processo de restauração digital pode ser obtido pela aproximação de $f(x,y)$ dada uma $g(x,y)$ e um conhecimento de degradação sobre a forma de H .

Compressão de Imagens

O principal objetivo desta técnica é tratar o problema de reduzir a quantidade de dados necessária para representar uma imagem digital, através da remoção de dados redundantes.

Do ponto de vista matemático, corresponde a transformar uma matriz de pixels de duas dimensões num conjunto de dados estatisticamente sem correlação. A imagem pode ser comprimida e descomprimida de acordo com a imagem original ou de uma aproximação dela.

O estudo desta técnica vem crescendo continuamente no campo científico, devido a tendência do uso de dados utilizados pela computação.

Segmentação da imagem

Esta técnica tem como objetivo extrair informação a partir de uma imagem pré-processada para, através de uma análise inicial das suas características, estabelecer num primeiro instante o reconhecimento de padrões.

A segmentação subdivide a imagem em partes ou objetos constituintes de acordo com a natureza do problema de análise a ser resolvido, podendo contribuir significativamente com o sucesso ou o fracasso dessa análise.

Representação e Descrição

Esta técnica tem como objetivo representar e descrever num formato apropriado os agrupamentos de pixels segmentados gerados a partir da técnica anterior para reconhecer e interpretar imagens.

O processo de representação pode ser feito a partir de características internas (os pixels que compõem a região) ou características externas da imagem (os pixels que formam a fronteira da região), para então descrever a imagem baseada na representação escolhida.

A descrição define que tipo de descritor a ser usado para “capturar” as características dos objetos ou classe de objetos segmentados.

Reconhecimento e Interpretação

Esta técnica está relacionada com a análise automática de imagens, ou seja, o processo de descobrimento, de identificação e de entendimento de padrões relevantes de acordo com a natureza da aplicação que será realizada com a imagem.

A meta da análise é tentar aproximar ao máximo o computador da capacidade humana de reconhecimento e interpretação automática do contexto. Resulta, por exemplo, em uma leitura automática de documentos datilografados com a finalidade de reconhecimento, no máximo possível, de caracteres como se fossem seres humanos realizando esta tarefa.

Uma forma de reconhecer automaticamente a imagem é projetar um certo grau de *inteligência*, através um sistema de análise automática, para extrair características e detalhes relevantes da imagem. Esse reconhecimento prévio da imagem proporcionaria a capacidade de aprender através de uma nova, ou diferente circunstância, habilidade de fazer inferências a partir de informações incompletas da imagem.

A técnicas de aquisição de imagens, melhoramento, restauração e compressão são comuns em aplicações envolvendo processamento de imagens. A técnica de reconhecimento abrange outras áreas afins de processamento digital de imagens e ainda é um desafio a estudiosos e pesquisadores da área.

É importante ressaltar que o resultado da aplicação de algumas dessas técnicas de processamento da imagem está relacionada a um objetivo específico. Portanto, o resultado dependerá da natureza da finalidade do imageamento.

5 ALGORITMOS DE REALCE

As técnicas de realce são utilizadas para processar imagens de forma que, o resultado de uma imagem processada seja melhor que sua forma original. É importante ressaltar que as técnicas de realce são utilizadas para diversos fins e são interpretadas de acordo com o conhecimento prévio do observador acerca das imagens analisadas.

As técnicas de realce proporcionam os mais diversos resultados. “Logo, não existem técnicas capazes de resolver 100% dos problemas que uma imagem digital possa apresentar” [MARQUES FILHO; VIEIRA NETO, 2000]. Assim, nem sempre uma técnica de realce produz os mesmos resultados de imagens processadas para serem usadas em aplicações diferentes.

Existem dois métodos utilizados pela técnica de realce para melhorar a aparência de uma imagem: as técnicas de filtragem no domínio espacial e as técnicas de filtragem no domínio da frequência.

Neste trabalho serão apresentadas algumas técnicas de realce utilizadas sobre a imagem digital após a sua digitalização por meio físico. A ideia básica é filtrar a imagem

digital original através do método de filtragem no domínio espacial para realçar e melhorar a aparência do visual da imagem resultando numa imagem digital filtrada para posterior uso de outras técnicas de processamento digital ou mesmo para ampliar as condições de observação.

5.1 TEORIA DOS FILTROS

As técnicas de filtragem de imagens no domínio espacial realizam operações diretamente sobre a imagem, através de convolução com máscaras. A convolução permite filtrar a imagem retirando dela algumas frequências espaciais. [SANTOS, 1995].

Por frequência espacial entende-se como o modo em que as alterações dos valores dos pixels são distribuídos ao longo da imagem. Áreas da imagem que apresentam tons ou cores constantes são consideradas de baixas frequências, pois há pouca ou nenhuma variação de cor. Áreas da imagem que apresentam grandes mudanças de tons ou cores são consideradas de alta frequência. [MARQUES FILHO; VIEIRA NETO, 2000]

Portanto, imagens com baixa frequência ou com frequência 0 são aquelas que representam objetos que possuem uma só cor e imagens com alta frequência são aquelas que possuem mudanças bruscas de tons, como por exemplo os contornos dos objetos [SANTOS, 1995].

5.1.1 PRINCÍPIO DE FUNCIONAMENTO DOS FILTROS

O princípio de funcionamento dos filtros que operam no domínio espacial baseia-se com relação às vizinhanças do pixel a ser filtrado pela máscara de convolução. O processo de filtragem é feito através das operações de convolução sobre do valor da cor de cada pixel da imagem original com a máscara de convolução. A máscara é preenchida com o pixel da imagem original e sua vizinhança, podendo variar de acordo com a vizinhança do um pixel. Pode ser direta de 4 elementos, de 8 elementos ou dos elementos diagonais que o circulam.

As imagens digitais são bidimensionais e discretas, logo chamamos de convolução discreta um somatório duplo, Equação 5.1, aplicado a cada ponto da imagem [SANTOS, 1995]:

$$g(i, j) = \sum_{k=i-w}^{i+w} \sum_{l=j-v}^{j+v} f(k, l)h(i-k, j-l) \quad (\text{Equação 5.1})$$

onde $\pm w$ e $\pm v$ são os limites do filtro espacial em cada dimensão.

A Figura 5.1 apresenta a aplicação de um filtro que utiliza uma máscara 3x3. O ponto central é $i(2,2)$ corresponde ao pixel da imagem a ser filtrada e resultará num pixel filtrado cujo o valor, após as operações de convolução, é demonstrado na equação 5.2 a seguir:

$$\begin{aligned} r_{2,2} = & h_{1,1} * i_{(1,1)} + h_{1,2} * i_{(1,2)} + h_{1,3} * i_{(1,3)} + \\ & + h_{2,1} * i_{(2,1)} + h_{2,2} * i_{(2,2)} + h_{2,3} * i_{(2,3)} + \\ & + h_{3,1} * i_{(3,1)} + h_{3,2} * i_{(3,2)} + h_{3,3} * i_{(3,3)} \end{aligned} \quad (\text{equação 5.2})$$

$$\text{Filtro } 3 \times 3 = \begin{bmatrix} h_{1,1} & h_{2,1} & h_{3,1} \\ h_{2,1} & h_{2,2} & h_{2,3} \\ h_{3,1} & h_{3,2} & h_{3,3} \end{bmatrix}$$

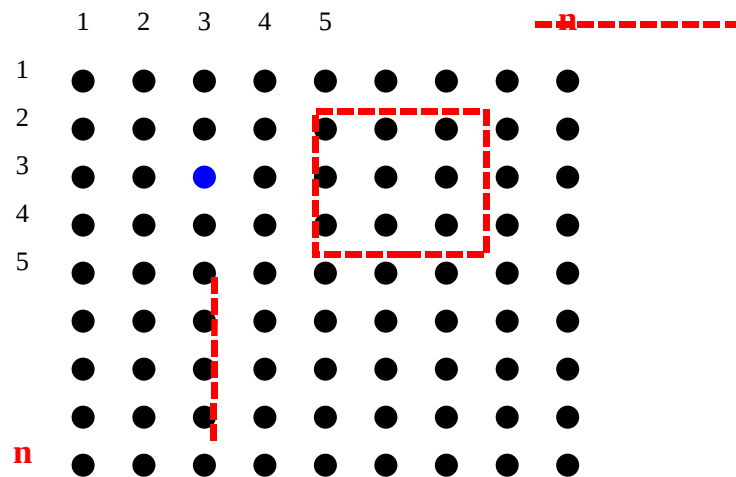


Fig. 5.1 – Imagem $n \times n$ – Adaptada de [SANTOS, 1995].

