

UNIVERSIDADE ESTADUAL DE MARINGÁ
CENTRO DE TECNOLOGIA
DEPARTAMENTO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

FERNANDA TAMY ISHII

Lucas-Kanade Tensorial: Aplicação da álgebra tensorial para a
estimativa de fluxo ótico em sequência de imagens coloridas

Maringá

2018

FERNANDA TAMY ISHII

Lucas-Kanade Tensorial: Aplicação da álgebra tensorial para a estimativa de fluxo ótico em sequência de imagens coloridas

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação do Departamento de Informática, Centro de Tecnologia da Universidade Estadual de Maringá, como requisito parcial para obtenção do título de Mestre em Ciência da Computação.

Orientador: Prof. Dr. Franklin César Flores

Maringá
2018

Dados Internacionais de Catalogação na Publicação (CIP)
(Biblioteca Central - UEM, Maringá, PR, Brasil)

I791 Ishii, Fernanda Tamy
Lucas-Kanade Tensorial: aplicação da álgebra tensorial para a estimativa de fluxo ótico em sequência de imagens coloridas / Fernanda Tamy Ishii. -- Maringá, 2018.
56 f. : il. color., figs., tabs.

Orientador: Prof. Dr. Franklin César Flores.
Dissertação (mestrado) - Universidade Estadual de Maringá, Centro de Tecnologia, Departamento de Informática, Programa de Pós-Graduação em Ciência da Computação, 2018.

1. Processamento de imagem colorida. 2. Processamento de vídeo. 3. Álgebra tensorial. 4. Estimativa de Fluxo Ótico. 5. Método de Lucas Kanade. 6. Morfologia (matemática). 7. Gradiente morfológico tensorial. I. Flores, Franklin César, orient. II. Universidade Estadual de Maringá. Centro de Tecnologia. Departamento de Informática. Programa de Pós-Graduação em Ciência da Computação. III. Título.

CDD 21.ed. 006.37

Mariza Nogami
CRB 9/1569

FOLHA DE APROVAÇÃO

FERNANDA TAMY ISHII

Lucas-Kanade Tensorial: aplicação da álgebra tensorial para a estimativa de fluxo ótico em sequência de imagens coloridas

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação do Departamento de Informática, Centro de Tecnologia da Universidade Estadual de Maringá, como requisito parcial para obtenção do título de Mestre em Ciência da Computação pela Banca Examinadora composta pelos membros:

BANCA EXAMINADORA



Prof. Dr. Franklin César Flores
Universidade Estadual de Maringá – DIN/UEM



Prof. Dr. Ademir Aparecido Constantino
Universidade Estadual de Maringá – DIN/UEM



Profª. Dra. Leticia Rittner
Universidade Estadual de Campinas – DCA/Unicamp

Aprovada em: 19 de fevereiro de 2018.

Local da defesa: Sala 101, Bloco C56, *campus* da Universidade Estadual de Maringá.

DEDICATÓRIA

Dedico este trabalho a Deus, a minha família e aos amigos, que me acompanham desde sempre e que me apoiaram nessa missão!

AGRADECIMENTOS

Inicialmente, agradeço a Deus por fortalecer-me de modo que cada novo dia seja mais um limite ultrapassado.

Também agradeço aos meus pais, Marli Akiko Ishii (*in memoriam*) e Ronaldo Choiti Ishii, pela atenção e dedicação empenhada para que eu me tornasse uma pessoa crítica, mas capaz de ponderar sobre os fatos.

Reafirmo minha gratidão a Maria Ioshie Yamada, minha sempre "tia Maria", que me acolheu em sua própria casa na época da graduação e não me abandonou durante o mestrado, mesmo quando eu lhe proferia palavras ingratas.

Ubiquamente agradeço aos meus tios e meus primos, que sempre me proporcionaram abençoadas reuniões, boas lições, calorosas discussões e um respeito memorável, mesmo que eu tomasse decisões que não os priorizassem.

Sem sentimentalismos, agradeço a minha irmã Francine Maky Ishii, minha caçulinha sete-anos-mais-nova que eu, que ainda me mostra uma resposta não errada, mas diferente da minha para um grande universo de situações.

