



UNIVERSIDADE ESTADUAL DE CAMPINAS  
FACULDADE DE CIÊNCIAS MÉDICAS

WEBER SHOITY RESENDE TAKAKI

DESENVOLVIMENTO E ANÁLISE DE MÉTODO PARA CODIFICAÇÃO E  
INDEXAÇÃO DE IMAGENS E VÍDEOS PARA EXAMES MÉDICOS COM  
APLICAÇÕES EM TELEMEDICINA

CAMPINAS

2020

WEBER SHOITY RESENDE TAKAKI

DESENVOLVIMENTO E ANÁLISE DE MÉTODO PARA CODIFICAÇÃO E  
INDEXAÇÃO DE IMAGENS E VÍDEOS PARA EXAMES MÉDICOS COM  
APLICAÇÕES EM TELEMEDICINA

Tese apresentada à Faculdade de Ciências Médicas da Universidade  
Estadual de Campinas como parte dos requisitos exigidos para a  
obtenção do título de Doutor em Ciências.

ORIENTADOR: PROFESSOR DOUTOR WU FENG CHUNG  
COORIENTADOR: PROFESSORA DOUTORA HUEI DIANA LEE

ESTE TRABALHO CORRESPONDE À VERSÃO  
FINAL DA TESE DEFENDIDA PELO ALUNO  
WEBER SHOITY RESENDE TAKAKI, E ORIENTADO  
PELO PROF. DR. WU FENG CHUNG.

CAMPINAS

2020

Ficha catalográfica  
Universidade Estadual de Campinas  
Biblioteca da Faculdade de Ciências Médicas  
Maristella Soares dos Santos - CRB 8/8402

T139d Takaki, Weber Shoity Resende, 1976-  
Desenvolvimento e análise de método para codificação e indexação de  
imagens e vídeos para exames médicos com aplicações em telemedicina /  
Weber Shoity Resende Takaki. – Campinas, SP : [s.n.], 2020.

Orientador: Wu Feng Chung.

Coorientador: Huei Diana Lee.

Tese (doutorado) – Universidade Estadual de Campinas, Faculdade de  
Ciências Médicas.

1. Colonoscopia. 2. Gastroscoopia. 3. Endoscopia. 4. Compressão de dados.  
5. Armazenamento e recuperação da informação. I. Wu, Feng Chung. II. Lee,  
Huei Diana. III. Universidade Estadual de Campinas. Faculdade de Ciências  
Médicas. IV. Título.

Informações para Biblioteca Digital

**Título em outro idioma:** Development and analysis of image and video coding and indexing  
method for medical examinations with applications in telemedicine

**Palavras-chave em inglês:**

Colonoscopy

Gastroscoopia

Endoscopia

Data compression

Information storage and retrieval

**Área de concentração:** Fisiopatologia Cirúrgica

**Titulação:** Doutor em Ciências

**Banca examinadora:**

Wu Feng Chung [Orientador]

Adriano Fregonesi

Orlando Petrucci Júnior

Jefferson Tales Oliva

Oscar Kenji Nihei

**Data de defesa:** 18-02-2020

**Programa de Pós-Graduação:** Ciências da Cirurgia

**Identificação e informações acadêmicas do(a) aluno(a)**

- ORCID do autor: <https://orcid.org/0000-0002-3540-5732>

- Currículo Lattes do autor: <http://lattes.cnpq.br/3007666905986319>

---

# **COMISSÃO EXAMINADORA DA DEFESA DE DOUTORADO**

**WEBER SHOITY RESENDE TAKAKI**

---

**ORIENTADOR: PROF. DR. WU FENG CHUNG**

**COORIENTADOR: PROF<sup>a</sup>. DR<sup>a</sup>. HUEI DIANA LEE**

---

## **MEMBROS:**

- 1. PROF. DR. WU FENG CHUNG.**
  - 2. PROF. DR. ADRIANO FREGONESI.**
  - 3. PROF. DR. ORLANDO PETRUCCI JÚNIOR.**
  - 4. PROF. DR. JEFFERSON TALES OLIVA.**
  - 5. PROF. DR. OSCAR KENJI NIHEI.**
- 

Programa de Pós-Graduação em Ciências da Cirurgia da Faculdade de Ciências Médicas da Universidade Estadual de Campinas.

A ata de defesa com as respectivas assinaturas dos membros encontra-se no SIGA/Sistema de Fluxo de Dissertação/Tese e na Secretaria do Programa da FCM.

**Data de Defesa: 18/02/2020**

---

## DEDICATÓRIA

À Patrícia, companheira querida, pelo suporte constante e providencial, e por esforçar-se junto comigo com alegria e persistência, em demonstração clara de amor e abnegação.

Aos pais, Neuza e Norio (em memória), pelas oportunidades desta vida, por ensinar-me o valor do trabalho, da educação e do esforço pessoal, e pelos exemplos de generosidade, honestidade, respeito e dedicação.

## **AGRADECIMENTOS**

Ao Prof. Dr. Wu Feng Chung pela oportunidade de estar neste caminho de aprendizado e crescimento acadêmico, científico e pessoal representado pelo doutorado em Ciências da Cirurgia da Unicamp. Sou profundamente grato pela confiança, pela orientação segura e precisa, pelo incentivo constante e justo – apesar dos meus erros e dificuldades, e pelos exemplos cotidianos de liderança, ética, competência, discernimento, generosidade e dedicação incansável à Ciência, à Educação, ao bem comum e ao Brasil.

À Prof<sup>a</sup>. Dr<sup>a</sup>. Huei Diana Lee por ter-me conduzido neste belo e enriquecedor caminho acadêmico desde antes do mestrado, ainda na condição de aluno especial, até este momento. Sou profundamente grato pelas oportunidades recebidas, pelos ensinamentos e pela experiência científica e acadêmica compartilhadas, pelos constantes exemplos de dedicação, zelo, ética, generosidade, competência, responsabilidade, e de esforço em favor da Educação, da Ciência, do bem comum e do Brasil.

Aos Amigos Huei e Wu, pela amizade, por compartilhar conosco suas ideias e ideais, sua cultura, suas experiências de vida e seu conhecimento, com alegria, cuidado e muita generosidade. Estes anos de convivência tem sido um verdadeiro privilégio.

Aos Amigos do LABI pela convivência profícua, produtiva e alegre, pelos conhecimentos e experiências compartilhadas, e por formar um ambiente enriquecedor, harmonioso e ético, tão importante para o desenvolvimento acadêmico e científico.

Aos Amigos Newton Spolaôr e Leandro Ensina, pelo suporte e disponibilidade constantes, pelo intercâmbio técnico e científico e pela amizade.

Ao Professor Cláudio Saddy Rodrigues Coy pelo apoio e pelas oportunidades oferecidas ao desenvolvimento da pesquisa.

À Inova Unicamp, e em particular à equipe de Propriedade Intelectual – Teresa Fukase, Janaína César e Raquel Barbosa, pelo suporte e pela condução, com profissionalismo e dedicação, do processo de patenteamento do método.

Ao Amarildo Stabile Júnior pelo apoio e pela disponibilidade constante e paciente em relação às questões administrativas do programa.

Ao Programa de Pós-Graduação em Ciências da Cirurgia, pela oportunidade de participação neste programa, apesar de minha formação técnica em Computação.

Ao Narco Afonso Maciejewski pelas ideias e contribuições dadas na fase inicial do desenvolvimento do método.

Aos colegas da Unila, por permitirem a redução da minha jornada de trabalho, dando-me cobertura nos períodos em que estive ausente para dedicar-me ao doutorado.

Aos familiares e amigos, pela torcida, pelo incentivo, e pelo respeito e compreensão em relação à minha ausência em diversos momentos.

## RESUMO

O surgimento da videoendoscopia digestiva alta e baixa e da telemedicina constituem importantes avanços tecnológicos para o diagnóstico e o treinamento em procedimentos de exames do trato gastrointestinal por métodos ópticos. Nesse cenário, a transmissão e o armazenamento de imagens e vídeos digitais demandam técnicas de compressão adequadas às características da aplicação. Além disso, com a formação de grandes bases de dados, a recuperação por conteúdo das imagens de modo efetivo é também necessária. Assim sendo, este trabalho propõe um método original de codificação, indexação e recuperação de imagens e vídeos. O método usa a decomposição em valores singulares como técnica algébrica de fatoração que permite a ordenação dos fatores constituintes por nível de importância na constituição da imagem. O método apresenta, também, uma análise estatística prévia de cores obtidas de amostras representativas, formando um mapa de probabilidade baseado na frequência de cores, e que resulta na representação compacta das mesmas. Por meio de experimentos realizados com 2.700 quadros de endoscopia digestiva alta, o método de codificação alcançou taxa de compressão média de 82,07%, com desvio padrão de 6,37%, e qualidade objetiva de 38,85 dB com desvio padrão de 1,06 dB. Em comparação com o padrão de codificação H.263, houve diferença estatisticamente significativa em termos da taxa de compressão e da qualidade objetiva ( $p \leq 0,05$ ). Para a recuperação, foram indexados 7.212 quadros e utilizadas 97 imagens de exemplo para a busca. Como resultado, a precisão média alcançada foi de 94,85% na configuração que obteve o melhor desempenho geral, tendo sido igual ou superior a outros métodos reportados na literatura. Por meio dos experimentos realizados conclui-se que o método original proposto foi eficaz em codificar e indexar automaticamente vídeos de exames médicos para aplicações de telemedicina.

Palavras-chaves: Colonoscopia; gastroscopia; endoscopia; compressão de dados; armazenamento e recuperação da informação.



## ABSTRACT

The emergence of upper and lower gastrointestinal (GI) video endoscopy and telemedicine are essential technological advancements for the diagnosis and training related to examination procedures of the digestive tract using optical methods. In this context, transmission and storage of digital images and videos require compression techniques that are suitable for the application characteristics. Moreover, as the media databases grow massive, effective search and retrieval by image content become necessary. Thus, this work proposes an original method for coding, indexing and retrieving images and videos. The proposed method uses the singular value decomposition as the algebraic technique for matrix factorization that allows sorting the components by their level of relevance in the image composition. The process also presents a statistical analysis of colors, obtained from typical image samples, to form a probability map based on the color frequency, that further leads to a compact representation. Through the experiments executed on 2,700 frames of upper GI endoscopy, the coding method achieved an average compression ratio of 82.07% with a standard deviation of 6.37%, and objective quality of 38.85 dB with a standard deviation of 1.06 dB. In comparison with the H.263 coding standard, it was observed a statistically significant difference in terms of compression ratio and objective quality ( $p \leq 0,05$ ). As for the retrieval, 7212 frames were indexed and 97 sample images were used as search queries. As a result, an average precision of 94,85% was achieved for the configuration with best overall performance, outperforming other methods reported in the literature. Based on the results of the experiments it is possible to conclude that the proposed original method was effectively able to encode and automatically index medical examination videos for telemedicine applications.

Keywords: Colonoscopy; gastroscopy; endoscopy; data compression; information storage and retrieval.

## LISTA DE ILUSTRAÇÕES

Figura 1 – Posição dos índices da matriz de pixels no sistema de coordenadas....	21
Figura 2 – Vista aproximada dos pixels de uma área da imagem. ....	22
Figura 3 – Imagem original decomposta em componentes R, G e B. ....	23
Figura 4 – Sensibilidade dos cones S, M e L segundo o comprimento de onda. ....	23
Figura 5 – Gráfico de mapeamento de cores da CIE. ....	24
Figura 6 – Espaço de cores YUV para um nível de luminância específico. ....	25
Figura 7 – Camadas do modelo YCbCr. ....	26
Figura 8 – Sequência de quadros de vídeo.....	26
Figura 9 – Interpretação geométrica da fatoração de A no valor singular $\sigma_{i,i}$ e nos vetores singulares $v_i$ e $u_i$ .....	28
Figura 10 – Interpretação geométrica da fatoração de A para obtenção do segundo valor singular $\sigma_{i,i}$ e vetores singulares $v_i$ e $u_i$ respectivos. ....	28
Figura 11 – Exemplo de indexação por rótulo.....	37
Figura 12 – Histograma de cores RGB do quadro nº 1852. ....	38
Figura 13 – Resultado da aplicação do filtro <i>Local Range</i> ao quadro nº 1852.....	39
Figura 14 – Resultado da aplicação do filtro Sobel ao quadro nº 1852.....	39
Figura 15 – Diagrama de blocos dos métodos .....	49
Figura 16 – Passos da Etapa Experimental 1 .....	52
Figura 17 – Árvore de cores mais frequentes.....	53
Figura 18 – Tabela de conversão para pré-codificação de cores.....	53
Figura 19 – Ilustração dos passos da fatoração do bloco. ....	54
Figura 20 – Codificação binária e serialização dos valores e vetores singulares....	56
Figura 21 – Efeito da sumarização sobre um autovetor. ....	57
Figura 22 – Atualização da tabela de índices.....	58
Figura 23 – Passos da Etapa Experimental 2. ....	58
Figura 24 – Fatoração e reconstrução gradual.....	66
Figura 25 – Imagem codificada e reconstruída com método proposto e H.263 .....	67
Figura 26 – Recuperação do quadro nº 696 .....	74
Figura 27 – Recuperação do quadro nº 1206 .....	74

Figura 28 – Recuperação do quadro nº 2240 .....	75
Figura 29 – Recuperação do quadro nº 2394 .....	75
Figura 30 – Recuperação do quadro nº 3370 .....	76

## LISTA DE TABELAS

Tabela 1 – Expressões de busca .....	51
Tabela 2 – Faixas de valores de $R^2$ consideradas para ordenação .....	60
Tabela 3 – Resumo das configurações aplicadas à recuperação .....	61
Tabela 4 – Resultados da busca ampla para a codificação, atualizada em maio de 2018 .....	65
Tabela 5 – Resultados da busca específica para a codificação, atualizada em maio de 2018 .....	65
Tabela 6 – Resultados da busca ampla para a indexação, realizada em junho e julho de 2018 .....	65
Tabela 7 – Resultados da busca específica para a indexação, realizada em julho de 2018 .....	66
Tabela 8 – Desempenho do método de codificação .....	67
Tabela 9 – Análise estatística dos resultados .....	67
Tabela 10 – Resultados dos experimentos de recuperação considerando os três primeiros quadros recuperados .....	68
Tabela 11 – Precisão da recuperação considerando os cinco e os dez primeiros quadros recuperados .....	69
Tabela 12 – Quadros recuperados para o Caso 1 .....	69
Tabela 13 – Quadros recuperados para o Caso 2 .....	72

## LISTA DE ABREVIATURAS E SIGLAS

ACM	<i>Association for Computing Machinery</i>
BPP	<i>Bits per pixel</i>
BRIEF	<i>Binary robust independent elementary features</i>
CBIR	<i>Content-based image retrieval</i>
CBVR	<i>Content-based video retrieval</i>
CBVIR	<i>Content-based video indexing and retrieval</i>
CEDD	<i>Color and edge directivity descriptor</i>
CIPO	<i>Canadian Intellectual Property Office</i>
DCT	<i>Discrete cosine transform</i>
DICOM	<i>Digital imaging and communication in medicine</i>
DPCM	<i>Differential pulse-code modulation</i>
DPI	<i>Dots per inch</i>
DVC	<i>Distributed video coding</i>
DVS	Decomposição em valores singulares
DWT	<i>Discrete wavelet transform</i>
FAST	<i>Features from accelerated segment test</i>
FLV	<i>Flash video</i>
FPO	<i>Free patents online</i>
HEVC	<i>High efficiency video coding</i>
HSV	<i>Hue saturation value</i>
INCA	Instituto Nacional de Câncer José Alencar Gomes da Silva
INPI	Instituto Nacional da Propriedade Industrial
JPEG	<i>Joint Photographic Experts Group</i>
JPEG-LS	<i>Lossless JPEG</i>
JPO	<i>Japan Patent Office</i>
KLT	<i>Karhunen-Loève transform</i>
MPEG	<i>Moving Picture Experts Group</i>
MSER	<i>Maximally stable external regions</i>

ORB	<i>Oriented FAST and rotated BRIEF</i>
PCA	<i>Principal component analysis</i>
PHOG	<i>Pyramid of histograms of orientation gradients</i>
PSNR	<i>Peak signal-to-noise ratio</i>
QPS	<i>Quadros por segundo</i>
RGB	<i>Red green blue</i>
RLE	<i>Run-length encoding</i>
RMS	<i>Root mean square</i>
SIFT	<i>Scale invariant feature transform</i>
SIMPLE	<i>Searching images with MPEG-7-like powered localized descriptors</i>
SPIHT	<i>Set partitioning in hierarchical trees</i>
SSIM	<i>Structural similarity index</i>
SURF	<i>Speeded up robust features</i>
USPTO	<i>United States Patent and Trademark Office</i>
VLC	<i>Variable length coding</i>
WIPO	<i>World Intellectual Property Organization</i>

## SUMÁRIO

1. INTRODUÇÃO .....	16
2. OBJETIVO .....	20
3. REVISÃO DA LITERATURA E BASE DE DADOS MUNDIAIS DE PATENTES ...	21
3.1. Imagens e vídeos digitais .....	21
3.2. Decomposição em valores singulares .....	27
3.3. Codificação de imagens e vídeos .....	29
3.4. Indexação e recuperação .....	37
4. MATERIAL E MÉTODOS.....	48
4.1. Material.....	48
4.2. Métodos.....	48
4.3. Busca de anterioridade.....	50
4.4. Delineamento experimental .....	52
4.4.1. Etapa 1: Codificação e indexação .....	52
4.4.2. Etapa 2: Recuperação .....	58
4.4.3. Etapa 3: Análise estatística .....	64
4.4.4. Etapa 4: Avaliação do resultado da recuperação .....	64
5. RESULTADOS.....	65
5.1. Busca de anterioridade.....	65
5.2. Etapa 1 – Codificação e indexação .....	66
5.3. Etapa 2 – Recuperação .....	68
6. DISCUSSÃO .....	77
6.1. Codificação.....	81
6.2. Indexação e recuperação .....	86
7. CONCLUSÃO .....	92
8. GLOSSÁRIO .....	93
9. REFERÊNCIAS.....	105
ANEXOS .....	121
A – Parecer do Comitê de Ética em Pesquisa .....	121
B – Depósito da patente do método proposto – primeira página .....	126
C – Tabelas de resultados dos experimentos de recuperação .....	127

## 1. INTRODUÇÃO

O câncer no trato digestório é considerado uma das enfermidades mais graves na medicina, pois exige acompanhamento e terapêutica complexa e por essa razão torna-se vital detectar precocemente essa patologia (1-3).

De acordo com o Instituto Nacional de Câncer José Alencar Gomes da Silva (INCA), a neoplasia maligna do cólon e do reto configura-se como a terceira modalidade de maior ocorrência entre homens, e segunda entre mulheres (3-5). Até o ano de 2013, o câncer colorretal consistia no terceiro tipo mais frequente entre as mulheres, porém a partir de 2014 passou a ocupar a segunda posição, superando o câncer de colo do útero em novos casos (4-6). Vale ressaltar que, para cada ano do biênio 2018-2019, essa instituição estima o surgimento de câncer nesses segmentos intestinais em aproximadamente 6,24% do total de novos casos (3).

Decorrente à evolução grave, a detecção precoce de lesões pré-malignas ou malignas constitui-se fator decisivo, pois a constatação dessas lesões aumenta significativamente a eficácia do tratamento, e, conseqüentemente, a cura (7-9).

O tipo mais frequente de neoplasia colorretal, o adenocarcinoma, quando diagnosticado ainda na fase inicial da enfermidade, apresenta percentual de sobrevida dos pacientes em cinco anos entre 90 a 100%, e essa condição tornou-se ainda mais viável com o advento dos exames colonoscópicos (10-13).

Sob esse aspecto, a investigação do cólon completo foi viabilizada na década de 1960 com o advento da fibra óptica, sendo esse componente utilizado na concepção do fibroendoscópio (14). Antes, os instrumentos para inspeção da região intraluminal eram rígidos, e por isso limitavam-se ao canal anal, reto e região terminal do cólon sigmoide (15).

Na década de 1980 e 1990, com o surgimento do videocolonoscópio e dos aparelhos com magnificação de imagem, respectivamente, foram alcançados avanços tecnológicos importantes que favoreceram, em muito, o diagnóstico médico, e, conseqüentemente, as indicações terapêuticas e de acompanhamento dos pacientes (16-18). Vale ressaltar que nesse mesmo período iniciou-se no Brasil o esforço de interconexão de redes universitárias e de pesquisa à rede mundial de computadores, bem como sua expansão para redes comerciais, marcando assim, o início da Internet no Brasil (19).



O advento da Internet provocou mudanças fundamentais nas tecnologias voltadas para o campo da saúde, e, especificamente, a área de telemedicina ganhou impulso nesse período devido à redução nos custos dos canais de comunicação de dados e aos avanços nas técnicas de compressão de imagens e de vídeos (20,21).

Historicamente, o primeiro relato de assistência remota considerada pioneira em telemedicina data de 1879, e foi realizada por via telefônica (21). Desde então a telemedicina tem demonstrado ser uma ferramenta útil na melhoria da prestação de assistência médica, especialmente em regiões isoladas ou distantes dos grandes centros urbanos. Em particular, a redução de custos e a celeridade no atendimento ao evitar deslocamentos, a disseminação da medicina especializada para regiões remotas e a possibilidade de aumento da cooperação entre pesquisadores e especialistas, são os principais benefícios obtidos com essa tecnologia (22-24).

Nesse cenário, a transmissão de imagens e de vídeos constitui a principal aplicação em telemedicina, e dentre as modalidades de assistência mediada por tecnologias descritas por Hoyt e Fleury (25) e por Starren et al. (21) apenas a teleconsulta por mensagens eletrônicas não utiliza comunicação audiovisual. No Brasil, o Conselho Federal de Medicina publicou a Resolução nº 2.227/2018 (26) autorizando novas modalidades de assistência por telemedicina apoiadas no compartilhamento síncrono e ou assíncrono de imagens e ou de vídeos. Dentre essas modalidades, apenas no telemonitoramento a comunicação audiovisual não era obrigatória. Essa resolução foi revogada um mês após sua publicação por meio da Resolução nº 2.228/2019 (27), porém indicou o caminho provável de evolução da telemedicina no Brasil, e reforçou o papel central das tecnologias de transmissão de imagens e de vídeo nesse contexto.

Além da transmissão, é também necessário nos sistemas de telemedicina a gravação e o armazenamento de todas as comunicações para posterior consulta ou auditoria. Nesse sentido, a transmissão e o armazenamento de imagens e de vídeos digitais tornaram-se viáveis graças ao desenvolvimento de técnicas de compressão, tais como os padrões *Joint Photographic Experts Group* (JPEG) e *Motion Picture Experts Group* (MPEG), respectivamente (28,29). Sem compressão, a transmissão, em tempo real, e o armazenamento seriam inviabilizados pelo enorme volume de dados gerados pelas imagens e vídeos (30,31).

Por esse motivo, a codificação de imagens e ou de vídeos digitais envolve a definição de algoritmos e técnicas de compressão de dados, mas seu principal

objetivo é estabelecer a sintaxe do fluxo de bits que o codificador e o decodificador devem obedecer (32,33). Apesar de existirem formatos sem compressão para imagens e vídeos, por padrão a codificação implica na compressão dos dados e, decorrente ao fato, os termos “compressão” e “codificação” são frequentemente usados como sinônimos no contexto de imagens e de vídeos digitais (34). Nesse sentido, a codificação está associada aos padrões e formatos de imagem e de vídeo, tais como JPEG e MPEG, enquanto a compressão trata das técnicas e algoritmos utilizados pelos codificadores, tais como a transformada discreta de cosseno (DCT), a transformada *wavelet* discreta (DWT) e a codificação de comprimento variável (VLC) (28-34).

Em se tratando de imagens médicas, o padrão *Digital imaging and communications in medicine* (DICOM) suporta a representação de imagens sem compressão. No entanto, a necessidade de compressão foi observada desde a sua concepção (35), sendo aceitos atualmente os formatos JPEG, *Run-length encoding* (RLE), ZIP, JPEG2000 e JPEG-LS (36).

Sob a ótica da engenharia, imagens e vídeos são sinais multidimensionais, e por esse motivo as técnicas de codificação aplicáveis foram, na maioria das vezes, herdadas das técnicas desenvolvidas para a análise e o processamento de sinais unidimensionais (31,33). Um sinal unidimensional é uma sequência ordenada de números que descrevem as tendências e variações de uma grandeza física, medida ao longo do tempo e ou do espaço (37).

Dentre os padrões de compressão de imagens bidimensionais já propostos até os dias atuais, o JPEG foi o primeiro esforço conjunto para elaboração de um padrão internacional de compressão de imagens digitais (28). A primeira versão foi lançada em 1992, mas ainda hoje é amplamente utilizada, sendo referência em compressão de imagens (28,38,39).

Em relação à compressão de vídeos, as mesmas técnicas utilizadas na compressão de imagens são também aplicadas para compactar quadros individuais. No entanto, as redundâncias temporais, presentes nas sequências de quadros do vídeo, são também eliminadas por meio da aplicação de técnicas de estimação e compensação de movimento. Nesse contexto, o padrão MPEG foi um marco e principal referência em codificação de vídeo, criado também por um esforço conjunto para elaboração de padrões internacionais para a indústria de áudio e de vídeo digital (29,40,41).

Esses padrões de codificação contribuíram para massificar a tecnologia de imagens e de vídeos digitais em diversas áreas, como exemplo a médica (30,36,42,43). Desse modo, com o acúmulo de vídeos e de imagens armazenadas nas bases de dados de hospitais e clínicas, torna-se também necessário adicionar a essas bases de dados, maneiras eficientes de recuperar o conteúdo. É digno de nota que a indexação e a recuperação de imagens e de vídeos baseados em conteúdo são áreas de pesquisa que buscam preencher essa lacuna (44-46).

O procedimento habitual de busca em bases de dados por expressões textuais, apesar de natural para os usuários, em diversos casos não permite expressar com precisão todo o conteúdo da imagem ou do vídeo desejado e, por esse motivo, abordagens alternativas tem sido investigadas e propostas. Essas abordagens são denominadas *Content-based image retrieval* (CBIR) e *Content-based video retrieval* (CBVR), ou ainda *Content-based video indexing and retrieval* (CBVIR), e consistem na recuperação de imagens e vídeos baseada em conteúdo (47-51).

Diferentes técnicas de indexação e recuperação baseada em conteúdo já foram propostas e avaliadas, bem como diferentes métodos de codificação de imagens e vídeos. Porém, as abordagens já relatadas tratam cada aspecto de maneira isolada e independente. Como resultado, a indexação é usualmente realizada *a posteriori*, depois da transmissão ou do armazenamento da imagem ou do vídeo codificado (44,47).

Com isso, este trabalho propõe um método original de codificação, indexação e recuperação de imagens e vídeos que trata essas tarefas de modo integrado, utilizando-se de diferentes técnicas matemáticas usualmente aplicadas a essas finalidades. Desse modo, a indexação é realizada simultaneamente à codificação, e o processo de recuperação será capaz de ser realizado imediatamente após o término do processo.

## **2. OBJETIVO**

Propor e analisar um método de codificação, indexação e recuperação automática de vídeos de exames médicos para aplicações de telemedicina.

### 3. REVISÃO DA LITERATURA E BASE DE DADOS MUNDIAIS DE PATENTES

#### 3.1. Imagens e vídeos digitais

Imagens digitais são compostas por um conjunto de pontos luminosos, denominados pixels, distribuídos espacialmente (52). Para obter ou produzir uma imagem digital, a intensidade e a cor da luz emitida em cada ponto da cena são amostradas e medidas quantitativamente, sendo os valores obtidos agrupados e organizados na forma de matrizes. Desse modo, os índices da matriz correspondem à posição do pixel na cena capturada pela imagem, sendo a origem, por convenção, posicionada no canto superior esquerdo (53), conforme ilustra a Figura 1.

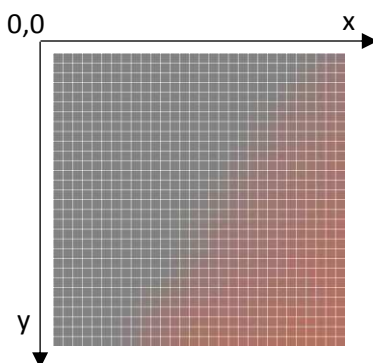


Figura 1 – Posição dos índices da matriz de pixels no sistema de coordenadas.

A distância entre cada pixel da imagem e a quantidade de linhas e colunas amostradas determinam a resolução espacial da imagem, sendo maior quanto mais amostras forem obtidas. Assim, a fidelidade de uma imagem digitalizada à cena real depende não somente da quantidade de linhas e colunas amostradas, mas também da densidade de pixels por unidade de área, usualmente indicada em pixels por polegada (DPI) (33,52).

Nos dispositivos e sensores de captura de imagens utilizados, por exemplo, em câmeras fotográficas, a resolução espacial é fixa, e a densidade de pixels por área é ajustada externamente por meio de lentes e ou pela distância física do objeto ou cena. Vale ressaltar que a resolução espacial ideal é determinada pela necessidade de cada aplicação. A Figura 2 ilustra uma ampliação gradativa (*zoom*) de uma imagem de exame endoscópico, na qual é possível observar os pixels de uma pequena área da figura (31,54).

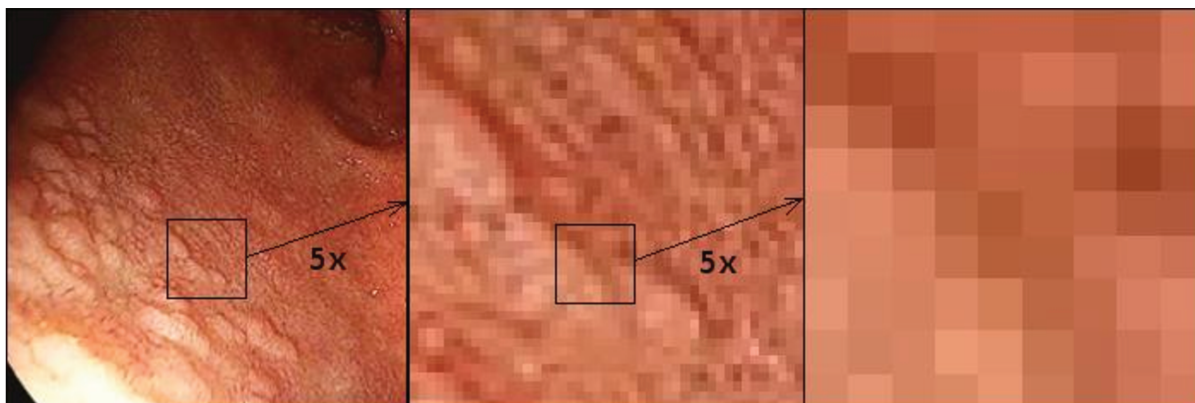


Figura 2 – Vista aproximada dos pixels de uma área da imagem.

Além da resolução espacial, a resolução da intensidade em cada ponto da imagem é um fator que deve ser considerado e ajustado conforme a aplicação, pois pode afetar a qualidade de imagem. A resolução da intensidade refere-se à menor variação perceptível na intensidade da luz e ou da cor do pixel, e deve permitir representar e reproduzir quantos níveis diferentes forem necessários. A quantidade de níveis de intensidade é usualmente definida em potências de dois devido ao sistema de representação binária empregado em sistemas digitais (53).

Nesse sentido, para imagens monocromáticas de aplicação genérica, é suficiente utilizar 256 (ou  $2^8$ ) níveis de intensidade para representar todas as variações possíveis observadas em cada ponto na cena. Entretanto, em imagens médicas de diagnóstico por imagem, como exemplo as radiográficas, ultrassonográficas e de ressonância magnética, é comum utilizar 65.536 (ou  $2^{16}$ ) níveis diferentes de intensidade a fim de preservar o máximo possível de detalhes das estruturas anatômicas observadas (35).

Vale ressaltar que todos os dispositivos e processos utilizados na captura, armazenamento e apresentação da imagem, tais como sensores, protocolos de codificação e monitores de vídeo, devem suportar a mesma resolução de intensidade, caso contrário os benefícios da resolução superior não serão aproveitados (37).

No caso de imagens coloridas, a cor de cada pixel é resultante da combinação de três cores primárias: vermelho, verde e azul. Essa modalidade de representação de imagens coloridas é denominada RGB, acrônimo para *Red green blue*. Portanto, em lugar de uma única matriz de intensidades, imagens coloridas são representadas por três matrizes que indicam, cada uma, quanto de cada cor primária existe em cada pixel (31). A Figura 3 exemplifica uma imagem decomposta em seus componentes R, G e B.

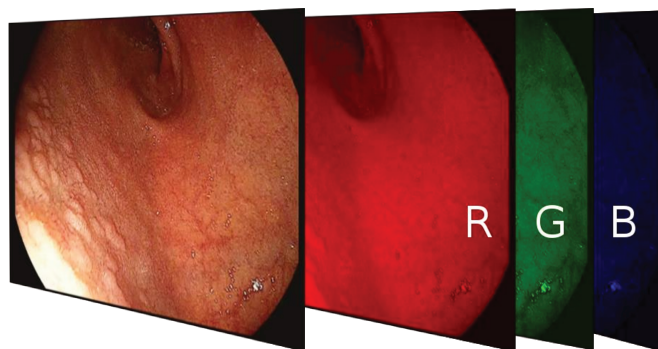


Figura 3 – Imagem original decomposta em componentes R, G e B.

Esse modelo de cores foi inspirado nas características do olho humano. As células presentes na retina responsáveis pela percepção de cores, os cones, possuem níveis de sensibilidade diferentes para cada espectro da luz, resultando em três tipos de cones específicos. Por esse motivo os cones são denominados conforme o comprimento de onda para o qual possuem maior sensibilidade, ou seja, S para ondas curtas, M para ondas médias, e L para ondas longas (52,55).

Os comprimentos de onda curtos, médios e longos correspondem, respectivamente, às cores azul, verde e vermelho. Na Figura 4 são apresentadas as respostas normalizadas de cada tipo de cone, em termos de sensibilidade, segundo o comprimento de onda da luz.

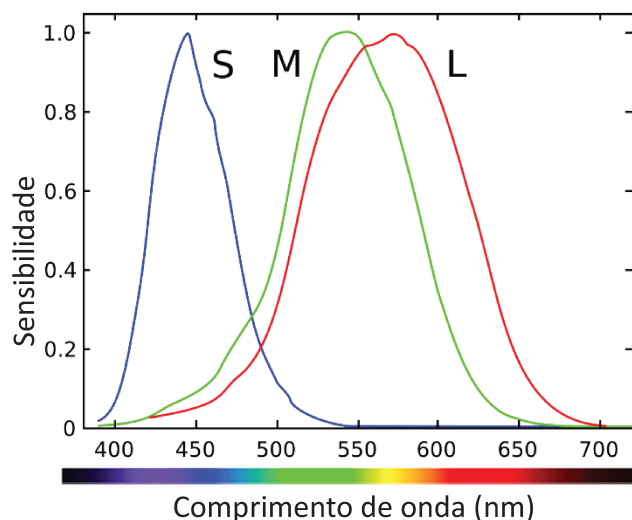


Figura 4 – Sensibilidade dos cones S, M e L segundo o comprimento de onda.

Fonte: Adaptado de LMS Color Space (56).