O formato da máscara pode variar de acordo com o tipo de filtro a ser aplicado. Dependendo do algoritmo de cada filtro, podem ser utilizadas submáscaras com formatos pré-determinados, tais como triangulares, pentagonais, hexagonais ou formatos irregulares. Até a máscara de convolução principal pode assumir formatos diferentes, como por exemplo cruz ou retângulo. A máscara percorrerá a matriz representando a imagem original tanto na horizontal quanto na vertical, ao longo de cada linha e entre as várias linhas, da esquerda para a direita, de cima para baixo [SANTOS, 1995].

A utilização das máscaras de convolução pode acarretar problemas nas bordas da imagem uma vez que para realizar as operações de convolução é necessário que seja conhecida toda a vizinhança de um dado pixel. Para contornar este problema, “pode-se restringir o posicionamento da máscara para que a mesma não seja sobreposta a pontos não pertencentes à imagem ou atribui-se aos pontos fora da imagem um valor pré-determinado” [ARECO, 1995].

Neste caso, a solução consiste em duas opções: a primeira pode-se preencher as bordas com um valor constante ou com o valor do ponto mais próximo; e a segunda considera-se a imagem “cíclicamente” fechada ou atribui 0 (zero) para pontos fora da imagem. [ARECO, 1995]

Uma seleção apropriada dos elementos das máscaras utilizadas pelos filtros pode resultar numa variedade de operações úteis tais como: a redução de ruído, a suavização, o refinamento e a detecção das características da imagem.

Mas, como desvantagem, uma operação de convolução com máscaras exige um grande esforço computacional. Por exemplo, “uma imagem de 512x512 pixels sendo processada com uma máscara 3x3 requer cerca de nove multiplicações e oito adições para localização de cada pixel resultando num total de 2.359.296 multiplicações e 2.097.152 adições” [MARQUES FILHO; VIEIRA NETO, 2000].

5.2 ALGORÍTMOS DE REALCE DE IMAGENS

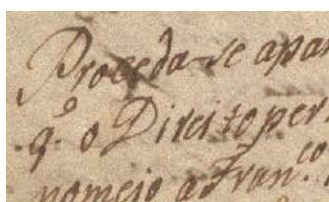
Nesta seção serão apresentadas as idéias básicas de alguns dos algoritmos dos filtros de suavização espacial mais interessantes, pesquisados no Núcleo de Processamento Digital de Imagens da Universidade Federal de Minas Gerais.

A seleção para implementação dos filtros foi feita baseando-se naqueles que possuísem aplicações na remoção de ruído, suavização e refinamento das características da imagem. Foram implementados, também, algoritmos cujas funções são de alterar frequências da imagem para atuar em conjunto com os algoritmos de realce para, juntos, obterem melhores resultados na visualização das imagens específicas utilizadas neste trabalho.

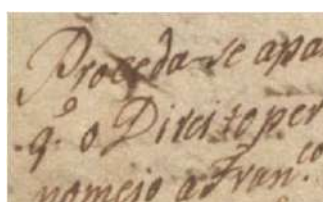
Os filtros são:

Média

Um dos filtros mais simples definidos no domínio espacial. A idéia é a aplicação da média aritmética dos elementos da máscara cujo centro é o pixel da imagem original. O objetivo desse filtro é a atenuação do ruído através da média pois cada pixel da imagem filtrada será função de todos os pixels dentro de uma janela ou máscara com pesos pré-determinados. A forma mais simples de utilização desse filtro é através de uma máscara 3x3 com peso iguais a 1 cujo resultado é o somatório de todos os níveis de cinza de cada elemento da máscara divididos por 9 (número total de elementos da máscara). Também é possível a utilização de dimensões maiores do tipo 5x5, 7x7 e 9x9 com ou sem pesos pré-determinados, mas sempre com o resultado da média aritmética de acordo com as suas dimensões. A figura 5.2 abaixo apresenta a aplicação deste filtro sobre uma imagem digitalizada. [MARQUES FILHO; VIEIRA NETO, 2000]



(a) Original com 20% de ruído



(b) Média 3x3

Fig. 5.2 – Filtro da Média 3x3

Média com K vizinhos mais próximos

Este filtro é uma ampliação do filtro da média tradicional. Seu funcionamento se baseia em máscara de formatos pré-determinados selecionada a partir de um grupo de elementos pertencente à vizinhança do pixel central com o próprio pixel. O cálculo realizado por este filtro é o mesmo da média tradicional e o resultado será aplicado ao respectivo pixel da imagem original. A aplicação desse filtro atinge uma redução do ruído com preservação de bordas e características da imagem de acordo com o valor de K. Quanto maior a escolha dos k vizinhos mais próximos maior a redução de ruído, mas menor a preservação de bordas e características da imagem. Quanto menor for k, menor será a redução de ruído e maior será a preservação das características da imagem. Os valores mais usuais de K são 2, 4, 6 e 8 para uma máscara 3x3. A figura 5.3 apresenta o resultado da aplicação do filtro da média com k = 2 vizinhos mais próximos [ARECO, 1995].

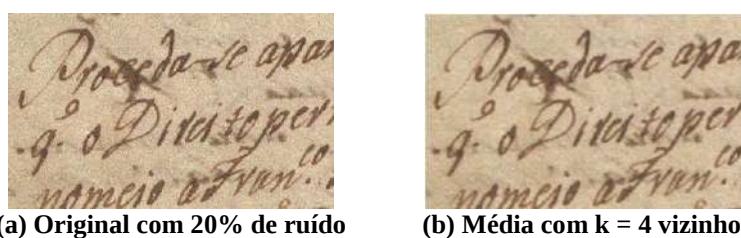


Fig. 5.3 – Filtro da Média 3x3 com k = 4 vizinhos mais próximos

Moda

O filtro da moda é uma aplicação bem simples e consiste na escolha do elemento que é mais freqüente na máscara. Este filtro apresenta melhor resultado se aplicado em conjunto com outros filtros como a média ou mediana. A figura 5.4 abaixo apresenta a aplicação deste filtro sobre uma imagem digitalizada [ARAÚJO, 1989].

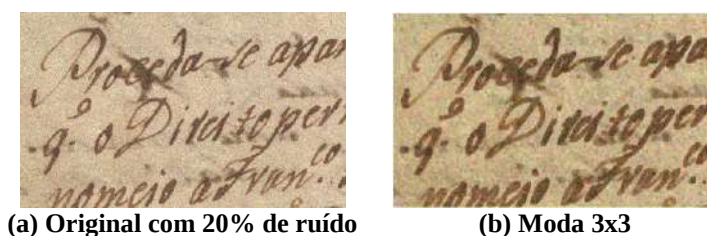


Fig. 5.4 – Filtro da Moda 3x3

Ordem

O filtro da ordem ordena os valores dos elementos da máscara e substitui o valor do pixel da imagem original pelo valor do menor dos elementos da máscara, filtro da ordem mínimo, e/ou o maior valor dos elementos da máscara, filtro da ordem máxima. Também pode ser utilizado o k-ésimo valor do nível de cinza de um dos elementos da máscara. Por exemplo, numa máscara 3x3 o elemento de número 8 é o penúltimo elemento da máscara com valores ordenados de seus elementos. A figura 5.5 abaixo apresenta a aplicação deste filtro sobre uma imagem digitalizada [ARAÚJO, 1989].



Fig. 5.5 – Filtro da Ordem 3x3

Mediana

O filtro da mediana é adaptação do filtro da ordem. Parte da idéia de que pontos ruidosos da imagem tendem a se posicionar nas extremidades dos elementos da máscara ordenados. Este filtro calcula o valor mediano dos elementos da máscara ordenados cujo resultado será aplicado sobre o pixel da imagem original. Possui uma tendência de “consolidar regiões de mesma intensidade”, mas apresenta um bom resultado visual suavizando discretamente a imagem. A figura 5.6 abaixo apresenta a aplicação deste filtro sobre uma imagem digitalizada [MARQUES FILHO; VIEIRA NETO, 2000].