Também destaco a participação nesta conquista do Sérgio Grijó de Sousa Filho, meu namorado desde o início do mestrado, meu parceiro de pesquisa e estudos, meu oponente nas partidas de tabuleiro e meu companheiro nos diversos momentos.

Indubitavelmente, agradeço ao meu orientador Franklin César Flores, pela atenção dispendida, pelas reuniões no laboratório misturadas a uma porção de conversas tangentes e, principalmente, pelo carinho e pela preocupação que causei nesse período.

Nominal e individualmente, agradeço aos amigos que fiz durante o processo de tornar-me mestre: João Choma, Juliana Gonçalves, André Felipe, Fernando Freire, Nilton Queiroz, Gustavo Bennemann, Roberto Soltoski, Ricardo Nakae, Mariane Medeiros, Victor França, Renan Haertel, Lucas Helal, Bruno Cavalcante. Sou grata pela companhia nas quartas-feira de sonho da igreja, nas idas ao fim da tarde no mercadinho, pelas partidas despretensiosas de tabuleiro, pelas conversas filosóficas e pelas expressões artísticas colaborativas da sala do mestrado, que me proporcionaram açaúcares, pensamentos, diversão, risadas, distração e uma grande parceria que já me causa nostalgia.

Meus agradecimentos à Maria Inês Davanço, secretária do programa de mestrado, pelo empenho e pela competência dedicados, sem os quais nosso trabalho seria muito complicado; e também pela atenção e pelas conversas no meio da tarde, que tornaram nossa jornada mais agradável.

E não poderia deixar de agradecer a o programa de pós graduação em ciência da computação da Universidade Estadual de Maringá (PCC - UEM), ao Conselho

Nacional de Desenvolvimento Científico e Tecnológico (CNPq), a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e à própria Universidade Estadual de Maringá (UEM), que me proporcionaram, nas devidas razões, a oportunidade e os recursos para tornar-me mestre, bem como as marcantes experiências e recordações que levarei para toda a vida.

*Muito Obrigada por permanecerem comigo durante esta jornada
e por tornar possível essa conquista!*

Maringá, 17 de Janeiro de 2018.

Lucas-Kanade Tensorial: Aplicação da álgebra tensorial para a estimativa de fluxo ótico em sequência de imagens coloridas

RESUMO

O aspecto multi-canal das imagens coloridas e as características de movimento oriundas do sequenciamento de imagens constituem dados fundamentais para a análise semântica de um vídeo colorido, bem como fatores que desafiam o seu processamento. O primeiro instiga pesquisadores a definir medidas de distância entre atributos de cor, uma vez que a ordenação desses não é natural; enquanto o segundo impõe desafios em sua detecção e sua estimação.

Por outro lado, a teoria da álgebra de tensores se apresenta bastante consistente e robusta para a modelagem matemática de diversos problemas, desde a manipulação de grandezas escalares, que independem de orientação; à estruturas maiores e mais complexas.

Rittner et al. (2010) apresentam uma abordagem satisfatória para o cálculo do gradiente aplicado à segmentação de imagens coloridas utilizando um tensor de segunda ordem. Este trabalho tem como objetivo propor uma extensão do trabalho de Rittner et al. (2010) para a estimativa de fluxo ótico para sequências de imagens coloridas. O método foi construído com base no método Lucas-Kanade (Lucas e Kanade, 1981).

O método foi comparado com o método Lucas-Kanade avaliando-se em relação ao padrão ouro por dois critérios de qualidade distintos (Erro angular e Erro absoluto). Resultados experimentais mostraram que o método proposto apresentaram erros médios menores que o método de referência.

Palavras-chave: Processamento de imagem colorida. Processamento de vídeo. Álgebra tensorial. Estimativa de fluxo ótico. Método de Lucas-Kanade. Morfologia matemática. Gradiente morfológico tensorial.

Tensorial Lucas Kanade Application of a tensorial algebra for estimating the optical flow for color image sequence

ABSTRACT

The multi-channel aspect of color images and the motion feature present in a color image sequence are important data for its semantic analysis, as well as challenging elements. While the previous instigates researchers to define distance or dissimilarity measures between two colors; the latter challenges in detection and its estimation.

On the other hand, the theory of tensor algebra is consistent and robust to model and solve several problems. Rittner et al. (2010) provides a good approach to calculating the tensorial morphological gradient (TMG) applied to the color image segmentation using a second-rank tensor.