Pela Figura 4 pode-se observar também o fato de a percepção visual de cores nos seres humanos ser resultante da estimulação simultânea dos três tipos de cones, assim como diferentes tonalidades e matizes podem ser reproduzidas a partir

da mistura dessas cores primárias (56). Como consequência, foi possível desenvolver telas de TV e monitores de vídeo capazes de reproduzir uma grande variedade de cores a partir de apenas três. Portanto, as diferentes cores reproduzidas pelos dispositivos eletrônicos mencionados não resultam da emissão da luz com diferentes comprimentos de onda no espectro visível, mas da mistura, em gradações variadas, de três cores fixas. Essa característica promove a estimulação dos cones na retina de modo igual ou compatível com a estimulação que cada comprimento de onda visível produz (57).

Para mapear a relação entre o comprimento de onda e a percepção às cores de um observador humano padrão, a Comissão Internacional de Iluminação (CIE) estabeleceu, em 1931, funções de mapeamento para as cores visíveis, definindo-as como combinações lineares das três cores primárias, nos comprimentos de onda de 435,8 (B), 546,1(G) e 700 nm (R) (58). O modelo de cores proposto, denominado XYZ, foi elaborado de tal maneira que as misturas de cores possíveis, com nível máximo de saturação e intensidade luminosa constante, fossem projetadas sobre um dos planos do espaço de cores tridimensional (57,58). A Figura 5 ilustra essa característica, com as cores projetadas sobre o plano  $xy$ . Na figura, os valores de  $x$  e  $y$  correspondem aos valores de  $X$  e  $Y$  normalizados de modo que sejam limitados entre 0 e 1, e podem ser calculados por meio das Equações 1 e 2 (33).

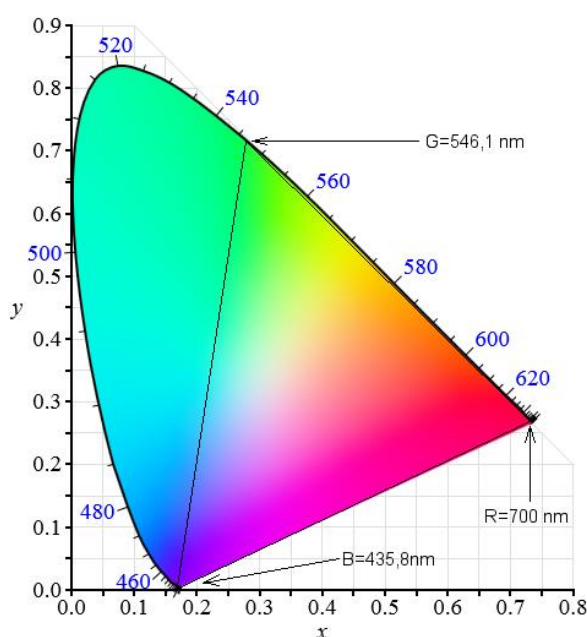


Figura 5 – Gráfico de mapeamento de cores da CIE.

Fonte: Adaptado de CIE 1931 Color Space (59).



$$x = X/(X + Y + Z) \quad (1)$$

$$y = Y/(X + Y + Z) \quad (2)$$

Adicionalmente, no modelo XYZ a luminância é representada separadamente. Esse modelo é utilizado como base para definir outros espaços de cor conhecidos, entre eles os modelos YUV para sinais de TV analógica, e YCbCr, utilizado na codificação de vídeos digitais (52,53). Em particular, os espaços de cor aplicados a sistemas de TV e ou vídeo utilizam essa separação entre luminância e cores, pois viabiliza a codificação de um fluxo único compatível com sistemas monocromáticos e coloridos, além de permitir comprimir os sinais de cor sem perda perceptível de qualidade (31,33). A Figura 6 ilustra o espaço de cores do modelo YUV para um nível de luminância específico.

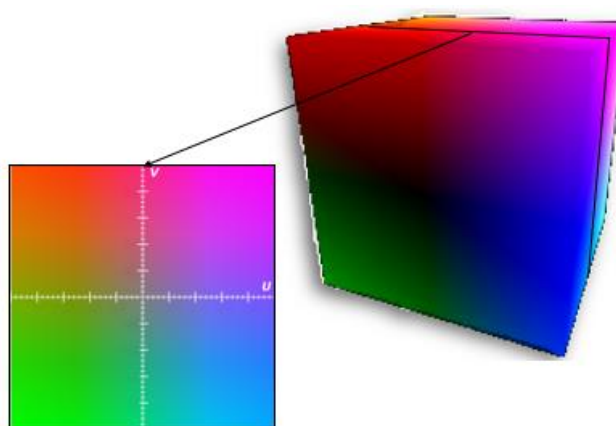


Figura 6 – Espaço de cores YUV para um nível de luminância específico.

Fonte: Adaptado de YUV e YUV Colorspace (60,61).

Outro aspecto relevante em relação à seleção do espaço de cores diz respeito ao fato de o olho humano ser mais sensível à percepção de variações de intensidade do brilho, quando comparado à percepção de cores (31). Por esse motivo a utilização de um modelo de cores que separa os componentes de cor do componente de luminância é mais vantajosa para aplicações de vídeo. Desse modo é possível alcançar melhor qualidade do vídeo dando maior ênfase à luminância (31,33).

O principal modelo de cor utilizado em vídeos digitais é o YCbCr, o qual baseia-se no modelo YUV. No modelo YUV os componentes de cor (U e V) são normalizados de modo que os valores mínimo e máximo sejam 0 e 1, enquanto no

modelo YCbCr os valores das matrizes são quantizados em 256 níveis, sendo, consequentemente, representados por números inteiros de oito bits (31).

A Figura 7 exemplifica as camadas do modelo YCbCr para uma imagem de endoscopia. A camada de luminância corresponde à imagem em escala de cinza, e favorece a identificação de características e a indexação.

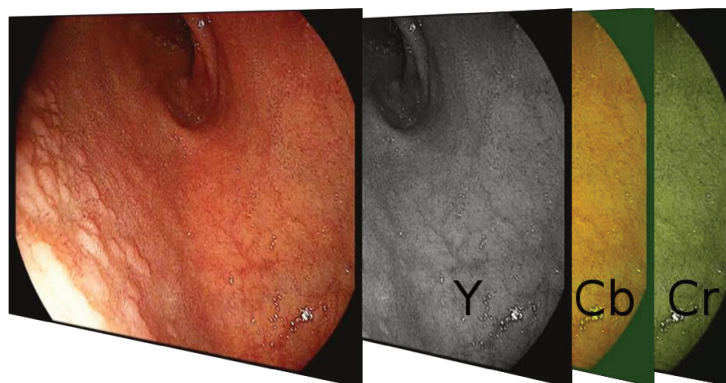


Figura 7 – Camadas do modelo YCbCr.

Estas mesmas características aplicadas a imagens são válidas também para vídeos, e um vídeo é formado por imagens sequenciais exibidas a intervalos breves e regulares. Nesse sentido, a rápida sucessão de imagens produz a sensação de movimento e a Figura 8 ilustra uma sequência de imagens de um fluxo de vídeo.

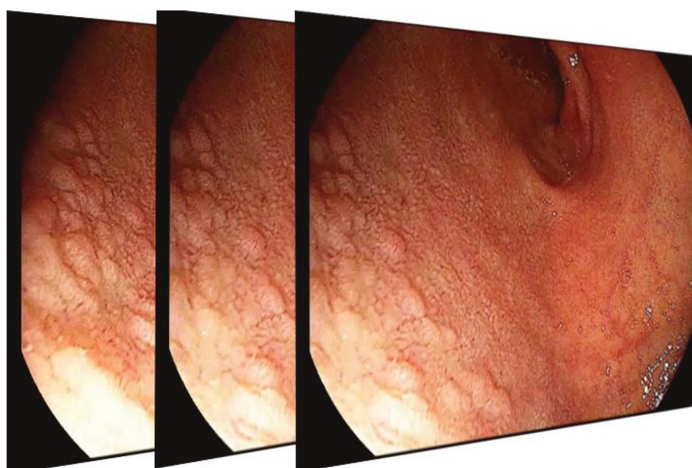


Figura 8 – Sequência de quadros de vídeo.

As imagens individuais que formam o vídeo são denominadas quadros ou *frames*. No processo de codificação de imagens e ou de quadros, busca-se explorar redundâncias espaciais, tais como as repetições de pixels em posições adjacentes,

de modo a representar grupos de pixels iguais ou semelhantes em conjunto. Desse modo, obtém-se uma representação compactada da informação visual (62).

No caso de vídeos, além das redundâncias espaciais presentes nos quadros, existem também redundâncias temporais, ou seja, semelhanças entre quadros subsequentes, que podem ser exploradas a fim de obter níveis de compressão mais elevados. As técnicas de codificação “inter-quadros” aplicam a codificação sobre a diferença numérica entre quadros subsequentes ou “resíduo”, obtendo uma redução da informação a ser codificada (33,62).

### 3.2. Decomposição em valores singulares

A decomposição em valores singulares (DVS) é uma técnica de fatoração de matrizes que obtém, como resultado, uma decomposição na forma:

$$A = U \times \Sigma \times V^T \quad (3)$$

Na Equação 3 (63,64),  $\Sigma$  é uma matriz diagonal de valores singulares, e  $U$  e  $V$  são matrizes ortogonais de vetores associados aos valores singulares.

O objetivo da DVS é encontrar um conjunto de vetores  $v_i$  que ao serem multiplicados pela matriz de entrada  $A$  se igualem a um único vetor de direção  $u_i$  e comprimento  $\sigma_{i,i}$ . Essa igualdade pode ser expressa na forma da Equação 4 (63):

$$A \times v_i = \sigma_{i,i} \times u_i \quad (4)$$

Essa operação pode ser interpretada geometricamente como uma projeção do vetor  $v_i$  orientada pela multiplicação pela matriz  $A$ , de tal modo que o vetor  $\sigma_{i,i} \times u_i$  represente a magnitude e a orientação da transformação produzida pela matriz  $A$  sobre o vetor  $v_i$  (63). Desse modo pode-se inferir em quais direções a matriz  $A$  concentra suas características mais relevantes. A Figura 9 ilustra esses procedimentos para a obtenção do primeiro valor singular e vetor  $u$  respectivo.

Esse procedimento é repetido em sequência até que todos os valores e vetores singulares da matriz  $A$  sejam encontrados. A Figura10 ilustra o segundo passo de obtenção dos valores e vetores singulares da matriz  $A$ .

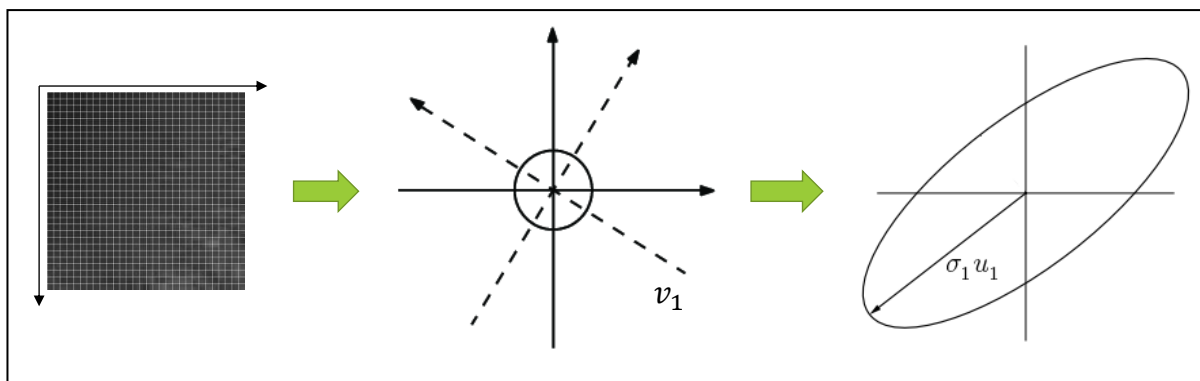


Figura 9 – Interpretação geométrica da fatoração de  $A$  no valor singular  $\sigma_{i,i}$  e nos vetores singulares  $v_i$  e  $u_i$ .

Fonte: Adaptado de Poole (63).

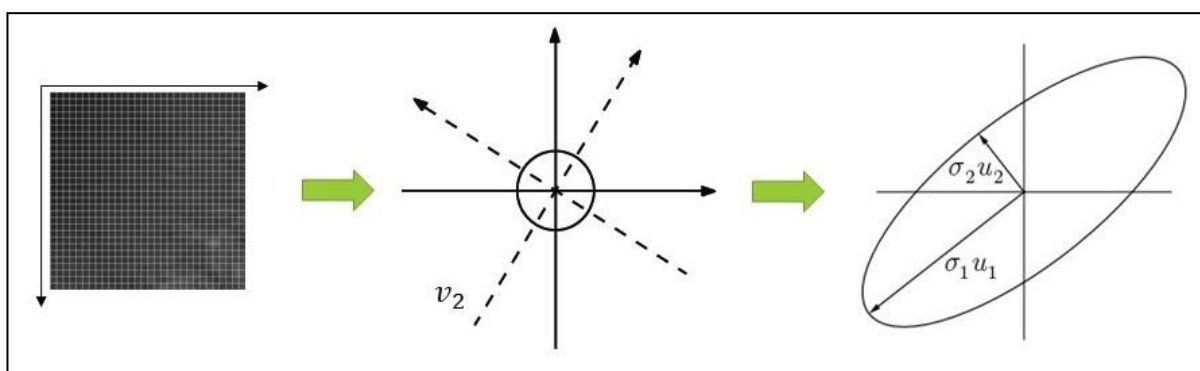


Figura 10 – Interpretação geométrica da fatoração de  $A$  para obtenção do segundo valor singular  $\sigma_{i,i}$  e vetores singulares  $v_i$  e  $u_i$  respectivos.

Fonte: Adaptado de Poole (63).

O conjunto de vetores e valores singulares podem ser ordenados de modo decrescente segundo a magnitude dos valores singulares  $\sigma_{i,i}$ . Desse modo, obtém-se uma sequência de vetores na qual os vetores mais representativos da matriz  $A$  são posicionados no início da relação. O conjunto de vetores  $u_i$  e  $v_i$  agrupados compõem as matrizes  $U$  e  $V$ , e são denominados vetores singulares à esquerda e à direita, respectivamente, enquanto os valores singulares  $\sigma_{i,i}$  formam a diagonal principal da matriz  $\Sigma$  (63,64).

O número de autovalores e autovetores da matriz  $A$  corresponde ao número de vetores linha linearmente independentes da matriz (65).

Os valores singulares  $\sigma_{i,i}$  são obtidos a partir das raízes quadradas dos autovalores da matriz resultante da multiplicação da matriz de entrada  $A$  por sua transposta, ou seja,  $A^T \times A$  (64,65). A multiplicação de uma matriz por sua transposta

produz uma matriz simétrica, e este fato possui a vantagem de evitar a presença de números complexos nos autovalores devido a todas as raízes do polinômio característico serem valores positivos, simplificando o cálculo e a representação desses valores (63).

A fatoração representada pela DVS é um procedimento reversível, ou seja, a matriz original  $A$  pode ser reconstruída a partir das matrizes fatoradas  $U$ ,  $\Sigma$  e  $V$ . Vale ressaltar que o produto  $U \times \Sigma \times V^T$  pode ser alcançado de maneira gradual por meio do somatório:

$$\sum_{i=1}^k u_i \times \sigma_{i,i} \times v_i^T \quad (5)$$

Desse modo, a cada parcela adicionada, a soma resultante se aproxima gradualmente da matriz original. Na Equação 5 (65), o valor de  $k$  é o posto da matriz  $A$ , e equivale ao número de vetores linearmente independentes na matriz.

No contexto de imagens digitais, a redundância espacial presente nas imagens se traduz em matrizes compostas de vetores semelhantes, e com frequência um subconjunto desses vetores são iguais. Em consequência dessa característica, o número de vetores linearmente independentes da matriz é menor que o total de vetores da matriz, resultando em um posto de menor valor. No somatório da Equação 5, isso significa que o valor de  $k$  será menor que as dimensões da matriz original, e, portanto, os vetores repetidos serão representados apenas uma vez (63).

Além disso, valores singulares de menor magnitude e seus respectivos vetores associados, possuem pouca ou nenhuma relevância na reconstrução da matriz original quando se trata de imagens digitais, e podem ser desconsiderados. Dessa forma é possível representar a matriz original com menos informações e alcançar uma compressão dos dados originais (63).

### 3.3. Codificação de imagens e vídeos

Para reduzir a demanda por espaço de armazenamento e por largura de banda dos canais de comunicação, imagens e vídeos são submetidos ao processo de compressão (33,40,41,66-68). Nesse cenário, as técnicas de compressão mais frequentes incluem as seguintes etapas:

1. Transformação da imagem do domínio espacial para o domínio de frequência por aplicação da transformada discreta de cosseno (DCT);

2. Quantização dos coeficientes da matriz transformada;
3. Codificação de símbolos utilizando o algoritmo de Huffman ou a codificação aritmética (28,52).

Vale ressaltar que exemplos de padrões que se baseiam nessas etapas e técnicas são o JPEG para imagens estáticas e MPEG para vídeos (28-30,33).

Depois do surgimento do MPEG, definiu-se o padrão H.263 pelo ITU-T SG15, grupo de estudo 15 do setor de normatização das telecomunicações da União Internacional de Telecomunicações. Esse padrão foi amplamente adotado em sistemas de videoconferência, e na distribuição de vídeos na Internet por meio da tecnologia Flash, tais como YouTube e MySpace (33,69-71).

Em termos de desempenho, o padrão H.263 é capaz de operar em diferentes taxas de bits, apresentando diferentes níveis de qualidade visual, e, como exemplo, aplicando-se resolução 352 x 288 a 15 quadros por segundo, a taxa de bits média, nessa modalidade de compressão, varia entre 70 e 1.152 Kbps, e a qualidade medida pelo PSNR entre 25 e 36 dB (70).

Em alternativa à codificação baseada na transformação de domínio, foram propostas técnicas de compressão baseadas na decomposição em valores singulares (DVS) (72-77).

Koshi et al. (78), na patente US 5615288A (*Singular Value Decomposition coding and decoding apparatuses*), reivindicaram um aparato de codificação e decodificação de imagens que usa DVS. O objetivo da invenção é reduzir a quantidade de cálculos necessários na codificação de imagens usando DVS. Trata-se, portanto, de método de codificação de imagens que utiliza a fatoração das matrizes de entrada. Por essa patente, a imagem de entrada é dividida em blocos e para cada bloco os valores singulares são calculados usando, por exemplo, o método de Jacobi (65). Os autovalores e os autovetores são codificados, e após o cálculo dos autovalores e autovetores, e dos autovetores normalizados, é determinada a força do último valor singular obtido, que é então acumulada. No momento em que a força acumulada dos valores singulares alcança taxa pré-determinada, ou quando a diferença da força acumulada e da força total do bloco de entrada alcança um valor fixo, não indicado no documento, o cálculo de valores singulares cessa. Os autovetores obtidos na decomposição são agrupados em classes de acordo com a magnitude dos valores singulares, e comparados com limiares pré-estabelecidos. A partir dessa comparação, os valores singulares e autovetores são agrupados em classes, e depois quantizados

linearmente de acordo com a faixa dinâmica da classe à qual pertencem. Em relação aos autovetores de uma segunda classe, estes são submetidos ao cálculo da diferença entre os elementos por meio de técnica denominada *differential pulse-code modulation* (DPCM) (79,77), sendo a diferença também quantizada linearmente.

Ainda, a patente avaliada faz referência aos resultados obtidos por Andrews e Patterson (72) e Garguir (80), e estabelece método para reduzir a quantidade de cálculos necessários na codificação, buscando manter os mesmos resultados. Andrews e Patterson (72) compararam a DVS com outras transformações, tais como as transformadas de cosseno e de Fourier, e demonstraram que a DVS era a melhor transformação ortogonal para compressão de energia de imagem. Porém, em termos do custo computacional, a DVS apresentou complexidade de tempo da ordem de  $N^3$ , enquanto a DCT, complexidade da ordem de  $4N^2 \cdot \log_2 2N$ .

Sob esse escopo, Garguir (80) propôs método de codificação de imagens monocromáticas baseado na DVS e comparou o desempenho do codificador com outra abordagem baseada na DCT. O autor concluiu que, em relação às taxas de compressão, as imagens reconstruídas usando o método DVS são ligeiramente melhores quando comparadas ao uso da DCT, e que a reconstrução pelo método DVS leva de 30 a 50 vezes menos tempo que a DCT. Porém, a DCT não possui limite teórico para a taxa de compressão, enquanto a DVS apresenta limite experimental de 0,6 bit/pixel.

O método mencionado é referenciado na patente US 8160368 B2 (81), a qual aplica a DVS sobre subdivisões da imagem original com objetivo de compressão e extração de característica. Nesse documento, os autores registraram que o método apresentou qualidade de imagem superior ao padrão JPEG para a mesma taxa de compressão.

Ashin et al. (82) propuserem um método híbrido de transformação e compressão das imagens que combina DVS e a transformada *wavelet* discreta (DWT) para obter melhores taxas de compressão ao mesmo tempo em que preserva a qualidade. Nesse trabalho, o método de compressão compreende três passos:

1. Transformação da imagem  $X$  na imagem  $X_1$  pela DWT de nível dois usando *wavelets* biortogonais 9/7;
2. Decomposição de  $X_1$  em blocos de tamanho  $2 \times 2$  com análise multi-resolução da DVS até o nível 6 para obter  $X_2$ ;

3. Compressão de  $X_2$  usando SPIHT – *set partitioning in hierarchical trees*, e compressão da imagem resultante com *gzip*.

Os autores realizaram experimentos comparando a qualidade da imagem reconstruída após codificação pelo método proposto e por outros métodos. As taxas de compressão alvo foram parametrizadas em 1,  $\frac{1}{2}$  e  $\frac{1}{4}$  bpp (bit por pixel), e os resultados apontaram o seguinte:

1. No caso de alta taxa de compressão ( $\frac{1}{4}$  bpp) o método proposto apresentou equilíbrio entre efeitos de bloco e imagens desfocadas. Vale ressaltar que, enquanto os codificadores baseados em DVS, *Karhunen-Loève Transform* (KLT) e JPEG apresentaram efeitos de bloco, o codificador *wavelet* com filtros longos apresentou imagens desfocadas;
2. Para a figura da impressão digital o desempenho do método proposto foi superior (PSNR = 40,22 dB);
3. Para a figura “Yogi” os codificadores baseados em DVS e KLT foram superiores decorrente ao fato da imagem possuir menos níveis de cinza (PSNR = 51,15 dB e 47,99 dB, respectivamente);
4. Para as demais imagens, o método proposto foi ligeiramente inferior ao codificador *wavelet*, mas superior aos demais (PSNR = 28,3 dB ~ 39,19 dB).

Gu et al. (77) propuseram um método de codificação de vídeo que utiliza a fatoração das matrizes de entrada por meio da DVS. O método descrito baseia-se na redução da redundância temporal presente nas imagens subsequentes de vídeo, alcançando taxa de compressão superior em comparação com a obtida sem a remoção. Nesse contexto, o processo algébrico empregado foi o 2-D DVS (75), que consiste em aplicar a fatoração sobre a média do grupo de  $N$  quadros e transmitir as matrizes de vetores singulares  $U$  e  $V$  do grupo, e os valores singulares de cada quadro. Assim, a quantidade de coeficientes transferidos a partir da aplicação do esquema proposto é reduzida para:

$$(2 + n) \cdot WH$$

Onde:

- $n$  é o número de quadros considerados;
- $W$  corresponde à largura do quadro;



- $H$  corresponde à altura do quadro.

De outro modo, sem o uso do método, o número de coeficientes necessários na transmissão equivale a:

$$nWH \cdot (2m^2 + m)/m^2$$

Onde:

- $n$ ,  $W$  e  $H$  referem-se, respectivamente, ao número de quadros considerados, à largura e à altura do quadro;
- $m$  corresponde ao tamanho do bloco.

Como resultado, o número de coeficientes necessários para a 2-D DVS corresponde, aproximadamente, à metade da quantidade necessária para a 1-D DVS. Com isso, o método proposto apresentou ganho médio em codificação de 6,15 dB para o mesmo *bit-rate* quando comparado à codificação baseada na DCT. Ainda, comparado aos codificadores H.264 sem estimação de movimento, codificador de vídeo distribuído (DVC) com predição intra-quadro, H.264 com predição intra-quadro e MJPEG DPCM, o método do documento foi superior em qualidade de imagem em 11 de 17 cenários para o mesmo *bit-rate*.

A patente US 8433148 (83) descreve um método para compressão de imagens texturizadas e particionadas em blocos de pixels. Nesse método, a matriz é decomposta em uma matriz-coluna e uma matriz-linha, de modo que a largura da matriz-coluna seja substancialmente menor que a altura da matriz-coluna, e a altura da matriz-linha seja consideravelmente inferior à largura da matriz-linha. A decomposição do método é caracterizada pela generalização do algoritmo K-médias ou *K-means*, que utiliza a DVS para calcular as matrizes coluna e linha, e por esse motivo é denominado *k-means SVD* (do inglês *singular value decomposition*) ou *K-SVD* (84). Cabe ressaltar que esse processo de decomposição busca minimizar o erro residual de modo que a norma de Frobenius de  $\|Y - DX\|_F^2$  seja mínima, onde  $Y$  é a imagem original, e  $D$  e  $X$  as matrizes-coluna e linha respectivamente. Para obter esse resultado, o método K-SVD usa um processo de busca que alterna entre codificação esparsa e atualização de tabela de códigos até convergir ou alcançar uma condição de terminação; o estágio de codificação esparsa determina os vetores  $x$  para cada pixel  $y$ , minimizando o resíduo  $\|y - Dx\|^2$ . Depois da decomposição, as linhas da

matriz-coluna são agrupadas usando um *bit-rate* alvo e um parâmetro de qualidade alvo. Em seguida, para obter maior compressão em imagens com padrões repetitivos, encontra-se a menor frequência dominante para a matriz-coluna usando uma transformada de Fourier. A partir do valor da frequência dominante seleciona-se uma submatriz da matriz-coluna que representa o padrão que se repete na imagem. Ainda, o método suporta forma alternativa de compressão na qual a imagem 2-D é particionada em vetores 1D, e sobre cada vetor são aplicadas a transformada *wavelet* ou DCT. Em seguida, os vetores transformados são comprimidos por codificação de entropia.

A patente US 8451906 (85) descreve um método e sistemas para reconstrução de sequências de quadros de vídeo recebidos de um dispositivo emissor num ambiente de codificação distribuída. O método descrito pressupõe uma arquitetura em que o vídeo codificado é transmitido pelo dispositivo cliente para um servidor, e do servidor para outro cliente. Segundo os autores, esse arranjo se justifica pelo fato de o servidor ter maior capacidade de processamento, e poder recodificar o vídeo de modo a aliviar a carga de processamento do cliente receptor. Nesse cenário, o dispositivo emissor do vídeo codifica quadros-chaves por completo utilizando um método projetado para demandar uma carga de processamento mínima. Por outro lado, para os quadros não-chave apenas uma parte dos blocos de pixels são codificados pelo emissor e, esse emissor identifica, por meio de cálculo, quais blocos podem ser substituídos com segurança por blocos correspondentes no quadro-chave. Nesse contexto, a decomposição em valores singulares é utilizada para codificar os blocos de pixels de quadros não-chave.

Na patente CN103237205 (86) é descrito um método de compressão de imagens para câmera digital baseado na observação da matriz de Toeplitz e em técnica de aprendizagem de dicionário. O método pretende resolver o problema relacionado à dificuldade de aplicar a matriz de observação aleatória gaussiana em sistemas reais. Por isso, aplica-se a matriz de observação de Toeplitz, a qual está relacionada ao corpo teórico de sensoriamento comprimido, cuja hipótese estabelece a possibilidade de obtenção de uma representação esparsa do sinal original por meio da construção de um sistema de observação não relacionado ao dicionário. Ainda, a partir da representação esparsa é possível reconstruir o sinal com alta probabilidade por meio do algoritmo de otimização. A técnica de aprendizado de dicionário aplicada pelos autores é a K-SVD. Nesse cenário, os autores compararam a qualidade da

reconstrução utilizando o dicionário DCT e a K-SVD, e verificaram que o desempenho da K-SVD é superior para todas as taxas de amostragem testadas. Após avaliações, observou-se que o PSNR variou de 24,32 dB a 36,14 dB para a imagem “Lena”, e de 24,94 dB a 38,37 dB para a imagem “House”. Cabe destacar que a qualidade das imagens reconstruídas pelo método depende do dicionário construído.

Pace (87) propõe, na patente US 8964835, método e sistema de processamento de vídeo para detecção de uma ou mais instâncias de uma característica candidata. Nessa patente, os quadros decodificados previamente são processados para identificar potenciais coincidências da característica candidata, e, no momento em que uma quantidade substancial de partes do quadro apresenta instâncias da característica candidata, estas são reunidas num conjunto. Vale ressaltar que cada conjunto de características candidatas é usado para criar um modelo baseado em característica, que considera a variação de deformação e a variação de aparência das instâncias. Os modelos de aparência e estruturais são criados usando a análise de componentes principais (PCA). A partir disso, a codificação do vídeo é feita com o protocolo MPEG, com exceção das regiões cuja taxa de compressão for maior usando a compressão baseada em características. Neste caso a região é codificada utilizando o método de compressão baseado em características, e as demais utilizando o codificador convencional (MPEG).

A patente US20160241861 (88) descreve uma unidade de predição intra-quadro cuja finalidade é melhorar a eficiência das predições entre planos de cor de uma imagem RGB para o protocolo de codificação *High efficiency video coding* (HEVC) (89). Esta unidade de predição remove componentes de alta frequência com um filtro passa-baixa, e separa o valor de luminância do pixel. Em seguida é gerado um bloco de predição para o qual estima-se o valor de cromaticidade usando um filtro passa-baixa de modo *LM chroma*. O modo *LM chroma* consiste em uma técnica para geração do bloco de predição a partir da correlação entre os planos de cor e de luminância, aplicado após remoção dos ingredientes de alta-frequência. O protocolo HEVC prevê a possibilidade de agregação de extensões, e nesse sentido a proposta dessa patente se aplica a imagens em formato RGB com amostragem uniforme dos componentes de cor (formato de pixel 4:4:4).

Também aplicada ao protocolo HEVC, a patente US20160261885 (90) propõe a análise da imagem de entrada com o objetivo de identificar o espaço de cores, a taxa de amostragem e a profundidade de bits (número de bits necessários

para representar cada pixel) que melhores níveis de compressão podem oferecer ao codificador. O método propõe estratégias de adaptação aplicáveis à mesma imagem, ou a uma sequência de quadros de um vídeo. Para selecionar a melhor configuração, o método propõe uma análise de taxa de distorção de diferentes estratégias, e favorece os parâmetros que apresentam a menor distorção. Considerando o espaço de cores, a estratégia proposta prioriza o modelo de cores YUV para cenas e RGB para telas de computadores.

A patente CN105763880 (91) descreve um método iterativo para cálculo de valores singulares e vetores correspondentes que objetiva reduzir ou eliminar o efeito de bloco de um processo de codificação de vídeo não descrito no documento. O método prevê a aplicação de um filtro baseado na DVS, e propõe técnica para decompor a matriz apenas o suficiente para as necessidades do método, com possibilidade de ajuste da precisão da decomposição. No trabalho em consideração, o processo de eliminação do efeito de bloco inicia-se com a obtenção de uma matriz de grupo de similaridade estrutural a partir do conteúdo do vídeo codificado. Essa matriz é usada como entrada para o processo iterativo de cálculo, que se desdobra em dois laços de repetição aninhados para que as linhas e as colunas da matriz sejam processadas gradualmente. Pelo laço interno, a matriz de grupo de similaridade estrutural é reconstruída utilizando-se o maior valor singular e respectivos vetores singulares à direita e à esquerda. Em seguida, a matriz de entrada é subtraída da matriz reconstruída, gerando uma matriz residual que serve de entrada para o laço externo. No laço externo, o valor singular final e os vetores calculados são novamente combinados, e ao ser concluído obtém-se uma matriz de grupo de similaridade estrutural na qual o efeito de bloco é eliminado.

A patente US 10027974 (92) descreve um método de codificação de imagens que aplica a DVS sobre bloco de predição para obter as matrizes de autovetores U e V, e realiza a codificação sobre os dados residuais da diferença entre o bloco predito e o bloco efetivamente a ser codificado. O método prevê o cálculo de uma segunda transformação sobre os dados residuais para obter um segundo conjunto de coeficientes de transformação. Após, o desempenho de ambas as transformações é comparado a fim de embasar a escolha da transformação e dos coeficientes correspondentes. Além da segunda transformação, o método ainda prevê a possibilidade de uma terceira, cujo desempenho também será confrontado e, caso seja superior, será selecionado.

### 3.4. Indexação e recuperação

A indexação é uma técnica utilizada para acelerar o acesso a um dado ou a um conjunto de dados armazenados (93). Este processo consiste no mapeamento de um conjunto de objetos (documentos, imagens, vídeos, etc.) por meio de símbolos representativos desses objetos, que podem ser palavras, números ou outras características, com o objetivo de tornar efetivo e eficiente o acesso aos objetos de interesse (44,45,47).

Nesse contexto, a busca de informações por meio de palavras-chaves constitui a maneira mais frequente e simples de obter acesso a documentos e ou objetos de interesse (47,94,95). Porém, no caso da busca por informações não textuais como imagens e vídeos, para que tais conteúdos possam ser recuperados por meio de uma pesquisa textual é necessário atribuir rótulos identificadores a cada objeto (44,47). A Figura 11 exemplifica a atribuição manual de rótulos ao quadro de nº 1852 do conjunto de entrada utilizado nos experimentos.



Figura 11 – Exemplo de indexação por rótulo.

O trabalho de atribuição de rótulos pode ser manual e ou executado automaticamente por meio de técnicas de reconhecimento de padrões. A atribuição manual tem a vantagem de ser acurada e precisa, porém, é custosa e torna-se inviável quando se trata de indexar vídeos completos e grandes bases de dados de imagens. Nesse caso, a atribuição de rótulos automaticamente por meio de técnicas

computacionais torna-se mais vantajoso, mesmo que apresentem perdas em precisão e acurácia dos resultados obtidos (44,45,96,97).

Alternativamente à atribuição de rótulos, a indexação pode ser feita com base em características extraídas das imagens. Os tipos de características utilizadas podem ser classificados em três tipos: cores, textura e características morfológicas (47,52,96). Na Figura 12 é apresentado o histograma de cores para a imagem da Figura 11, exemplificando uma característica de cor comumente utilizada na indexação de imagens. O histograma de cores apresenta a frequência absoluta de ocorrência, na imagem, de cada nível dos componentes R, G e B dos pixels.