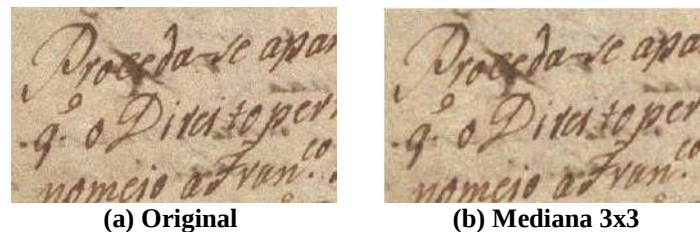


Fig. 5.6 – Filtro da Mediana 3x3

Transformada de Aguçamento Extremo - TAEX

O filtro TAEX tem como funcionamento o cálculo da diferença absoluta entre o ponto central e os valores extremos dos elementos da máscara. Aquele que possuir menor diferença absoluta substituirá o valor do pixel da imagem original. É um filtro bom para ser utilizado em conjunto com outros filtros dependendo da aplicação. Por exemplo, utilizado em conjunto com o filtro da mediana possui um bom resultado em imagens biomédicas. A figura 5.7 abaixo apresenta a aplicação deste filtro. [ARAÚJO, 1989].

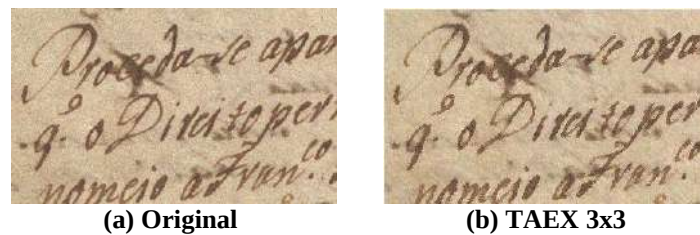


Fig. 5.7 – Filtro da TAEX 3x3

Passa Alta

O filtro Passa Alta parte da idéia de que são serão filtradas as frequências baixas. A figura 5.8 abaixo apresenta a aplicação deste filtro [MASCARENHAS & VELASCO, 1984].

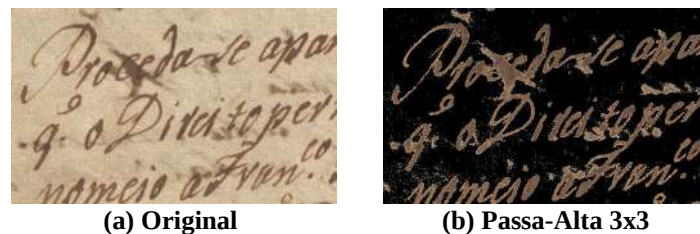


Fig. 5.8 – Filtro Passa-Alta

Passa Baixa

O filtro Passa Baixa parte da idéia de que são serão filtradas as frequências altas. v A figura 5.9 abaixo apresenta a aplicação deste filtro [MASCARENHAS & VELASCO, 1984].

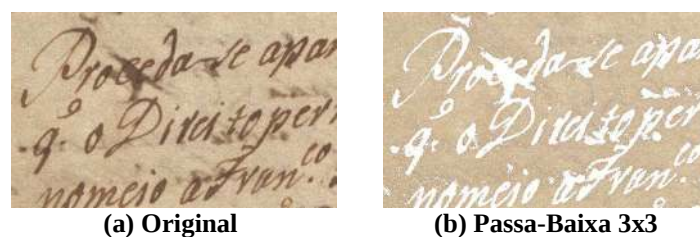


Fig. 5.9 – Filtro Passa-Baixa

Alguns filtros foram adaptados a partir do filtro Passa-Faixa. Este filtro atenua tanto frequências altas quanto baixas, preservando, contudo, frequências médias. Para realizar as operações de aumento e diminuição de frequências espaciais, foram gerados testes para cada região característica para definir as frequências médias de cada uma. Contudo, estes testes possibilitaram um realce na aparência da imagem que, aplicado com os filtros tradicionais de realce, resultaram num resultado satisfatório nas imagens quanto ao seu visual [MASCARENHAS & VELASCO, 1984].

Os filtros adaptados foram:

Escurecer

O filtro Escurecer é uma adaptação do filtro passa-alta. Utilizado como recurso para escurecer as regiões mais homogêneas. Este filtro atenua ruídos e contornos. A figura 5.10 abaixo apresenta a aplicação deste filtro sobre uma imagem digitalizada

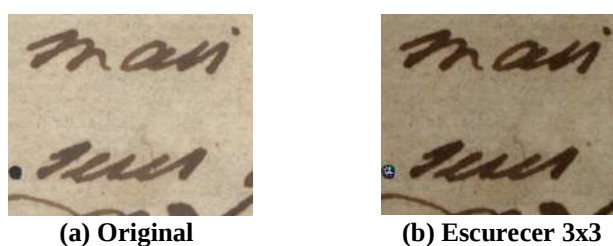


Fig. 5.10 – Filtro da Escurecer 3x3

Clarear

O filtro Clarear é uma adaptação do filtro passa-baixa. Utilizado como recurso para clarear as regiões claras e homogêneas da imagem. A figura 5.11 abaixo apresenta a aplicação deste filtro sobre uma imagem digitalizada

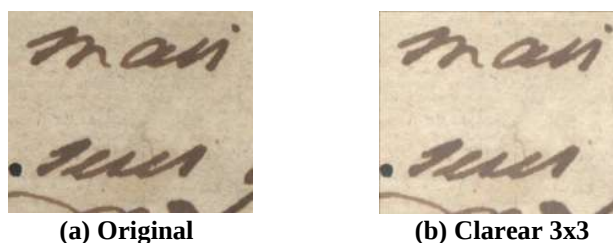


Fig. 5.11 – Filtro da Clarear 3x3

TrataClara

O filtro TrataClara é uma adaptação do filtro passa-alta básico. Utiliza um máscara 3x3 aplicando pesos negativos na vizinhança e um peso positivo no ponto central. Uma adaptação do operador Lapaciano, que é utilizado para detectar bordas na imagem. A figura 5.12 abaixo apresenta a aplicação deste filtro [ARAÚJO, 1989].

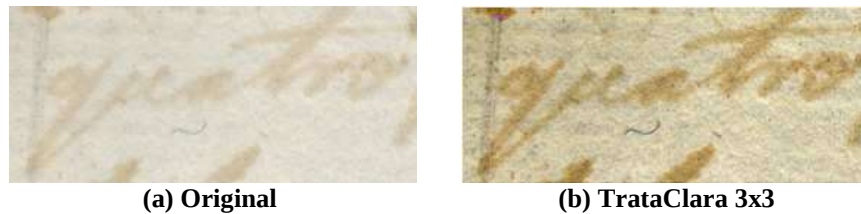


Fig. 5.12 – Filtro da TrataClara 3x3

MédiaClara

O filtro MédiaClara é uma adaptação do filtro da média com o filtro passa-baixa. Aplica o filtro da média e aumenta as frequências baixas da imagem resultando numa suavização com efeito de clareamento. A figura 5.13 abaixo apresenta a aplicação deste filtro.

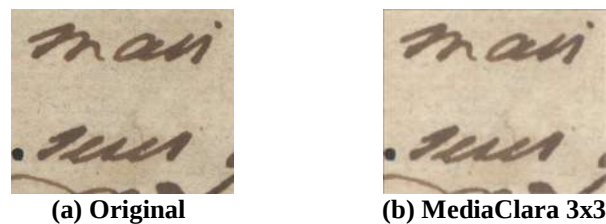


Fig. 5.13 – Filtro da MédiaClara 3x3

RealçarBorda

O filtro RealçarBorda é uma adaptação do filtro passa-faixa. Ao aplicar o filtro é analisado as frequências altas da imagem com o objetivo de aumentar uma faixa específica dessas frequências. A figura 5.14 abaixo apresenta a aplicação deste filtro.



Fig. 5.14 – Filtro RealçarBorda 3x3

5.2.1 APLICAÇÃO DOS FILTROS

No universo de documentos históricos digitalizados é necessário que quanto melhor for a visualização das imagens melhor será a análise do observador visual, pois este é um dos objetivos de tais imagens.

A implementação dos filtros de realce tem como objetivo possibilitar que sua utilização seja uma das ferramentas fundamentais na obtenção de um melhor resultado visual da imagem. Aplicados isolados ou em conjunto, para este universo específico de imagens digitais.

A partir de testes realizados com a aplicação dos filtros sobre as imagens, procurou-se observar como aumentar a visibilidade das imagens através da remoção de ruídos, preservação e aguçamento de bordas, preservação de pequenos detalhes da imagem, tomando o cuidado em manter as características originais dos documentos.

As imagens foram digitalizadas com resolução de 600 dpi coloridas em 24bits e os algoritmos foram implementados em Delphi 5.0.

A degradação dos documentos demonstra a urgência da sua digitalização, como meio para preservar os originais do manuseio. As folhas desmancham pelo simples fato de serem movidas a uma pequenina distância e também pela quantidade de orifícios e rasgos ocasionados pelo tempo.