This document proposes to extend the TMG by presenting a optical flow estimation method, called tensorial Lucas Kanade method, based on Lucas Kanade method and using a second order tensor to manipulate color information.

The optical flow estimated by proposed method was compared to a ground truth OF and evaluated by two different criteria. Experimental results show that proposed method have had smaller average error than the LK method.

Keywords: Color image processing. Video processing. Tensor algebra. Optical flow estimation. Lucas Kanade Method. Mathematical Morphology. Tensorial morphological gradient (TMG).

LISTA DE FIGURAS

Figura 2.1	Exemplos de elementos estruturantes	19
Figura 2.2	Sistema de cores RGB	21
Figura 2.3	Representação gráfica do sistema de cores HSL	22
Figura 2.4	Representação elíptica de um tensor de segunda ordem	25
Figura 2.5	Representação tensorial de cores	26
Figura 2.6	Sequência de imagens e ilustração do campo de movimento	28
Figura 2.7	Movimento de alguns <i>pixels</i> em sequência de imagens coloridas	29
Figura 2.8	Problema de abertura	30
Figura 2.9	Técnica de estimativa de fluxo ótico Lucas-Kanade	32
Figura 3.1	Técnica de estimativa de fluxo ótico Lucas-Kanade Tensorial	34
Figura 3.2	Elementos estruturantes utilizados no método LK tensorial	34
Figura 3.3	Filtro Laplaciano com $w = 4$	35
Figura 4.1	Exemplos de pares de imagens utilizadas nos experimentos	43

LISTA DE TABELAS

Tabela 2.1	Funções de distância para o gradiente morfológico tensorial . .	27
Tabela 4.1	Síntese dos experimentos	44
Tabela 4.2	Resultados do primeiro experimento	46
Tabela 4.3	Resultados do segundo experimento	47
Tabela 4.4	Resultados do terceiro experimento	48
Tabela 4.5	Resultados do quarto experimento	48
Tabela 4.6	Resultados do quinto experimento	49
Tabela 4.7	Resultados do sexto experimento	49
Tabela 4.8	Resultados do sétimo experimento	50

LISTA DE SIGLAS E ABREVIATURAS

CG: Computação Gráfica

EA: Erro Angular

EAbs: Erro Absoluto

FO: Fluxo Ótico

GMT: Gradiente Tensorial Morfológico (*Tensorial Morphological Gradient*)

HSL: Sistema de cores HSL (*Hue/Matiz, Saturation/Saturação, Lightness/Luminosidade*)

LK: Método Lucas-Kanade

LKT: Método Lucas-Kanade Tensorial

LUV: Sistema de cores LUV (ou CielUV)

PB: Imagens em preto e branco

MM: Morfologia Matemática

PDI: Processamento Digital de Imagens

PNL: Programação Não Linear

RGB: Sistema de cores RBG (*Red/Vermelho, Green/Verde, Blue/Azul*)

VC: Visão Computacional

LISTA DE SÍMBOLOS E OPERAÇÕES

Símbolos Gerais

\mathbb{N} : números naturais

\mathbb{Z} : números inteiros

\mathbb{Q} : números racionais

\mathbb{R} : números reais

$[\cdot]^T$: Matriz Transposta

$[\cdot]^{-1}$: Matriz Inversa

I : Matriz Identidade

T_2 : Tensor de segunda ordem ou tensor bidimensional

T_3 : Tensor de terceira ordem ou tridimensional

$D(T_i, T_j)$: Medida de dissimilaridade entre os tensores T_i e T_j

$\vee(\cdot)$: Função supremo, função máximo

Tensor bidimensional

λ_i : i-ésimo autovalor do tensor

λ : conjunto de autovalores do tensor

θ : medida angular do autovalor principal em relação ao eixo de referência x .