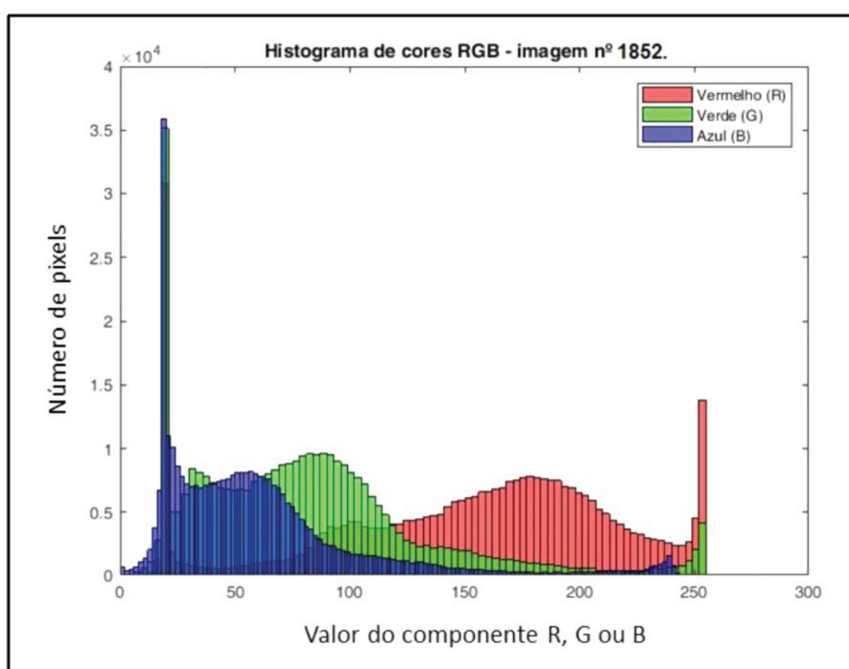


Figura 12 – Histograma de cores RGB do quadro nº 1852.

Para a extração de características de textura e morfológicas costuma-se submeter a imagem indexada a um filtro, de modo a revelar ou ressaltar as características desejadas (52). A Figura 13 exemplifica o resultado da aplicação do filtro *Local Range* para análise de textura da imagem da Figura 11.

Após a aplicação do filtro *Local Range* obtém-se uma matriz cujos valores indicam a variabilidade dos pixels para uma área delimitada em torno de cada pixel. A partir dos coeficientes da matriz resultante, após a aplicação do filtro, é possível obter um valor representativo da característica analisada para toda ou parte da imagem. Esse valor, situado em uma escala específica para a característica, permite realizar comparações posteriores entre imagens (52,53).

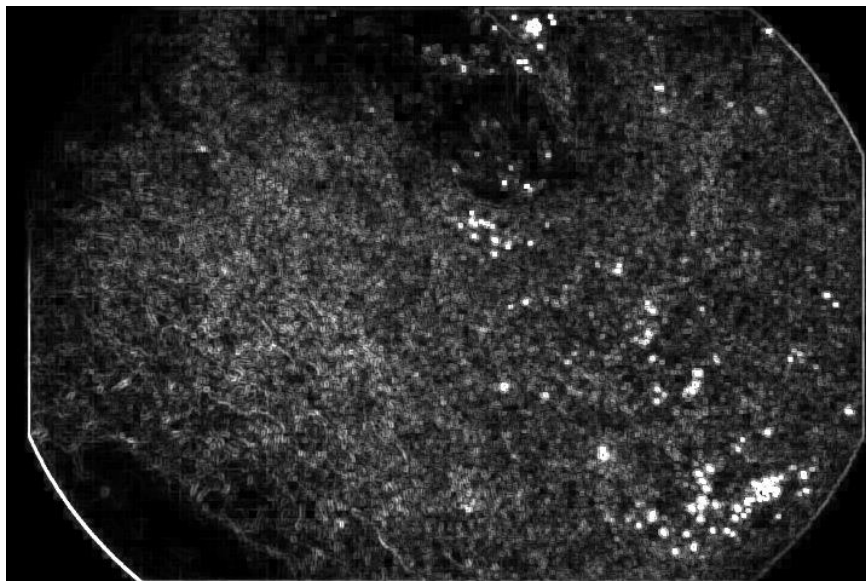


Figura 13 – Resultado da aplicação do filtro *Local Range* ao quadro nº 1852.

Características morfológicas em imagens referem-se à forma, quantidade, localização e aspecto dos objetos presentes na cena, e podem incluir, dentre as características extraídas, informações a respeito da geometria e topologia dos objetos detectados, além de cor e textura. O processo de detecção de objetos é denominado segmentação, e pode envolver etapas de filtragem, detecção de bordas, limiarização e identificação de regiões (53). Em seguida, as características morfológicas são então extraídas dos segmentos identificados. A Figura 14 ilustra o resultado da aplicação do filtro Sobel, comumente utilizado para auxiliar na tarefa de detecção de bordas e segmentação de imagens (52).

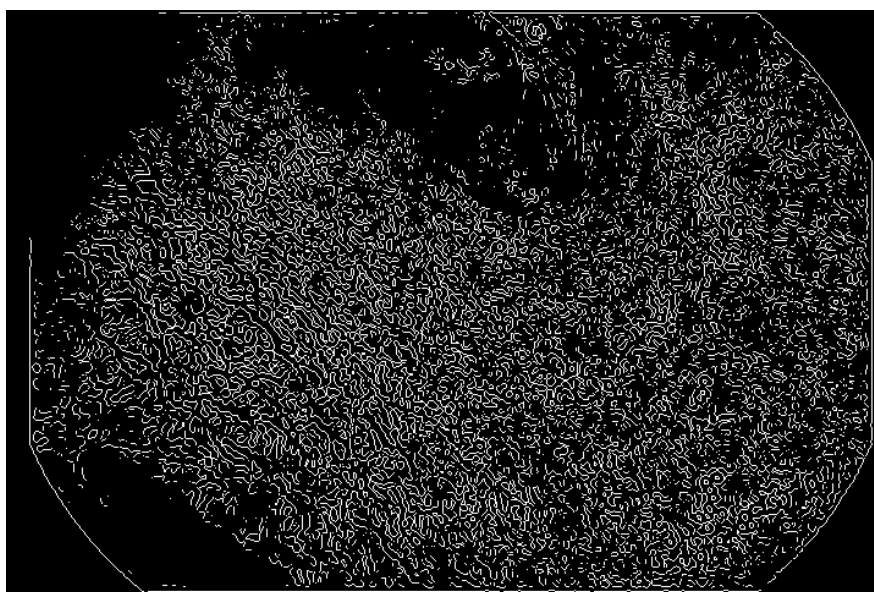


Figura 14 – Resultado da aplicação do filtro Sobel ao quadro nº 1852.

As características de cor, textura e forma são ditas de “baixo nível”, pois são processamentos algébricos e estatísticos que resultam em dados de baixo nível de abstração. Sem o contexto, características de baixo nível são insuficientes para identificar o conteúdo das imagens. Por outro lado, rótulos textuais são características ditas de “alto nível”, pois possuem alto nível de abstração e permitem a identificação do conteúdo das cenas (45).

As características de baixo nível são agrupadas de modo a formar vetores de características. Em um passo posterior, a recuperação das imagens ocorre por comparação entre os vetores de características da imagem buscada e das imagens indexadas previamente, sendo o grau de semelhança entre imagens determinado por uma medida de distância entre os vetores (44,47,96).

A obtenção de características de alto nível de modo automático envolve a utilização de técnicas de aprendizado de máquina e reconhecimento de padrões. Essas técnicas permitem, entre outros aspectos, atribuir rótulos automaticamente e ou identificar novas classes segundo o conteúdo das imagens (45).

Vale ressaltar que a maneira de buscar e recuperar as imagens depende do tipo de indexação realizada, sendo comuns a busca textual e a busca por exemplo. Na busca por exemplo uma imagem é fornecida como critério de busca, e obtém-se como resposta um conjunto de imagens considerados mais semelhantes e ou de vídeos contendo quadros similares à imagem buscada, ou apenas as imagens e ou vídeos que alcançarem um nível de similaridade mínimo previamente estabelecido. Portanto, a busca pode ocorrer por meio da combinação de palavras e ou de características extraídas das imagens, ou ainda por meio de exemplos (96,98).

Dentre as características ditas de “baixo nível” mais frequentes pode-se destacar: histograma de cores, características de textura tais como aspereza, contraste e direcionalidade, e características morfológicas tais como as dimensões e a posição dos objetos presentes na imagem (53,95).

Além da indexação baseada em características, foram propostas e descritas na literatura técnicas que buscam identificar características específicas nas imagens e descrevê-las em termos dos aspectos detectados, e por isso são chamados de detectores e descritores de características. Exemplos de detectores e descritores incluem: *Scale invariant feature transform* (SIFT), *Maximally stable extremal regions* (MSER), *Features from accelerated segment test* (FAST), *Speeded up robust features* (SURF), *Binary robust independent elementary features* (BRIEF), *Oriented FAST and*



*rotated BRIEF* (ORB), *Color and edge directivity descriptor* (CEDD), *Color layout descriptor* (CLD) e *Pyramid histogram of orientation gradients* (PHOG) (99-101).

Indexação e recuperação são usualmente tratadas de modo independente da codificação, mesmo nos casos em que a indexação é realizada sobre imagens e vídeos codificados (95). No entanto, Vasconcelos e Lippman (102) exploraram a relação entre modelagem probabilística e compressão de dados para propor um método de codificação e de recuperação de vídeos codificados. A proposta dos autores consiste em aplicar um tipo de codificação denominada *library-based coding* ao método MPEG, a qual estabelece um conjunto de classes de probabilidades calculado para cada quadro processado, adicionando-o ao fluxo de dados codificados. Desse modo, a biblioteca de classes substitui as tabelas de códigos da quantização, e por esse motivo não prejudica a taxa de compressão do método. Além disso, por serem obtidas a partir de análise probabilística dos quadros, são utilizadas como índice para a recuperação. De acordo com os experimentos realizados pelos autores, os resultados indicaram acurácia de 99,2% para a primeira imagem recuperada, e 87,7% para a quarta imagem recuperada. Para o cálculo destes resultados, a imagem idêntica à utilizada na busca não foi considerada.

Na patente CA 2155901 (103) é descrito um método para identificação de rostos e extração de características com a finalidade de permitir a recuperação de fotos em que determinada pessoa se encontra registrada na cena. Nessa patente, o processo é constituído de quatro módulos ou unidades:

1. Uma unidade de aquisição de imagens em formato digital;
2. Uma unidade de pré-processamento para normalização da geometria e intensidade de valores;
3. Uma unidade para codificação das imagens e atualização da base de dados;
4. Uma unidade de interface gráfica para facilitar o uso do sistema.

O documento relacionado se aplica exclusivamente à recuperação de imagens de pessoas em arquivos fotográficos.

A patente US 5991028 (104) descreve aparato e sistema para classificação automática ou semi-automática de células, cujas imagens são obtidas a partir da técnica de espectroscopia. A espectroscopia e a espectrometria constituem a principal característica tecnológica dessa patente, pois determinam os procedimentos e

análises posteriores à captura das imagens. Pelo referido trabalho, observou-se que cada tipo de célula apresenta resposta diferente a um mesmo espectro de frequência, e por isso são obtidas diversas imagens da mesma amostra de células ou tecido biológico, sendo cada uma correspondente a um espectro em particular. Em seguida, estas imagens são comparadas com um grupo de referência, gerando um mapa de similaridades, que pode ser apresentado na forma de uma imagem em escala de cinza para ser analisado por técnicas de processamento de imagens e ou visão computacional. A partir disso, as amostras são classificadas de acordo com a quantidade e os tipos de células presentes nas imagens/amostras.

Em experimento realizado com 21 amostras de células de carcinoma de mama, diversas imagens da mesma amostra, obtidas a partir da espectroscopia, foram agrupadas na forma de cubos espectrais, conforme determina o método. Após redução da dimensionalidade, os cubos tiveram seus pixels corretamente classificados em seis grupos por uma rede neural de dois níveis, conforme o tipo de resposta espectral obtida. Em seguida, os grupos de pixels já classificados no passo anterior, foram novamente submetidos a uma rede neural que classificou as amostras em oito grupos, sendo cada qual relativo a um tipo de célula maligna. Como resultado, o desempenho das duas redes em cascata foi considerado excelente pelos autores, pois classificou corretamente 20 das 21 amostras (95%). Nesse experimento, apenas uma amostra do tipo lobular *in situ* foi incorretamente classificada como sendo do tipo ductal *in situ* cribriforme. Além da classificação automática, o método produz imagens de classificação morfométricas determinadas espectralmente que apresentam mais claramente as características de células cancerosas, em comparação com imagens no padrão RGB.

A patente US 6574378 (105) descreve um processo para indexação e recuperação de imagens usando “palavras-chaves visuais”, cujo método prevê a formação de uma espécie de dicionário de palavras-chaves visuais a partir de um conjunto de imagens pré-selecionadas, consideradas como documentos. Para realizar a indexação do conteúdo e assim viabilizar a busca de modo otimizado, as palavras-chaves visuais de amostras dos documentos são extraídas por meio de processos de aprendizado de máquina supervisionado e ou não supervisionado. Uma imagem ou quadro-chave de uma sequência de vídeo é então descrita e indexada por uma assinatura que registra a distribuição espacial das palavras-chaves visuais presentes

na cena. A recuperação é feita por semelhança entre as assinaturas da consulta e dos documentos presentes na base de dados.

O método descrito na patente US 7158692 B2 (106) possui diversas funções complexas que envolvem a obtenção, o tratamento, a segmentação de imagens aplicando várias técnicas possíveis, o alinhamento de imagens correlacionadas, a extração de características, o armazenamento em banco de dados e a busca/recuperação. Nessa patente, a finalidade principal é auxiliar no registro e acompanhamento do resultado de testes clínicos, em particular na avaliação de novas drogas. Por esse método, o acompanhamento é realizado por meio da comparação entre imagens obtidas em momentos diferentes, para um mesmo paciente ou para múltiplos pacientes. Desse modo, as imagens de um conjunto são alinhadas semi-automaticamente e em seguida segmentadas com base em características de forma e textura. De cada segmento são extraídas as seguintes características: volume total, coordenadas do centroide, coordenadas dos limites do cubo, extensão das linhas, colunas e lâminas, produto cruzado dos diâmetros, coordenadas dos pontos finais dos diâmetros dos eixos principais, e média, desvio padrão, mínima e máxima intensidade dos pixels.

Em trabalho anterior desenvolvido pelos autores e citado no documento de patente (107), a segmentação do parênquima cerebral e seus tecidos constitutivos foi realizada automaticamente com dois algoritmos diferentes. O algoritmo de segmentação morfológica obteve índice de similaridade média de 91,8% em relação à segmentação manual, e o algoritmo de alinhamento das imagens de ressonância magnética do cérebro de um atlas com as de um sujeito obteve média de 95,3% de similaridade. Ainda, o algoritmo de classificação do tipo de tecido obteve um índice de similaridade de 83,3% para matéria cinza, e 76,6% para matéria branca. Como resultado, o desempenho alcançado foi considerado aceitável pelos usuários finais em termos de precisão e tempo de processamento. Ao final do processo, as características extraídas e as imagens são armazenadas em banco de dados e podem ser consultadas utilizando recursos do próprio sistema gerenciador de bases de dados.

As reivindicações apresentadas na patente US 7689544 B2 (108) estabelecem um método geral para indexação e recuperação de documentos baseados em conteúdo, e de imagens referenciadas pelos documentos. No referido trabalho, os textos armazenados em um sistema de gerenciamento de documentos

são processados com base em uma gramática livre de contexto, e submetidos à extração de características. A gramática permite apenas a análise da estrutura sintática das frases, a partir de um vocabulário predefinido. Nesse contexto, as características são termos relevantes no cenário de aplicação do método, que depois de selecionadas são usadas para construir um índice binário reverso. Arquivos de imagens são também associados às características detectadas nos documentos, e referenciadas no índice reverso. Além do sistema gerenciador de documentos e de imagens, o método inclui uma arquitetura cliente-servidor para permitir a busca de modo distribuído, enquanto os dados e índices permanecem centralizados. Trata-se, portanto, de um método genérico para organização de documentos e imagens, que prevê a indexação, a busca e a recuperação de imagens baseada em conteúdo, porém não especifica técnicas utilizadas na indexação.

O método descrito na patente US 8682093 (109) inclui a captura de duas imagens formadas por dados de espectros eletromagnéticos visíveis e não visíveis, como exemplo as provenientes de ressonância magnética, e os combina de maneira que contornos e elementos de interesse para o diagnóstico sejam ressaltados. O método permite ainda adequar a apresentação da imagem final de acordo com a necessidade da aplicação, por exemplo, substituindo determinada cor por outra de modo a favorecer a visão de uma pessoa portadora de daltonismo. Nesse trabalho, a maneira como os dados visuais são combinados passa pela separação das imagens de entrada em dados espectrais e de contorno, que são obtidos após calcular os gradientes das imagens de entrada. A decomposição em valores singulares (DVS) é aplicada sobre os gradientes das imagens e não sobre as imagens originais, de modo que as informações espectrais possam ser combinadas, e os contornos preservados. Os autovetores da decomposição dessa matriz  $Z$  de gradientes, chamada de matriz Di Zeno, correspondem às direções em  $x$  e  $y$  das mudanças máxima e mínima do gradiente relativo a cada pixel, e os valores singulares representam a magnitude das mudanças.

Roldan-Carlos et al. (100) propõem um método cujo objetivo é permitir a recuperação do vídeo e do momento no vídeo em que uma imagem foi capturada, durante procedimentos cirúrgicos, bem como a recuperação de vídeos com conteúdo semelhante. O método possui aplicação particular em grandes bases de vídeos médicos cujo conteúdo não foi indexado durante a realização dos procedimentos. Para o processamento dos vídeos, foram extraídos dos quadros três descritores

globais (CEDD, *Auto Color Correlogram* e PHOG) e um descritor local (CEDD localizado) (110-113). Os descritores foram então combinados por técnica de fusão tardia de modelos, a fim de gerar um descritor unificado. Para os experimentos foram indexados 593.446 quadros, e realizadas 600 buscas a partir de imagens capturadas por cirurgiões. O tempo médio de recuperação foi de 30 segundos, e a precisão média alcançada, na melhor configuração, foi de 79,8% considerando apenas o primeiro resultado, e 84,6% considerando os três primeiros resultados.

Beecks et al. (113), partem do trabalho de Roldan-Carlos et al. (100) e propõem um método para cálculo do grau de correspondência entre uma imagem estática e trechos de vídeo baseado em assinatura. Os autores propõem a combinação entre uma assinatura de características baseada em categorização adaptativa e uma medida de distância da correspondência entre assinaturas. O delineamento experimental foi o mesmo de Roldan-Carlos et al., e o desempenho obtido, na melhor configuração, em termos de sensibilidade foi de 88,2% considerando apenas o primeiro resultado, e de 90,5% considerando os três primeiros resultados.

Na patente JP 5734460B (114) é descrito um método que busca objetos em imagens, e realiza a comparação entre os objetos ou cenas encontradas a fim de permitir a recuperação. Inicialmente são identificados pontos chaves nas imagens de busca e de referência. Em seguida são comparadas características locais da área ao redor dos pontos-chaves em ambas as imagens, de modo a determinar se existe correspondência entre os pontos. Nesse cenário, os pontos que possuem correspondência entre si são chamados *inliers*, enquanto os que não possuem são chamados *outliers*. Num primeiro e principal estágio a compatibilidade geométrica entre os *inliers* é avaliada a fim de determinar se as visões dos objetos existentes na cena são correspondentes, e em caso positivo o grau de similaridade é calculado. O segundo e terceiro estágios podem ainda ser executados, nos quais as distâncias entre os pontos chaves são calculadas e resumidas na forma de “taxas de distância”, e posteriormente analisadas estatisticamente.

No trabalho de Varytimidis et al. (101), os autores também buscam detectar correspondências entre pontos em imagens diferentes, porém com a finalidade de filtrar características que sejam irrelevantes para a segmentação de vídeos de cirurgias laparoscópicas. As características que se repetem por cinco ou mais quadros subsequentes são utilizadas na identificação de pontos ideais para a segmentação

automática do vídeo, sendo a presença de instrumentos cirúrgicos na cena o principal elemento de diferenciação entre trechos. As características dos quadros são extraídas e sistematizadas por meio de descritores SIFT, que por sua vez são fundidos com os descritores que se repetem em quadros subsequentes, e depois agrupados utilizando a técnica *k-means* a fim de obter um descritor global do trecho. Ainda, os autores utilizam um modelo baseado em características “profundas”, obtidas após a submissão do vídeo a uma rede neural convolucional (CNN), para avaliar a recuperação de trechos de vídeo similares. Para a segmentação automática do vídeo a precisão foi de 78,3% no melhor caso. Na recuperação, o melhor desempenho foi de 81,7%.

A patente CN 107330950A (115) apresenta um método utilizado na reconstrução de imagens de ressonância magnética utilizando uma técnica modificada de decomposição em valores singulares chamada *Non-Local SVD*, adequada para matrizes tridimensionais, cujo objetivo é melhorar a qualidade das imagens reconstruídas. Pela técnica descrita, a imagem obtida é dividida em blocos, e para cada bloco um conjunto de blocos similares é buscado na mesma imagem. Os blocos similares são fatorados usando a DVS para obter um coeficiente de valor singular correspondente. A partir disso, um valor singular é estimado utilizando como critério o menor erro quadrático médio linear, e a variância do ruído para o conjunto de blocos similares é estimada a fim de aumentar a precisão do autovalor estimado. Com o ajuste do coeficiente do valor singular do bloco, os autores esperam melhorar a qualidade da imagem.

A patente CN 106651930 A (116) descreve um método de análise de imagens médicas denominado “método de percepção de cores” cuja finalidade é colorir regiões de interesse em imagens monocromáticas de exames médicos, a fim de ressaltar estruturas ou áreas relevantes para o diagnóstico. Para alcançar o objetivo, o método prevê a aplicação de um filtro Sobel em quatro orientações diferentes ( $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  e  $135^\circ$ ) que identifica os gradientes de variação de cores nessas quatro direções e, assim, resalta os contornos de elementos presentes na imagem. Após a eliminação de ruídos por meio de filtro gaussiano, as quatro matrizes de gradientes são combinadas para identificar áreas da imagem que pertencem às mesmas estruturas anatômicas, e cores que devem ser aplicadas em cada área. No processo de combinação dos gradientes, são geradas múltiplas camadas a partir das comparações entre os gradientes, sendo que cada camada possui dimensões

superiores em relação às camadas inferiores. A decomposição em valores singulares (DVS) é aplicada nesse momento para reduzir a dimensionalidade da matriz da camada mais externa, e identificar os valores de RGB para o pixel correspondente, a partir dos três autovalores mais significativos.

## 4. MATERIAL E MÉTODOS

Este trabalho foi submetido à avaliação da Comissão de Ética em Pesquisa da Unicamp sob o número CAAE 53887216.3.0000.5404, tendo sido aprovado no parecer de número 1.477.557 (Anexo A).

### 4.1. Material

Os materiais utilizados nesse trabalho foram:

1. Gravação de videoendoscopia digestiva alta de paciente mulher, com idade de 40 anos, realizada com videoendoscópio Pentax EPKi e obtida por meio de software de captura Zscan, com as seguintes características:
  - a. Duração: 9 minutos;
  - b. Resolução do vídeo: 720 x 576 pixels;
  - c. Taxa de quadros: 15 quadros por segundo (qps);
  - d. Formato de pixel: YUV 4:2:0;
  - e. Total de quadros não repetidos: 7.212.
2. Memória *flash* USB (*pen drive*) da marca Kingston, com capacidade de 16GB;
3. Computador pessoal *laptop* com as seguintes características:
  - a. Marca e modelo: Lenovo Ideapad;
  - b. Processador: Intel Core i3 – 3217U 1,8GHz;
  - c. Memória: 4 GB;
  - d. Sistema operacional: Microsoft Windows 8.
4. Ambiente e linguagem de programação para cálculo numérico MathWorks MATLAB versão 2017a;
5. Linguagem R versão 3.4.3 para cálculo dos testes estatísticos;
6. Ferramenta FFmpeg para codificação e decodificação de vídeos em formato FLV (H.263), versão N-81234-ge1be80a para sistema operacional Windows.

### 4.2. Métodos

Os métodos de codificação e de indexação/recuperação possuem uma etapa comum, referente à fatoração das matrizes que formam as imagens de entrada. A técnica de fatoração utilizada pôde ser aplicada à compressão e à indexação de



imagens por permitir a ordenação dos vetores mais significativos que caracterizam as matrizes. Além disso, conforme pode ser observado na Figura 15, os processos de indexação e recuperação foram representados juntos pelo fato de formarem único ciclo, estando separados apenas temporalmente. A indexação ocorre antes, simultaneamente à codificação, enquanto a busca e recuperação das imagens ocorre em momento posterior, tendo como base os índices gerados na indexação.

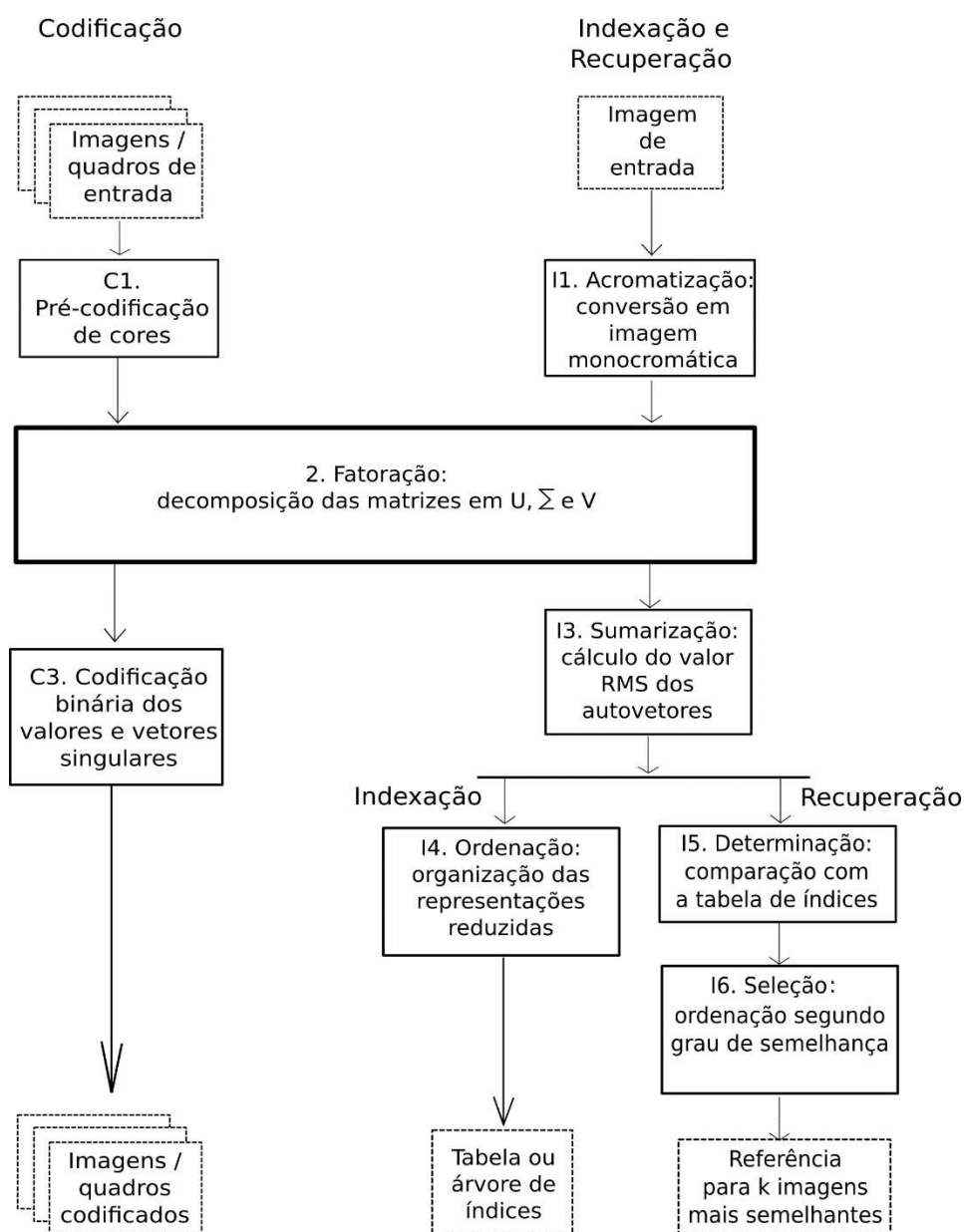


Figura 15 – Diagrama de blocos dos métodos.

A descrição dos passos do diagrama é apresentada a seguir.

- C1 – Pré-codificação de cores: envolve a conversão da imagem para o modelo YCbCr, e a compressão dos componentes Cb e Cr para um único byte. A saída deste passo são imagens compostas por apenas duas matrizes, diferente das imagens de entrada compostas por três matrizes;
- 2 – Fatoração: decomposição das matrizes de entrada em  $U$ ,  $\Sigma$  e  $V$ . As matrizes são fatoradas por meio da decomposição em valores singulares;
- C3 – Codificação binária dos valores e vetores singulares: os vetores e valores singulares mais importantes obtidos após a decomposição são codificados em formato binário;
- I1 – Acromatização: as imagens de entrada são convertidas para tons de cinza;
- I3 – Sumarização: os vetores singulares são sintetizados por meio do cálculo do valor RMS (média quadrática);
- I4 – Ordenação: os vetores resumidos são dispostos em uma tabela de modo ordenado;
- I5 – Determinação: comparação com a tabela de índices. Os vetores singulares obtidos após a decomposição da imagem de exemplo são comparados com o índice;
- I6 – Seleção: as imagens mais semelhantes são selecionadas e ordenadas segundo o grau de similaridade.

### **4.3. Busca de anterioridade**

Para a avaliação da patenteabilidade do método proposto, foi realizada uma busca de anterioridade em bases mundiais de patentes (CIPO, Esp@cenet, FPO, INPI, JPO, USPTO, WIPO), bem como nas principais bases de artigos científicos relacionadas ao tema do trabalho (ACM, IEEEExplore, Scopus) no período entre 26 de junho de 2017 a 15 de julho de 2018. Foram realizadas duas buscas: uma por trabalhos relacionados à codificação de imagens e vídeos, e outra por trabalhos relacionados à indexação e recuperação.

Para a busca relacionada à codificação de imagens e vídeos foram utilizadas duas expressões diferentes, sendo uma específica e outra ampla. A expressão ampla foi utilizada para a busca de patentes, enquanto a específica foi aplicada na busca por patentes e artigos científicos.

De modo análogo, para a busca relacionada à indexação e recuperação foram utilizadas duas expressões, sendo a ampla utilizada na busca por patentes, e a específica utilizada na busca por artigos científicos relacionados.

No Tabela 1 são apresentadas as expressões de busca ampla e específica utilizadas nesta etapa.

Tabela 1– Expressões de busca.

Método	Expressão de busca
Codificação	<p>Busca ampla:</p> <p>((image OR picture OR video) AND (("singular value decomposition" OR "factor analysis" OR "principal component decomposition" OR "empirical orthogonal function analysis") AND "diagonal matrix") AND (compress* OR coding))</p> <p>Busca específica:</p> <p>(color*) AND (analysis OR statistic* OR frequency) AND ("singular value decomposition" OR "factor analysis" OR "principal component decomposition" OR "empirical orthogonal function analysis") AND (compress* OR coding))</p>
Indexação e recuperação	<p>Busca ampla:</p> <p>(image OR picture OR video) AND ("diagonal matrix" OR "singular value" OR eigenvalue) AND (index OR retriev* OR label) AND ("coefficient of determination" OR "R squared" OR "determination coefficient")</p> <p>Busca específica:</p> <p>(image OR picture OR video) AND ("diagonal matrix" OR "singular value" OR eigenvalue) AND (index OR retriev* OR label) AND ("coefficient of determination" OR "determination coefficient")</p>

## 4.4. Delineamento experimental

Os experimentos foram realizados em duas etapas, sendo uma primeira etapa de codificação e indexação, e uma segunda de recuperação.

### 4.4.1. Etapa 1: Codificação e indexação

Os quadros originais foram submetidos, um a um em sequência, aos passos ilustrados na Figura 16, e detalhados a seguir.

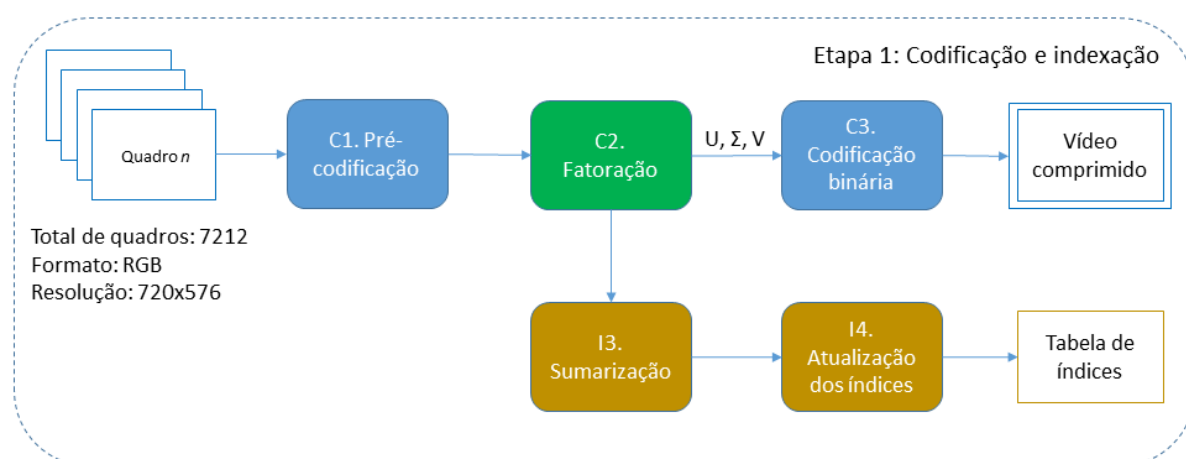


Figura 16 – Passos da Etapa Experimental 1.

#### C1 – Pré-codificação

Inicialmente os quadros de entrada foram convertidos para o formato YCbCr. Com base em uma tabela de conversão previamente estabelecida, cujo cálculo é demonstrado na etapa C1.a, as matrizes correspondentes aos componentes Cb e Cr foram convertidas a uma única matriz CbCr'. Na nova matriz CbCr' os valores de Cb e Cr correspondentes a cada pixel são combinados em um único valor. Vale ressaltar que a matriz relativa ao componente Y foi mantida inalterada.

#### C1.a) Cálculo da tabela de pré-codificação de cores

Os quadros que compõem o vídeo de referência foram representativos do domínio de aplicação do método. Assim, sobre eles foi realizada análise estatística descritiva a fim de obter a média e o desvio padrão das cores mais frequentes para cada nível de luminância. Esta etapa, considerada de treinamento, objetivou criar uma tabela de conversão para servir de suporte para a pré-codificação das cores, sem prejudicar a qualidade da imagem.

No processo de análise das imagens ou vídeos de treinamento, uma estrutura de dados auxiliar do tipo árvore foi construída para armazenar os valores mais frequentes encontrados na análise, conforme ilustrado na Figura 17. Os valores mínimo e máximo de Cb e Cr foram determinados a partir da média e desvio padrão encontrados. Além dos valores dos componentes Y, Cb e Cr, cada nó da árvore também armazena a média e o desvio padrão para permitir a decodificação.