A Figura 5.15 apresenta uma imagem típica dos documentos históricos, os quais são objetos deste trabalho.

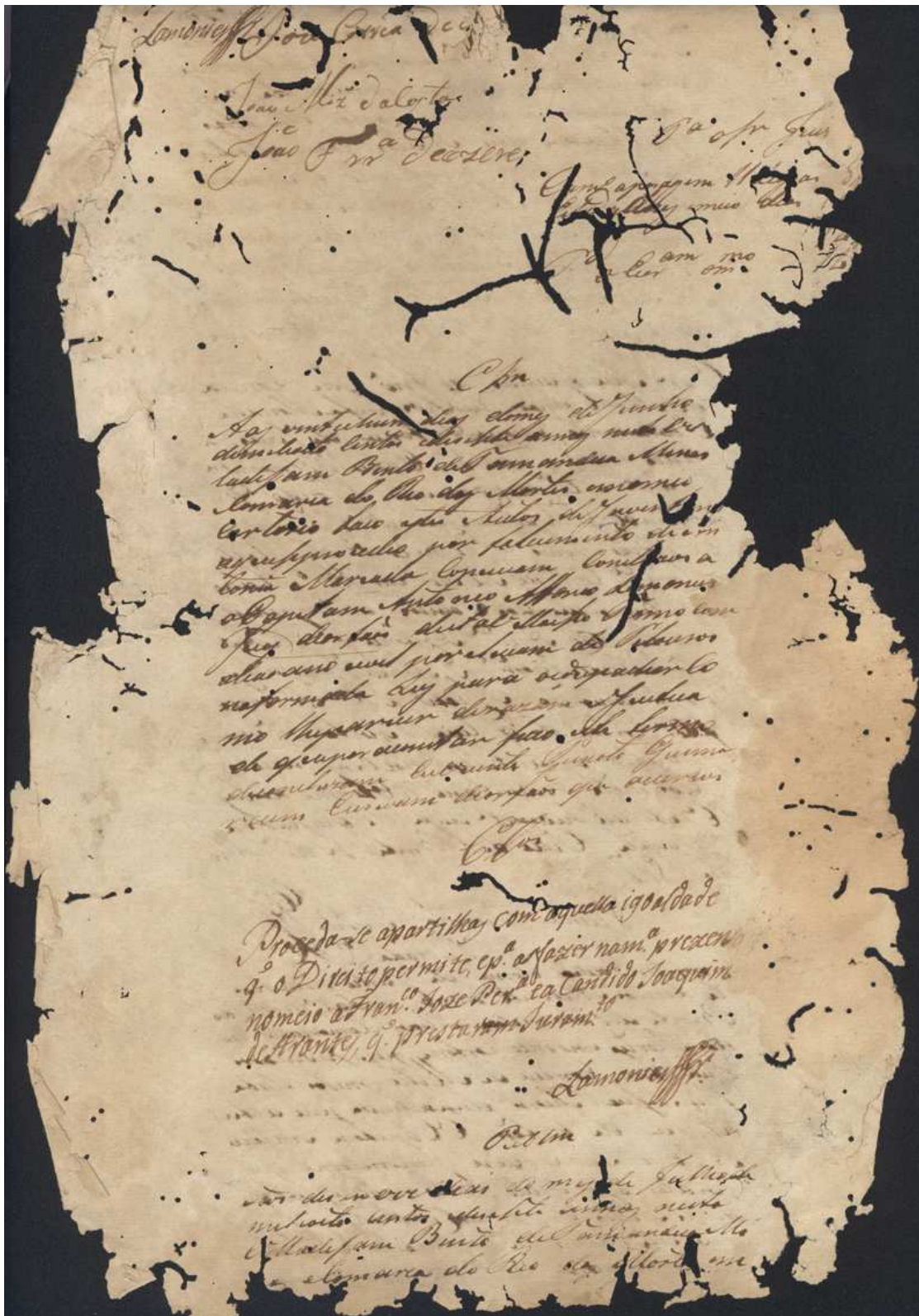


Fig. 5.15 – Imagem de um Documento Histórico em Processo de Deterioração

Nas imagens de documentos históricos a degradação apresenta-se através de imagens muito envelhecidas amareladas e escurecidas, normalmente muito borradas, apagadas, com texto variando com espessura da tinta fina e grossa. Na maioria dos documentos pesquisados foram detectados que a maior parte dessas características estão presentes num único documento. Assim, para aplicação dos filtros a idéia foi de clarear o máximo possível tentando não perder demais as características da imagem utilizando, em seguida, os filtros de suavização para remoção dos ruídos e detecção das bordas.

A partir das características encontradas nos documentos, segue abaixo as demonstrações das aplicações dos filtros de acordo com cada uma delas.

5.2.2 IMAGENS APAGADAS



Fig. 5.16 – Imagem com conteúdo apagado

Figura 5.2(a)(b)(c)(d) são imagens cujo conteúdo apresenta-se muito claro e com pouca visibilidade objetiva-se destacar o texto do fundo para facilitar a leitura do pesquisador. O conjunto de filtros aplicados foi um algoritmo denominado *TrataClara* que é uma adaptação do operador Laplaciano que tem como finalidade realçar as bordas (mudanças ou descontinuidade local na luminosidade) da imagem. Em seguida, foi aplicado uma adaptação do filtro da média que além da remoção do ruído aumenta a luminosidade dos pontos da imagem, um algoritmo denominado *Clarear* para realçar os pontos mais claros na imagem junto com o filtro da mediana para suavizar a aparência da imagem. A dificuldade no tratamento desta imagem foi não deixar com que os pontos mais claros saturassem, ou seja, apresentarem como pontos cintilantes na imagem.

5.2.3 IMAGENS MUITO ENVELHECIDAS



Fig. 5.17– Imagem envelhecida com amarelado característico

Figura 5.3(a)(b)(c)(d) são imagens cujo conteúdo apresenta-se mais ou menos legível com um leve escurecimento ocasionado pelo tempo objetiva-se clarear o fundo destacando o texto para facilitar a leitura do pesquisador. O conjunto de filtros aplicados foi um algoritmo denominado *RealçarBorda* que é um algoritmo com a finalidade de destacar as partes com maior frequência. Em seguida, foi aplicado o filtro *Clarear* para realçar os pontos mais claros na imagem com o filtro da *MédiaClara* para suavizar a aparência da imagem. A desvantagem nesse conjunto de aplicações de filtros foi de realçar junto com o texto da imagem as partes borradas.

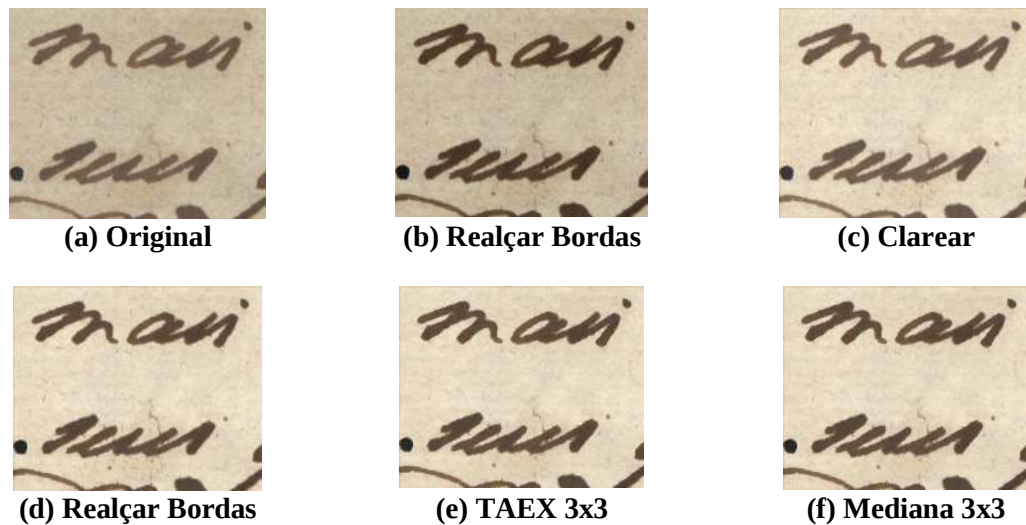


Fig. 5.18– Imagem envelhecida com escurecimento característico

Figura 5.4(a)(b)(c)(d)(e)(f) são imagens cujo conteúdo apresenta-se mais legível com um leve escurecimento ocasionado pelo tempo objetiva-se clarear o fundo destacando o texto para facilitar a leitura do pesquisador. O conjunto de filtros aplicados foi um algoritmo denominado *RealçarBorda* que é um algoritmo com a finalidade de destacar as partes com maior frequência. Em seguida, foi aplicado o filtro *Clarear* para realçar os pontos mais claros na imagem acompanhados dos filtros da *TAEX 3X3* e *Mediana 3x3* para remover ruídos e suavizar a aparência da imagem.