ϕ : medida angular do autovalor em relação ao eixo de referência y .

e_i : i-ésimo autovetor do tensor

e : conjunto de autovetores do tensor

t_{xy} : elemento da matriz posicionado na x -ésima linha e y -ésima coluna

T_i : Tensor de i -ésima ordem

SUMÁRIO

1	Introdução	15
1.1	Motivação e justificativa	16
1.2	Objetivos e contribuições esperadas	17
1.3	Organização do documento	17
2	Conceitos Preliminares	18
2.1	Imagem digital	18
2.2	Cor	20
2.3	Modelo de Cores	20
2.3.1	<i>RGB</i>	21
2.3.2	<i>HSL</i>	22
2.3.3	<i>LUV</i>	23
2.4	Tensor	24
2.4.1	<i>Definição matemática</i>	24
2.4.2	<i>Representação Tensorial de Cor</i>	25
2.5	Gradiente morfológico tensorial	26
2.6	Processamento de vídeo e fluxo ótico	28
2.6.1	Método Lucas-Kanade	30
3	Método Lucas-Kanade Tensorial	33
3.1	Definição do método	33
4	Resultados experimentais	37
4.1	Ambiente de execução dos experimentos	37
4.2	Imagens utilizadas no experimento	38
4.3	Critérios de qualidade	38
4.4	Filtro de realce	39
4.5	Procedimento experimental	40
4.5.1	<i>Experimento I</i>	40
4.5.2	<i>Experimento II</i>	40
4.5.3	<i>Experimento III</i>	41
4.5.4	<i>Experimento IV</i>	41
4.5.5	<i>Experimento V</i>	41
4.5.6	<i>Experimento VI</i>	41
4.5.7	<i>Experimento VII</i>	42

4.6	Análise dos resultados experimentais	45
5	Considerações Finais	51
5.1	Principais Resultados e Contribuições	51
5.2	Limitações deste trabalho	52
5.3	Trabalhos futuros	52
	REFERÊNCIAS	54

Introdução

Com a popularização de equipamentos de registro de vídeo mais acessíveis, tais como câmeras digitais e aparelhos celulares do tipo *smartphone*, uma quantidade notável e crescente de material audiovisual tem sido produzido e distribuído por meio da Internet.

Concomitantemente a popularização de equipamentos de vídeo, o avanço tecnológico reduziu também os custos para armazenamento de dados, tornando possível e viável a adoção de outros formatos de mídia diferentes das imagens em preto e branco.

Nesse cenário, estudos em processamento digital de imagens (PDI) e visão computacional (VC) têm ganhado notoriedade, uma vez que estes se propõem a, dentre outros objetivos, extrair e avaliar atributos e características de elementos do mundo físico e a compreendê-los com base em informações visuais, tal como imagens e sequências de imagens (vídeos).

No processamento digital de imagens e, mais especificamente, em processamento de vídeos coloridos, ou sequência de imagens coloridas, um dos grandes desafios em destaque é a segmentação adequada de vídeo, que consiste em dividir o vídeo em regiões semânticas compatíveis para a compreensão do mundo físico.

Isso é decorrente da dificuldade de mensurar a diferença entre as cores de dois pontos do vídeo e o significado dessa diferença. Quando os dois pontos pertencem a mesma imagem e apresentam distinção de cores humanamente percebíveis, é possível que estes pertençam a regiões que representem elementos distintos. Quando os pontos pertencem a uma mesma localização em quadros distintos do mesmo vídeo, pode ser indicador de movimento.

No processamento de imagens coloridas, parece adequada a representação de cores por meio de sistemas que particionam o dado da cor em partições ou canais específicos, pois a relação de ordem entre cores não é tão intuitiva e mensurável como na medição da intensidade do tom de cinza em imagens não coloridas, isto é, a definição de um critério de diferença entre duas cores não é tão natural quanto a aferição da diferença de intensidade de luz existente entre duas tonalidades de cinza.

Isso é um desafio para a construção de um gradiente para imagens coloridas, visto que sua própria definição considera a diferença entre os tons de cinza da vizinhança de um pixel de uma imagem e a equivalência dessa operação para cores não é tão simples.

Algumas outras características, tais como a existência de ruídos, iluminação não uniforme, sombreamento e textura, agravam ainda mais a segmentação de imagem pois podem inserir nas imagens informações que dificultam a separação dos objetos.

Já para a avaliação de movimento em sequência de imagens, um dos maiores desafios é encontrado na estimativa do fluxo ótico, que é a velocidade da imagem. Por meio de padrões de espaço e tempo de intensidade das imagens, é possível determinar uma aproximação do campo de movimento, isto é, especificar quanto cada pixel de uma imagem se moveu entre um quadro e seu subsequente em uma projeção bidimensional.