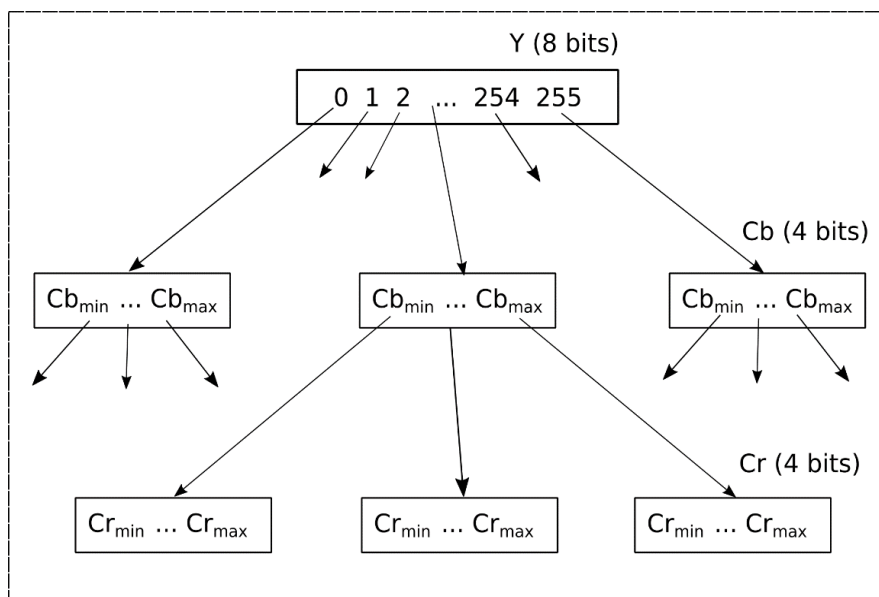


Figura 17 – Árvore de cores mais frequentes.

A partir da árvore de dados foi criada uma tabela de conversão, conforme ilustrado na Figura 18.

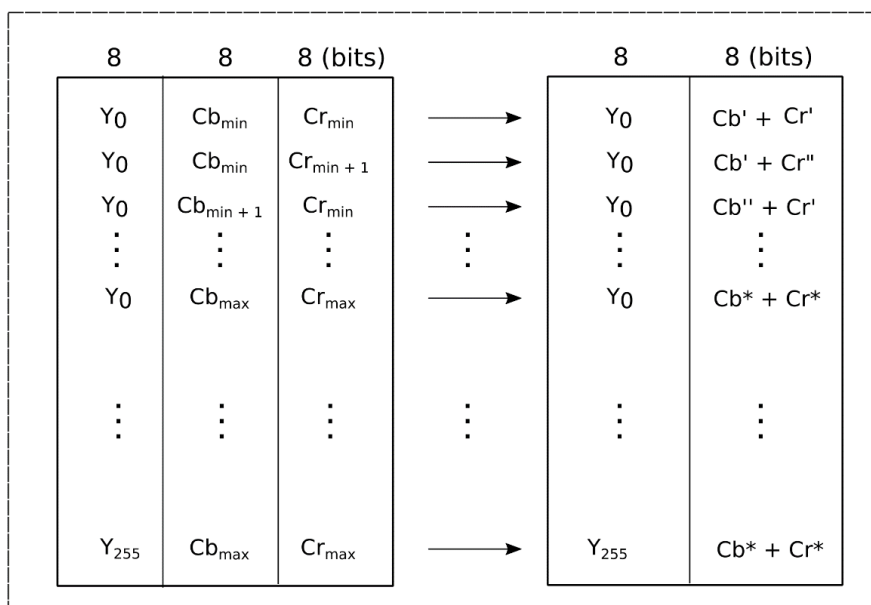


Figura 18 – Tabela de conversão para pré-codificação de cores.

A abordagem utilizada na construção da tabela consiste em codificar o novo valor a partir da distância em relação à média. Desse modo, para encontrar o valor de  $Cb'$  e  $Cr'$  foi necessário aplicar uma regra de três. Após a conversão, os valores de  $Cb'$  e  $Cr'$  foram concatenados, dando origem a um número de oito bits.

O procedimento de reversão de  $Cb'$  e  $Cr'$  para os valores originais utilizou a mesma tabela de conversão, portanto o decodificador deve conhecer a mesma tabela de cores utilizada na codificação.

## C2 – Fatoração

As matrizes de entrada foram fatoradas nas matrizes  $U$ ,  $\Sigma$  e  $V$  pela DVS. Antes de serem fatoradas, as matrizes foram subdivididas em blocos menores de  $60 \times 60$  pixels. A Figura 19 ilustra os passos da fatoração para o caso hipotético de uma matriz  $2 \times 2$ .

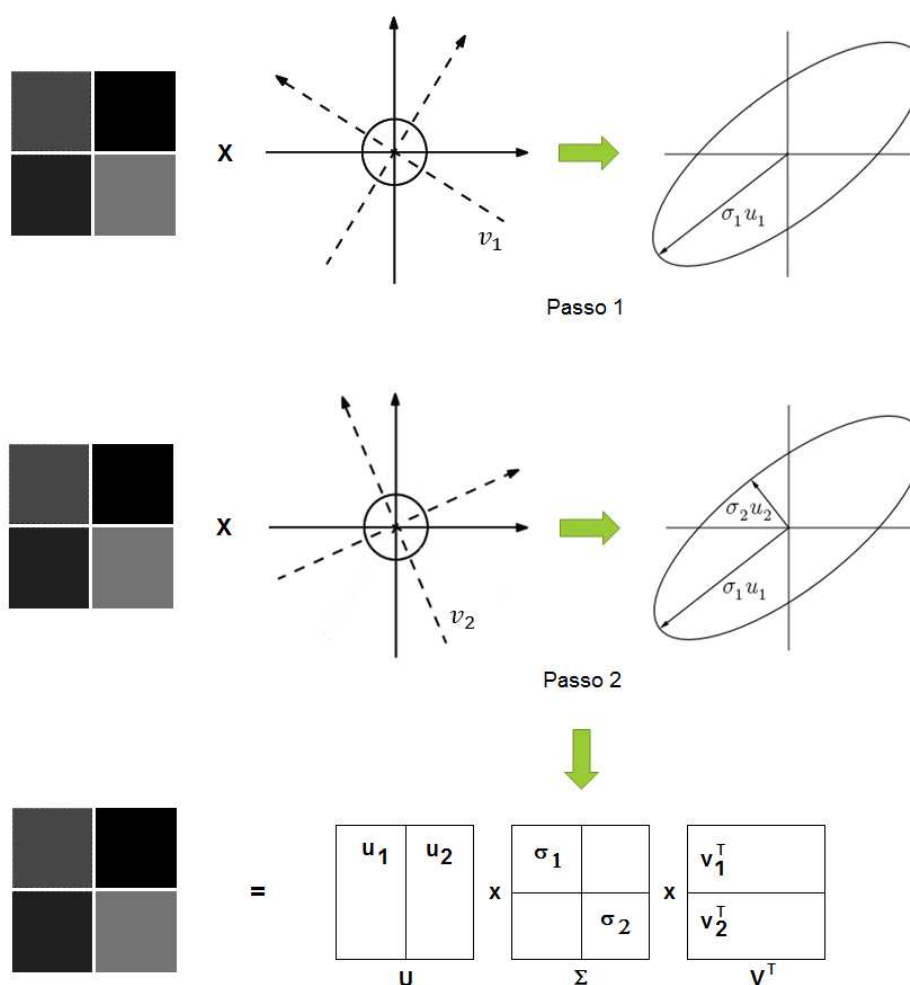


Figura 19 – Ilustração dos passos da fatoração do bloco.

O objetivo da DVS é encontrar duas matrizes ortogonais  $U$  e  $V$  que satisfazem à igualdade:

$$A = U \times \Sigma \times V^T$$

Onde  $\Sigma$  é uma matriz diagonal de valores singulares.

Para simplificar o cálculo da DVS partiu-se da condição de que a matriz de entrada era simétrica. Uma matriz simétrica pode ser obtida a partir da multiplicação da matriz de entrada  $A$  transposta ( $A^T$ ) por ela própria:

$$A^T \times A$$

Os elementos da diagonal principal de  $\Sigma$  foram unicamente determinados e correspondem às raízes quadradas dos autovalores de  $A^T \times A$ . Estes satisfazem à condição da Expressão 4:

$$\sigma_1, \sigma_2, \dots, \sigma_p, 0 \quad (\text{onde } p = \min \{m, n\}, \text{ e } \sigma_1 \geq \sigma_2 \dots \geq \sigma_p) \quad (6)$$

Onde  $m$  e  $n$  correspondem, respectivamente, ao número de linhas e colunas de  $A$ .

Desse modo, os elementos diagonais de  $\Sigma$  são os valores singulares, as colunas de  $U$  são os vetores singulares à esquerda, e as colunas de  $V$  são os vetores singulares à direita.

Após a fatoração, apenas os  $k$  primeiros fatores de cada bloco foram aproveitados. O valor de  $k$  foi obtido a partir do cálculo da qualidade após a reconstrução da imagem, sendo  $k$  menor que a largura e a altura do bloco. A qualidade da reconstrução foi medida pela relação sinal-ruído de pico (PSNR), e é dada pela fórmula abaixo:

$$\text{PSNR} = 10 \cdot \log_{10} \left( \frac{As_{\max}^2}{\sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} (e[x,y])^2} \right) \quad (7)$$

Onde:

- $X$  e  $Y$  são as dimensões da imagem (largura e altura);
- $A$  é a “área” da imagem, ou seja,  $X \cdot Y$ ;
- $s_{\max}$  é o valor máximo do pixel, neste caso 256;

- $e[x, y]$  é a diferença entre o valor original do pixel em  $(x, y)$  e seu correspondente após a reconstrução.

Na indexação, o valor de  $k$  é fixo pois a matriz original não precisa ser reconstruída, sendo menor ou igual a dez.

### C3. Codificação binária

A codificação das imagens ou dos vídeos utilizando as matrizes resultantes da decomposição em valores singulares foi expressa na forma de combinação linear, a qual consiste na soma dos  $k$  fatores  $u_i$ ,  $\sigma_i$  e  $v_i^T$  mais significativos da decomposição.

A combinação linear permite que os valores sejam serializados, e o acréscimo de fatores acontece até o nível em que a reconstrução do quadro alcance um valor mínimo de qualidade objetiva. A Figura 20 ilustra a serialização dos vetores e valores singulares.

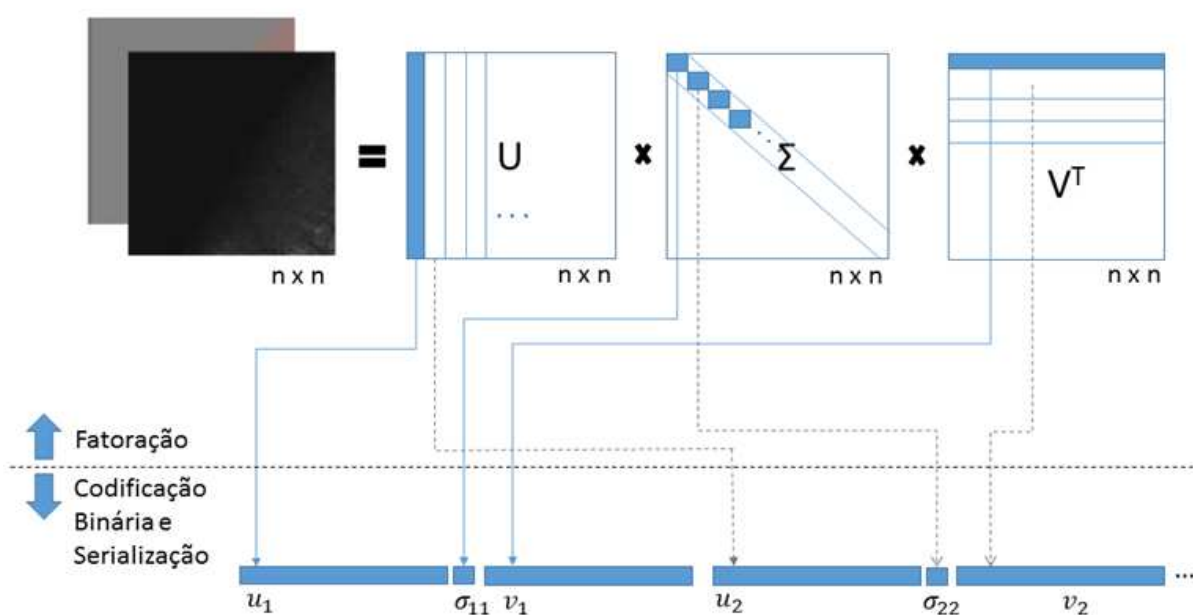


Figura 20 – Codificação binária e serialização dos valores e vetores singulares.

Os valores singulares, bem como os componentes dos vetores singulares  $U$  e  $V$  são números reais. Desse modo, para simplificar a representação, os valores foram convertidos a números inteiros de 16 bits. Os valores que compõem os vetores singulares variam entre -1 e 1, e por isso sua conversão a inteiro é obtida multiplicando-os pelo número 32.768 ( $2^{15}$ ). Após convertida a inteiro, a sequência de valores foi gravada em arquivo.



### I3 – Sumarização

Os dez vetores singulares à direita mais significativos foram resumidos aplicando-se o cálculo da média quadrática (RMS) sobre os valores de cada autovetor. O número de vetores singulares considerados pode variar conforme a aplicação, não sendo superior a 10. O cálculo do valor RMS é dado pela Equação 8.

$$Vi_{RMSj} = \sqrt{\frac{1}{n}(v_{ij_1}^2 + v_{ij_2}^2 + \dots + v_{ij_n}^2)} \quad (8)$$

Onde:

- $i$  é o índice do autovetor, e varia de 1 a 10,
- $j$  é o índice do valor RMS que resume  $n$  valores do autovetor  $i$ , e
- $n$  é o número de elementos do autovetor tomados em grupo para calcular o RMS.

O valor de  $n$  na Equação 8 corresponde ao divisor do tamanho do autovetor original, e foi estipulado em 20. A Figura 21 ilustra o efeito da sumarização pela média quadrática sobre um autovetor, apresentado na forma gráfica.

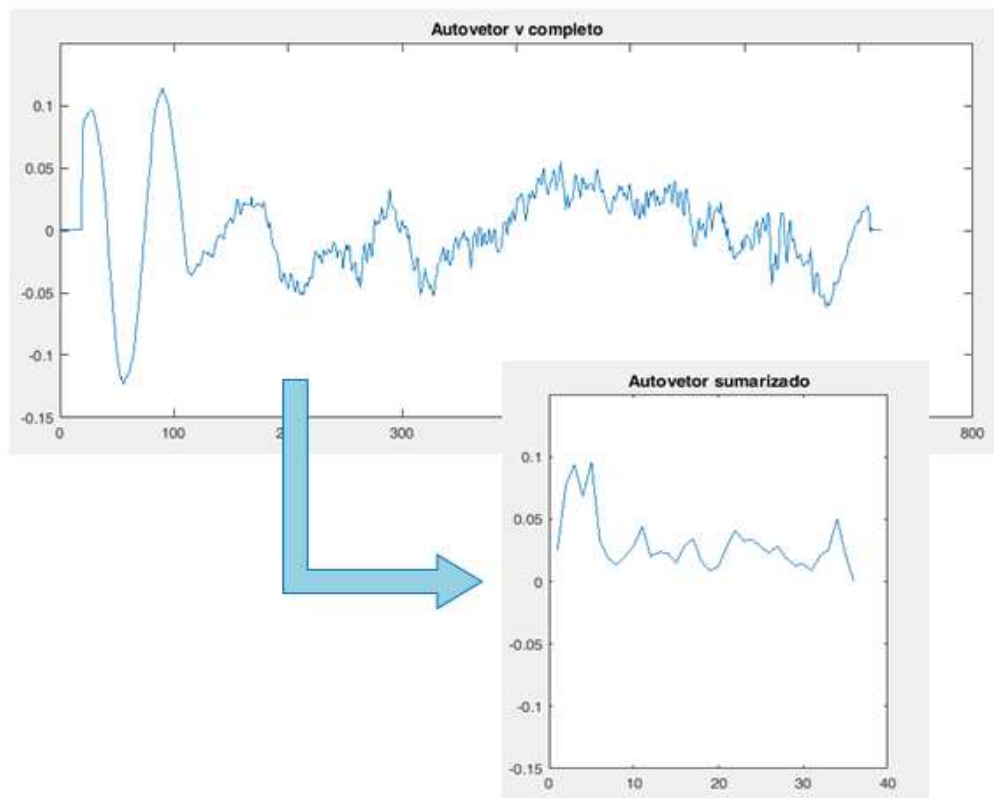



Figura 21 – Efeito da sumarização sobre um autovetor.

#### I4 – Atualização dos índices

As sequências de autovetores sumarizados foram concatenadas e adicionadas à tabela de índices em ordem, com a referência para o quadro correspondente, conforme ilustrado na Figura 22.



Identificador	Quadro
[0.0227 0.0149 0.0289 0.0343 ... ]	523
[0.0256 0.0779 0.0937 0.0679 ... ]	1617
[0.0283 0.0408 0.0324 0.0338 ... ]	1101
[0.0287 0.0188 0.0125 0.0143 ... ]	41

Figura 22 – Atualização da tabela de índices.

#### **4.4.2. Etapa 2: Recuperação**

Os passos da Etapa 2 estão representados graficamente na Figura 23, e detalhados em seguida.

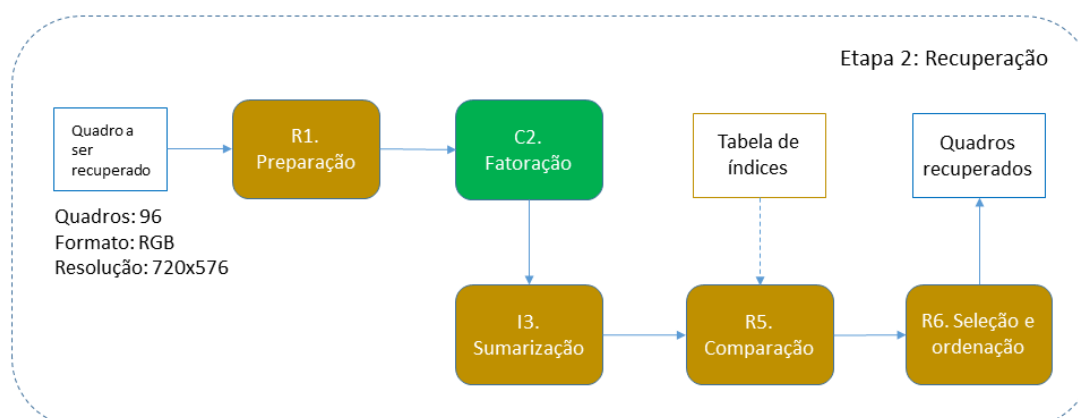


Figura 23 – Passos da Etapa Experimental 2.

#### R1 – Preparação

A imagem de exemplo foi convertida para uma representação monocromática em tons de cinza. No caso de imagens com formato de pixel YCbCr, o componente Y corresponde à imagem original em tons de cinza, e não requer outra conversão.

#### C2 – Fatoração

O procedimento de fatoração foi o mesmo descrito no passo C2 da Subseção 4.4.1 – Etapa 1.

### I3. Sumarização

Nesta etapa, a sumarização também envolveu os mesmos procedimentos do passo I3 da Subseção 4.4.1 – Etapa 1.

### R5. Determinação

A determinação é um passo exclusivo do processo de recuperação. Nele os vetores singulares resumidos extraídos da imagem de entrada ( $V_E$ ) foram comparados com os vetores singulares resumidos presentes na tabela de índices ( $V_T$ ) um a um. Para esta comparação utilizou-se o coeficiente de determinação ( $R^2$ ), cujo cálculo é dado pela Equação 9.

$$R^2 = 1 - \frac{SQ_e}{SQ_e + SQ_r} \quad (9)$$

Onde:

- $SQ_e$  ou a *variação não explicada* é o somatório do quadrado dos erros, ou seja, das diferenças entre o valor de  $v_T$  e o valor correspondente de  $v_T$  previsto pelo modelo (Equação 9.1);
- $SQ_r$  ou a *variação explicada* é o somatório do quadrado das diferenças entre o valor de  $v_T$  previsto e a média de  $v_T$  (Equação 9.2).

$$SQ_e = \sum_i (v_{T_i} - \hat{v}_T)^2 \quad (9.1)$$

$$SQ_r = \sum_i (\hat{v}_T - \bar{v}_T)^2 \quad (9.2)$$

Onde:

- $v_{T_i}$  corresponde aos valores de cada vetor singular resumido  $V_T$ ;
- $\hat{v}_T$  corresponde ao valor de  $v_T$  previsto usando o modelo de regressão linear derivado dos dados de  $V_E$  e  $V_T$ . O valor de  $\hat{v}_T$ , por sua vez, é determinado pela Equação 9.1.1;
- $\bar{v}_T$  é a média dos valores do vetor singular resumido  $V_T$ .

$$\hat{v}_T = \hat{\beta}_0 + \hat{\beta}_1 \cdot v_{E_i} \quad (9.1.1)$$

Onde:

- $\hat{\beta}_0$  e  $\hat{\beta}_1$  são, respectivamente, o intercepto e a inclinação do modelo de regressão linear, estimados a partir dos valores de  $V_E$  e  $V_T$  utilizando o método de mínimos quadrados;
- $v_{E_i}$  corresponde aos valores de cada vetor singular resumido  $V_E$ .

## 16. Seleção e ordenação

Para a realização dos experimentos, 97 quadros relevantes para o diagnóstico foram selecionados por especialista médico. As imagens escolhidas foram utilizadas como exemplos a serem recuperados.

No processo de seleção e ordenação, após o passo anterior (R5. Determinação), foram selecionados os quadros da tabela de índice cujo valor de  $R^2$  relativo ao primeiro vetor singular foi igual ou superior ao limiar de correspondência. Nesse sentido, foram avaliados três valores para o limiar de correspondência: 0,9, 0,95 e 0,975.

Após a seleção, os quadros foram ordenados segundo dois critérios diferentes. No primeiro caso, os resultados foram ordenados, de modo decrescente, pela soma dos valores de  $R^2$  maiores que 0,9.

No segundo caso, os valores de  $R^2$  foram agrupados por faixa de valores, e ordenados, de maneira decrescente, pela quantidade de valores em cada faixa, priorizando as faixas de maior valor. Assim, foram definidas dez faixas no intervalo de 0 a 1, conforme ilustrado na Tabela 2.

Os resultados foram ordenados segundo a contagem de cada faixa de valores de  $R^2$ , considerando primeiro a Faixa 1 e depois as seguintes, até a Faixa 10.

Tabela 2 – Faixas de valores de  $R^2$  consideradas para ordenação.

Faixa	Valores de $R^2$
1	$0,9 < R^2 \leq 1,0$
2	$0,8 < R^2 \leq 0,9$
3	$0,7 < R^2 \leq 0,8$
4	$0,6 < R^2 \leq 0,7$

*Continua na próxima página*

Conclusão da Tabela 2.

Faixa	Valores de $R^2$
5	$0,5 < R^2 \leq 0,6$
6	$0,4 < R^2 \leq 0,5$
7	$0,3 < R^2 \leq 0,4$
8	$0,2 < R^2 \leq 0,3$
9	$0,1 < R^2 \leq 0,2$
10	$0,0 < R^2 \leq 0,1$

As configurações experimentais para os Casos 1 e 2 estão resumidas na Tabela 3.

Tabela 3 – Resumo das configurações aplicadas à recuperação.

Caso	Configuração	Seleção ( $R^2$ para o primeiro autovetor)	Ordenação
1	a	$> 0,9$	Soma dos valores de $R^2 > 0,9$
	b	$> 0,95$	
	c	$> 0,975$	
2	a	$> 0,9$	Quantidade de valores de $R^2$ nas faixas de valores superiores.
	b	$> 0,95$	
	c	$> 0,975$	

Os Algoritmos 1 e 2 detalham os passos de seleção e ordenação para os Casos 1 e 2 respectivamente.

Por meio desses procedimentos, foram retornadas dez referências para as imagens ou vídeos com maior semelhança à imagem de entrada.

Algoritmo 1 – Seleção e ordenação para o Caso 1, formatado segundo padrão IEEE.

**Descrição** Algoritmo para seleção e ordenação dos resultados da busca para o Caso 1.

**Entrada:** Autovetores resumidos da imagem de exemplo para a busca, e tabela de índices.

**Saída:** Lista das dez imagens mais semelhantes.

*Definição de parâmetros constantes:*

*p: quantidade de autovalores considerados na comparação*

1:  $p = 10$

*Continua na próxima página.*

### Conclusão do Algoritmo 1.

---

*limiar\_correspondencia*: valor mínimo de  $R^2$  a ser aplicado na seleção

2:  $limiar\_correspondencia = 0,9, 0,95$  ou  $0,975$   
*k*: quantidade de imagens a ser retornada na recuperação

3:  $k = 10$

*Definição de variáveis:*

*contador*: quantidade de autovetores com  $R^2$  maior que  $0,9$

4:  $contador = 0$   
*soma\_R2*: valor acumulado de  $R^2$  para imagem avaliada

5:  $soma\_R2 = 0$

6: **Para** cada imagem indexada, **faça**:

7:     **Faça**  $i$  variando de 1 a  $p$ :

8:          $resumo\_ref \leftarrow$  autovetor\_ $i$  resumido da imagem de exemplo

9:          $resumo\_ind \leftarrow$  autovetor\_ $i$  resumido da imagem indexada

10:          $R2 \leftarrow$  coeficiente de determinação entre  $resumo\_ref$  e  $resumo\_ind$

11:         **Se**  $i == 1$  e  $R2 \geq limiar\_correspondencia$

12:             **Selecione** a imagem indexada

13:         **Senão**

14:             **Desconsidere** a imagem indexada e salte para a próxima

15:         **Fim-Se**

16:         **Se**  $R2 \geq 0,9$

17:              $contador \leftarrow contador + 1$

18:              $soma\_R2 \leftarrow soma\_R2 + R2$

19:         **Fim-Se**

20:     **Fim\_Faça**

21: **Fim-Para**

22: **Ordene** a lista selecionada pelos valores de  $contador$  e de  $soma\_R2$

23: **Retorna** as  $k$  primeiras imagens da lista

---

Algoritmo 2 - Seleção e ordenação para o Caso 2, formatado segundo padrão IEEE.

---

**Descrição** Algoritmo para seleção e ordenação dos resultados da busca para o Caso 2.

---

*Continua na próxima página.*

*Continuação do Algoritmo 2.*

---

**Entrada:** Autovetores resumidos da imagem de exemplo para a busca, e tabela de índices.

**Saída:** Lista das dez imagens mais semelhantes.

*Definição de parâmetros constantes:*

*p: quantidade de autovalores considerados na comparação*

1:  $p = 10$

*limiar\_correspondencia: valor mínimo de  $R^2$  a ser aplicado na seleção*

2:  $\text{limiar\_correspondencia} = 0,9, 0,95 \text{ ou } 0,975$

*k: quantidade de imagens a ser retornada na recuperação*

3:  $k = 10$

*Definição de variáveis:*

*contadores: vetor de contadores de autovetores, agrupados por faixa de valor de  $R^2$*

4:  $\text{contadores}[1..10] = 0$

5: **Para** cada imagem indexada, **faça:**

6:     **Faça**  $i$  variando de 1 a  $p$ :

7:          $\text{resumo\_ref} \leftarrow$  autovetor\_ $i$  resumido da imagem de exemplo

8:          $\text{resumo\_ind} \leftarrow$  autovetor\_ $i$  resumido da imagem indexada

9:          $R2 \leftarrow$  coeficiente de determinação entre  $\text{resumo\_ref}$  e  $\text{resumo\_ind}$

10:        **Se**  $i == 1$  e  $R2 \geq \text{limiar\_correspondencia}$

11:            **Selecione** a imagem indexada

12:            **Senão**

13:                **Desconsidere** a imagem indexada e salte para a próxima

14:            **Fim-Se**

15:     **Escolha**  $R2$

16:         **caso**  $> 0,9$  e  $\leq 1$  **faça**  $\text{contadores}[1] \leftarrow \text{contadores}[1]+1$

17:         **caso**  $> 0,8$  e  $\leq 0,9$  **faça**  $\text{contadores}[2] \leftarrow \text{contadores}[2]+1$

18:         **caso**  $> 0,7$  e  $\leq 0,8$  **faça**  $\text{contadores}[3] \leftarrow \text{contadores}[3]+1$

19:         **caso**  $> 0,6$  e  $\leq 0,7$  **faça**  $\text{contadores}[4] \leftarrow \text{contadores}[4]+1$

20:         **caso**  $> 0,5$  e  $\leq 0,6$  **faça**  $\text{contadores}[5] \leftarrow \text{contadores}[5]+1$

21:         **caso**  $> 0,4$  e  $\leq 0,5$  **faça**  $\text{contadores}[6] \leftarrow \text{contadores}[6]+1$

22:         **caso**  $> 0,3$  e  $\leq 0,4$  **faça**  $\text{contadores}[7] \leftarrow \text{contadores}[7]+1$

23:         **caso**  $> 0,2$  e  $\leq 0,3$  **faça**  $\text{contadores}[8] \leftarrow \text{contadores}[8]+1$

24:         **caso**  $> 0,1$  e  $\leq 0,2$  **faça**  $\text{contadores}[9] \leftarrow \text{contadores}[9]+1$

25:         **caso**  $\leq 0,1$  **faça**  $\text{contadores}[10] \leftarrow \text{contadores}[10] + 1$

26:     **Fim-Escolha**

---

*Continua na próxima página.*

### *Conclusão do Algoritmo 2.*

---

```
27:   Fim_Faça
28:   Fim_Para
29:   Ordene a lista selecionada pelos valores de contadores[1],
      em seguida pelos valores de contadores[2], e assim
      sucessivamente
30:   Retorne as k primeiras imagens da lista
```

---

#### **4.4.3. Etapa 3: Análise estatística**

O resultado da compressão (Etapa 1) foi avaliado pela taxa de compressão obtida para cada quadro, e pela qualidade objetiva da reconstrução da imagem, dada pelo cálculo de PSNR.

Após, o desempenho obtido foi comparado estatisticamente com o desempenho do codificador H.263 para o mesmo grupo de imagens de entrada e as mesmas configurações.

Os grupos de dados comparados referem-se à taxa de compressão e à qualidade objetiva calculada pelo PSNR. Cada grupo foi submetido ao teste de normalidade Shapiro-Wilk, com rejeição da hipótese nula para  $p\text{-valor} \leq 0,05$ .

Com o resultado do teste de normalidade, foi definido o teste estatístico para dados pareados fixando-se o intervalo de confiança de 95%.

#### **4.4.4. Etapa 4: Avaliação do resultado da recuperação**

A recuperação dos quadros foi considerada adequada sempre que o próprio quadro foi reconhecido e ordenado em primeiro lugar. Os demais quadros recuperados deveriam estar próximos ou adjacentes à cena buscada e serem visualmente semelhantes.

A verificação da semelhança visual foi feita com a supervisão de especialista da área médica.



## 5. RESULTADOS

Neste trabalho, os resultados da busca de anterioridade, bem como os alcançados por meio da avaliação experimental estão apresentados a seguir.

### 5.1. Busca de anterioridade

Nas Tabelas 4 a 7 estão apresentados os resultados das quatro buscas realizadas. Na primeira e na segunda tabelas encontram-se os resultados da busca relacionada à codificação, enquanto na terceira e na quarta tabelas os resultados relativos à indexação e recuperação. Em cada tabela, na primeira coluna estão identificadas as bases consultadas, e nas colunas que seguem à direita constam o número de resultados retornados (segunda coluna), e o número de trabalhos selecionados para comparação (terceira coluna) após a leitura dos resumos e ou das reivindicações.

Tabela 4 – Resultados da busca ampla para a codificação atualizada em maio de 2018.

Base	Resultados	Documentos selecionados
Esp@cenet	10	0
USPTO	332	3

Tabela 5 – Resultados da busca específica para a codificação atualizada em maio de 2018.

Base	Resultados	Documentos selecionados
CIPO	4	0
Esp@cenet	0	0
JPO	57	0
USPTO	6	0

Tabela 6 – Resultados da busca ampla para a indexação realizada em junho e julho de 2018.

Base	Resultados	Documentos selecionados
CIPO	4	1

*Continua na próxima página.*

Conclusão da Tabela 6.

Base	Resultados	Documentos selecionados
Esp@cenet	0	0
FPO (todas as bases selecionadas)	198	0
INPI	0	0
JPO	2	0
USPTO	13	1
WIPO	1	1

Tabela 7 – Resultados da busca específica para a indexação realizada em julho de 2018.

Base	Resultados	Documentos selecionados
ACM	1	0
IEEEExplore	33	0
Scopus	1	0

## 5.2. Etapa 1 – Codificação e indexação

A Figura 24 ilustra, na forma de imagem monocromática, a decomposição em matrizes e a reconstrução utilizando três, dez e 20 valores e vetores singulares.

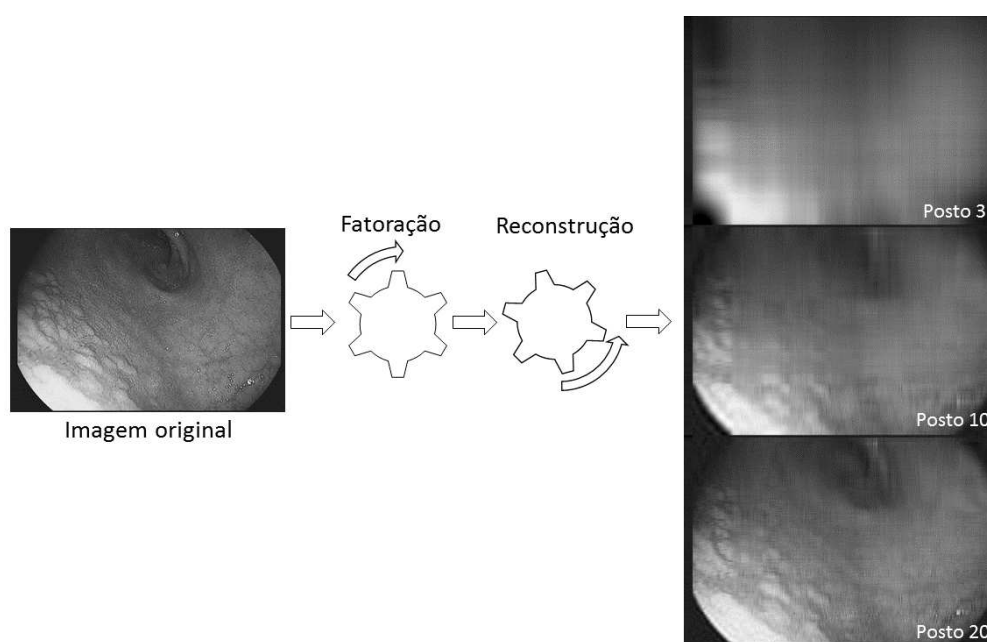


Figura 24 – Fatoração e reconstrução gradual.

Os resultados da codificação em termos da taxa de compressão alcançada e a qualidade obtida estão expressos na Tabela 8. Nas colunas do lado esquerdo da tabela estão indicados os desempenhos alcançados por ambos os codificadores para a taxa de compressão. Por outro lado, nas colunas do lado direito, constam os desempenhos alcançados para a qualidade de imagem, medidas em decibéis pelo PSNR.

Tabela 8 – Desempenho do método de codificação.

	Taxa de compressão		PSNR	
	Método proposto	H.263	Método proposto	H.263
<b>Média</b>	82,07%	90,28%	38,85 dB	53,93 dB
<b>Desvio padrão</b>	6,37%	2,44%	1,06 dB	0,50 dB

O resultado do teste de normalidade indicou que os grupos de dados não apresentaram distribuição normal. Desse modo, foi aplicado o teste de Wilcoxon, o qual demonstrou diferença estatisticamente significativa entre os grupos ( $p$ -valor  $< 0,0001$ ). A Tabela 9 sintetiza estes resultados.

Tabela 9 – Análise estatística dos resultados.