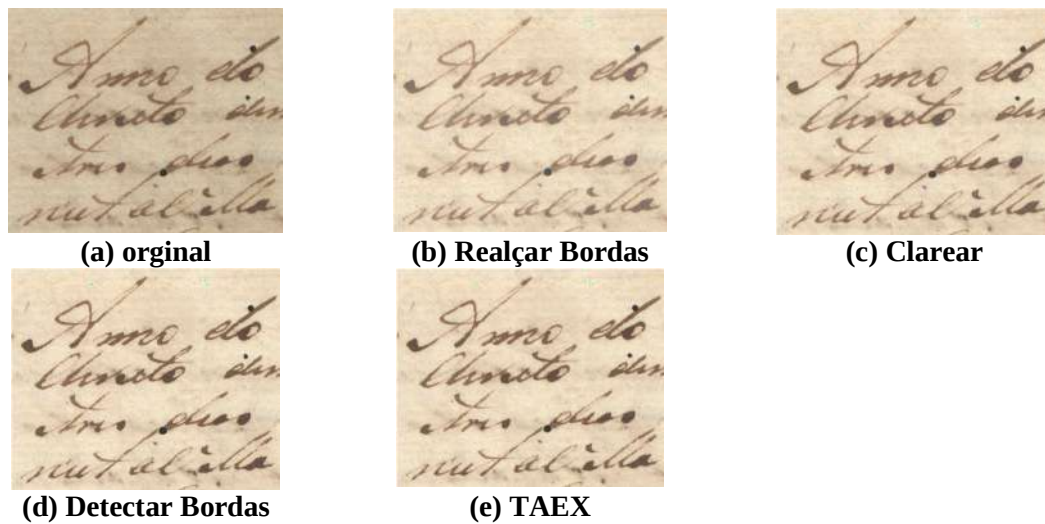


Fig. 5.19– Imagem envelhecida com escurecimento característico

Figura 5.5(a)(b)(c)(d) são imagens cujo conteúdo apresenta-se mais legível com um leve escurecimento ocasionado pelo tempo objetiva-se clarear o fundo destacando o texto para facilitar a leitura do pesquisador. O conjunto de filtros aplicados foi um algoritmo denominado *RealçarBorda* que é um algoritmo com a finalidade de destacar as partes com maior frequência. Em seguida, foi aplicado o filtro *Clarear* para realçar os pontos mais claros. Nesta imagem, como a escrita é de uma espessura fina, houve a necessidade de clarear mais o fundo tomando o cuidado para que não comprometesse as bordas limítrofes do texto. O filtros *TAEX 3X3* e *Mediana 3x3* foram utilizados para remover ruídos e suavizar a imagem.

5.2.4 IMAGENS BORRADAS E BASTANTE DETERIORADAS

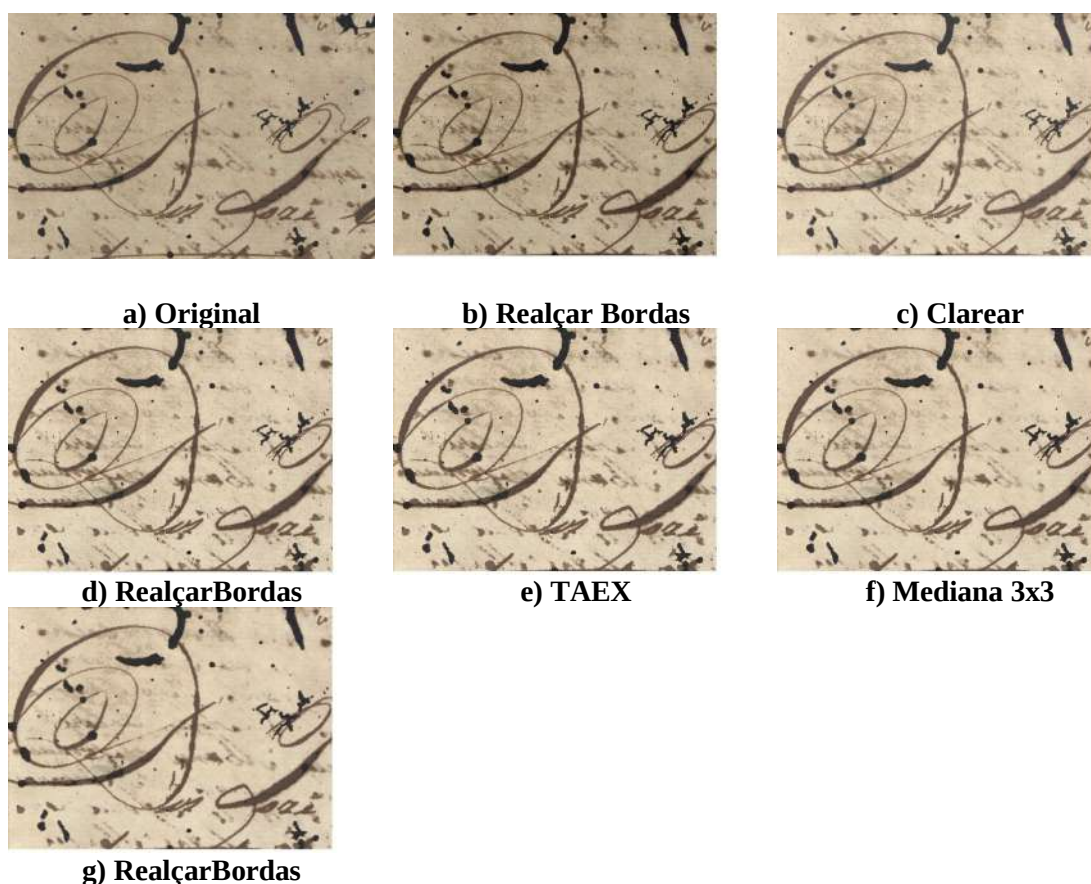


Fig. 5.20 – Imagem envelhecida e com muitos buracos e rasgos

Figura 5.6(a)(b)(c)(d)(e)(f)(g) são imagens cujo conteúdo apresenta-se muito pouco legível, escurecida e cheias de furos e rasgos objetiva-se clarear o fundo destacando o texto para facilitar a leitura do pesquisador, preocupando-se em não perder as características originais. O conjunto de filtros aplicados foi um algoritmo denominado *RealçarBorda* que é um algoritmo com a finalidade de destacar as partes com maior frequência. Em seguida, foi aplicado o filtro *Clarear* para realçar os pontos mais claros e novamente o *RealçarBorda*. Nesta imagem, como a escrita é de uma espessura fina, houve a necessidade de clarear mais o fundo de forma bastante discreta, tomando o cuidado para que não comprometesse as bordas limítrofes do texto. O filtros *TAEX 3X3* e *Mediana 3x3* foram utilizados para remover ruídos e suavizar a imagem.