Ainda que a segmentação de vídeo compartilhe com a segmentação de imagens algumas dificuldades, tais como a existência de ruídos e/ou iluminação não uniforme, ela também apresenta suas dificuldades próprias, como a existência de movimentos bruscos, superfícies oclusas e objetos deformáveis.

1.1 Motivação e justificativa

Muito esforço tem sido dispendido para desenvolver novas técnicas e estratégias de segmentação de vídeo com a intenção de superar as dificuldades inerentes e atender às necessidades dos diversos sistemas que aplicam a segmentação.

Dentre as variadas aplicações válidas para a segmentação de vídeo, são destacadas: a recuperação e o reconhecimento de informações existentes em vídeos, na criação de animações e renderizações não realistas, em sistemas para compressão de vídeos.

Sob a perspectiva morfológica, tensores são estruturas algébricas consistentes e bastante robustas para trabalhar com grandezas escalares e grandezas vetoriais, que requerem informações de orientação e sentido. Sua aplicação para a segmentação de imagens coloridas, conforme o gradiente morfológico tensorial (GMT) proposto por Rittner et al. (2010), que utilizou uma representação de cores baseada em tensores,

permitiu-lhes estabelecer um critério de semelhança entre cores de *pixels* vizinhos, ensejando-lhe uma solução compatível ao esperado.

Nesse sentido, uma extensão do GMT apresenta-se como uma conveniente possibilidade para definir critérios de segmentação de vídeo, considerando as informações de cor.

1.2 Objetivos e contribuições esperadas

Como proposta para a dissertação, almeja-se neste trabalho, estender o GMT, proposto por Rittner et al. (2010), de modo a ser aplicado na estimativa de fluxo ótico de sequências de imagens coloridas, por meio da adoção de uma estrutura tensorial; e avaliar o método construído.

Dessa maneira, é esperado neste trabalho que um método baseado no GMT seja elaborado e avaliado para permitir estimativa de FO em sequências de imagens coloridas.

1.3 Organização do documento

Este documento foi estruturado conforme o desenvolvimento da pesquisa, partindo-se de uma visão mais ampla (contexto da pesquisa) até a proposta da pesquisa.

Dessa maneira, o texto foi organizado como segue: A seção 2 apresenta uma revisão de literatura sobre os principais assuntos abordados, mostrando os conceitos preliminares e algumas definições necessárias para a compreensão da proposta deste trabalho. Na seção 3 é apresentado o método proposto neste trabalho. Os experimentos conduzidos, os resultados obtidos e a análise realizada são apresentados na seção 4. As considerações finais, considerando as contribuições deste trabalho, as limitações encontradas e trabalhos futuros possíveis são apresentadas na seção 5.

Conceitos Preliminares

Esta seção objetiva apresentar os conceitos e definições fundamentais de processamento de imagens, bem como buscando apresentar os fundamentos essenciais que permeiam a teoria de Tensores.

Conceitos como o de imagem e cores, naturalmente compreendida pelas pessoas, mas com complexa descrição, formam parte do alicerce usado no desenvolvimento do trabalho produzido nesta dissertação e, por este motivo, são apresentados nas seções que seguem.

2.1 Imagem digital

Uma imagem digital monocromática é uma discretização de um objeto ou cena real, que pode ser definida como sendo um função bidimensional $f(x, y) \in \mathbb{Z}$, tal que $x \in Z \subset \mathbb{Z}$ e $y \in Z \subset \mathbb{Z}$ são coordenadas espaciais e a grandeza de f em cada par de coordenadas (x, y) é a intensidade ou nível de cinza da imagem naquele ponto. (Gonzalez e Woods, 2008)

O menor elemento de uma imagem digital, isto é, cada elemento do contra-domínio da função $f(x, y)$ localizada na coordenada $((x, y))$ é denominada pixel.

Os *pixels* de coordenadas $(x - 1, y - 1)$, $(x - 1, y)$, $(x - 1, y + 1)$, $(x, y - 1)$, $(x, y + 1)$, $(x + 1, y - 1)$, $(x + 1, y)$ e $(x + 1, y + 1)$ são chamados de vizinhos de primeira ordem ou adjacentes. Os $(x - 1, y)$, $(x + 1, y)$, $(x, y - 1)$ e $(x, y + 1)$ são chamados de vizinhos mais próximos e os pontos $(x - 1, y - 1)$, $(x - 1, y + 1)$, $(x + 1, y - 1)$ e $(x + 1, y + 1)$ são ditos vizinhos mais distantes.