Teste estatístico		Taxa de compressão	PSNR
<b>Wilcoxon</b>	Método proposto vs. H.263 (P-valor)	$< 0,0001$	$< 0,0001$

Na Figura 25 estão demonstradas, lado a lado, uma imagem de exemplo (esquerda) e as versões da mesma imagem após terem sido comprimidas e reconstruídas utilizando o método proposto (centro) e o codificador H.263 (direita).



Figura 25 – Imagem codificada e reconstruída com método proposto e H.263.

### 5.3. Etapa 2 – Recuperação

Na Tabela 10 estão inseridos os resultados alcançados nos experimentos em termos de precisão, sensibilidade, especificidade e tempo médio de busca, considerando os três primeiros quadros recuperados para cada imagem de exemplo.

Os dados relativos ao Caso 1 constam nas colunas 2, 3 e 4, cada qual correspondendo a um critério específico de seleção e ordenação dos resultados, conforme detalhado na Tabela 3. De modo similar, os resultados para o Caso 2 figuram nas colunas 5, 6 e 7.

Cada linha da tabela corresponde a uma medida de desempenho do recuperador. As medidas de precisão, sensibilidade e especificidade foram calculadas considerando os três primeiros quadros de cada busca. Na precisão foram considerados relevantes os quadros com semelhança visual da mesma região anatômica, e com distância máxima de 100 quadros (antes ou depois) em relação ao quadro buscado. Para o cálculo de sensibilidade, a distância em relação ao quadro de referência foi desconsiderada, e apenas a compatibilidade visual e anatômica foi levada em consideração.

Tabela 10 – Resultados dos experimentos de recuperação considerando os três primeiros quadros recuperados.

Configuração	Caso 1			Caso 2		
	a	b	c	a	b	c
<b>Precisão</b>	94,85%	94,85%	94,85%	94,50%	94,50%	95,19%
<b>Sensibilidade</b>	96,22%	96,22%	96,22%	94,85%	94,85%	95,88%
<b>Especificidade</b>	99,99%	99,99%	99,99%	99,98%	99,98%	99,98%
<b>Tempo médio (s)</b>	1,70	1,30	1,10	59,56	42,37	20,80

Na Tabela 11 são apresentados os resultados em termos de precisão considerando os cinco e os dez primeiros quadros recuperados para cada imagem de exemplo.

Tabela 11 – Precisão da recuperação considerando os cinco e os dez primeiros quadros recuperados.

Configuração	Caso 1			Caso 2		
	a	b	c	a	b	c
<b>Precisão (Cinco primeiros quadros)</b>	89,69%	89,69%	89,69%	90,93%	90,93%	90,93%
<b>Precisão (10 primeiros quadros)</b>	77,42%	77,42%	77,42%	80,41%	80,41%	80,62%

Na Tabela 12 são apresentados os resultados da recuperação para os primeiros 20 quadros, referentes ao Caso 1. A tabela completa contendo os resultados para todas as imagens buscadas encontra-se no Anexo C, Tabela C.1. Na primeira coluna à esquerda estão identificados os quadros buscados, e na segunda coluna os quadros recuperados, listados em ordem decrescente de similaridade. Embora a recuperação tenha retornado dez quadros para cada busca, na tabela foram apresentados apenas os três quadros melhor classificados.

Nas colunas restantes estão listados os valores que embasam a ordenação dos resultados. A terceira coluna apresenta a quantidade de vetores singulares para os quais o cálculo de  $R^2$  resultou maior que o limiar de correspondência. Similarmente, na quarta coluna são apresentados os resultados da soma dos  $R^2$  que foram maiores que 0,9.

Tabela 12 – Quadros recuperados para o Caso 1.

Nº do quadro buscado	Quadros retornados	Vetores com $R^2 > 0,9$	Soma dos $R^2 > 0,9$
9	9	10	10,0000
	7228	1	0,9927
	7229	1	0,9922
10	10	10	10,0000
	6984	1	0,9839
	6988	1	0,9838
216	216	10	10,0000
	215	1	0,9958
	217	1	0,9956

*Continua na próxima página.*

Continuação da Tabela 12.

<b>Nº do quadro buscado</b>	<b>Quadros retornados</b>	<b>Vetores com <math>R^2 &gt; 0,9</math></b>	<b>Soma dos <math>R^2 &gt; 0,9</math></b>
231	231	10	10,0000
	230	1	0,9993
	228	1	0,9991
232	232	10	10,0000
	229	1	0,9992
	228	1	0,9990
236	236	10	10,0000
	239	1	0,9992
	269	1	0,9991
290	290	10	10,0000
	289	1	0,9999
	291	1	0,9999
339	339	10	10,0000
	338	2	1,9375
	340	1	1,0000
340	340	10	10,0000
	338	2	1,9211
	341	2	1,9019
342	342	10	10,0000
	341	2	1,9110
	343	1	0,9999
346	346	10	10,0000
	347	1	0,9998
	345	1	0,9997
406	406	10	10,0000
	397	1	0,9995
	398	1	0,9995
407	407	10	10,0000
	409	1	0,9997
	410	1	0,9996
439	439	10	10,0000
	440	2	1,9500
	438	1	0,9998

*Continua na próxima página.*

Conclusão da Tabela 12.

<b>Nº do quadro buscado</b>	<b>Quadros retornados</b>	<b>Vetores com <math>R^2 &gt; 0,9</math></b>	<b>Soma dos <math>R^2 &gt; 0,9</math></b>
527	527	10	10,0000
	526	1	0,9999
	529	1	0,9997
566	566	10	10,0000
	565	1	0,9990
	530	1	0,9986
624	624	10	10,0000
	631	1	0,9991
	632	1	0,9987
691	691	10	10,0000
	692	1	0,9976
	693	1	0,9965
696	696	10	10,0000
	697	2	1,9115
	706	1	0,9992
702	702	10	10,0000
	703	3	2,8458
	698	1	0,9997

Na Tabela 13 são apresentados os resultados da recuperação para as primeiras 20 imagens referentes ao Caso 2, na configuração que obteve melhor desempenho. A tabela completa encontra-se no Anexo C, Tabela C.2.

Semelhantemente à Tabela 12, na primeira coluna à esquerda estão identificados os quadros buscados, na segunda coluna os quadros recuperados, listados em ordem decrescente de similaridade, e nas duas colunas restantes constam os valores que embasam a ordenação dos resultados. A terceira coluna apresenta a quantidade de vetores singulares para os quais o cálculo de  $R^2$  resultou maior que 0,9, correspondendo à Faixa 1 (Tabela 2). Ademais, na quarta coluna consta a quantidade de vetores comparados cujo  $R^2$  ficou entre 0,8 e 0,9, correspondendo à Faixa 2 de valores. As colunas referentes às faixas de valores seguintes foram omitidas por não terem influência sobre os resultados apresentados.

Tabela 13 – Quadros recuperados para o Caso 2.

<b>Nº do quadro buscado</b>	<b>Quadros retornados</b>	<b>Nº de vetores com <math>R^2 &gt; 0,9</math></b>	<b>Nº de vetores com <math>0,9 \geq R^2 &gt; 0,8</math></b>
9	9	460	0
	6864	90	26
	7	86	20
10	10	460	0
	7026	92	10
	7015	89	18
216	216	460	0
	217	83	44
	215	74	45
231	231	460	0
	230	154	56
	229	125	51
232	232	460	0
	227	162	56
	228	134	46
236	236	460	0
	247	212	21
	235	207	34
290	290	460	0
	289	333	19
	288	294	34
339	339	460	0
	338	212	47
	340	186	36
340	340	460	0
	341	279	43
	338	274	56
342	342	460	0
	343	294	35
	341	287	36
346	346	460	0
	345	282	38
	347	257	48

*Continua na próxima página.*



Conclusão da Tabela 13.

<b>Nº do quadro buscado</b>	<b>Quadros retornados</b>	<b>Nº de vetores com <math>R^2 &gt; 0,9</math></b>	<b>Nº de vetores com <math>0,9 \geq R^2 &gt; 0,8</math></b>
406	406	460	0
	407	240	32
	405	229	40
407	407	460	0
	406	240	32
	411	229	26
439	439	460	0
	440	352	26
	441	305	31
527	527	460	0
	528	249	33
	529	245	28
566	566	460	0
	565	290	29
	564	249	16
624	624	460	0
	623	145	16
	631	134	7
691	691	460	0
	690	139	36
	692	133	42
696	696	460	0
	697	188	59
	695	145	33
702	702	460	0
	703	222	63
	704	160	51

As Figuras 26 a 30 exemplificam resultados da recuperação para quadros representativos de diferentes regiões anatômicas, incluindo esôfago, estômago e duodeno. No canto esquerdo de cada figura consta a imagem de exemplo utilizada na busca, seguida à direita pelas duas imagens melhor classificadas. As imagens da fileira superior correspondem ao Caso 1, e as da fileira inferior ao Caso 2.

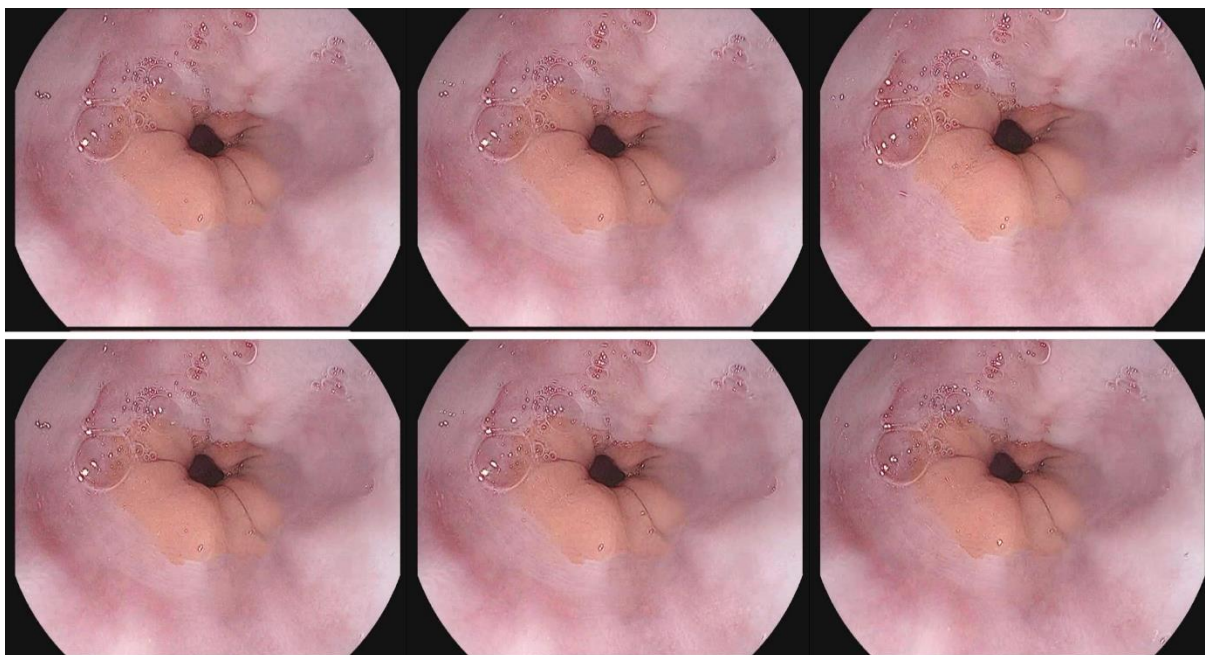


Figura 26 – Recuperação do quadro nº 696 (busca nº 19).



Figura 27 – Recuperação do quadro nº 1206 (busca nº 28).

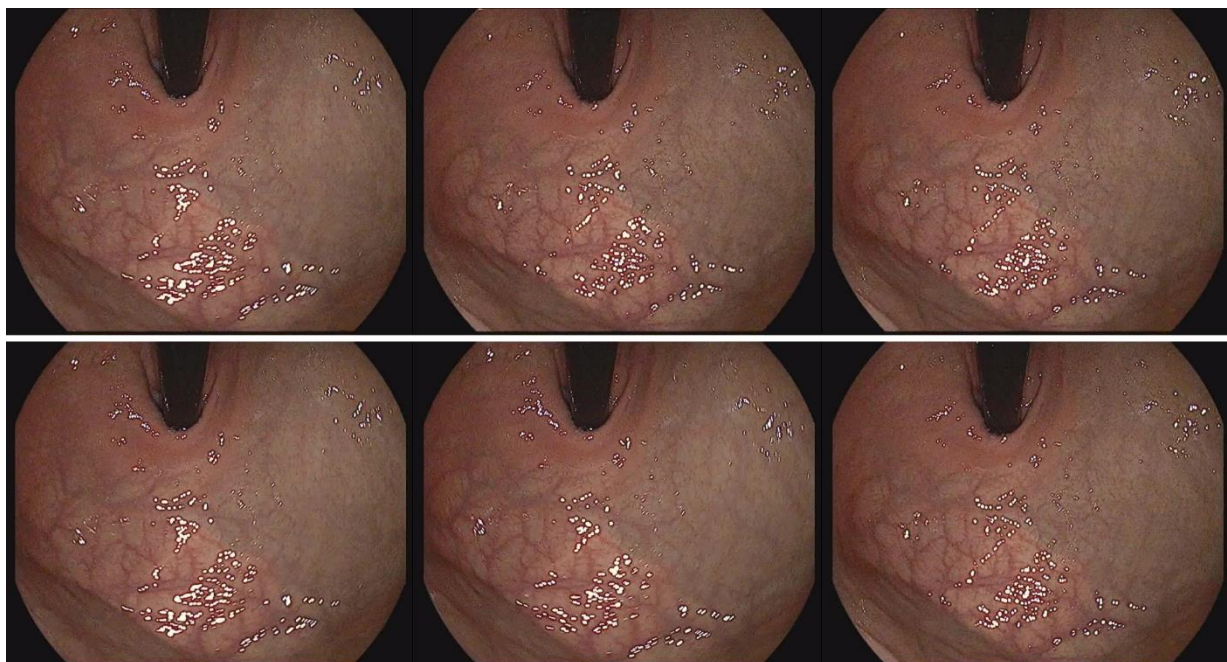


Figura 28 – Recuperação do quadro nº 2240 (busca nº 35).

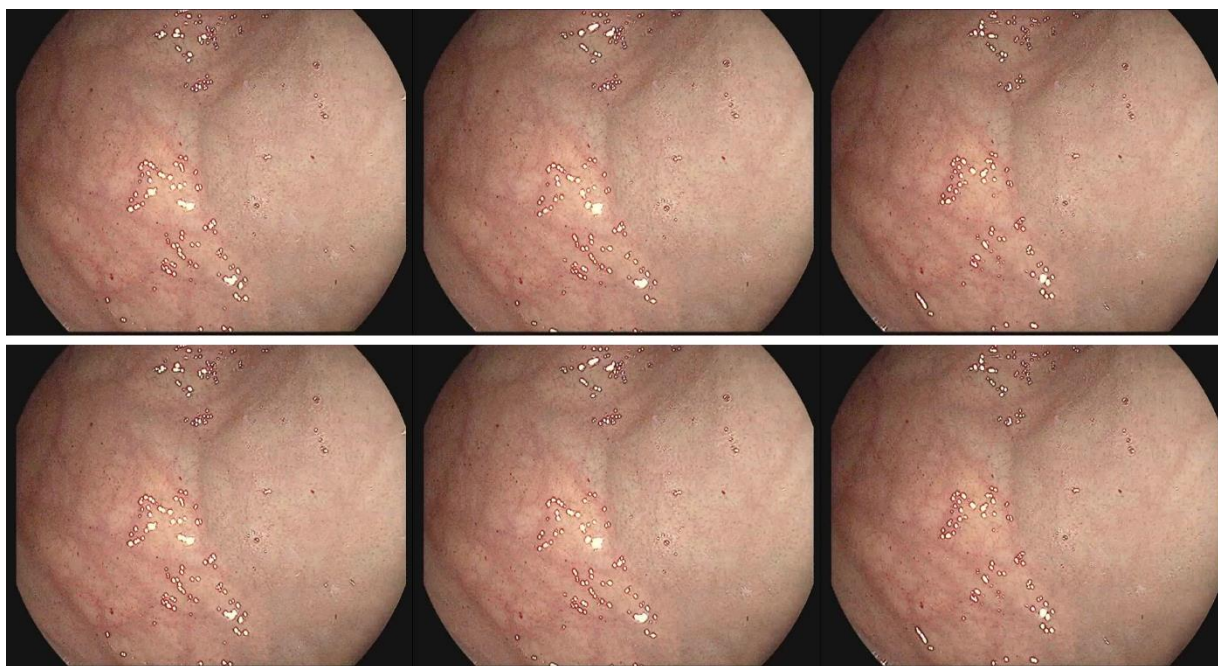


Figura 29 – Recuperação do quadro nº 2394 (busca nº36).

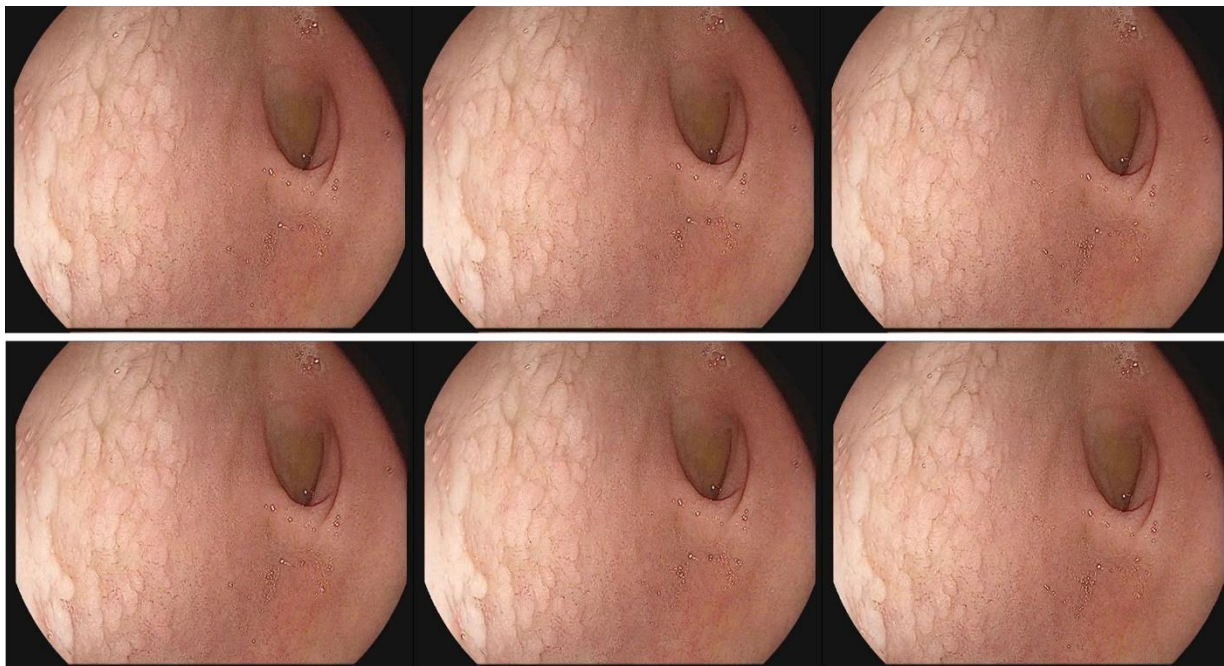


Figura 30 – Recuperação do quadro nº 3370 (busca nº 50).

## 6. DISCUSSÃO

A compressão de dados aplicada a informações multimídia foi um fator preponderante na massificação de tecnologias de imagem, áudio e vídeo digital (28-33). A possibilidade de armazenar e transmitir grandes volumes de imagens e vídeos médicos deve-se, em grande parte, aos avanços alcançados pelas técnicas de codificação (35,37). Atualmente, aplicações de vídeo-chamada e videoconferência tornaram-se comuns no cotidiano, e podem ser realizadas a partir de aparelhos pessoais de telefonia celular, com qualidade e segurança (117,118).

Nesse cenário, a busca de anterioridade e a revisão de literatura objetivaram identificar trabalhos relacionados à codificação de imagens e de vídeos que tivessem finalidades e campos de aplicação similares, além de executarem técnicas semelhantes às utilizadas neste trabalho. Por outro lado, devido à característica de originalidade do método foram utilizadas duas expressões de busca complementares, sendo uma ampla e outra específica. Pela busca ampla foram identificados trabalhos assemelhados em termos dos objetivos e ou técnicas utilizadas, enquanto na busca específica foram identificados possíveis trabalhos com grau de semelhança alto o suficiente para caracterizar a ausência de novidade no método proposto. Os resultados das buscas relacionadas à codificação estão resumidos nas Tabelas 4 e 5.

O método proposto neste trabalho difere dos métodos usuais de compressão nos seguintes aspectos:

1. A transformação algébrica não tem por objetivo a mudança da representação das imagens do domínio espacial para o domínio de frequência, mas a fatoração em componentes que possam ser ordenados por relevância;
2. A compressão ocorre como resultado do descarte de componentes irrelevantes;
3. Os componentes de cor são pré-codificados a fim de aumentar a taxa de compressão;
4. A indexação é realizada simultaneamente à codificação, pois ambos os processos utilizam a mesma técnica de fatoração das matrizes de entrada.

Quanto ao método algébrico de fatoração das matrizes utilizado (DVS), este difere de outras transformações algébricas no aspecto do pré-condicionamento

da matriz de entrada (33,64,119,120). Por este pré-condicionamento a matriz de entrada é multiplicada por sua transposta, resultando em uma matriz quadrada simétrica que apresenta duas vantagens principais:

- A matriz de entrada (imagem ou bloco) pode ter linhas e colunas de tamanhos diferentes (matriz retangular), sendo mais flexível que outros codificadores que requerem blocos quadrados;
- Evita a presença de números complexos na determinação dos valores singulares de  $\Sigma$ .

Após a fatoração, os vetores e valores singulares são ordenados de modo decrescente. Desse modo, obtém-se uma decomposição na qual os componentes mais representativos da matriz original são dispostos nas posições iniciais das matrizes  $U$ ,  $\Sigma$  e  $V$ . Assim, tanto a compressão quanto a indexação podem tirar proveito desse fato, concentrando-se nos primeiros componentes das matrizes, pois neles estão acumuladas as informações mais relevantes para a reconstrução e a indexação das imagens (65,72,80). Conforme já indicado por Andrews e Patterson (72), a DVS é a melhor transformada ortogonal para compressão da energia de imagens.

Portanto, o método de codificação, indexação e recuperação de imagens e de vídeos apresentado neste trabalho trata essas tarefas como procedimento integrado, e por isso utiliza um processo algébrico comum.

Como resultado, a indexação é realizada simultaneamente à codificação, eliminando-se a necessidade de reprocessamento da imagem ou do vídeo, como acontece nas abordagens tradicionais (45-51). Por esse motivo, a maioria dos documentos selecionados na revisão de literatura foram obtidos na pesquisa de anterioridade para a análise de patenteabilidade do método. Os temas codificação e indexação são tradicionalmente tratados separadamente, tendo sido encontradas apenas duas referências que abordassem os dois aspectos no mesmo método, porém utilizando abordagens e estratégias diferentes (81,102).

Em Nakamura et al. (81) os autores apresentaram método integrado de compressão de imagens e de extração de características, tendo sido o trabalho com maior semelhança ao método proposto em termos do delineamento metodológico. Porém, os autores buscaram obter taxas de compressão mais elevadas em comparação com o padrão JPEG. Como consequência, o custo computacional se

eleva e inviabiliza a codificação de vídeos em tempo real, diferente do método proposto neste trabalho que tem por objetivo viabilizar a codificação e a indexação em tempo real. Além disso, o processo metodológico do trabalho relacionado não especifica quais as características são extraídas para a finalidade, portanto não estabelece um método de indexação e recuperação, mas apenas indica de modo superficial o processo de compressão. Ainda, em tarefas de recuperação em que são utilizados atributos de baixo nível, além da construção do índice, é necessário definir uma medida de distância a fim de mensurar o grau de similaridade entre os objetos, grandeza essa não especificada no documento relacionado. Sob esse aspecto, a medida de distância utilizada neste trabalho foi o coeficiente de determinação ( $R^2$ ), o qual permite comparar dois vetores numéricos com precisão (121).

Em outra abordagem similar, Vasconcelos e Lippman (102) sugeriram a adaptação do padrão MPEG para que a indexação fosse realizada simultaneamente à codificação. Os autores exploraram formas de calcular a biblioteca de códigos, e reconheceram que, dependendo da precisão desejada, é necessário reprocessar os quadros de entrada, tornando o processo computacionalmente custoso. Adicional a este cenário, a técnica proposta ainda apresentou menor capacidade discriminatória em relação a imagens homogêneas devido à semelhança existente entre os blocos que as constituem. Cabe ressaltar que imagens endoscópicas possuem alto grau de homogeneidade, ou seja, os pixels que compõem a imagem se distribuem em uma faixa ampla de valores, porém a variação ocorre de modo gradual em relação aos pixels da vizinhança (122). Este aspecto pode ser representado numericamente por um valor entre 0 e 1, constituindo, assim, uma característica útil na comparação entre imagens (52). Apesar disso, graus elevados de homogeneidade dificultam as tarefas de indexação por características visuais, conforme relatado também por Beecks et al. (113) e Munzer et al. (123).

Além do aspecto de homogeneidade presente em vídeos e imagens médicas em geral, os procedimentos endoscópicos possuem particularidades adicionais que dificultam a aplicação de técnicas convencionais de visão computacional a vídeos, exigindo abordagens específicas (113,123-125). Dentre essas peculiaridades estão a ausência de cortes ou mudanças abruptas de cena, o deslocamento rápido e de grande amplitude devido à magnificação, e a ocorrência de quadros borrados ou desfocados (124,125). Tais características não favorecem a

segmentação temporal em vídeos, nem a segmentação espacial em imagens ou quadros, para fins de indexação e recuperação (124-126).

Beecks et al. (113) abordaram tal dificuldade em lidar com indexação e recuperação de imagens endoscópicas propondo sintetizar diversas características em uma assinatura adaptável à faixa de variação das cores presentes nas imagens, cujo objetivo foi diferenciar mais claramente imagens com cores homogêneas. Além disso, os autores propuseram uma medida de distância para tais assinaturas que permite o pareamento parcial, e a comparação entre assinaturas de tamanhos diferentes. Estes aspectos contrastam com o método proposto neste trabalho, o qual não é afetado pela homogeneidade cromática pois as características são extraídas das imagens convertidas a tons de cinza. Além disso, as assinaturas são um resumo dos vetores que mais contribuem para a conformação das imagens, promovendo a identificação pelo seu aspecto global mais relevante. Por outro lado, diferentemente da medida de distância proposta por Beecks et al. (113) que permitem assinaturas de tamanhos diferentes, neste trabalho os vetores devem necessariamente possuir o mesmo tamanho.

Primus (124) e Primus et al. (127) também abordaram o problema da indexação e recuperação de imagens endoscópicas, e reconheceram a necessidade de desenvolver e aplicar técnicas específicas, que considerassem as características de homogeneidade cromática e ausência de cortes ou mudanças de cenas nos vídeos. Nesse sentido, os autores analisaram as imagens codificadas no modelo de cores HSV (53) a fim de detectar a presença e identificar os instrumentos presentes nos vídeos. Com isso, observaram que os valores de saturação (S) são baixos nos pixels que constituem os instrumentos cirúrgicos, enquanto nas regiões de tecido biológico os pixels possuem saturação alta e predominância do matiz vermelho. A mesma diferenciação entre as cores predominantes nos instrumentos e em tecidos do corpo foi observada no modelo de cores CIE  $L^*a^*b^*$  (52). Porém, a diferenciação entre regiões de tecidos sem a presença de instrumentos não foi abordada, e a indexação e a recuperação de segmentos foi reportada como trabalho futuro.

Cao et al. (125,126) também descreveram os aspectos desafiadores de imagens endoscópicas, e propuseram um método de segmentação de vídeos específico para exames colonoscópicos. Para apoio à segmentação do vídeo, os autores estabeleceram um modelo no qual os seis segmentos do intestino grosso são representados (ceco, restante do cólon ascendente, cólon transversal, cólon



descendente, sigmoide e reto), e buscaram detectar as transições entre os trechos a partir da presença de quadros borrados devido à aproximação da câmera em relação à mucosa nas curvaturas que demarcam os segmentos do cólon. Os autores, no entanto, não mencionaram outros aspectos que demarcam transições sem curvas entre os segmentos, tais como saliências ou depressões na parede interna das alças intestinais, e por isso não deixam claro como o método detecta as transições entre ceco e restante do cólon ascendente, e entre sigmoide e reto (126). Além disso os autores tampouco reportam a indexação e recuperação dos vídeos por semelhança de imagens, diferentemente deste trabalho que visa permitir a recuperação de vídeos a partir de imagens de exemplo.

Ainda neste trabalho, a identificação do conteúdo presente nas imagens não considera características de cor, pois conforme discutido anteriormente, o alto grau de homogeneidade cromática de imagens endoscópicas não favorece a diferenciação baseada em características de cor. Essa característica foi evidenciada de outro modo por Haneishi et al. (128) em estudo sobre o fator de reflexão espectral da mucosa gastrointestinal, no qual ficou demonstrado que a resposta de maior intensidade da mucosa se concentra na região entre as cores vermelha e púrpura.

Contudo, a homogeneidade cromática pode ser aproveitada na codificação de imagens e de vídeos endoscópicos para obter melhores taxas de compressão, conforme proposto neste trabalho. Na etapa de pré-codificação de cores buscou-se explorar a maior ocorrência de pixels de tonalidade vermelha estabelecendo uma tabela de conversão por meio da qual os códigos das cores foram reduzidos pela metade. Os códigos foram calculados com base na frequência de ocorrência das cores, conforme detalhado no passo C.1 da etapa de Codificação e Indexação do Delineamento Experimental (Seção 4.4.1).

Além disso, a pré-codificação de cores não afeta as características de luminância das imagens, as quais são mantidas com maior resolução, resultando em maior qualidade visual da imagem reconstruída após a decodificação. A preservação da luminância na pré-codificação também favorece a indexação das imagens, pois estas são submetidas à extração das assinaturas em seu formato original (31).

## **6.1. Codificação**

No cenário de codificação de cores, Choe e Choi (88) apresentaram método que realiza análise cromática para identificar correlações entres os

componentes de luminância e de cor, cujo objetivo é melhorar a eficiência da compressão. A abordagem dos autores é uma extensão aplicada ao padrão HEVC (41,89) na qual é proposto um método de predição de cores que explora a correlação existente entre os componentes de baixa frequência de luminância e de cor. Com o aumento da eficácia na predição, o resíduo a ser codificado torna-se menor, resultando em melhores taxas de compressão. Desse modo, no que concerne à codificação de cores, o trabalho mencionado não considera a frequência de ocorrência das cores no cálculo de predição, o que resulta, em última instância, em códigos de cores de tamanho variável. Além disso o método citado não trata do processo completo de codificação, mas apenas de uma pequena parte associada à predição ou estimação de pixels de blocos adjacentes. No método proposto neste trabalho os códigos de cores possuem tamanhos fixos, e são calculados com base na frequência de ocorrência observada nas imagens de treinamento.

De maneira semelhante, Li et al. (90) também apresentaram uma técnica direcionada ao padrão HEVC de codificação cuja finalidade é permitir que a codificação de cores seja adaptável em termos do espaço de cores, da taxa de amostragem de cores, e da quantidade de bits usados na representação das cores. Os autores estabelecem a possibilidade de alternar os critérios de codificação de cores mencionados entre quadros de um vídeo, ou mesmo entre blocos de um mesmo quadro. No método proposto neste trabalho os parâmetros de codificação de cores são fixos, e o espaço de cores deve representar a luminância separadamente da cor para favorecer a indexação concomitante à codificação, sem a necessidade de processamento extra para uniformizar a imagem. Além disso, o trabalho relacionado não consiste em um algoritmo de codificação completo como o atual método proposto.

Outros trabalhos relacionados são comparáveis ao método proposto na tarefa de codificação propriamente dita, ou em alguma parte desse processo. Neste sentido, Zhang et al. (91) propuseram a aplicação da DVS com finalidade específica de eliminação do efeito de bloco, não indicando seu uso para codificação ou indexação do conteúdo. Ainda, o documento não faz distinção a respeito dos componentes de cor, nem menciona o modelo de cores adotado. O enfoque do trabalho citado é no método de cálculo da DVS com nível de precisão apenas o suficiente para os propósitos do método, porém os autores não especificam o valor. Contrariamente ao trabalho citado, neste trabalho não foi estabelecida restrição em relação ao método de cálculo da DVS, porém sabe-se que métodos numéricos iterativos, tais como o

método de Jacobi, são mais apropriados que métodos algébricos exatos por tornarem a DVS de grandes matrizes um problema tratável. Apesar de serem aproximados, os métodos numéricos iterativos apresentam aproximações adequadas aos propósitos do método (65).

Sob esse escopo, no método apresentado por Cao et al. (92), também focado em codificação, a escolha dos coeficientes de transformação pela DVS implica em até três repetições do cálculo algébrico por bloco de predição. Conseqüentemente, as repetições resultam em processamento extra e maior tempo de execução. Além disso, a patente indica que os valores codificados são a diferença entre os autovetores do resíduo e os autovetores do bloco predito. Os valores dos autovetores singulares são números reais que variam entre -1 e 1, e por isso os valores residuais podem sofrer perdas de acurácia em consequência dos arredondamentos decorrentes da representação binária compactada, caso o processo de quantização não leve esses aspectos em consideração (129).

O método da patente US 8964835 (87) não substitui o método de codificação convencional, mas o estende com objetivo de obter melhores taxas de compressão utilizando técnicas de visão computacional. Por aplicar técnicas de extração de características, o método é empregado em quadros completos, o que pode levar a atrasos que prejudicam sua utilização em aplicações de tempo real.

Beecks et al. (130) corroboram com essa afirmação ao relatar tempos de processamento variando desde segundos até minutos para o cálculo de assinaturas de imagens utilizadas em tarefas de recuperação por conteúdo. Embora sejam aplicações diferentes da patente US 8964835 (87), o uso de técnicas de visão computacional pressupõe a extração de características e a identificação de padrões nas imagens, sendo, por isso, comparável com os experimentos relatados por Beecks et al.

Também em Porikli (83) o método descrito possui complexidade de tempo e custo computacional alto, pois aplica a DVS para obter duas matrizes de dimensões menores que a matriz original, e posteriormente emprega a transformada de Fourier com o objetivo de alcançar taxas de compressão maiores. A complexidade equivalente a esses passos é da ordem de  $N^3 + 2N^2 \log_2 N$  conforme reportado por Andrews e Patterson (72). A patente mencionada prevê ainda a possibilidade de realizar o pré-processamento da imagem utilizando a técnica PCA, o que elevaria a complexidade do método para ordem de  $2N^3 + 2N^2 \log_2 N$ . Tais características de

complexidade podem tornar o método inapropriado para aplicações de tempo real, particularmente nos cenários em que os dispositivos decodificadores possuem menor poder computacional. Nesse sentido, o método proposto neste trabalho executa a DVS apenas uma vez, e da própria matriz decomposta extrai as características para a indexação das imagens. Desse modo, a complexidade de tempo para ambas as tarefas de codificação e indexação em conjunto é da ordem de  $N^3$ .