5.2.5 IMAGENS MUITOS BORRADAS

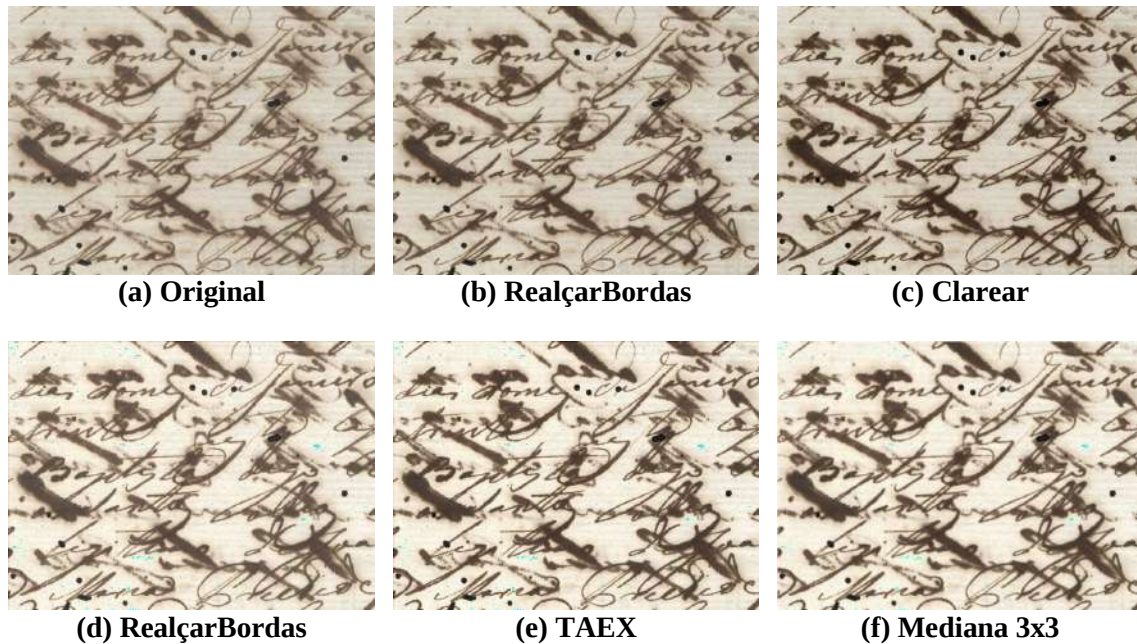


Fig. 5.21– Imagem muito borrada

Figura 5.7(a)(b)(c)(d)(e)(f)(g) são imagens cujo conteúdo apresenta-se muito pouco legível devido a excesso de borramento causado pelo tempo objetiva-se clarear o fundo discretamente destacando o texto para facilitar a leitura do pesquisador, preocupando-se em não perder as características originais. O conjunto de filtros aplicados foi um algoritmo denominado *RealçarBorda* que é um algoritmo com a finalidade de destacar as partes com maior frequência. Em seguida, foi aplicado o filtro *Clarear* para realçar os pontos mais claros e novamente o *RealçarBorda*. Nesta imagem, como a escrita é de uma espessura fina, houve a necessidade de clarear mais o fundo de forma bastante discreta, tomando o cuidado para que não compromettesse as bordas limítrofes do texto. O filtros *TAEX 3X3* e *Mediana 3x3* foram utilizados para remover ruídos e suavizar a imagem.

Avaliação dos Testes

Aplicando-se os filtros de realce conjuntamente, pôde-se observar que a medida que vai se clareando o fundo da imagem os ruídos vão aparecendo ocasionados pela oxidação da tinta no papel causando um amarelado característico de documentos antigos.

Estas foram as aplicações iniciais e consideradas interessantes, pois focalizaram partes características diferentes dos documentos, ampliando e melhorando a sua visualização.

Como já foi abordado, a utilização dos filtros de realce tem como objetivo melhorar a aparência das imagens digitais. A avaliação da eficiência dos filtros utilizados em conjunto dependerá do observador e da sua necessidade de visualização.

5.2.6 OUTROS TESTES

A seguir, será apresentado outros testes realizados com as imagens de documentos históricos. Para cada testes, será apresentado o conjunto de filtros utilizados no realce das imagens.

Os efeitos são similares aos anteriores, porém foram escolhidas outras variações de filtros para atingir resultados diferentes.

TESTE 1:



Fig. 5.22– Imagem clara

Figura 5.22(a)(b)(c)(d) são imagens cujo conteúdo apresenta-se muito claro e com pouca visibilidade objetiva-se destacar o texto do fundo para facilitar a leitura do pesquisador. O conjunto de filtros aplicados neste teste foram *RealçarBordas*, *TAEX* e *MediaClara*. A utilização de um conjunto diferente de filtros nesta imagem teve como objetivo atingir um resultado parecido ou melhor dos obtidos em testes anteriores. A dificuldade no tratamento de imagem clara é não deixar com que os pontos mais claros saturassem, ou seja, apresentarem como pontos cintilantes na imagem.

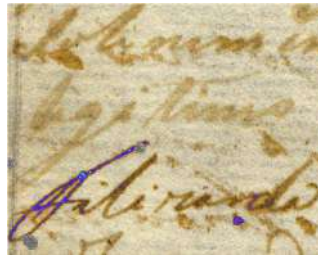
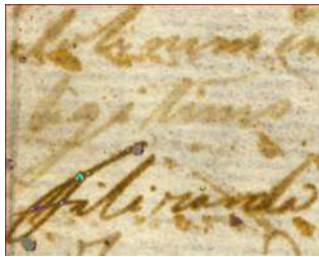
TESTE 2:**(a) original****(b) TrataClara****(c) Ordem Mínimo****(d) MediaClara****Fig. 5.23– Imagem clara**

Figura 5.23(a)(b)(c)(d) são imagens cujo conteúdo apresenta-se muito claro e com pouca visibilidade objetiva-se destacar o texto do fundo para facilitar a leitura do pesquisador. O conjunto de filtros aplicados neste teste foram *TrataClara*, *Ordem Mínimo* e *MediaClara*. A utilização de um conjunto diferente de filtros nesta imagem teve como objetivo atingir um resultado parecido ou melhor dos obtidos em testes anteriores. A dificuldade no tratamento de imagem clara é não deixar com que os pontos mais claros saturassem, ou seja, apresentarem como pontos cintilantes na imagem.

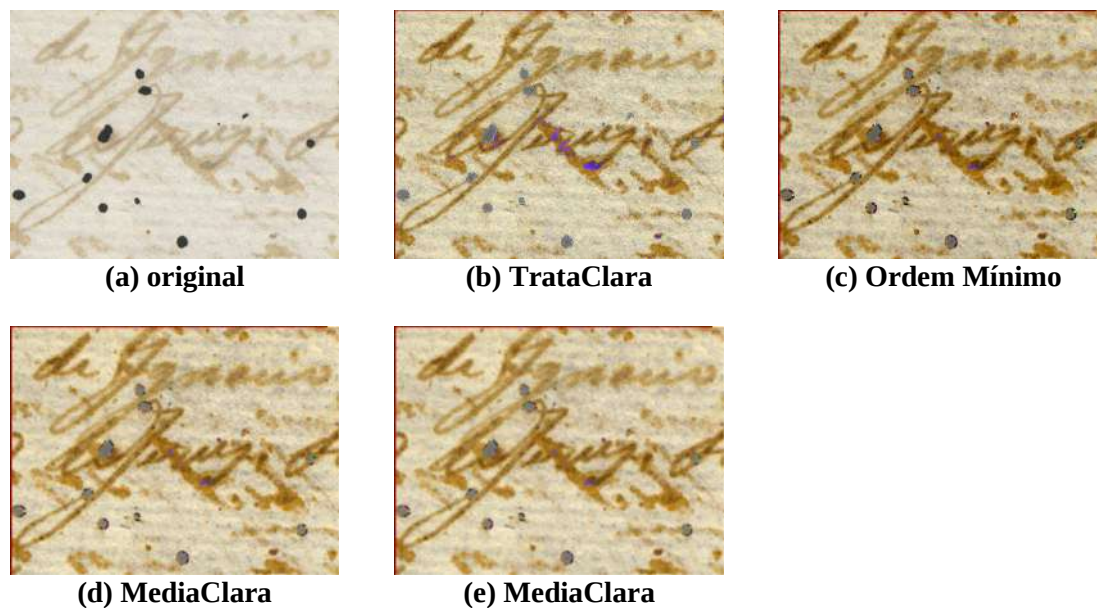
TESTE 3:**Fig. 5.24**– Imagem clara

Figura 5.24(a)(b)(c)(d)(e) são imagens cujo conteúdo apresenta-se muito claro e com pouca visibilidade objetiva-se destacar o texto do fundo para facilitar a leitura do pesquisador. O conjunto de filtros aplicados neste teste foram *TrataClara*, *Ordem Mínimo*, *MediaClara*, este último aplicado duas vezes. A utilização de um conjunto diferente de filtros nesta imagem teve como objetivo atingir um resultado parecido ou melhor dos obtidos em testes anteriores. A dificuldade no tratamento de imagem clara é não deixar com que os pontos mais claros saturassem, ou seja, apresentarem como pontos cintilantes na imagem.