Imagens digitais binárias, também conhecidas como imagens em preto e branco, são as imagens mais simples que existem e são definidas pela função $f(x, y) \in \{0, 1\} \subset \mathbb{Z}$.

Imagens em tons de cinza podem ser definidas com uma simples expansão da definição das imagens binárias, pela função $f(x, y) \in [0, 2^k)$, em que k é a resolução radiométrica do sensor que estabelece a quantidade de bits necessárias para armazenar a informação de cinza do pixel.

Imagens coloridas, por sua vez, pode ser definida pela extensão da imagem em tons de cinza com o aumento da dimensão do contra-domínio. Isso significa que a função que a define é do tipo $f(x, y) \in [0, 2^k)^n$ em que n indica o número de bandas requeridas para armazenar a informação de cor do pixel. Usualmente, os modelos de representação de cores utilizam $n = 3$.

Em morfologia matemática, Najman e Talbot (2013) definem um elemento estruturante como sendo um conjunto de pixels, de tamanho definido e conhecido, comumente representado por uma imagem binária ou em tons de cinza, que podem definir uma relação de vizinhança ou conectividade entre os pixels.

A Figura 2.1 mostra alguns exemplos de elementos estruturantes de tamanho 3×3 . A Figura 2.1(a) é o elemento estruturante quadrado, também conhecido como *flat*, em que todos os pontos apresentam valores de intensidade iguais a 1; a Figura 2.1(b) é chamado de elemento estruturante X, em que apenas as diagonais principais e secundárias da imagem recebem o valor 1 e Figura 2.1(c) é chamado de diamante por ter esse aspecto semelhante a um losango.

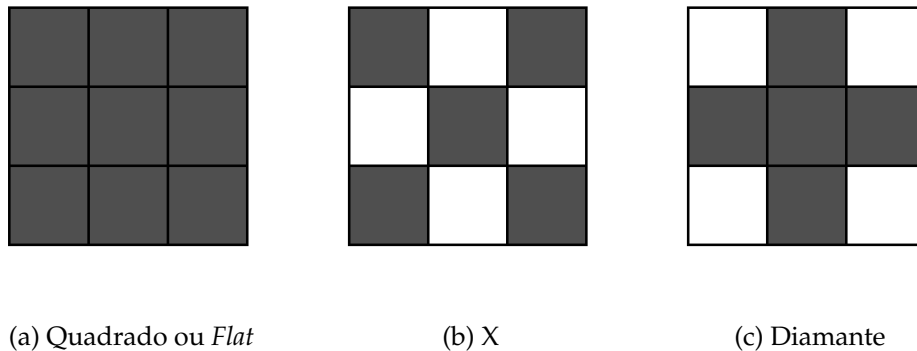


Figura 2.1: Exemplos de elementos estruturantes

Comumente, uma maneira para representar uma imagem é por meio de uma matriz bidimensional discreta e finita, de ordem $N \times M$, cujos elementos representam

o pixel, isto é, apresentam valores numéricos quando referenciam imagens binárias ou monocromáticas; e apresentam uma tupla de valores, geralmente de tamanho 3, ao tratar-se de imagens coloridas.

$$f(x, y) = \begin{pmatrix} f(1,1) & f(1,2) & \cdots & f(1,M) \\ f(2,1) & f(2,2) & \cdots & f(2,M) \\ \vdots & \vdots & \ddots & \vdots \\ f(N,1) & f(N,2) & \cdots & f(N,M) \end{pmatrix} \quad (2.1)$$

Gonzalez e Woods (2008) afirmam que dois *pixels* adjacentes são conexos se apresentam o mesmo nível de cinza ou se satisfazem a algum outro critério de similaridade.

2.2 Cor

Trabalhar com cores têm se mostrado uma atividade árdua, principalmente no que relaciona à visão computacional, dados os diversos significados e/ou sentimentos que as cores despertam nos seres humanos. Em processamento de imagens, uma das maiores dificuldades é encontrada na representação computacional de cores, que apresenta uma modelagem difícil.