Em Rajagopalan et al. (85), o método descrito baseou-se em um paradigma de compressão de vídeo que desloca a maior carga de processamento do emissor para o receptor (131), por isso os autores especificaram um componente servidor intermediário a fim de absorver parte da carga de trabalho e reduzir o custo computacional da decodificação. Porém, ao fazê-lo, adicionaram ao processo um ponto suscetível a falhas representado pelo servidor, do qual depende o funcionamento do método. Além disso, a tecnologia descrita pelos autores propôs uma codificação híbrida que não favorece a indexação direta a partir das imagens decompostas, pois a DVS foi aplicada na codificação de blocos residuais de quadros não-chave. Este aspecto ocorreu devido ao pressuposto de que as matrizes que correspondem a esses blocos possuem *rank* baixo e, conseqüentemente, geram menos valores e vetores singulares significativos na decomposição (64). Esta abordagem pode ser vantajosa para a compressão, porém prejudica a indexação, pois seria necessário reconstruir os quadros chaves para então poder reconstruir quadros não chave, e então proceder à extração de características, resultando em processamentos extras (95,132,133).

A técnica aplicada por Yang et al. (86) de aprendizado de dicionário K-SVD obteve bom desempenho em termos de qualidade da reconstrução (PSNR entre 24,31 dB e 38,37 dB, e SSIM entre 0,92 e 0,99) conforme relatado pelos autores, porém é susceptível a um treinamento inadequado, e sua aplicação à codificação de vídeo em tempo real é limitada (134,135). Ainda, o paradigma de amostragem das imagens utilizado no documento relacionado difere do paradigma convencional. Por aplicar sub-amostragem na origem para obter maior compressão, o método estabelece um limite *a priori* de qualidade da imagem reconstruída. Desse modo, o método descrito não permite aumentar a qualidade das imagens sem que seja modificado, e por isso possui aplicabilidade restrita.

O método descrito em Gu et al. (77) embora tenha conseguido, com qualidade superior, taxa de compressão compatível com outros codificadores

baseados na DCT, possui custo computacional mais alto em comparação com a DCT (72). Além disso, devido à necessidade de agrupar quadros para o cálculo dos vetores singulares do grupo, o tempo de processamento e transmissão pode ultrapassar o limite máximo de 150 ms de atraso desejável para aplicações de vídeo telefonia (29). De fato, os resultados reportados pelos autores foram obtidos considerando grupos de 24 quadros, o que em muitos casos corresponde a um segundo de vídeo. Portanto, aplicações de tempo real, ou que envolvam conversação por vídeo ficam impossibilitadas, diferindo do método inédito proposto. Em particular, a DVS é mais vantajosa que a DCT no processo de decodificação, pois conforme reportado por Garguir (80) a reconstrução pela DVS pode ser até 50 vezes mais rápida que a DCT. Tal aspecto pode favorecer aplicações nas quais os dispositivos decodificadores apresentem menor poder computacional.

Também inadequado para a codificação de vídeos em tempo real é o método proposto por Ashin et al. (82), pois requer a aplicação de dois processos algébricos diferentes (DWT e DVS), o que torna o método mais complexo e de elevado custo computacional. Além disso, em comparação com o codificador baseado na DWT, o método proposto teve desempenho ligeiramente inferior, aproximadamente 1% em quatro de seis imagens avaliadas. Portanto sua aplicação seria justificável apenas nos casos envolvendo altas taxas de compressão, pois os outros métodos apresentaram qualidade de imagem inferior, ou para a codificação de impressões digitais.

O método proposto por Koshi et al. (78) possui diferenças relevantes em relação ao critério de seleção da quantidade de valores singulares e autovetores que são aproveitados, e em relação à quantização e à codificação dos valores utilizados neste trabalho. Em relação ao critério de parada, a patente em questão utiliza um parâmetro de qualidade indireto, ou seja, a relação da energia acumulada nos autovalores com a energia total do bloco, diferentemente deste trabalho de tese apresentado, que utilizou um parâmetro direto obtido a partir do cálculo da qualidade alcançável da reconstrução.

Além disso, no documento relacionado, Koshi et al. (78) empregam estratégia de quantização em dois níveis, e a técnica DPCM na codificação. A técnica DPCM pode favorecer a compressão, permitindo que o método alcance taxas mais elevadas, porém requer processamento adicional em aplicações de indexação e recuperação, pois a obtenção dos valores originais da DVS somente é possível após

a decodificação. Ainda, os autores não reportaram as taxas de compressão obtidas com a aplicação do método. No método proposto no trabalho de tese, a quantização e a codificação são simples e diretas, pois visam favorecer a busca e a comparação com imagens de referência sem a necessidade de decodificação. Ademais, priorizou-se a pré-codificação da informação de cores, e com esse processo ocorreu a produção de compressão preliminar em até 33% em relação à imagem original, e uma taxa de compressão de 82,07% no desempenho geral.

## **6.2. Indexação e recuperação**

A indexação de imagens e vídeos torna-se mais importante à medida que grandes bases de dados de imagens e vídeos são formadas. Em particular, a recuperação de imagens e vídeos de procedimentos médicos por semelhança pode ser útil na formação e treinamento de profissionais, no apoio ao diagnóstico, e na revisão de procedimentos realizados (100,101,113,124,126).

Nesse cenário, a busca de anterioridade e a revisão de literatura relacionada à indexação e recuperação objetivaram identificar trabalhos semelhantes em termos das técnicas utilizadas neste trabalho. Por outro lado, de modo análogo à codificação foram também utilizadas duas expressões de busca complementares, sendo uma ampla e outra específica, devido à característica de originalidade do método. Pela busca ampla foram identificados trabalhos assemelhados em termos dos objetivos e ou técnicas utilizadas. Por outro lado, na busca específica não foram identificados trabalhos com grau de semelhança alto o suficiente para caracterizar a ausência de novidade no método proposto. Os resultados das buscas relacionadas à indexação e recuperação estão resumidos nas Tabelas 6 e 7.

Nos trabalhos mais semelhantes do ponto de vista das técnicas aplicadas, a DVS foi utilizada para reduzir a dimensionalidade de vetores de características e, conseqüentemente, agilizar o processo de comparação entre vetores necessário à recuperação.

Na patente US 6574378 (105) a DVS é aplicada às assinaturas relativas à distribuição espacial das palavras-chaves visuais na obtenção de uma descrição codificada e de menores dimensões, e como consequência tornar a comparação mais rápida. A técnica apresentada indexa e recupera imagens com base em características extraídas da imagem, atuando num nível de abstração intermediário entre os dados brutos e o conteúdo semântico (45), enquanto no método proposto

neste trabalho, este atua sobre os dados brutos, com menor perspectiva de perdas de dados reais.

Em outro método, no qual é utilizada a técnica de espectroscopia para análise de tecidos biológicos (104), o uso de DVS foi mencionado como técnica de análise estatística alternativa à PCA para descorrelacionar grandes volumes de dados. Como resultado da PCA obtém-se também representações compactas das imagens, no entanto a transformação reversa para reconstrução dos dados originais não é possível pelo método apresentado. Nesse sentido, embora o método não tenha como objetivo indexar e recuperar imagens por conteúdo, a aplicação da PCA à indexação e recuperação é factível, porém inviabiliza a codificação pelo mesmo processo matemático.

Também na patente CA 2155901 (103) os autores indicaram a fatoração das imagens segmentadas para obter uma representação compactada, e ao mesmo tempo permitir que fossem reconstruídas. Entretanto há passos não detalhados que dificultam a avaliação do processo de análise realizado. Ainda, o documento não menciona a análise ou recuperação de vídeos, e é restrito à análise com extração de características de rostos de pessoas para fins de identificação, não sendo, portanto, um método genérico.

A patente CN 106651930 proposta por Zeng e He (116) também utiliza a DVS em seu processo com objetivo de reduzir a dimensionalidade de matrizes geradas pelo próprio método, porém não prevê a indexação das imagens e a recuperação com base nos índices gerados. Nessa patente a redução no tamanho das matrizes tem por objetivo permitir o alinhamento de imagens médicas e a determinação de cores a serem aplicadas às regiões anatômicas delineadas pelo método.

Com esses processos posicionados, no método de indexação proposto neste trabalho, a DVS tem a função de extrair das imagens as características mais relevantes, o que é obtido após a ordenação dos autovetores de acordo com a magnitude dos valores singulares correspondentes (64). A redução no tamanho dos vetores é promovida pela aplicação da média quadrática, conforme detalhado no Passo I.3 da etapa de codificação e indexação do método (Seção 4.4.1).

Na patente CN 107330950A, Liu et al. (115) propuseram melhorar a qualidade da reconstrução de imagens de ressonância magnética por meio do ajuste dos coeficientes dos valores singulares de blocos da imagem com base em blocos

similares. No entanto, os parâmetros utilizados no processo não foram detalhados. Por isso, não foi possível avaliar se o método de fato aperfeiçoa a imagem em termos visuais e ou diagnósticos, ou se apenas uniformiza a qualidade. Além disso, a fatoração empregada neste documento não é aplicada à indexação ou recuperação de imagens.

O método proposto na patente JP5734460B (114) permite correlacionar objetos presentes em imagens diferentes mesmo na presença de alterações de escala e perspectiva, sendo uma vantagem do método. Porém o processamento torna-se computacionalmente custoso, em particular se comparado ao método proposto neste trabalho que não requer processamento extra, além do realizado na codificação, para obter as assinaturas das imagens. Os autores citaram o uso de valores singulares para condicionar os valores de taxa de distância no processo de cálculo da correspondência entre pontos chaves. Contudo, a DVS não é aplicada às imagens brutas como ocorre no presente trabalho.

Em Connah et al. (109) – patente US 8682093, a aplicação da DVS tem a finalidade de permitir que as informações/dados espectrais de contraste de uma imagem sejam “transportados” para a outra, gerando uma nova imagem colorida automaticamente, e sem a remoção de bordas e contornos necessários para a avaliação diagnóstica. A patente descrita não tem por finalidade indexar imagens para permitir a recuperação posterior.

A patente US 7689544 B2 (108) não detalha as técnicas utilizadas para indexar o conteúdo das imagens, apenas indica, de modo superficial, o processamento de documentos textuais associados às imagens. Apesar de o documento mencionar a busca e recuperação com base em conteúdo e extração de características de documentos, o autor não deixou claro se características de imagens foram também extraídas. Ainda, o autor limitou-se a estabelecer que o índice foi gerado a partir de pelo menos duas características extraídas dos arquivos, sem especificar quais. Além disso, as técnicas para buscar e recuperar as imagens a partir dos índices tampouco foram detalhadas.

Também na patente US 7158692 B2 (106) o método não detalha quais técnicas utiliza em cada passo, apenas define um processo que envolve a extração de um conjunto de características de imagens previamente segmentadas, armazenando-as em bancos de dados para consulta posterior. Os métodos de indexação e recuperação são disponibilizados pelo sistema gerenciador de banco de



dados, e não apresentam novidade. Além disso, o documento não cita o uso de fatoração de matrizes por DVS, diferentemente deste trabalho.

No método proposto neste trabalho de tese, após a aplicação do processo algébrico de fatoração pela DVS, as imagens ou quadros são representados de maneira condensada e ordenada, facilitando a seleção/recuperação de imagens semelhantes em conteúdo, ou mesmo a seleção baseada em rótulos previamente atribuídos. A representação condensada alcançada pelo cálculo do RMS reduz o tamanho dos autovetores a um percentual do comprimento original do vetor, por exemplo 5%, que ao serem tomadas em conjunto sintetizam as características mais significativas da imagem.

Cabe destacar que uma vantagem importante do método de decomposição em valores singulares é o fato dos autovetores serem normalizados, ou seja, o vetor possui comprimento igual a 1 (64,65). Como efeito a imagem também sofre uma normalização em relação à luminância, e desse modo duas imagens da mesma cena resultarão em autovetores idênticos ou quase idênticos, mesmo que obtidas com níveis de iluminação diferentes. Para os procedimentos de indexação e recuperação essa característica torna o método mais robusto em relação às diferenças de luminosidade nas imagens e apresenta-se como uma das características originais do método.

Além disso, apesar do custo computacional da DVS ser teoricamente superior ao da DCT (72), o processo de reconstrução das imagens não requer a aplicação da transformada inversa, mas de operações simples de multiplicação e adição de matrizes. Essa característica do método proposto torna a reconstrução menos custosa, e por isso pode favorecer aplicações que tenham como receptores/decodificadores dispositivos com baixo poder computacional (80).

Em relação à recuperação, o nível de semelhança pode ser ajustado, bem como a quantidade de resultados considerados significativos. O método permite ainda o tratamento do quadro por inteiro ou subdividido em blocos de pixels de dimensões variáveis, ou mesmo de apenas parte do quadro, de modo a melhor ajustar-se às aplicações e tipos de conteúdo visual.

A partir dos experimentos realizados, observou-se que a precisão média obtida em relação aos três primeiros quadros foi de 94,85% para o Caso 1, e de 95,19% para o Caso 2. As buscas que tiveram desempenho baixo em termos de precisão foram de imagens sem foco ou com pouca definição, e por isso irrelevantes

do ponto de vista diagnóstico. Se essas imagens forem desconsideradas, a precisão alcança 100% para o conjunto de imagens utilizado.

Apesar do Caso 2 ter obtido desempenho ligeiramente superior em termos de precisão, em termos do tempo de processamento o desempenho alcançado foi significativamente inferior. Portanto, na avaliação geral, o melhor resultado foi obtido com os parâmetros utilizados no Caso 1.c, conforme pode ser observado na Tabela 10.

Ao comparar com outros trabalhos reportados na literatura, constatou-se que o desempenho foi compatível ou superior. Ahmed (51) relata precisão média em relação aos dez primeiros quadros entre 59,7% e 100% para dez bases de dados diferentes. Em Lux (136), a precisão alcançada foi de 58,6%, enquanto em Barra et al. (137), a precisão foi de 23%, apesar das diferenças metodológicas. Por outro lado, Howarth e Ruger (138) analisaram dois conjuntos de características diferentes, tendo obtido 14,3% de precisão com uma combinação de três características padrões de Tamura (139), e 12,2% com características de homogeneidade. Neste trabalho, a precisão média em relação aos 10 primeiros quadros recuperados foi de 77,42% para o Caso 1, e 80,62% para o Caso 2, conforme pode ser observado na Tabela 11.

Roldan-Carlos et al. (100) relataram a aplicação de descritores de características de imagens em conjunto com o software LIRE (136) de indexação e recuperação de imagens, com o objetivo de recuperar vídeos de cirurgias laparoscópicas a partir de “fotos” obtidas durante o procedimento. Os autores utilizaram três abordagens, duas baseadas em características globais das imagens e técnicas de fusão de características, e outra baseada em características locais de áreas de interesse identificadas nas imagens. Ainda, estes calcularam e reportaram a precisão alcançada nos experimentos para apenas os três primeiros resultados. O melhor desempenho obtido utilizou o descritor SIMPLE local (140). Nesse caso, os autores relataram que em 79,8% das consultas o vídeo-fonte da imagem de busca foi o primeiro da lista de resultados, e em 84,6% o vídeo-fonte estava entre os três primeiros resultados. Os autores avaliaram ainda outros tipos de descritores, e o melhor desempenho foi obtido utilizando um descritor local. Diferentemente do trabalho citado, o método proposto neste trabalho utiliza características globais da imagem, e obteve desempenho de 94,85% para os três primeiros quadros recuperados.

Beecks et al. (113) partiram do trabalho de Roldan-Carlos et al. (100) e utilizam o mesmo delineamento experimental e a mesma base de dados, porém propuseram uma forma de caracterização do conteúdo das imagens baseada em assinaturas de características com agrupamento adaptativo, e avaliaram diferentes medidas de distância baseadas em assinaturas. Os autores reportaram um desempenho superior em relação ao trabalho de Roldan-Carlos, indicando a recuperação correta de até 88,2% para o primeiro resultado. As características que formam a assinatura combinaram aspectos globais e locais de cor, posição e textura. Conforme mencionado anteriormente, neste trabalho de tese, o método proposto utiliza características globais da imagem, e obteve desempenho de 94,85% para os três primeiros quadros recuperados.

No cenário de aplicação médica, é frequente a captura de imagens de pontos importantes durante exames ou procedimentos cirúrgicos, de modo que possam compor o laudo ou relatório final e, ainda, em circunstâncias de aperfeiçoamento de conhecimento (141).

Assim sendo, com o método proposto é possível recuperar o ponto exato do vídeo em que a foto foi capturada, e ainda indicar outros trechos de conteúdos semelhantes no próprio ou em outros vídeos, contribuindo assim e de modo positivo, para a finalidade proposta.

## **7. CONCLUSÃO**

Neste trabalho, por meio dos resultados alcançados, conclui-se a efetividade do método original proposto para codificação e indexação automática de vídeos de exames médicos para aplicações de telemedicina.

## 8. GLOSSÁRIO

**Acurácia:** proximidade entre o valor obtido experimentalmente e o valor verdadeiro na medição de uma grandeza física (142). Medida estatística que envolve tanto erros sistemáticos (tendência) como aleatórios (dispersão) (143).

**Alinhamento de imagens:** técnica de processamento de imagens usada para alinhar automaticamente duas ou mais imagens da mesma cena. Também denominado “registro de imagens” (52).

**Analógico:** diz-se de sinal representativo de uma grandeza física cuja variação ocorre de modo contínuo, sem lacunas ou hiatos (142).

**Aprendizado de máquina supervisionado:** tipo de aprendizado de máquina no qual o conjunto de dados utilizado como treinamento inclui a classe. Desse modo, o esquema opera sob “supervisão” considerando que a classificação correta de cada exemplo de treinamento é conhecida (134).

**Aprendizado de máquina não supervisionado:** tipo de aprendizado de máquina no qual o conjunto de dados de treinamento não inclui a classe (134).

**Aprendizagem de dicionário:** técnica aplicada no processamento de sinais que consiste na determinação de vetores representativos, denominados átomos, que são selecionados por meio de vetores esparsos na codificação de sinais. O sinal codificado é, portanto, compactado pois é constituído de vetores esparsos, que representam uma combinação linear de átomos do dicionário (144).

**Aspereza:** aspecto de textura de uma imagem ou região da imagem, caracterizada por baixa uniformidade e alta entropia (52).

**Autovalor:** o autovalor de uma matriz  $A$  é um escalar  $\lambda$  que, multiplicado pelo autovetor associado  $x$ , torna verdadeira a igualdade  $Ax = \lambda x$  (64).

**Autovetor:** o autovetor  $x$  de uma matriz  $A$  é um vetor não nulo que, multiplicado pela matriz  $A$ , mantém sua direção inalterada, e sua magnitude expandida ou reduzida pelo fator  $\lambda$ , denominado autovalor (64). Pode ser entendido como um operador linear que projeta e resume a matriz  $A$  em um único vetor de comprimento  $\lambda$ .

**Bit:** dígito binário. Menor parcela de informação processada por um computador. Algoritmo do sistema binário que somente pode assumir os valores 0 ou 1 (142).

**Bit-rate:** ver “Taxa de bits”.

**BRIEF:** “*Binary robust independent elementary features*” é um descritor de características extraídas de pontos de interesse em imagens que utiliza cadeias de bits para representar os pontos. É calculado utilizando testes simples de diferença de intensidade. Diferentemente de outros descritores que utilizam técnicas para redução da dimensionalidade dos vetores de características, o BRIEF calcula as sequências reduzidas de bits diretamente a partir das imagens (145,99).

**Byte:** conjunto de oito bits adjacentes (142).

**CEDD:** ver “*Color and edge directivity descriptor*”.

**Centroide:** centro geométrico de uma figura plana que corresponde à posição média de todos os pontos da figura. Baricentro (142).

**Cliente:** em sistemas distribuídos, cliente é o processo ou computador que requisita a execução de operações de um servidor interligado em rede (146).

**Codificação de comprimento variável:** técnica de codificação de entropia que consiste em atribuir códigos de menor tamanho aos símbolos mais frequentes, utilizando uma estratégia que garante que nenhum código seja igual ao prefixo de outro código (41,147).

**Codificação de entropia:** processo reversível que permite a compressão adicional por meio da codificação dos elementos sintáticos (coeficientes de transformação, modos de predição, vetores de movimento, etc.) em um arquivo final de saída (41). Entropia, no contexto da teoria da informação, tem relação direta com informação e incerteza (148), ou seja, quanto maior a incerteza, e conseqüentemente maior a surpresa na ocorrência de um símbolo, maior a informação. Neste caso, a entropia é sinônimo de informação. Desse modo, a codificação de entropia leva em consideração a probabilidade de ocorrência dos símbolos de entrada a fim de obter representações compactadas como saída.

**Color and edge directivity descriptor (CEDD):** descritor de características de baixo nível extraídas de imagens, que pode ser usado para indexação e recuperação. O método incorpora informações de textura e cor em um histograma para descrever imagens, sendo o tamanho do descritor limitado a 54 bytes. O atributo mais importante considerado pelos autores é a baixa complexidade computacional do algoritmo para extração das características (110).

**Color layout descriptor:** descritor de características cromáticas de imagens, utilizado em aplicações de indexação e recuperação de vídeos e imagens. O método especifica a distribuição espacial das cores a partir de alguns coeficientes DCT quantizados de modo não linear, obtidos a partir de cores médias, e organizados na forma matricial. O descritor ocupa apenas 63 bits (aproximadamente 8 bytes) por imagem (111).

**Combinação linear:** expressão matemática construída a partir da soma de um conjunto de termos  $(x_1, x_2, \dots, x_n)$ , cada qual multiplicado por uma constante  $(ax_1 + bx_2 + \dots + kx_n)$  (64).

**Compressão:** em Ciência da Computação, compressão de dados consiste em reduzir a representação da informação de modo que ocupe menos espaço de armazenamento, sem perda de conteúdo. A compressão de imagens é a técnica de reduzir a quantidade de dados necessários para representar uma imagem (52).

**Crominância:** refere-se às propriedades da luz que definem sua cor, caracterizada pelo comprimento de onda predominante na combinação de ondas de luz (matiz), e pelo grau de pureza ou mistura com a luz branca (saturação) (53).

**Contraste:** variação nas tonalidades de luz e sombra, claro e escuro, zonas opacas e transparentes, numa imagem ou vídeo (142).

**Decomposição:** ver “Fatoração”.

**Direcionalidade:** propriedade global de uma imagem, ou região da imagem, que indica se o padrão de textura apresenta direção. Envolve aspectos relativos à forma dos elementos e à regra de posicionamento desses elementos, podendo ser direcional ou não direcional (139).

**Domínio de frequência:** domínio de uma função periódica cujos valores de entrada são valores de frequência. Sinais periódicos podem ser representados como um conjunto de todas as frequências que os constituem, sendo expressos como uma soma de senos e cossenos. Quando representados dessa forma, diz-se que o sinal está expresso no domínio de frequência (148).

**Domínio espacial:** domínio de uma função cujos valores de entrada são a posição no espaço. Imagens digitais podem ser interpretadas como funções cujo domínio são as coordenadas nos eixos horizontal (x) e vertical (y). O valor da função em cada ponto corresponde ao valor do pixel, que tomados em conjunto, reconstituem a imagem em sua totalidade. Desse modo, o domínio espacial refere-se ao próprio plano da imagem (53).

**Efeito de bloco:** tipo de distorção perceptível em imagens digitais caracterizado por um efeito quadriculado na imagem, resultante da codificação em blocos. Em decorrência da aplicação de níveis de quantização diferentes em blocos adjacentes de uma mesma imagem, e, conseqüentemente, do nível de qualidade visual, os limites entre um bloco e outro podem tornar-se visíveis, produzindo o efeito visual indesejado da divisão artificial da imagem em blocos (33,149).

**Entropia:** medida da desordem ou da imprevisibilidade da informação (142). Na teoria da informação, a entropia está diretamente relacionada a incerteza e surpresa, ou seja, um valor ou símbolo com alta probabilidade de ocorrência possui baixa incerteza e causa pouca surpresa, por isso porta pouca informação, e conseqüentemente apresenta baixa entropia (148).

**Erro quadrático médio:** é a média das diferenças, elevadas ao quadrado, entre os valores estimados e os valores efetivamente observados. O erro quadrático médio enfatiza o efeito de erros maiores, independentemente de serem positivos ou negativos em relação ao valor estimado, pois os valores são elevados ao quadrado. Desse modo, quando aplicado à análise da qualidade de imagens, quanto menor a métrica, mais a imagem distorcida se aproxima da imagem original (52,53,64).

**Espectroscopia:** conjunto de métodos para análise de substâncias baseados na produção e interpretação de seus espectros de emissão ou absorção de radiações



eletromagnéticas (p.ex., nas regiões do infravermelho, ultravioleta, raios X, visível etc.) (142).

**Fatoração:** a fatoração de matrizes consiste na decomposição da matriz original em um produto de duas ou mais matrizes. Há diferentes tipos de decomposição matricial, com propriedades e finalidades distintas, como a solução de sistemas de equações lineares e a resolução de sistemas de equações diferenciais. As fatorações favorecem a análise e interpretação de características dos sistemas representados na forma matricial (64).

**FAST:** ver “*Features from accelerated segment test*”.

***Features from accelerated segment test (FAST)*:** é um método para detecção de quinas em imagens que utiliza técnicas de aprendizado de máquina, cujo objetivo é permitir a detecção de características em tempo real, de modo que possa ser utilizado no processamento de vídeos sem perdas de quadros. Por esse método, os critérios para detecção são decisões acerca da classe dos pixels, os quais podem pertencer ou não a uma quina. Os critérios são organizados na forma de regras de uma árvore de decisão capaz de classificar corretamente todos as quinas presentes nas imagens de treinamento, sendo posteriormente aplicadas a outras imagens (150).

**Filtro:** dispositivo ou técnica aplicada sobre um sinal de entrada para separar o sinal ou componente desejado de outros sinais modulados associados, com o objetivo de remover ou atenuar componentes e ou características indesejadas, ou ressaltar aspectos de interesse (53,148).

**Filtro gaussiano:** em sistemas de comunicação e ou de processamento de sinais, é um filtro cuja função de resposta ao impulso se aproxima de uma função gaussiana. No processamento de imagens, a aplicação de filtros no domínio espacial é realizada por meio de matrizes denominadas máscaras, as quais são aplicadas sobre cada pixel da imagem, considerando os pixels da vizinhança. No filtro gaussiano, os coeficientes da máscara são derivados dos pixels da vizinhança, a partir de uma função gaussiana bidimensional. Desse modo, o filtro gaussiano produz uma suavização da imagem por meio da substituição de cada pixel por uma média ponderada dos pixels vizinhos, de modo que o peso atribuído a um vizinho decresce gradualmente com a distância do pixel central (53,148).

**Filtro passa-baixa:** filtro que atenua as altas frequências, selecionando as frequências baixas e médias. Na análise e processamento de imagens, altas frequências estão relacionadas com a informação de detalhes, portanto os efeitos visuais de um filtro passa-baixa são a suavização da imagem e a redução de ruídos (53).

**Filtro passa-faixa:** filtro que seleciona um intervalo de frequências do sinal a ser realçado, removendo frequências muito baixas e muito altas. O efeito visual obtido caracteriza-se pelo realce de bordas e limites em consequência da remoção de frequências baixas, e ao mesmo tempo a eliminação de ruídos como consequência da remoção de altas frequências (53).

**Filtro Sobel:** filtro utilizado para ressaltar transições em uma determinada direção (vertical, horizontal, etc.) por meio da ampliação das diferenças entre pontos localizados em lados opostos de uma fronteira. Portanto, este filtro favorece a detecção de bordas em imagens (37).

**Frame:** ver “quadro”.

**Gradiente:** valor indicativo da direção e inclinação de uma reta. O gradiente é dado pela razão entre a variação vertical e a variação horizontal da reta. Quando aplicado ao cálculo vetorial, o gradiente de uma função de múltiplas variáveis é um vetor cujos componentes são as derivadas parciais da função em um ponto (151).

**Gramática livre de contexto:** é uma especificação para a estrutura sintática de uma linguagem. A denominação “livre de contexto” deve-se ao fato da substituição das classes de elementos sintáticos pelos símbolos denominados terminais (palavras reconhecidas pela linguagem, pontuação e outros sinais gráficos) poder ser realizada em qualquer contexto, ou seja, sem depender de qualquer análise quanto às regras de substituição ou símbolos que antecedem ou sucedem a regra em questão (152-154).

**HSV:** modelo de cores alternativo ao RGB que representa separadamente as informações de cor e de luminância. Nesse modelo, as cores são mapeadas em valores de matiz e saturação, sendo o matiz a cor essencial, e a saturação o grau de pureza do matiz (53).

**Indexação:** técnica ou procedimento que conecta um símbolo para um tópico (palavra, número, característica ou imagem) com o material pertinente àquele tópico, com objetivo de tornar efetivo e eficiente o acesso à informação (155).

**Índice reverso:** estrutura de dados que mapeia uma palavra a um conjunto de documentos que contém a palavra, ou um item de busca mínimo a um conjunto de unidades indexadas que contém o item. Semelhante a um índice remissivo, porém com referência a múltiplos documentos por palavra ou item de busca. A maioria dos métodos e ferramentas de busca e recuperação de informações baseadas em texto livre são implementadas utilizando índices reversos (156).

**Largura de banda:** é uma propriedade física do meio de transmissão, medida em hertz (Hz), que consiste na faixa de frequências que passam pelo meio com atenuação mínima. Essa característica do meio limita a quantidade de dados que o sinal aceito pelo canal pode transportar. Desse modo, em comunicação de dados, a largura de banda do meio físico está diretamente relacionada à taxa de dados que o canal é capaz de transmitir. Por esse motivo, a taxa de bits suportada pelo canal é frequentemente (e equivocadamente) usada para indicar a largura de banda (157).

**Luminância:** relação entre a intensidade luminosa de uma superfície e a área da superfície projetada na direção de observação. A luminância corresponde ao brilho medido em um ponto da superfície, e é medida em candelas/m<sup>2</sup> (53).

**Matriz de co-ocorrência:** representação, na forma de matriz, das relações espaciais entre pares de valores de pixels em uma imagem digital. As matrizes de co-ocorrência descrevem com que frequência pares de pixels com determinados níveis de cinza ou de cores, separados por uma distância específica e alinhados em certa direção, ocorrem na imagem. A partir da matriz de co-ocorrência são calculadas medidas estatísticas que descrevem a textura da imagem (53).

**Matriz de Toeplitz:** é uma matriz cujos valores das diagonais descendentes da esquerda para a direita são iguais, ou seja, uma matriz  $T_n$  tal que  $t_{k,j} = t_{k-j}$ . A matriz de Toeplitz é utilizada na análise de sinais porque a multiplicação por um polinômio trigonométrico, condensado para um espaço finito, pode ser representado utilizando essa matriz (158).

**Média quadrática:** a média quadrática é dada pela raiz quadrada da média aritmética dos quadrados dos valores (142).

***Maximally stable extremal regions (MSER):*** trata-se de um método para detecção de objetos binários em imagens que tem por objetivo estabelecer correspondências entre pares de imagens cujas cenas foram obtidas de perspectivas diferentes. O método propõe o conceito de regiões extremas, que são regiões que possuem propriedades desejáveis para o objetivo do método. Estas propriedades dizem respeito ao conjunto de dados que compõe essas regiões que é fechado sob duas condições: (a) transformações contínuas e, portanto, projetivas de coordenadas, e (b) transformações monotônicas dos níveis de intensidade da imagem (variações da iluminação). Desse modo, o método se torna independente de escala e rotação. As regiões são detectadas usando uma análise de conectividade e calculando regiões conectadas de intensidade máxima e mínima (159,99).

**MSER:** ver “*Maximally stable external regions*”.

**Norma de Frobenius:** norma de matrizes dada pela raiz quadrada da soma dos quadrados dos elementos da matriz. A norma é considerada geometricamente como o comprimento de um vetor. Neste caso, a norma da matriz é uma norma de vetor no espaço vetorial cujos elementos são matrizes (65).

**ORB:** ver “*Oriented FAST and rotated BRIEF*”.

***Oriented FAST and rotated BRIEF (ORB):*** detector e descritor de características de imagens proposto como alternativa mais rápida ao SURF. O método prevê a adição de um componente de orientação rápido e preciso ao FAST, o cálculo eficiente de características BRIEF orientadas, a análise de variância e correlação das características BRIEF orientadas, e um método de aprendizagem para descorrelacionar as características BRIEF sob rotação da imagem (160,99).

**PHOG:** ver “*Pyramid of histograms of orientation gradients*”.

**Pixel:** cada ponto de uma imagem digital. Acrônimo do inglês *Picture Element* (53).

***Pyramid of histograms of orientation gradients:*** descritor de características de imagens cujo objetivo é representar o conteúdo da imagem por suas características

de forma e layout espacial. A forma localizada é capturada pela distribuição sobre as orientações de bordas dentro de uma região, enquanto o layout espacial é capturado pela subdivisão da imagem em blocos em múltiplas resoluções. O descritor consiste em um histograma dos gradientes de orientação em cada sub-região da imagem, em cada nível de resolução, sendo, por isso, denominado pirâmide de histogramas de gradientes de orientação (112).

**Precisão:** grau de consistência da grandeza medida com sua média. Medida estatística que envolve apenas erros aleatórios (143).

**Predição intra-quadro:** técnica utilizada na codificação de vídeos pela qual os pixels de um quadro são estimados a partir de pixels adjacentes do mesmo quadro, sendo codificadas a posição da referência e a diferença entre os pixels estimados e os observados. Esta técnica explora a correlação espacial existente entre blocos de pixels adjacentes, contribuindo para a obtenção de taxas de compressão maiores (33).

**Profundidade de bits:** corresponde ao número de bits necessários para representar todos os valores possíveis de cada pixel. Por exemplo, no caso de uma imagem monocromática, com até 256 tons de cinza, a profundidade de bits é dada por  $\log_2 256 = 8$ . De modo análogo, para imagens coloridas em formato RGB, a profundidade de bits equivale a  $3 \times \log_2 256 = 24$  (53).

**Quadro:** cena completa obtida em um instante de tempo. O conjunto de quadros agrupados em sequência constituem o vídeo digital (31).

**Quadro-chave:** em algoritmos de compressão de vídeo que aplicam predição de movimento entre quadros, é o quadro utilizado como referência. Quadros-chaves são codificados por completo sem dependência em relação a outros quadros (31,33).

**Quantização:** procedimento que visa enquadrar os valores de uma função a um conjunto finito e predeterminado de valores discretos. Portanto, a quantização mapeia um valor de entrada a um valor de saída com objetivo de obter uma representação discreta para um sinal contínuo, ou para reduzir a faixa dinâmica de valores de um sinal já digitalizado (31,33).

**Rank:** quantidade de colunas ou de linhas linearmente independentes de uma matriz. Mesmo que posto (64,65).

**Rede neural:** rede neural artificial é um método de aprendizagem de máquina inspirado no modelo biológico, cujo sistema de aprendizado é construído sobre complexas redes de neurônios artificiais interconectados (161).

**Redundância espacial:** correlação existente entre pixels ou blocos de pixels adjacentes de um mesmo quadro ou imagem (33).

**Redundância temporal:** correlação existente entre quadros, temporalmente adjacentes, de uma sequência de vídeo (33).

**Resolução:** dimensões espaciais de uma imagem ou quadro, indicado pelo número de colunas e linhas da imagem (33).

**Scale invariant feature transform (SIFT):** detector e descritor de características de imagens independente de escala e rotação. Pelo método de detecção, pontos de interesse são extraídos das imagens em dois passos. No primeiro, a imagem de entrada é repetidamente suavizada utilizando filtros gaussianos, e dela são extraídas sub-imagens em escalas menores. Desse modo uma pirâmide de imagens é construída tendo a imagem de entrada posicionada na base. No segundo passo, os pontos de interesse são descobertos considerando uma vizinhança de tamanho igual a três pixels (vizinhança  $3 \times 3 \times 3$ ). Os pontos são obtidos onde a diferença dos valores gaussianos estiverem dentro de um limite máximo, tanto no domínio espacial quanto no nível da escala da pirâmide. Assim, os pontos de interesse extraídos independem da escala e de rotação. O método prevê ainda um descritor para os pontos de interesse que se baseia no histograma das direções dos gradientes da imagem local em torno dos pontos de interesse (99).