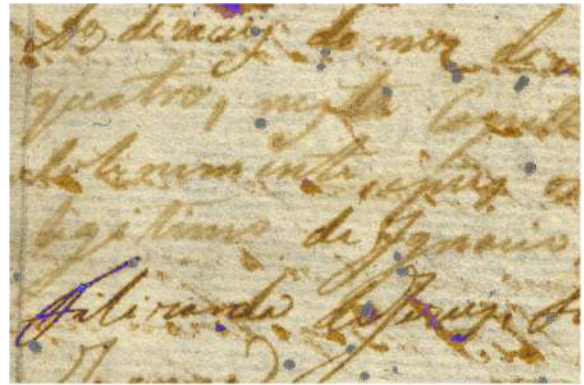
TESTE 4:**(a) Original****(b) TrataClara****(c) Ordem Mínimo****(d) MediaClara****(e) Mediana****Fig. 5.25— Imagem clara**

Figura 5.25(a)(b)(c)(d)(e) são imagens cujo conteúdo apresenta-se muito claro e com pouca visibilidade objetiva-se destacar o texto do fundo para facilitar a leitura do pesquisador. O conjunto de filtros aplicados neste teste foram *TrataClara*, *Ordem Mínimo*, *MediaClara* e *Mediana*. A utilização de um conjunto diferente de filtros nesta imagem teve como objetivo atingir um resultado parecido ou melhor dos obtidos em testes anteriores. A dificuldade no tratamento de imagem clara é não deixar com que os pontos mais claros saturassem, ou seja, apresentarem como pontos cintilantes na imagem.

TESTE 5:



Fig. 5.26– Imagem Homogênea

Figura 5.26(a)(b)(c)(d)(e) são imagens cujo conteúdo apresenta-se envelhecido e escurecida apesar de se apresentar como uma região homogênea. O conjunto de filtros aplicados neste teste foram *RealçarBorda*, *Moda*, *Mediana*, *MediaClara* e *RealçarBorda*. A utilização de um conjunto diferente de filtros nesta imagem teve como objetivo atingir um resultado parecido ou melhor dos obtidos em testes anteriores. A dificuldade no tratamento de

imagem envelhecida é não deixar com que a diferença na escrita, na aplicação do filtro, possibilite a perda de características da imagem original.

TESTE 6:

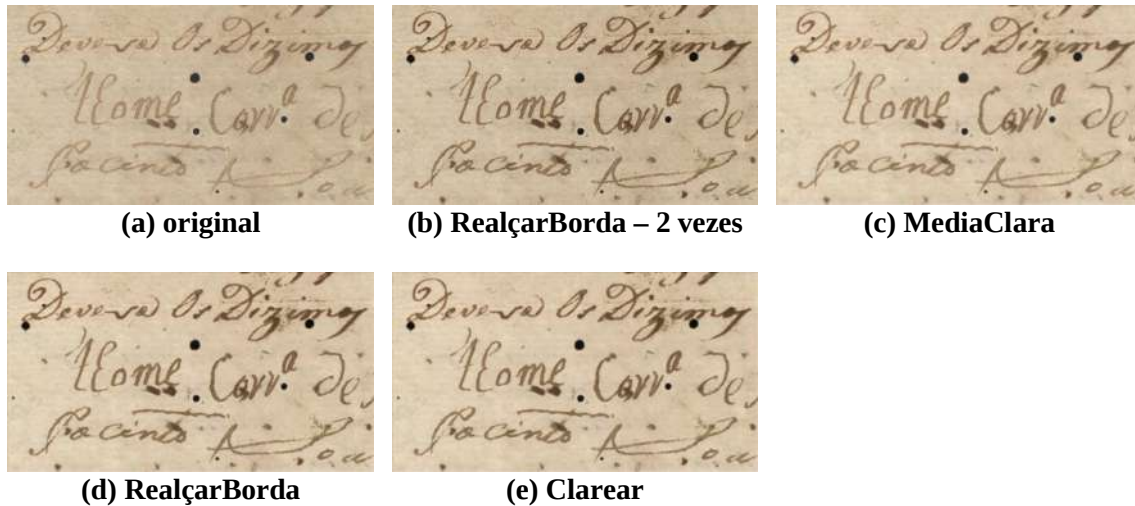


Fig. 5.27– Imagem Homogênea

Figura 5.26(a)(b)(c)(d)(e) são imagens cujo conteúdo apresenta-se envelhecido e escurecida apesar de se apresentar como uma região homogênea. O conjunto de filtros aplicados neste teste foram *RealçarBorda* aplicado duas vezes, *MediaClara*, *RealçarBorda* e *Clarear*. A utilização de um conjunto diferente de filtros nesta imagem teve como objetivo atingir um resultado parecido ou melhor dos obtidos em testes anteriores. A dificuldade no tratamento de imagem envelhecida é não deixar com que o realce na escrita possibilite a perda de características da imagem original.

6 CONCLUSÃO E TRABALHOS FUTUROS

O presente trabalho propôs metodologias que são utilizadas na preservação das imagens digitais. Também, apresenta uma proposta da utilização dos algoritmos de realce como passo inicial no tratamento dos documentos digitalizados. O objeto de pesquisa utilizado foi um documento histórico pertencente ao Arquivo Eclesiástico da Diocese de São João del-Rei.

Este trabalho tem como intuito iniciar discussões das questões acerca da digitalização dos documentos históricos do Arquivo Eclesiástico. Porém, a partir do que foi realizado, observa-se que outras atividades são necessárias para que se possa disponibilizar de forma efetiva as imagens dos documentos históricos.

Muito ainda precisa ser feito para realizar a conversão do acervo tradicional para o universo digital. Inicialmente, pode-se gerar uma estrutura de recuperação das imagens digitalizadas com a finalidade de acesso junto a base textual existente para ampliar as pesquisas e consultas realizadas pela comunidade interessada. Pode-se utilizar um recurso

mais especializado que faça a ligação da imagem e texto, como por exemplo, um banco de dados com suporte ao gerenciamento de dados complexos (imagens, textos, som, etc.).

Foi possível observar com os testes realizados, que num mesmo documento pode-se encontrar regiões degradadas com diferentes características, o que dificulta a aplicação de uma única técnica de melhoramento capaz de tratar o documento como um todo. Assim, torna-se necessário o desenvolvimento de um software capaz de detectar as diferentes regiões da imagem e tratá-las individualmente, através das técnicas propostas pelo processamento digital de imagem.

Para as atividades futuras propostas, seria necessário um estudo mais aprofundado, que fogue ao escopo deste trabalho, para possibilitar a realização da informatização do acervo documental do Arquivo Eclesiástico da Diocese de São João del-Rei..

O trabalho foi desenvolvido perante a necessidade em divulgar as preciosidades históricas contidas no Arquivo Eclesiástico da Diocese de São João del-Rei e a possibilidade de aplicar, na prática, muitos dos conceitos abordados.

REFERÊNCIAS BIBLIOGRÁFICAS

[GONZALEZ; WOODS, 2000] GONZALEZ, Rafael C.; WOODS, Richard E.. **Processamento Digital de Imagens**. Rio do Sul, SC: Editora Edgard Blücher Ltda, 2000.

[MASCARENHAS; VELASCO, 1982] MASCARENHAS, N. D. A. & VELASCO, F. R. D. **Processamento digital de Imagens**. São Paulo, Instituto de Matemática e Estatística, 1984

[MARQUES FILHOS; VIEIRA NETO, 1989] MARQUES FILHO, Ogê; VIEIRA NETO Hugo. **Processamento Digital de Imagens**. Editora Brasport, 1999

[ARAÚJO, 1989] ARAÚJO, Arnaldo de Albuquerque. **Técnica de Realce para Imagem**. 1989. Apostila do IX Congresso da SBC – UFU, Uberlândia.

[ARECO, 1995] ARECO, Eduardo Roberto. **Comparação de Algoritmos de Suavização Espacial**. 1995. Tese de Mestrado - UFMG, Belo Horizonte.

[VALLE JR.; 2003] VALLE JR., Eduardo Alves do. **Sistemas de Informação Multimídia na Preservação de Acervos Permanentes**. 2003. Tese de Mestrado - UFMG, Belo Horizonte.

[LEONE, 2002] SANT'ANA, Marcelo Leone. **Digitalização de Documentos de Arquivo: O Caso das Plantas de Parcelamento de Belo Horizonte**. 2002. Tese de Mestrado - UFMG, Belo Horizonte.

[BUSQUIM, 2002] BUSQUIM, Rodney. **Projeto de Implantação de um Robô Móvel**. Projeto de Mestrado em andamento - 2002. USP, São Paulo

[MARÇAL, 2001] MARÇAL, Joaquim. **Digitalização para Preservação**. Projeto Conservação Preventiva em Bibliotecas e Arquivos - CPBA – Fundação Biblioteca Nacional. 2001. Rio de Janeiro.

[SANTOS, 1995] SANTOS, Luís Paulo Peixoto dos. **Visão por Computador**. Texto de Apoio. Universidade do Ninho. 1995. Lisboa, Portugal.

[SCURI, 1999] SCURI, Antonio Escãno. **Fundamento da Imagem Digital**. TecGraf. 1999. PUC, Rio de Janeiro.