**Segmentação:** é a subdivisão da imagem em suas regiões ou objetos constitutivos, podendo ser realizada manual ou automaticamente. No caso de segmentação automática, o nível de detalhes que orienta a subdivisão depende do tipo de problema a ser resolvido (52).

**Sensoriamento comprimido:** técnica de aquisição de dados de modo comprimido que tem por objetivo obter diretamente apenas a porção dos dados relevantes para a reconstrução do sinal ou da informação original. Considerando que no processo de compressão de sinais, parte dos dados é descartada sem perda perceptível de

informação, a abordagem de sensoriamento comprimido busca adquirir diretamente a parte dos dados que não é descartada pelos processos de compressão (162).

**Servidor:** em sistemas distribuídos, servidor refere-se a um programa em execução (um processo) em um computador interligado em rede, que aceita pedidos de programas em execução em outros computadores para efetuar um serviço e responder apropriadamente (146).

**SIFT:** ver “*Scale invariant feature transform*”.

**SIMPLE:** acrônimo para “*searching images with MPEG-7-like powered localized descriptors*”, é um descritor de características de imagens que combina o detector de pontos locais do SURF para identificação de pontos de interesse, e três descritores globais do padrão MPEG-7, juntamente com o descritor global CEDD. Os descritores obtidos são combinados em um esquema denominado *Bag-of-Visual-Words*, que busca descrever o conteúdo de imagens como “palavras visuais”, sendo utilizado em aplicações de recuperação de imagens baseada em conteúdo (140).

**Sinal:** é uma sequência ordenada de números que descrevem as tendências e variações de uma grandeza física, medida ao longo do tempo e ou do espaço (37). Um sinal pode ser representado como uma função de uma ou mais variáveis independentes (53).

**Speeded up robust features (SURF):** algoritmo para detecção e descrição de características de imagens independente de escala e rotação. O método é similar ao SIFT, porém baseado em matriz hessiana gerada a partir da convolução da derivada gaussiana de segunda ordem. Os pontos de interesse são extraídos da mesma maneira que o SIFT. O detector SURF é encontrado em cada ponto de interesse por meio de atribuição de orientação e da análise do componente descritor. A orientação do ponto é calculada pela resposta *wavelet* de Haar nas direções x e y, na vizinhança circular de cada ponto de interesse, e a orientação dominante é encontrada a partir da soma das orientações. As respostas *wavelet* na região quadrada, direcionada na orientação dominante produz os descritores SURF (163,99).

**SURF:** ver “*Speeded up robust features*”.

**Taxa de bits:** em comunicação de dados, é a quantidade de informação transmitida por um canal por unidade de tempo, sendo igual ao número de símbolos transmitidos por segundo multiplicado pelo número de bits necessários para representar cada símbolo. No caso de vídeos digitais, a taxa de bits corresponde à quantidade de bits usados para representar um segundo de vídeo (33,157).

**Taxa de compressão:** é a razão entre o tamanho original da imagem ou vídeo e o tamanho alcançado após compressão (33).

**Transformada:** transformada ou Transformação é uma função ( $f$ ) que relaciona um espaço vetorial de domínio ( $V$ ) a um espaço vetorial de contradomínio ( $W$ ), ou seja,  $f$  é uma transformação de  $V$  em  $W$  (64).

**Variable length coding:** ver “codificação de comprimento variável”.

**Wavelet:** transformação algébrica baseada em pequenas ondas de frequências variáveis e duração limitada. Diferente da transformada de Fourier, na qual a informação relativa à duração é perdida no processo de transformação, a transformada *wavelet* provê informações sobre frequência e informações espaciais da imagem ou quadro de entrada (33,52).

**XYZ:** modelo de cores teórico que utiliza três valores para representar todas as cores do espectro visível. Cada valor, identificado individualmente como X, Y ou Z, representa uma cor primária. Nesse modelo, a variação das cores, percebidas visualmente, é resultado da combinação aditiva das cores primárias em quantidades diferentes (52,53).

**YCbCr:** modelo de cores mais utilizado em vídeos digitais no qual as informações de luminância e de cor são representados separadamente. Nesse modelo, o componente Y representa a luminância, enquanto o valor de Cb representa a diferença entre o componente azul da cor e um valor de referência, e o valor de Cr representa a diferença entre a cor vermelha e um valor de referência (31,53).

**YUV:** modelo de cores utilizado em vídeo analógico dos padrões PAL (acrônimo para *Phase Alternation by Line*) e SECAM (acrônimo para *Séquentiel Couleur à Mémoire*). De modo similar ao modelo YCbCr, o componente Y representa a luminância, e os valores de U e V correspondem aos componentes de cor dos pixels (31,53).



## 9. REFERÊNCIAS

- [1] American Cancer Society. Cancer Facts and Figures 2018. Atlanta, EUA: American Cancer Society; 2018.
- [2] Quilici FA, Oliveira LAR. Câncer colorretal. In: Quilici FA, Grecco EC, editors. Colonoscopia. São Paulo: Lemos Editorial;2000. p.121-32.
- [3] Instituto Nacional de Câncer José Alencar Gomes da Silva. Estimativa 2018: Incidência de câncer no Brasil. Rio de Janeiro, RJ: Coordenação de Prevenção e Vigilância; 2017.
- [4] Instituto Nacional de Câncer José Alencar Gomes da Silva. Estimativa 2014: Incidência de câncer no Brasil. Rio de Janeiro, RJ: Coordenação de Prevenção e Vigilância; 2014.
- [5] Instituto Nacional de Câncer José Alencar Gomes da Silva. Estimativa 2016: Incidência de câncer no Brasil. Rio de Janeiro, RJ: Coordenação de Prevenção e Vigilância; 2015.
- [6] Instituto Nacional de Câncer José Alencar Gomes da Silva. Estimativa 2012: Incidência de câncer no Brasil. Rio de Janeiro, RJ: Coordenação Geral de Ações Estratégicas, Coordenação de Prevenção e Vigilância; 2011.
- [7] Mendonça R, Valadão M, Couto AC, Koifman S. Tendência de mortalidade por câncer colorretal em cinco capitais brasileiras de 1980 a 2009. Caderno de Saúde Coletiva. 2012;20(3):329-35.
- [8] Assis RVBF. Rastreamento e vigilância do câncer colorretal: guidelines mundiais. GED Gastroenterologia Endoscopia Digestiva. 2011;30(2):62-74.
- [9] Kudo SE, Lambert R, Allen JI, Fujii H, Fujii T, Kashida H, et al. Nonpolypoid neoplastic lesions of the colorectal mucosa. Gastrointestinal Endoscopy. 2008;68(4):S3-S47.
- [10] Quilici FA, Grecco EC. Colonoscopia. São Paulo: Lemos Editorial; 2000.
- [11] Ratto C, Sofo L, Ippoliti M, Merico M, Doglietto GB, Crucitti F. Prognostic factors in colorectal cancer. Diseases of the Colon & Rectum. 1998;41(8):1033-49.
- [12] Huwe FG, Anton EM, Eisenhardt MF, Foletto E, Jackisch S, Severo B, et al. Avaliação das características clínicas e epidemiológicas e sobrevida global de

- pacientes portadores de câncer colorretal. *Revista de Epidemiologia e Controle de Infecção*. 2013;3(4):112-6.
- [13] Siegel R, DeSantis C, Jemal A. Colorectal cancer statistics, 2014. *CA A Cancer Journal for Clinicians*. 2014;64(2):104-17.
- [14] Edmonson JM. History of the instruments for gastrointestinal endoscopy. *Gastrointestinal Endoscopy*. 1991;37(Supp.2):S27-S56.
- [15] Staut Junior JE, Staut GCC. Equipamento e instrumental. In: Quilici FA, Grecco EC, editors. *Colonoscopia*. São Paulo: Lemos Editorial; 2014. p. 27-30.
- [16] Sivak MV, Fleischer DE. Colonoscopy with a VideoEndoscope: preliminary experience. *Gastrointestinal Endoscopy*. 1984;30(1):1-5.
- [17] Sivak M. Video endoscopy. *Clinics in gastroenterology*. 1986 April;15(2):205-34.
- [18] Demling L, Hagel HJ. Video Endoscopy: fundamentals and problems. *Endoscopy*. 1985;17:167-9.
- [19] Rede Nacional de Pesquisa. Relatório de atividades 1997/1998. Ministério da Ciência e Tecnologia; 1998 [Acesso em 12 de março de 2019]. Disponível em: <https://memoria.rnp.br/arquivo/documentos/rel-rnp98.pdf>.
- [20] Wolff WI. The JPEG still picture compression standard. *Communications of the ACM*. 1989;84(9):1017-25.
- [21] Starren JB, Nesbitt TS, Chiang MF. Telehealth. In: Shortliffe EH, Cimino JJ, editors. *Biomedical Informatics: computer applications in health care and biomedicine*. 4th ed. London, United Kingdom: Springer; 2014. p. 541-60.
- [22] Pereira BMT, Calderan TRA, Silva MTN, Silva AC, Junior ACM, Fraga GP. Initial experience at a university teaching hospital from using telemedicine to promote education through video conferencing. *São Paulo Medical Journal*. 2012;130(1):32-6.
- [23] Machado RB. Método computacional para acompanhamento e interação remota em tempo real para videocolonoscopia [tese]. Universidade Estadual de Campinas, Faculdade de Ciências Médicas. Campinas, SP; 2013.

- [24] Ribeiro Filho JL, Messina LA, de Lima Lopes PR, editors. RUTE100: as 100 Primeiras Unidades de Telemedicina no Brasil e o Impacto da Rede Universitária de Telemedicina (RUTE). Rio de Janeiro: e-Papers; 2014.
- [25] Hoyt RE, Fleury R. Telemedicine. In: Hoyt RE, Yoshihashi A, editors. Health Informatics - Practical Guide for Healthcare and Information Technology Professionals. 6ed. USA:Informatics Education; 2014. p.416-39.
- [26] Conselho Federal de Medicina. Resolução nº 2.227/2018. Disciplina a telemedicina como forma de prestação de serviços médicos mediados por tecnologia; 2019. Publicada no D.O.U de 6 de fevereiro de 2019, Seção I, p.58.
- [27] Conselho Federal de Medicina. Resolução nº 2.228/2019. Revoga a Resolução CFM nº 2.227, publicada no D.O.U. de 6 de fevereiro de 2019, Seção I, p.58, a qual define e disciplina a telemedicina como forma de prestação de serviços médicos mediados por tecnologia, e restabelece expressamente a vigência da Resolução CFM nº 1.643/2002, publicada no D.O.U. de 26 de agosto de 2002, Seção I, p.205; 2019. Publicada no D.O.U de 6 de março de 2019, Seção I, p.91.
- [28] Wallace GK. The JPEG still picture compression standard. Communications of the ACM. 1991;34(4):33-44.
- [29] Gall DJ. The MPEG video compression algorithm. Signal Processing Image Communications. 1992;4(2):129-40.
- [30] Bergwall B. MPEG technology: bringing video up to speed. Tech Directions. 1995;55(4):33-6.
- [31] June F. An introduction to video compression in C/C++. USA: CreateSpace; 2010.
- [32] Turaga D, Chen T. Fundamentals of video compression: H.263 as an example. In: Sun MT, Reibman AR, editors. Compressed Video over Networks. USA: Marcel Dekker; 2001. p. 3-34.
- [33] Bull DR. Communicating picture: a course in image and video coding. USA: Elsevier; 2014.
- [34] Richardson IEG. H.264 and MPEG-4 video compression. Wiltshire, Great Britain: Jon Wiley & Sons; 2003.

- [35] Pianykh OS. Digital Imaging and Communications in Medicine (DICOM): a practical introduction and survival guide. Boston, MA: Springer; 2008.
- [36] Mustra M, Delac K, Grgic M. Overview of the DICOM standard. In: 2008 50th International Symposium ELMAR. Vol. 1. Zadar, Croatia; 2008. p.39-44.
- [37] Najarian K, Splinter R. Biomedical Signal and Image Processing. Boca Raton, FL: CRC Press; 2006.
- [38] Chand D, Lakha B, Kumar H. Study of JPEG compression with different DCT methods. In: Fifth Innovative Conference on Embedded Systems, Mobile Communication and Computing ICEMC2. Goa; 2010.
- [39] Mukhopadhyay A, Raj A, Shaji RP. LRJPEG: A Luminance Reduction based Modification for JPEG Algorithm to Improve Medical Image Compression. In: 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI); 2018. p. 617-23.
- [40] Sikora T. MPEG digital video-coding standards. IEEE Signal Processing Magazine. 1997;14(5):82-100.
- [41] Bing B. Next-generation video coding and streaming. New Jersey, USA: Jon Wiley & Sons; 2015.
- [42] Nouri N, Abraham D, Moureaux JM, Dufaut M, Hubert J, Perez M. Subjective MPEG2 compressed video quality assessment: application to telesurgery. In: 2010 IEEE International Symposium on Biomedical Imaging: From Nano to Macro. Rotterdam, The Netherlands; 2010. p. 764-7.
- [43] Paul M, Sorwar G. Encoding and decoding techniques for medical video signal transmission and viewing. In: 6th IEEE/ACIS International Conference on Computer and Information Science (ICIS 2007). Melbourne, Australia; 2007. p. 750-6.
- [44] Kumar A, Kim J, Cai W, Fulham M, Feng D. Content-based medical image retrieval: a survey of applications to multidimensional and multimodality data. Journal of Digital Imaging. 2013;26:1025-39.
- [45] Priya R, Shanmugam TN. A comprehensive review of significant researches on content based indexing and retrieval of visual information. Frontiers of Computer Science. 2013;7(5):782-99.

- [46] Zhou W, Li H, Tian Q. Recent Advance in Content-based Image Retrieval: A Literature Survey. CoRR. 2017;arXiv/1706.06064.
- [47] Oussalah M. Content based image retrieval: review of state of art and future directions. In: 2018 First Workshops on Image Processing Theory, Tools and Applications. Sousse, Tunisia: IEEE; 2008.
- [48] Gupta A, Agarwal D, Veenu. Performance Analysis of Content Based Image Retrieval Systems. 2018 International Conference on Computing, Power and Communication Technologies (GUCON). Greater Noida, UP, India: IEEE; 2018. p. 899-902. Sep 28-29.
- [49] Kulkarni SP, Goudar RH. Design and Analysis of Content Based Video Retrieval System. 2017 International Conference On Smart Technology for Smart Nation. Bangalore, India: IEEE; 2017. p. 421-4. Aug. 17-19.
- [50] Mahdi M, Ahmad AR, Ismail R. Similarity Search Techniques in Exploratory Search: A Review. In: Proceedings of TENCON 2018 – 2018 IEEE Region 10 Conference. Jeju, Korea: IEEE; 2018. p. 2193-8.
- [51] Ahmed KT, Ummesafi S, Iqbal A. Content based image retrieval using image features information fusion. Information Fusion. 2019;51:76-99.
- [52] Gonzalez RC, Woods RE. Digital Image Processing. 3rd ed. Upper Saddle River, NJ: Pearson; 2008.
- [53] Pedrini H, Schwartz WR. Análise de imagens digitais: princípios, algoritmos e aplicações. São Paulo: Thomson Learning; 2008.
- [54] Poon TC, Liu JP. Fundamentals of optics. In: Cristóbal G, Schelkens P, Thienpont H, editors. Optical and digital image processing - fundamentals and applications. Weinheim: Wiley-VCH; 2011. p. 1-24.
- [55] Guyton AC, Hall JE. Guyton & Hall - Tratado de fisiologia médica. 13ed. Rio de Janeiro, RJ: Elsevier; 2017.
- [56] LMS Color Space [2019]. [Acesso em 18 de maio de 2019]. Disponível em: [https://en.wikipedia.org/wiki/LMS\\_color\\_space](https://en.wikipedia.org/wiki/LMS_color_space).

- [57] Abraham C. A beginner's guide to (CIE) colorimetry; 2016. [Acesso em 7 de outubro de 2019]. Disponível em: <https://medium.com/hipster-color-science/a-beginners-guide-to-colorimetry-401f1830b65a>.
- [58] Fairman HS, Brill MH, Hemmendinger H. How the CIE 1931 color matching functions were derived from Wright-Guild data. *Color Research & Application*. 1997;22(1):11-23.
- [59] CIE 1931 Color Space [2019]. [Acesso em 7 de junho de 2019]. Disponível em: [https://en.wikipedia.org/wiki/CIE\\_1931\\_color\\_space](https://en.wikipedia.org/wiki/CIE_1931_color_space).
- [60] YUV [2019]. [Acesso em 18 de maio de 2019]. Disponível em: <https://en.wikipedia.org/wiki/YUV>.
- [61] YUV Colorspace [2019]. [Acesso em 18 de maio de 2019]. Disponível em: <http://softpixel.com/~cwright/programming/colorspace/yuv/>.
- [62] Shen Y. *Image and video coding - Fundamentals, standards and perspectives of image and video coding*. USA: LAP Lambert Academic Publishing; 2010.
- [63] Poole D. *Álgebra linear - Uma introdução moderna*. 2ed. São Paulo,SP: Cengage Learning; 2017.
- [64] Anton H, Rorres C. *Álgebra Linear com Aplicações*. 10ed. Porto Alegre, RS: Bookman; 2012.
- [65] Watkins DS. *Fundamentals of Matrix Computations*. 3rd ed. New Jersey: John Wiley & Sons; 2010.
- [66] Aravind R, Cash GL, Duttweiler DL, Hang H, Haskell BG, Puri A. Image and video coding standards. *AT&T Technical Journal*. 1993 Jan;72(1):67-89.
- [67] Ostermann J, Bormans J, List P, Marpe D, Narroschke M, Pereira F, et al. Video coding with H.264/AVC: tools, performance and complexity. *Circuits and Systems Magazine*. 2004;4(1):7-28.
- [68] Sikora T. Trends and Perspectives in Image and Video Coding. *Proceedings of the IEEE*. 2005 Jan;93(1):6-17.
- [69] Ohm J, Sullivan GJ, Schwarz H, Tan TK, Wiegand T. Comparison of the Coding Efficiency of Video Coding Standards – Including High Efficiency Video Coding

- (HEVC). IEEE Transactions on Circuits and Systems for Video Technology. 2012 Dec;22(12):1669-84.
- [70] Rijkse K. H.263: video coding for low-bit-rate communication. IEEE Communications Magazine. 1996 Dec;34(12):42-5.
- [71] International Telecommunication Union. Video Coding for Low Bitrate Communication – ITU-T Recommendation H.263. Telecommunication Standardization Sector of ITU; 2005.
- [72] Andrews HC, Patterson CL. Singular Value Decomposition (SVD) image coding. IEEE Transactions on Communications. 1976;24(4):425-32.
- [73] Yang JF, Lu CL. Combined techniques of singular value decomposition and vector quantization for image coding. IEEE Transactions on Image Processing. 1995;4(8):1141-6.
- [74] Ochoa H, Rao KR. A hybrid DWT-SVD image-coding system (HDWTSVD) for monochromatic images. WSEAS Transactions on Circuits and Systems. 2003;4:419-24.
- [75] Ding C, Ye J. Two-dimensional singular value decomposition for 2D maps and images. In: Kargupta H, Srivastava J, Kamath C, Goodman A, editors. Proceedings of the 2005 SIAM International Conference on Data Mining. Philadelphia, PA: Society for Industrial and Applied Mathematics; 2005. p. 32-43.
- [76] Guaragnella C, Politi T. Constant bit rate Video Coding based on SVD video decomposition. In: 2007 IEEE Symposium on Virtual Environments, Human-Computer Interfaces and Measurement Systems; 2007. p. 113-6.
- [77] Gu Z, Lin W, Lee B, Lau C. Low-Complexity Video Coding Based on Two-Dimensional Singular Value Decomposition. IEEE Transactions on Image Processing. 2012;21(2):674-87.
- [78] Koshi Y, Kimura S, Yokose T, Kamizawa K, Minami T, Nakamura O, inventors; Fuji Xerox Co. Ltd., assignee. Singular value decomposition coding and decoding apparatuses. USA Patent US 5615288A. 1997 March 25.

- [79] O'Neal JB. Predictive quantizing systems (differential pulse code modulation) for the transmission of television signals. The Bell System Technical Journal. 1966;45(5):689-721.
- [80] Garguir N. Comparative performance of SVD and adaptive cosine transform in coding images. IEEE Transactions on Communications. 1979;27(8):1230-4.
- [81] Nakamura Y, Iwasaki M, Obata M, Kondo K, Sasada S, inventors; Japan Science and Technology Agency, The Doshisha, assignee. Image feature extraction method and image compression method. USA Patent US 8160368 B2. 2012 April 17.
- [82] Ashin R, Morimoto A, Nagase M, Vaillancourt R. Image compression with multiresolution Singular Value Decomposition and other methods. Mathematical and Computer Modelling. 2005;41(6-7):773-90.
- [83] Porikli F, inventor; Mitsubishi Electric Research Laboratories, Inc, assignee. Method for compressing textured images. USA Patent US 8433148. 2013 April 30.
- [84] Aharon M, Elad M, Bruckstein A. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. IEEE Transactions on Signal Processing. 2006;54(11):4311-22.
- [85] Rajagopalan SS, Pai D, Revankar SV, Sood A, Chaitanya PK, inventors; Adobe Systems Incorporated, assignee. Reconstructing efficiently encoded video frames in a distributed video coding environment. USA Patent US 8451906. 2013 May 28.
- [86] Yang S, Jiao L, Qi Z, Liu F, Ma W, Ma J, inventors; University of Xidian, assignee. Compress imaging method for digital camera based on Toeplitz matrix observation and dictionary learning. China Patent CN103237205. 2013 August 7.
- [87] Pace CP, inventor; Euclid Discoveries LLC., assignee. Feature-based video compression. USA Patent US 8964835. 2015 February 24.
- [88] Choe YS, Choi JW, inventors; University-Industry Foundation (UIF), Yonsei University, assignee. Method, device and system for encoding and decoding image. USA Patent US 20160241861A1. 2016 August 18.



- [89] International Telecommunication Union. High Efficiency Video Coding – ITU-T Recommendation H.265. Telecommunication Standardization Sector of ITU; 2016.
- [90] Li B, Xu J, Wu F, inventors; Microsoft Technology Licensing, LLC, assignee. Encoding strategies for adaptive switching of color spaces, color sampling rates and/or bit depths. USA Patent US 20160261885 A1. 2016 September 8.
- [91] Zhang J, Jia C, Ma S, inventors; University of Beijing, assignee. Method for eliminating block effect of video coding images. China Patent CN105763880. 2017 July 13.
- [92] Cao X, He Y, Zheng X, Zheng J, inventors; Huawei technologies Co., Ltd., Tsinghua University, assignee. Image coding/decoding method, device, and system. USA Patent US 10027974. 2018 July 17.
- [93] Tow D. SQL Tuning. USA: O'Reilly; 2004.
- [94] Westerveld T. Image retrieval: content versus context. In: Proceedings RIAO 2000 Content-Based Multimedia Information Access - Volume 1. Paris, France; 2000. p. 276-84.
- [95] Mandal MK, Idris F, Panchanathan S. A critical evaluation of image and video indexing techniques in the compressed domain. *Image and Vision Computing*. 1999;17:513-29.
- [96] Hu W, Xie N, Li L, Zeng X, Maybank S. A Survey on visual content-based video indexing and retrieval. *IEEE Transactions on Systems, Man and Cybernetics - Part C: Applications and Reviews*. 2011;41(6):797-819.
- [97] Soek CGM, Worring M. Multimodal video indexing: a review of the State-of-the-Art. *Multimedia Tools and Applications*. 2005;25:5-35.
- [98] Faloutsos C. Efficient and effective querying by image content. *Journal of Intelligent Information Systems*. 1994; 3:231-62.
- [99] Mukherjee D, Wu QMJ, Wang G. A comparative experimental study of image feature detectors and descriptors. *Machine Vision and Applications*. 2015: 26:443-66.

- [100] Roldan-Carlos J, Lux M, Giro-i-Nieto X, Munoz P, Anagnostopoulos N. Visual information retrieval in endoscopic video archives. In: Proceedings of the 13<sup>th</sup> International Workshop on Content-Based Multimedia Indexing (CBMI). June 1-12. Prague, Czech Republic: IEEE; 2015. p.1-6.
- [101] Varytimidis C, Konstantinos R, Loukas C, Kollias S. Surgical video retrieval using deep neural networks. In: Proceedings of the 19<sup>th</sup> International Conference on Medical Image Computing and Computer-Assisted Intervention – MICCAI. Athens, Greece. 2016. p4-14.
- [102] Vasconcelos N, Lippman A. Library-based coding: a representation for efficient video compression and retrieval. In: Proceedings DCC '97. Data Compression Conference. Snowbird, UT, USA: IEEE; 1997. p. 121-30. Mar. 25-27.
- [103] Brunelli R, Mich O, inventors; Instituto Trentino di Cultura (Italy), assignee. Method of storing and retrieving images of people, for example, in photograph archives and for the construction of identikit images. Canada, Patent CA 2155901. 1995 August 11.
- [104] Calib D, Buckwald RA, Malik Z, Ben-Yosef N, inventors; Applied Spectral Imaging Ltd., assignee. Spectral bio-imaging methods for cell classification. USA Patent US 5991028. 1999 November 23.
- [105] Lim JH, inventor; Kent Ridge Digital Labs (Singapore, SG), assignee. Method and apparatus for indexing and retrieving images using visual keywords. USA Patent US 6574378. 2003 June 3.
- [106] Chalana V, Fogarasi S, NG L, Oelund J, Pathak S, Racki S, Sparks B, Wyman B, inventors; Insightful Corporation, assignee. System and method for mining quantitative information from medical images. USA Patent US 7158692 B2. 2007 January 2.
- [107] Chalana V, NG L, Rystrom LR, Gee JC, Haynor DR. Validation of brain segmentation and tissue classification algorithm for T1-weighted MR images. SPIE Proceedings, Medical Imaging 2001: Image Processing. 2001;4322.
- [108] Koenig H, inventor; Siemens Aktiengesellschaft, assignee. Automatic indexing of digital image archives for content-based, context-sensitive searching. USA Patent US 7689544 B2. 2010 March 30.

- [109] Connah D, Drew MS, Finlayson G, inventors University of East Anglia (GB), assignee. Method and system for generating accented image data. Patent US 8682093. 2014 March 25.
- [110] Chatzichristofis AS, Boutalis YS. CEDD: color and edge directivity descriptor: a compact descriptor for image indexing and retrieval. In: Gasteratos A, Vincze M, Tsotsos JK, editors. Computer Vision Systems – 6<sup>th</sup> International Conference on Computer Vision Systems, ICVS 2008. Santorini, Greece: Springer-Verlag: 2008. p.312-22.
- [111] Kasutani E, Yamada A. The MPEG-7 color layout descriptor: a compact image feature description for high-speed image/video segment retrieval. In: Proceedings of 2001 International Conference on Image Processing. October 7-10. Thessaloniki, Greece: IEEE; 2001. p.674-7.
- [112] Bosch A, Zisserman A, Munoz X. Representing shape with a spatial pyramid kernel. In: Proceedings of the 6<sup>th</sup> ACM International Conference on Image and Video Retrieval – CIVR '07. July 9-11. Amsterdam, The Netherlands: ACM, 2007. p.401-8.
- [113] Beecks C, Schoeffmann K, Lux M, Uysal MS, Seidl T. Endoscopic video retrieval: a signature-based approach for linking endoscopic images with video segments. In: Proceedings of the 2015 IEEE International Symposium on Multimedia. Miami, FL, 2015; pp.33-8.
- [114] Cordara G, Francini G, Lepsoy S, Gusmao PB, inventors; Telecom Italia, assignee. Method and system for comparing images. Japan Patent JP5734460B. 2015 June 17.
- [115] Liu S, Cao J, Shen X, Song J, Zhang X, Zhang K, inventors. MRI image reconstruction method based on nonlocal singular value decomposition and estimation; 2017. China Patent Application CN107330950A.
- [116] Zeng X, He S, inventors; University Chongqing Posts & Telecom, assignee. Multi-level manifold learning medical image color perceiving method. China Patent CN106651930. 2017 May 10.
- [117] Antoniou ZC, Panayides AS, Pantzaris M, Constantinides AG, Pattichis CS, Pattichis MS. Real-time adaptation to time-varying constraints for medical video

- communications. *IEEE Journal of Biomedical and Health Informatics*. 2018;22(4):1177-88.
- [118] Demaerschalk BM, Vegunta S, Vargas BB, Wu Q, Channer DD, Hentz JG. Reliability of real-time video smartphone for assessing national institutes of health stroke scale scores in acute stroke patients. *Stroke*. 2012;43(12):3271-7.
- [119] Goyal VK. Theoretical foundations of transform coding. *IEEE Signal Processing Magazine*. 2001;18(5):9-21.
- [120] Bairagi VK, Sapkal AM, Gaikwad MS. The role of transforms in image compression. *Journal of The Institution of Engineers (India): Series B*. 2013;94(2):135-40.
- [121] Larson R, Farber B. *Estatística aplicada*. 6ed. São Paulo,SP: Pearson Education do Brasil; 2016.
- [122] Takaki WSR. Proposta de um sistema embarcado para transmissão de vídeos em tempo real com aplicação em telemedicina [Dissertação de Mestrado]. Universidade Estadual do Oeste do Paraná. Foz do Iguaçu, PR; 2015.
- [123] Munzer B, Schoeffmann K, Boszormenyi L. Content-based processing and analysis of endoscopic images and videos: a survey. *Multimedia Tools and Applications*. 2018; 77:1323-62.
- [124] Primus MJ. Segmentation and indexing of endoscopic videos. In: *Proceedings of the 22<sup>nd</sup> ACM International Conference on Multimedia*. Orlando, FL, 2014. Pp 659-62.
- [125] Cao Y, Tavanapong W, Kim K, Wong J, Oh JH, de Groen PC. A framework for parsing colonoscopy videos for semantic units. In: *2004 IEEE International Conference on Multimedia and Expo (ICME)*. Taipei, Taiwan: IEEE; 2004. p.1879-82.
- [126] Cao Y, Tavanapong W, Li D, Oh JH, de Groen PC, Wong J. A Visual Model Approach for parsing colonoscopy. In: Enser P, Kompatsiaris Y, O'Connor NE, Smeaton AF, Smeulders AWM, editors. *Image and Video Retrieval. CIVR 2004*.

Lecture Notes in Computer Science, vol.3115. Berlin, Heidelberg: Springer; 2004. p. 160-9.

- [127] Primus MJ, Schoeffmann K, Boszormenyi L. Segmentation of recorded endoscopic videos by detecting significant motion changes. In: Proceedings of the 11th International Workshop on Content-Based Multimedia Indexing (CBMI). Veszprém, Hungary: IEEE; 2013. p. 223-8.
- [128] Haneishi H, Shiobara T, Miyake Y. Color correction for colorimetric color reproduction in an electronic endoscope. *Optics Communications*. 1995;114(1-2):57-63.
- [129] Gersho A, Gray RM. Vector quantization and signal compression. New York, USA: Springer Science + Business Media; 1992.
- [130] Beecks C, Kletz S, Schoeffmann K. Large-scale endoscopic image and video linking with gradient-based signatures. In: Proceedings of the IEEE Third International Conference on Multimedia Big Data. Laguna Hills,CA,USA; 2017. p. 17-21.
- [131] Girod B, Aaron AM, Rane S, Rebollo-Monedero D. Distributed Video Coding. *Proceedings of the IEEE*. 2005 Jan;93(1):71-83.
- [132] Chang SF. Compressed-domain techniques for image/video indexing and manipulation. In: Proceedings of the IEEE International Conference on Image Processing. Washington,DC,USA; 1995. p. 314-7.
- [133] Javed M, Nagabhushan P, Chaudhuri BB. A review on document image analysis techniques directly in the compressed domain. *Artificial Intelligence Review*. 2018;50(4):539-68.
- [134] Witten IH, Frank E, Hall MA. Data mining: practical machine learning tools and techniques. 3ed. Burlington, MA: Morgan Kaufmann; 2011.
- [135] Sahoo S, Makur A. Dictionary Training for Sparse Representation as Generalization of K-Means Clustering. *IEEE Signal Processing Letters*. 2013; 20:587-90.

- [136] Lux M. LIRE: Open-source image retrieval in Java. In: Proceedings of the 21st ACM International Conference on Multimedia; 2013. Barcelona, Spain: ACM; 2013. p. 843-6.
- [137] Barra GO, Lux M, Giro-i-Nieto X. Large scale content-based video retrieval with LlvRE. In: 14th International Workshop on Content-Based Multimedia Indexing; CBMI 2016; June 15-17. Bucarest, Romania: IEEE; 2016.
- [138] Howarth P, Rüger S. Evaluation of texture features for content-based image retrieval. In: Enser P, Kiompatsiaris Y, O'Connor NE, Smeaton AF, Smeulders AWM, editors. Image and Video Retrieval. CIVR: International Conference on Image and Video Retrieval. Dublin, Ireland: Springer; 2004. p. 326-34.
- [139] Tamura H, Mori S, Yamawaki T. Textural features corresponding to visual perception. IEEE Transactions on Systems, Man and Cybernetics. 1978;8:460-72.
- [140] Iakovidou C, Anagnostopoulos N, Kapoutsis AC, Boutalis Y, Chatzichristofis AS. Searching images with MPEG-7 (& MPEG-7-like) powered localized descriptors: the SIMPLE answer to effective content based image retrieval. In: Proceedings of the 12<sup>th</sup> International Workshop on Content-Based Multimedia Indexing (CBMI). June 18-20. Klagenfurt, Austria: IEEE, 2014. p.1-6.
- [141] Vargas C. Relatório padrão. In: Quilici FA, Grecco EC, editors. Colonoscopia. São Paulo: Lemos Editorial; 2014. p. 63-5.
- [142] Houaiss A, de Salles Villar M. Dicionário eletrônico Houaiss da língua portuguesa [CD-ROM]. Rio de Janeiro, RJ: Objetiva; 2007.
- [143] Monico JFG, Póz APD, Galo M, Santos MC, Oliveira LC. Acurácia e precisão: revendo os conceitos de forma acurada. Boletim de Ciências Geodésicas. 2009;15(3):469-83.
- [144] Skretting K. Dictionary learning tools for Matlab [Internet]. Stavanger; 2018. [Cited 2019 May 26]. Available from: <http://www.ux.uis.no/~karlsk/dle/>.
- [145] Calonder M, Lepetit V, Strecha C, Fua P. BRIEF: binary robust independent elementary features. In: Daniilidis K, Maragos P, Paragios N, editors. Computer

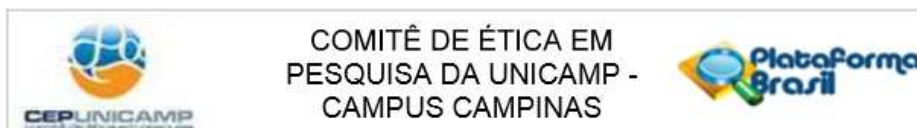
- Vision ECCV 2010 – 11<sup>th</sup> European Conference on Computer Vision. Greece: Springer-Verlag; 2010. p.778-92.
- [146] Coulouris G, Dollimore J, Kindberg T. Sistemas distribuídos: conceitos e projeto. 4ed. Porto Alegre, RS: Bookman; 2007.
- [147] Cormen TH, Leiserson CE, Rivest RL, Stein C. Introduction to algorithms. 3ed. Cambridge, MA: MIT Press; 2009.
- [148] Haykin S. Sistemas de comunicação: analógicos e digitais. 4ed. Porto Alegre, RS: Bookman; 2004.
- [149] Liu S, Bovik AC. Efficient DCT-domain blind measurement and reduction of blocking artifacts. IEEE Transactions on Circuits and Systems for Video Technology. 2002;12(12):1139-49.
- [150] Rosten E, Drummond T. Machine learning for high-speed corner detection. In: Leonardis A, Bischof H, Pinz A, editors. Computer Vision ECCV 2006 – 9<sup>th</sup> European Conference on Computer Vision. Austria: Springer-Verlag; 2006. p.430-43.
- [151] Stewart J. Cálculo: volume II. São Paulo: Cengage Learning; 2009.
- [152] Price AMA, Toscani SS. Implementação de linguagens de programação: compiladores. 2ed. Porto Alegre: Sagra Luzzatto; 2001.
- [153] Hopcroft JE, Ullman JD, Motwani R. Introdução à teoria de autômatos, linguagens e computação. Rio de Janeiro, RJ: Campus; 2003.
- [154] Loudon KC. Compiladores: princípios e práticas. São Paulo: Thomson Learning; 2004.
- [155] Cleveland DB, Cleveland AD. Introduction to indexing and abstracting. 4ed. Santa Barbara, CA: Libraries Unlimited; 2013.
- [156] Cutting D, Pedersen J. Optimizations for Dynamic Inverted Index Maintenance. SIGIR '90: Proceedings of the 13th annual international ACM SIGIR conference on research and development in information retrieval; Sep 5-7. Brussels, Belgium; 1990. p. 405-11.