[WATERS, 2001] WATERS, Donald J. **Do Microfilme à imagem digital: Como executar um projeto para estudo dos meios, custos e benefícios de conversão para imagens digitais de grandes quantidades de documentos preservados em microfilme**. Projeto Conservação Preventiva em Bibliotecas e Arquivos – CPBA. 2001. Rio de Janeiro.

[CONWAY, 2001] CONWAY, Paul. **Preservação no Universo Digital**. Projeto Conservação Preventiva em Bibliotecas e Arquivos – CPBA. 2001. Rio de Janeiro.

[WILLIS, 2001] WILLIS, Don. **Uma abordagem de Sistemas Híbridos para a preservação de materiais impressos**. Projeto Conservação Preventiva em Bibliotecas e Arquivos – CPBA. 2001. Rio de Janeiro.

[KENNEY, CHAPMAN; 2001] KENNEY, Anne R.; CHAPMAN, Stephen. **Requisitos de resolução digital para textos: métodos para o estabelecimento de critérios de qualidade de imagem**. Projeto Conservação Preventiva em Bibliotecas e Arquivos – CPBA. 2001. Rio de Janeiro.

[OGDEN e OUTROS; 2001] OGDEN, Sherelyn; JONES, LeeC.; WALKER, Gay; JONES, Norvel M.N. **Reformatação**. Projeto Conservação Preventiva em Bibliotecas e Arquivos – CPBA. 2001. Rio de Janeiro.

ANEXO A

PASSOS NECESSÁRIOS AOS ALGORITMOS DE REALCE

Algoritmo da Média

- 1º) Coletar o valor da cor do pixel da imagem original e da sua vizinhança e armazenar em uma matriz quadrática. A matriz deverá ser preenchida da seguinte forma: elemento central é o pixel da imagem original e os outros elementos a respectiva vizinhança do pixel da imagem original;
- 2º) Calcular o valor da média aritmética dos elementos da matriz;
- 3º) Substituir no valor do pixel da imagem original o resultado obtido.
- 4º) Repetir o processo para todos os pixels da imagem original;

Algoritmo da Média com K vizinhos mais próximos

- 1º) Coletar o valor da cor do pixel da imagem original e da sua vizinhança e armazenar em uma matriz quadrática. A matriz deverá ser preenchida da seguinte forma: elemento central é o pixel da imagem original e os outros elementos a respectiva vizinhança do pixel da imagem original;

- 2º) Calcular o valor da média aritmética do elemento central da matriz com os k elementos definidos por uma variável cujo valor é dado pelo usuário, de acordo com o tamanho da matriz;
- 3º) Substituir no valor do pixel da imagem original o resultado obtido;

Algoritmo da Moda

- 1º) Coletar o valor da cor do pixel da imagem original e da sua vizinhança e armazenar em uma matriz quadrática. A matriz deverá ser preenchida da seguinte forma: elemento central é o pixel da imagem original e os outros elementos a respectiva vizinhança do pixel da imagem original;
- 2º) Encontrar o elemento cujo valor é mais freqüente na matriz
- 3º) Substituir no valor do pixel da imagem original pelo valor do elemento mais freqüente na matriz.
- 4º) Repetir o processo para todos os pixels da imagem original.

Algoritmo da Ordem

- 1º) Coletar o valor da cor do pixel da imagem original e da sua vizinhança e armazenar em uma matriz quadrática. A matriz deverá ser preenchida da seguinte forma: elemento central é o pixel da imagem original e os outros elementos a respectiva vizinhança do pixel da imagem original;
- 2º) Ordenar os valores dos elementos da matriz;
- 3º) Para o algoritmo da ordem mínimo, substituir o valor do pixel da imagem original pelo elemento da matriz de menor valor; para o algoritmo da ordem máxima, substituir o valor do pixel da imagem original pelo elemento da matriz de maior valor;
- 4º) Repetir o processo para todos os pixels da imagem original.

Algoritmo da Mediana

- 1º) Coletar o valor da cor do pixel da imagem original e da sua vizinhança e armazenar em uma matriz quadrática. A matriz deverá ser preenchida da seguinte forma: elemento central é o pixel da imagem original e os outros elementos a respectiva vizinhança do pixel da imagem original;
- 2º) Ordenar os valores dos elementos da matriz;
- 3º) Substituir o valor do pixel da imagem original pelo valor do elemento mediano da matriz ordenada;
- 4º) Repetir o processo para todos os pixels da imagem original.

Algoritmo Transformada de Aguçamento Extremo - TAEX

- 1º) Coletar o valor da cor do pixel da imagem original e da sua vizinhança e armazenar em uma matriz quadrática. A matriz deverá ser preenchida da seguinte forma: elemento central é o pixel da imagem original e os outros elementos a respectiva vizinhança do pixel da imagem original;
- 2º) Ordenar os valores dos elementos da matriz;
- 3º) Calcular o valor absoluto dos valores do elemento médio com o elemento de menor valor e dos valores do elemento médio com o elemento de maior valor da matriz ordenada;
- 4º) Substituir o valor do pixel da imagem original pelo valor do elemento que resultou na menor diferença absoluta;
- 5º) Repetir o processo para todos os pixels da imagem original.

Algoritmo Passa-alta

- 1º) Coletar o valor da cor do pixel da imagem original;
- 2º) Se o valor da cor do pixel for menor do que 128 então substituir o valor do pixel da imagem original para 0;
- 3º) Repetir o processo para todos os pixels da imagem original.

Algoritmo Passa-Baixa

- 1º) Coletar o valor da cor do pixel da imagem original;
- 2º) Se o valor da cor do pixel for maior do que 128 então substituir o valor do pixel da imagem original para 255;
- 3º) Repetir o processo para todos os pixels da imagem original.

Algoritmo Escurecer

- 1º) Coletar o valor da cor do pixel da imagem original;
- 2º) Se o valor da cor do pixel estiver na faixa entre 0 e 180 então diminuir 50 o valor do pixel da imagem original;
- 3º) Repetir o processo para todos os pixels da imagem original.

Algoritmo Clarear

- 1º) Coletar o valor da cor do pixel da imagem original;
- 2º) Se o valor da cor do pixel estiver na faixa entre 170 e 250 então aumentar 30 o valor do pixel da imagem original;

- 3º) Se o valor da cor do pixel estiver na faixa entre 0 e 100 então substituir para 3 o valor do pixel da imagem original;
- 4º) Se o valor da cor do pixel estiver na faixa maior que 255 então atribuir 245 ao valor do pixel da imagem original;
- 5º) Se o valor da cor do pixel estiver na faixa menor que 0 então atribuir 0 ao valor do pixel da imagem original;
- 6º) Repetir o processo para todos os pixels da imagem original.

Algoritmo TrataClara

- 1º) Coletar o valor da cor do pixel da imagem original e de sua vizinhança e armazenar numa matriz com peso de 27 no elemento central e -1 nos outros elementos;
- 2º) Calcular a média aritmética dos elementos da matriz;
- 3º) Se o valor resultante da média for maior que 245 então diminuir 5 do valor resultante e substituir no valor do pixel da imagem original;
- 4º) Se o valor resultante da média for menor que 0 então atribuir 0 do valor resultante e substituir no valor do pixel da imagem original;
- 5º) Repetir o processo para todos os pixels da imagem original.

Algoritmo MediaClara

- 1º) Coletar o valor da cor do pixel da imagem original e da sua vizinhança e armazenar em uma matriz quadrática. A matriz deverá ser preenchida da seguinte forma: elemento central é o pixel da imagem original e os outros elementos a respectiva vizinhança do pixel da imagem original;
- 2º) Calcular o valor da média aritmética dos elementos da matriz;

- 4º) Se o valor resultante da média for menor que 0 então atribuir 0 do valor resultante e substituir no valor do pixel da imagem original;
- 5º) Se o valor resultante da média for maior que 255 então atribuir 245 do valor resultante e substituir no valor do pixel da imagem original;
- 6º) Se o valor resultante da média não estiver em nenhuma das faixas estipuladas anteriormente então aumentar 30 no valor resultante da média e substituir no valor do pixel da imagem original;
- 7º) Repetir o processo para todos os pixels da imagem original.

Algoritmo RealçarBorda

- 1º) Coletar o valor da cor do pixel da imagem original;
- 2º) Se o valor da cor do pixel estiver na faixa maior que 130 e menor que 180 então diminuir 15 ao valor do pixel da imagem original senão atribuir 0 ao valor do pixel da imagem original;
- 3º) Repetir o processo para todos os pixels da imagem original.