- [157] Tanenbaum AS. Computer Networks. 4ed. New Jersey: Pearson Education; 2003.
- [158] Gray RM. Toeplitz and circulant matrices: a review. *Foundations and trends in communications and information theory*. 2006;2(3):155-239.
- [159] Matas J, Chum O, Urban M, Pajdla T. Robust wide-baseline stereo from maximally stable extremal regions. *Image and Vision Computing*. 2004; 22(10):761-7.
- [160] Rublee E, Rabaud V, Konolige K, Bradski G. ORB: an efficient alternative to SIFT or SURF. In: *Proceedings of IEEE International Conference on Computer Vision; ICCV '11; November 6-13. Washington DC, USA: IEEE; 2011. p.2564-71.*
- [161] Mitchell TM. *Machine Learning*. Portland, Oregon: McGraw-Hill; 1997.
- [162] Donoho DL. Compressed sensing. *IEEE Transactions on Information Theory*. 2006;52(4):1289-306.
- [163] Bay H, Tuytelaars T, Gool LV. SURF: speeded up robust features. In: Leonardis A, Bischof H, Pinz A, editors. *Computer Vision ECCV 2006 – 9<sup>th</sup> European Conference on Computer Vision*. Austria: Springer-Verlag; 2006.p.404-17.



## ANEXOS

### A – Parecer do Comitê de Ética em Pesquisa



#### PARECER CONSUBSTANCIADO DO CEP

##### DADOS DO PROJETO DE PESQUISA

**Título da Pesquisa:** MÉTODO PARA TRANSMISSÃO DE ÁUDIO E DE VÍDEO DE PROCEDIMENTOS MÉDICOS EM TEMPO REAL PARA SISTEMAS EMBARCADOS E MÓVEIS

**Pesquisador:** Weber Shoity Resende Takaki

**Área Temática:**

**Versão:** 2

**CAAE:** 53887216.3.0000.5404

**Instituição Proponente:** Faculdade de Ciências Médicas - UNICAMP

**Patrocinador Principal:** Financiamento Próprio

##### DADOS DO PARECER

**Número do Parecer:** 1.477.557

##### Apresentação do Projeto:

"Esta pesquisa apresenta como finalidade desenvolver um sistema embarcado para transmissão de áudio e de vídeo de procedimentos médicos em tempo real, compatível com computadores de propósito geral e com dispositivos móveis tais como tablets e smartphones. Este sistema permitirá que estudantes ou profissionais de saúde acompanhem "ao vivo" a execução de eventos como procedimentos cirúrgicos ou exames por imagens realizados em locais distantes. Com isso, possibilita a interação entre os profissionais envolvidos por meio de chat, de áudio e de vídeo. Essas modalidades de interação proporcionam a redução de custos com deslocamentos de profissionais ou de pacientes. Vale ressaltar que sistemas embarcados apresentam custos de aquisição em geral mais baixos que computadores pessoais, repercutindo diretamente na redução dos custos da solução. Para a avaliação do quesito funcionalidade do sistema proposto, serão realizados experimentos com transmissão de fluxos de áudio e vídeo digital, com conteúdo capaz de simular um cenário real, de maneira que permita análise precisa da aplicabilidade da solução desenvolvida. As variáveis de desempenho medidas nos experimentos serão submetidas à análise estatística a fim de avaliar a estabilidade e a confiabilidade da solução proposta. Espera-se obter como resultado um protótipo funcional que atenda aos requisitos de desempenho, com interface de operação intuitiva e simples."

**Endereço:** Rua Tessália Vieira de Camargo, 126  
**Bairro:** Barão Geraldo **CEP:** 13.083-887  
**UF:** SP **Município:** CAMPINAS  
**Telefone:** (19)3521-8938 **Fax:** (19)3521-7187 **E-mail:** cep@fcm.unicamp.br



Continuação do Parecer: 1.477.557

**Objetivo da Pesquisa:**

"Objetivo Primário: Avaliar a capacidade de sistemas embarcados na codificação e transmissão de áudio e de vídeo em tempo real, bem como a compatibilidade destes sistemas com dispositivos móveis."

**Avaliação dos Riscos e Benefícios:**

"Riscos: O uso do vídeo gravado não envolve riscos nem desconfortos físicos previsíveis ao voluntário. Benefícios: Os benefícios previstos são indiretos, pois a pesquisa pretende contribuir com a melhoria da prestação dos serviços médicos por meio da cooperação entre especialistas, da qualificação da educação médica, e do apoio à descoberta de novos conhecimentos em Medicina."

**Comentários e Considerações sobre a Pesquisa:**

Estudo experimental, com amostra prevista de um voluntário, no qual deverá ceder um vídeo de endoscopia digestiva já realizada anteriormente. O pesquisador analisa que apenas um vídeo seria suficiente na projeto de pesquisa uma vez que as variáveis observadas são relativas ao desempenho técnico do equipamento na transmissão de áudio e vídeo.

O Termo de Consentimento Livre e Esclarecido (TCLE) deve ser adequado para estar em conformidade com a Resolução CNS nº466 de 2012.

**Considerações sobre os Termos de apresentação obrigatória:**

1. Protocolo de Pesquisa gerado pela Plataforma Brasil preenchido, inclusive Cronograma de execução da pesquisa e Orçamento, indicando financiamento pelo próprio pesquisador com custo estimado em R\$3.500,00 e início da coleta de dados em 04/07/2016;
2. Folha de Rosto preenchida e assinada pelo pesquisador responsável e pelo responsável legal pela instituição;
3. Termo de Consentimento Livre e Esclarecido (TCLE) anexado à Plataforma Brasil;
4. Projeto de Pesquisa, com finalidade de doutorado, anexado à Plataforma Brasil;

**Recomendações:**

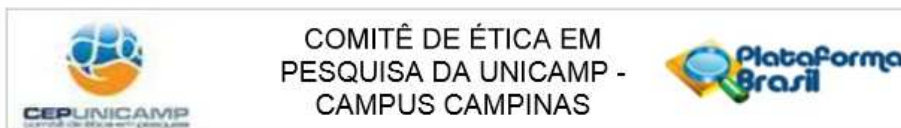
**Conclusões ou Pendências e Lista de Inadequações:**

1. Sobre o Termo de Consentimento Livre e Esclarecido (TCLE):

1.a. Solicitamos ao pesquisador trocar a palavra "cópia" no TCLE pela palavra "via", no seguinte trecho: "Cópia. Uma cópia deste termo será entregue ao voluntário.";

1.b. Solicitamos ao pesquisador esclarecer no TCLE que o CEP está disponível para atendê-los

**Endereço:** Rua Tessália Vieira de Camargo, 126  
**Bairro:** Barão Geraldo **CEP:** 13.083-887  
**UF:** SP **Município:** CAMPINAS  
**Telefone:** (19)3521-8938 **Fax:** (19)3521-7187 **E-mail:** cep@fcm.unicamp.br



Continuação do Parecer: 1.477.557

sobre questões éticas da pesquisa. Sugestão de texto: "Em caso de denúncias ou reclamações sobre sua participação e sobre questões éticas do estudo, você pode entrar em contato com a secretaria do Comitê de Ética em Pesquisa (CEP) da UNICAMP das 08:30hs às 13:30hs e das 13:00hs as 17:00hs na Rua: Tessália Vieira de Camargo, 126; CEP 13083-887 Campinas – SP; telefone (19) 3521-8936; fax (19) 3521-7187; e-mail: cep@fcm.unicamp.br"

1.c. O pesquisador responsável também deve assinar o TCLE. Solicitamos ainda ao pesquisador expressar o cumprimento das exigências na Resolução 466/12 CNS/MS (Sugestão: ver modelo de TCLE proposto pelo CEP/UNICAMP em [https://www.prp.unicamp.br/sites/default/files/3\\_estrutura\\_basica\\_tcleversao\\_09outubro2015-disponivel\\_site.docx](https://www.prp.unicamp.br/sites/default/files/3_estrutura_basica_tcleversao_09outubro2015-disponivel_site.docx))

RESPOSTA: O pesquisador apresentou nova versão do TCLE, através do documento "TermoDeConsentimentoLivreEsclarecido\_corrigido.pdf", anexado na Plataforma Brasil em 29/03/16 às 05h36. O novo documento contempla as modificações sugeridas pelo CEP/UNICAMP.

ANÁLISE: Pendência atendida.

**Considerações Finais a critério do CEP:**

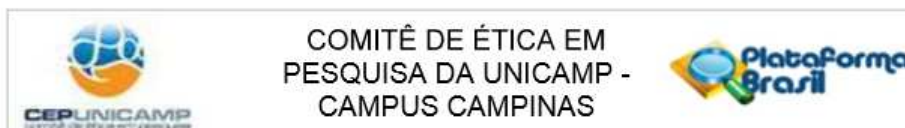
- O sujeito de pesquisa deve receber uma via do Termo de Consentimento Livre e Esclarecido, na íntegra, por ele assinado (quando aplicável).

- O sujeito da pesquisa tem a liberdade de recusar-se a participar ou de retirar seu consentimento em qualquer fase da pesquisa, sem penalização alguma e sem prejuízo ao seu cuidado (quando aplicável).

- O pesquisador deve desenvolver a pesquisa conforme delineada no protocolo aprovado. Se o pesquisador considerar a descontinuação do estudo, esta deve ser justificada e somente ser realizada após análise das razões da descontinuidade pelo CEP que o aprovou. O pesquisador deve aguardar o parecer do CEP quanto à descontinuação, exceto quando perceber risco ou dano não previsto ao sujeito participante ou quando constatar a superioridade de uma estratégia diagnóstica ou terapêutica oferecida a um dos grupos da pesquisa, isto é, somente em caso de necessidade de ação imediata com intuito de proteger os participantes.

- O CEP deve ser informado de todos os efeitos adversos ou fatos relevantes que alterem o curso normal do estudo. É papel do pesquisador assegurar medidas imediatas adequadas frente a evento

**Endereço:** Rua Tessália Vieira de Camargo, 126  
**Bairro:** Barão Geraldo **CEP:** 13.083-887  
**UF:** SP **Município:** CAMPINAS  
**Telefone:** (19)3521-8936 **Fax:** (19)3521-7187 **E-mail:** cep@fcm.unicamp.br



Continuação do Parecer: 1.477.557

adverso grave ocorrido (mesmo que tenha sido em outro centro) e enviar notificação ao CEP e à Agência Nacional de Vigilância Sanitária – ANVISA – junto com seu posicionamento.

- Eventuais modificações ou emendas ao protocolo devem ser apresentadas ao CEP de forma clara e sucinta, identificando a parte do protocolo a ser modificada e suas justificativas. Em caso de projetos do Grupo I ou II apresentados anteriormente à ANVISA, o pesquisador ou patrocinador deve enviá-las também à mesma, junto com o parecer aprovatório do CEP, para serem juntadas ao protocolo inicial.

- Relatórios parciais e final devem ser apresentados ao CEP, inicialmente seis meses após a data deste parecer de aprovação e ao término do estudo.

- Lembramos que segundo a Resolução 466/2012, item XI.2 letra e, "cabe ao pesquisador apresentar dados solicitados pelo CEP ou pela CONEP a qualquer momento".

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

Tipo Documento	Arquivo	Postagem	Autor	Situação
Informações Básicas do Projeto	PB_INFORMAÇÕES_BÁSICAS_DO_PROJETO_662388.pdf	29/03/2016 05:38:08		Aceito
Outros	CartaResposta_assinada.pdf	29/03/2016 05:36:30	Weber Shoity Resende Takaki	Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	TermoDeConsentimentoLivreEsclarecido_corrigido.pdf	29/03/2016 05:36:02	Weber Shoity Resende Takaki	Aceito
Outros	RelatorioFinalMatricula_Takaki.pdf	03/03/2016 08:33:58	Weber Shoity Resende Takaki	Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	TermoDeConsentimentoLivreEsclarecido.pdf	01/03/2016 21:50:12	Weber Shoity Resende Takaki	Aceito
Folha de Rosto	folhaDeRosto_assinada.pdf	22/02/2016 00:10:33	Weber Shoity Resende Takaki	Aceito
Projeto Detalhado / Brochura	ProjetoDePesquisa_ComiteDeEtica.pdf	22/02/2016 00:10:20	Weber Shoity Resende Takaki	Aceito

Endereço: Rua Tessália Vieira de Camargo, 126  
 Bairro: Barão Geraldo CEP: 13.083-887  
 UF: SP Município: CAMPINAS  
 Telefone: (19)3521-8936 Fax: (19)3521-7187 E-mail: cep@fem.unicamp.br



COMITÊ DE ÉTICA EM  
PESQUISA DA UNICAMP -  
CAMPUS CAMPINAS

Continuação do Parecer: 1.477.557

Investigador	ProjetoDePesquisa_ComiteDeEtica.pdf	22/02/2016 00:10:20	Weber Shoity Resende Takaki	Aceito
--------------	-------------------------------------	------------------------	--------------------------------	--------

**Situação do Parecer:**

Aprovado

**Necessita Apreciação da CONEP:**

Não

CAMPINAS, 05 de Abril de 2016

---

**Assinado por:**  
**Renata Maria dos Santos Celeghini**  
(Coordenador)

**Endereço:** Rua Tessália Vieira de Camargo, 126  
**Bairro:** Barão Geraldo **CEP:** 13.083-887  
**UF:** SP **Município:** CAMPINAS  
**Telefone:** (19)3521-8936 **Fax:** (19)3521-7187 **E-mail:** cep@fom.unicamp.br

**B – Depósito da patente do método proposto – primeira página**

27/12/2018 870180168196  
15:30  
29409161812000757

**Pedido nacional de Invenção, Modelo de Utilidade, Certificado de Adição de Invenção e entrada na fase nacional do PCT**

Número do Processo: BR 10 2018 077265 1

**Dados do Depositante (71)**

---

**Depositante 1 de 1**

**Nome ou Razão Social:** UNIVERSIDADE ESTADUAL DE CAMPINAS - UNICAMP

**Tipo de Pessoa:** Pessoa Jurídica

**CPF/CNPJ:** 46068425000133

**Nacionalidade:** Brasileira

**Qualificação Jurídica:** Órgão Público

**Endereço:** Cidade Universitária Zeferino Vaz

**Cidade:** Campinas

**Estado:** SP

**CEP:** 13084-971

**País:** Brasil

**Telefone:**

**Fax:**

**Email:** [patentes@inova.unicamp.br](mailto:patentes@inova.unicamp.br)

---

**PETICIONAMENTO  
ELETRÔNICO**

Esta solicitação foi enviada pelo sistema Petição Eletrônica em 27/12/2018 às 15:30, Petição 870180168196

## C – Tabelas de resultados dos experimentos de recuperação

Tabela C.1 – Resultados da recuperação para o Caso 1.

Nº do quadro buscado	Quadros retornados	Vetores com $R^2 > 0,9$	Soma dos $R^2 > 0,9$
9	9	10	10,0000
	7228	1	0,9927
	7229	1	0,9922
10	10	10	10,0000
	6984	1	0,9839
	6988	1	0,9838
216	216	10	10,0000
	215	1	0,9958
	217	1	0,9956
231	231	10	10,0000
	230	1	0,9993
	228	1	0,9991
232	232	10	10,0000
	229	1	0,9992
	228	1	0,9990
236	236	10	10,0000
	239	1	0,9992
	269	1	0,9991
290	290	10	10,0000
	289	1	0,9999
	291	1	0,9999
339	339	10	10,0000
	338	2	1,9375
	340	1	1,0000
340	340	10	10,0000
	338	2	1,9211
	341	2	1,9019
342	342	10	10,0000
	341	2	1,9110
	343	1	0,9999
346	346	10	10,0000
	347	1	0,9998
	345	1	0,9997

*Continua na próxima página.*

Continuação da Tabela C.1.

<b>Nº do quadro buscado</b>	<b>Quadros retornados</b>	<b>Vetores com <math>R^2 &gt; 0,9</math></b>	<b>Soma dos <math>R^2 &gt; 0,9</math></b>
406	406	10	10,0000
	397	1	0,9995
	398	1	0,9995
407	407	10	10,0000
	409	1	0,9997
	410	1	0,9996
439	439	10	10,0000
	440	2	1,9500
	438	1	0,9998
527	527	10	10,0000
	526	1	0,9999
	529	1	0,9997
566	566	10	10,0000
	565	1	0,9990
	530	1	0,9986
624	624	10	10,0000
	631	1	0,9991
	632	1	0,9987
691	691	10	10,0000
	692	1	0,9976
	693	1	0,9965
696	696	10	10,0000
	697	2	1,9115
	706	1	0,9992
702	702	10	10,0000
	703	3	2,8458
	698	1	0,9997
707	707	10	10,0000
	706	1	0,9997
	708	1	0,9993
797	797	10	10,0000
	798	1	0,9990
	799	1	0,9985
798	798	10	10,0000
	799	1	0,9991
	802	1	0,9990

*Continua na próxima página.*



Continuação da Tabela C.1.

<b>Nº do quadro buscado</b>	<b>Quadros retornados</b>	<b>Vetores com <math>R^2 &gt; 0,9</math></b>	<b>Soma dos <math>R^2 &gt; 0,9</math></b>
942	942	10	10,0000
	941	1	0,9990
	943	1	0,9980
1047	1047	10	10,0000
	1049	1	0,9969
	1070	1	0,9953
1155	1155	10	10,0000
	1154	2	1,9018
	1156	1	0,9998
1188	1188	10	10,0000
	1187	1	0,9989
	1186	1	0,9966
1206	1206	10	10,0000
	1205	1	0,9996
	1207	1	0,9995
1338	1338	10	10,0000
	1336	1	0,9999
	1340	1	0,9994
1477	1477	10	10,0000
	1479	1	0,9992
	1480	1	0,9986
1497	1497	10	10,0000
	1495	1	0,9993
	1493	1	0,9959
1852	1852	10	10,0000
	1849	1	0,9947
	1856	1	0,9923
1944	1944	10	10,0000
	1943	1	0,9979
	1942	1	0,9941
2198	2198	10	10,0000
	2195	5	4,6886
	2197	3	2,8553
2240	2240	10	10,0000
	2247	1	0,9985
	2248	1	0,9984

Continua na próxima página.

Continuação da Tabela C.1.

<b>Nº do quadro buscado</b>	<b>Quadros retornados</b>	<b>Vetores com <math>R^2 &gt; 0,9</math></b>	<b>Soma dos <math>R^2 &gt; 0,9</math></b>
2394	2394	10	10,0000
	2395	1	0,9998
	2393	1	0,9998
2414	2414	10	10,0000
	2413	2	1,9186
	2412	1	0,9998
2472	2472	10	10,0000
	2473	1	0,9990
	2478	1	0,9983
2511	2511	10	10,0000
	2513	1	0,9995
	2510	1	0,9993
2637	2637	10	10,0000
	2639	1	0,9994
	2636	1	0,9990
2658	2658	10	10,0000
	2663	1	0,9995
	2662	1	0,9988
2759	2759	10	10,0000
	2761	1	0,9992
	2762	1	0,9972
2957	2957	10	10,0000
	2956	1	0,9999
	2955	1	0,9996
3022	3022	10	10,0000
	3026	1	0,9962
	3027	1	0,9959
3047	3047	10	10,0000
	3048	1	0,9994
	3046	1	0,9993
3189	3189	10	10,0000
	3190	1	0,9998
	3191	1	0,9995
3280	3280	10	10,0000
	3281	1	0,9999
	3279	1	0,9998

*Continua na próxima página.*

Continuação da Tabela C.1.

<b>Nº do quadro buscado</b>	<b>Quadros retornados</b>	<b>Vetores com <math>R^2 &gt; 0,9</math></b>	<b>Soma dos <math>R^2 &gt; 0,9</math></b>
3333	3333	10	10,0000
	3332	1	0,9997
	3334	1	0,9997
3350	3350	10	10,0000
	3349	1	0,9999
	3351	1	0,9998
3370	3370	10	10,0000
	3369	1	0,9998
	3371	1	0,9997
3376	3376	10	10,0000
	3377	3	2,8285
	3375	2	1,9523
3515	3515	10	10,0000
	3517	1	0,9996
	3518	1	0,9977
3568	3568	10	10,0000
	3569	3	2,8454
	3571	3	2,8253
3597	3597	10	10,0000
	3596	3	2,8709
	3598	1	0,9999
3609	3609	10	10,0000
	3610	1	0,9996
	3608	1	0,9995
3677	3677	10	10,0000
	3200	1	0,9982
	3202	1	0,9982
3696	3696	10	10,0000
	3691	1	0,9985
	1305	1	0,9975
3820	3820	10	10,0000
	3819	1	0,9985
	787	1	0,9965
3893	3893	10	10,0000
	3892	9	8,3323
	3894	3	2,8781

Continua na próxima página.

Continuação da Tabela C.1.

<b>Nº do quadro buscado</b>	<b>Quadros retornados</b>	<b>Vetores com <math>R^2 &gt; 0,9</math></b>	<b>Soma dos <math>R^2 &gt; 0,9</math></b>
3945	3945	10	10,0000
	3944	2	1,9331
	3946	1	0,9999
3992	3992	10	10,0000
	3993	10	9,7122
	3994	10	9,5157
4041	4041	10	10,0000
	4040	1	0,9999
	4039	1	0,9990
4098	4098	10	10,0000
	4097	1	0,9998
	4099	1	0,9996
4194	4194	10	10,0000
	4195	1	0,9996
	4192	1	0,9995
4230	4230	10	10,0000
	4231	2	1,9392
	4228	2	1,9232
4531	4531	10	9,9950
	4532	1	0,9995
	4530	1	0,9991
4567	4567	10	10,0000
	4566	1	0,9997
	4568	1	0,9995
4600	4600	10	10,0000
	4599	1	0,9998
	4570	1	0,9996
4674	4674	10	10,0000
	4675	1	0,9980
	4677	1	0,9972
4934	4934	10	10,0000
	4935	5	4,6660
	4933	3	2,8457
4970	4970	10	10,0000
	4971	1	0,9958
	4972	1	0,9949

*Continua na próxima página.*

Continuação da Tabela C.1.

<b>Nº do quadro buscado</b>	<b>Quadros retornados</b>	<b>Vetores com <math>R^2 &gt; 0,9</math></b>	<b>Soma dos <math>R^2 &gt; 0,9</math></b>
5122	5122	10	10,0000
	5683	1	0,9884
	5121	1	0,9878
5240	5240	10	10,0000
	5238	2	1,9245
	5237	2	1,9170
5335	5335	10	10,0000
	5333	1	0,9987
	5336	1	0,9986
5479	5479	10	10,0000
	5478	10	9,4410
	5480	3	2,8601
6014	6014	10	10,0000
	6013	1	0,9975
	6015	1	0,9940
6097	6097	10	10,0000
	6096	1	0,9996
	6098	1	0,9995
6385	6385	10	10,0000
	6384	3	2,8378
	6386	1	0,9956
6513	6513	10	10,0000
	6514	1	0,9998
	6512	1	0,9996
6861	6861	10	10,0000
	6860	3	2,9100
	6859	2	1,9281
6864	6864	10	10,0000
	6865	2	1,9085
	6866	1	1,0000
6941	6941	10	10,0000
	6942	2	1,9398
	6943	1	0,9992
6949	6949	10	10,0000
	6950	1	0,9984
	6951	1	0,9982

*Continua na próxima página.*

Continuação da Tabela C.1.

<b>Nº do quadro buscado</b>	<b>Quadros retornados</b>	<b>Vetores com <math>R^2 &gt; 0,9</math></b>	<b>Soma dos <math>R^2 &gt; 0,9</math></b>
6985	6985	10	10,0000
	6987	3	2,8565
	6986	2	1,9527
7122	7122	10	10,0000
	1181	1	0,9933
	3512	1	0,9929
7183	7183	10	10,0000
	7184	2	1,9667
	7186	2	1,9074
7243	7243	10	10,0000
	7244	4	3,7324
	7242	2	1,9164
7310	7310	10	10,0000
	7311	3	2,8576
	7308	1	0,9999
7325	7325	10	10,0000
	7323	1	0,9996
	7326	1	0,9996
7405	7405	10	10,0000
	7404	1	0,9999
	7406	1	0,9995
7424	7424	10	10,0000
	3769	1	0,9950
	135	1	0,9945
7504	7504	10	10,0000
	7503	1	0,9984
	7502	1	0,9973
7550	7550	10	10,0000
	7551	1	0,9938
	7549	1	0,9932
7610	7610	10	10,0000
	7611	2	1,9214
	7609	1	0,9995
7671	7671	10	10,0000
	7473	1	0,9917
	2337	1	0,9916

*Continua na próxima página.*

*Conclusão da Tabela C.1.*

<b>Nº do quadro buscado</b>	<b>Quadros retornados</b>	<b>Vetores com <math>R^2 &gt; 0,9</math></b>	<b>Soma dos <math>R^2 &gt; 0,9</math></b>
8022	8022	10	10,0000
	8021	2	1,9061
	8023	1	0,9983
8075	8075	10	10,0000
	8076	2	1,9257
	8077	1	0,9996

Tabela C.2 – Resultados da recuperação para o Caso 2.

<b>Nº do quadro buscado</b>	<b>Quadros retornados</b>	<b>Nº de vetores com <math>R^2 &gt; 0,9</math></b>	<b>Nº de vetores com <math>0,9 \geq R^2 &gt; 0,8</math></b>
9	9	460	0
	6864	90	26
	7	86	20
10	10	460	0
	7026	92	10
	7015	89	18
216	216	460	0
	217	83	44
	215	74	45
231	231	460	0
	230	154	56
	229	125	51
232	232	460	0
	227	162	56
	228	134	46
236	236	460	0
	247	212	21
	235	207	34
290	290	460	0
	289	333	19
	288	294	34
339	339	460	0
	338	212	47
	340	186	36
340	340	460	0
	341	279	43
	338	274	56
342	342	460	0
	343	294	35
	341	287	36
346	346	460	0
	345	282	38
	347	257	48
406	406	460	0
	407	240	32
	405	229	40

*Continua na próxima página.*



Continuação da Tabela C.2.

<b>Nº do quadro buscado</b>	<b>Quadros retornados</b>	<b>Nº de vetores com <math>R^2 &gt; 0,9</math></b>	<b>Nº de vetores com <math>0,9 \geq R^2 &gt; 0,8</math></b>
407	407	460	0
	406	240	32
	411	229	26
439	439	460	0
	440	352	26
	441	305	31
527	527	460	0
	528	249	33
	529	245	28
566	566	460	0
	565	290	29
	564	249	16
624	624	460	0
	623	145	16
	631	134	7
691	691	460	0
	690	139	36
	692	133	42
696	696	460	0
	697	188	59
	695	145	33
702	702	460	0
	703	222	63
	704	160	51
707	707	460	0
	706	148	54
	708	136	34
797	797	460	0
	798	140	38
	799	124	44
798	798	460	0
	799	183	57
	800	170	53
942	942	460	0
	943	89	14
	941	85	19

Continua na próxima página.

Continuação da Tabela C.2.

<b>Nº do quadro buscado</b>	<b>Quadros retornados</b>	<b>Nº de vetores com <math>R^2 &gt; 0,9</math></b>	<b>Nº de vetores com <math>0,9 \geq R^2 &gt; 0,8</math></b>
1047	1047	460	0
	1049	60	32
	7328	26	5
1155	1155	460	0
	1154	104	65
	1156	102	49
1188	1188	460	0
	1187	84	7
	6998	82	8
1206	1206	460	0
	1207	103	84
	1205	96	83
1338	1338	460	0
	1336	52	36
	1340	44	22
1477	1477	460	0
	1479	36	25
	1328	34	4
1497	1497	460	0
	1495	56	63
	1498	33	12
1852	1852	460	0
	1856	51	27
	1849	39	26
1944	1944	460	0
	1943	59	39
	1942	45	24
2198	2198	460	0
	2195	109	86
	2197	100	78
2240	2240	460	0
	2238	48	21
	2248	47	11
2394	2394	460	0
	2395	72	35
	2393	60	34

Continua na próxima página.

Continuação da Tabela C.2.

<b>Nº do quadro buscado</b>	<b>Quadros retornados</b>	<b>Nº de vetores com <math>R^2 &gt; 0,9</math></b>	<b>Nº de vetores com <math>0,9 \geq R^2 &gt; 0,8</math></b>
2414	2414	460	0
	2413	78	36
	2412	63	23
2472	2472	460	0
	2477	49	19
	2476	48	21
2511	2511	460	0
	2513	75	39
	2510	70	32
2637	2637	460	0
	2639	43	14
	2641	40	5
2658	2658	460	0
	2659	106	34
	2657	91	21
2759	2759	460	0
	2761	50	26
	2767	34	8
2957	2957	460	0
	2958	109	45
	2956	102	47
3022	3022	460	0
	3033	45	19
	490	44	22
3047	3047	460	0
	3048	94	31
	3046	84	22
3189	3189	460	0
	3190	81	45
	3188	67	22
3280	3280	460	0
	3281	53	15
	3279	44	20
3333	3333	460	0
	3332	72	35
	3334	64	35

*Continua na próxima página.*

Continuação da Tabela C.2.

<b>Nº do quadro buscado</b>	<b>Quadros retornados</b>	<b>Nº de vetores com <math>R^2 &gt; 0,9</math></b>	<b>Nº de vetores com <math>0,9 \geq R^2 &gt; 0,8</math></b>
3350	3350	460	0
	3351	89	38
	3349	65	28
3370	3370	460	0
	3369	73	48
	3371	62	34
3376	3376	460	0
	3377	93	43
	3375	68	42
3515	3515	460	0
	3517	44	23
	3691	37	2
3568	3568	460	0
	3567	92	45
	3569	79	62
3597	3597	460	0
	3596	141	84
	3598	120	62
3609	3609	460	0
	3610	116	59
	3608	106	68
3677	3677	460	0
	520	91	15
	3678	90	11
3696	3696	460	0
	2257	51	9
	2357	50	9
3820	3820	460	0
	3819	86	26
	3821	79	13
3893	3893	460	0
	3894	93	55
	3892	89	44
3945	3945	460	0
	3946	105	39
	3944	90	38

Continua na próxima página.

Continuação da Tabela C.2.

<b>Nº do quadro buscado</b>	<b>Quadros retornados</b>	<b>Nº de vetores com <math>R^2 &gt; 0,9</math></b>	<b>Nº de vetores com <math>0,9 \geq R^2 &gt; 0,8</math></b>
3992	3992	460	0
	3993	237	50
	3991	229	45
4041	4041	460	0
	4040	160	56
	4039	105	51
4098	4098	460	0
	4099	132	39
	4097	117	33
4194	4194	460	0
	4195	108	24
	4196	88	12
4230	4230	460	0
	4228	60	46
	4231	35	10
4531	4531	460	0
	4532	112	43
	4530	110	55
4567	4567	460	0
	4568	79	63
	4566	78	64
4600	4600	460	0
	4599	86	66
	4596	81	62
4674	4674	460	0
	4672	54	29
	4675	53	16
4934	4934	460	0
	4933	125	49
	4935	114	66
4970	4970	460	0
	4971	96	39
	4972	86	34
5122	5122	460	0
	5121	75	29
	5123	71	21

Continua na próxima página.

Continuação da Tabela C.2.

<b>Nº do quadro buscado</b>	<b>Quadros retornados</b>	<b>Nº de vetores com <math>R^2 &gt; 0,9</math></b>	<b>Nº de vetores com <math>0,9 \geq R^2 &gt; 0,8</math></b>
5240	5240	460	0
	5239	78	47
	5238	74	44
5335	5335	460	0
	5336	71	58
	5333	68	35
5479	5479	460	0
	5480	128	86
	5478	102	80
6014	6014	460	0
	6015	182	75
	6016	132	64
6097	6097	460	0
	6096	122	80
	6098	112	85
6385	6385	460	0
	6384	115	50
	6386	95	46
6513	6513	460	0
	6514	243	100
	6512	232	87
6861	6861	460	0
	6860	264	67
	6859	227	53
6864	6864	460	0
	6865	291	21
	6866	241	18
6941	6941	460	0
	6942	165	65
	6940	118	42
6949	6949	460	0
	6950	101	33
	6951	87	29
6985	6985	460	0
	6987	185	44
	6986	177	64

Continua na próxima página.

Continuação da Tabela C.2.

<b>Nº do quadro buscado</b>	<b>Quadros retornados</b>	<b>Nº de vetores com <math>R^2 &gt; 0,9</math></b>	<b>Nº de vetores com <math>0,9 \geq R^2 &gt; 0,8</math></b>
7122	7122	460	0
	7123	175	9
	7032	173	3
7183	7183	460	0
	7184	202	45
	7182	157	36
7243	7243	460	0
	7244	85	67
	7242	75	75
7310	7310	460	0
	7311	106	61
	7308	86	55
7325	7325	460	0
	7326	63	32
	7327	48	18
7405	7405	460	0
	7404	69	30
	7406	62	36
7424	7424	460	0
	695	62	9
	698	57	7
7504	7504	460	0
	7503	51	15
	7505	43	9
7550	7550	460	0
	7549	100	23
	2938	100	9
7610	7610	460	0
	7609	94	60
	7611	87	30
7671	7671	460	0
	7670	155	53
	7672	130	52
8022	8022	460	0
	8021	128	62
	8023	89	30

Continua na próxima página.

Conclusão da Tabela C.2.

<b>Nº do quadro buscado</b>	<b>Quadros retornados</b>	<b>Nº de vetores com <math>R^2 &gt; 0,9</math></b>	<b>Nº de vetores com <math>0,9 \geq R^2 &gt; 0,8</math></b>
8075	8075	460	0
	8076	242	44
	8074	216	37