De maneira intuitiva, é fácil perceber que o estímulo causado por uma luz sem cor, ou acromática, é bem mais modesta que o causado pela luz colorida. Por outro lado, das teorias de ótica e ondulatória, é sabido que a cor de um dado objeto é determinado pelo comprimento de onda eletromagnética que emana deste, isto é, pela luz que é emitida ou refletida por ele.

A cor é associada aos diferentes comprimentos de onda do espectro eletromagnético, sendo que a cor de um material é determinada pelo comprimento de onda dos raios luminosos que suas moléculas constituintes refletem. Um objeto terá determinada cor se não absorver justamente os raios correspondentes à frequência de onda desta cor. (Gonzalez e Woods, 2008)

Na próxima seção serão apresentados alguns sistemas de cores e suas características.

2.3 Modelo de Cores

Gonzalez e Woods (2008) afirmam que um modelo de cores, também chamado de sistema de cores, ou ainda espaço de cores, é um subespaço formado por um sistema de coordenadas em que cada cor é representado por um único ponto neste subespaço; de

modo a permitir a especificação de uma cor em um determinado formato padronizado, específico e aceito por todos.

Busin et al. (2008) afirmam que um espaço de cores é uma representação geométrica das cores que comumente permitem a sua identificação por meio de uma estrutura composta por três componentes.

Existem diversos modelos de cores na literatura, conforme Gonzalez e Woods (2008) e Marques Filho e Vieira Neto (1999), sendo RGB, CMY, CMYK, HSI, HSV e HSL os mais utilizados no processamento digital de imagens. Para este trabalho, serão apresentados os modelos de cores: RGB, HSL e LUV, por serem modelos bastante conhecidos na literatura e por apresentarem características específicas de cada sistema.

2.3.1 RGB

O modelo RGB é baseado num sistema de três coordenadas cartesianas, em que cada componente representa a intensidade de uma cor primária. Neste sistema, as cores primárias são: o vermelho (*red*), o verde (*green*) e o azul (*blue*). A Figura 2.2 ilustra o modelo RGB em um sistema cartesiano em que cada eixo representa cada uma das três cores primárias do sistema.

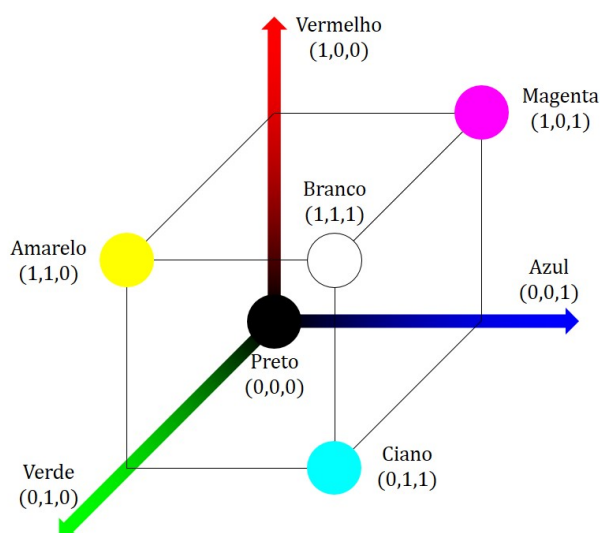


Figura 2.2: Sistema de cores RGB

Algumas vezes, por conveniência, assume-se que os valores máximos de R, G e B estão normalizados entre 0 e 1. (Marques Filho e Vieira Neto, 1999). Quando os valores de R, G e B são nulos, a cor resultante é preta; e, quando todos os valores são máximos, a cor resultante é branca.

O modelo RGB é bastante utilizado por câmeras, monitores de computador e telas de televisão.

2.3.2 HSL

O sistema de cores HSL classifica uma determinada cor em três componentes: o matiz (*hue*), a saturação (*saturation*) e a luminosidade (*lightness*). Isto é interessante pois essa classificação é mais natural que a do RGB, em comparação com a percepção do olho humano. (Gonzalez e Woods, 2008).

O HSL utiliza uma coordenada cilíndrica formada por um eixo que mensura a luminosidade, outro que mede a saturação e um valor angular que define o matiz, como ilustrado na Figura 2.3. (Busin et al., 2008)

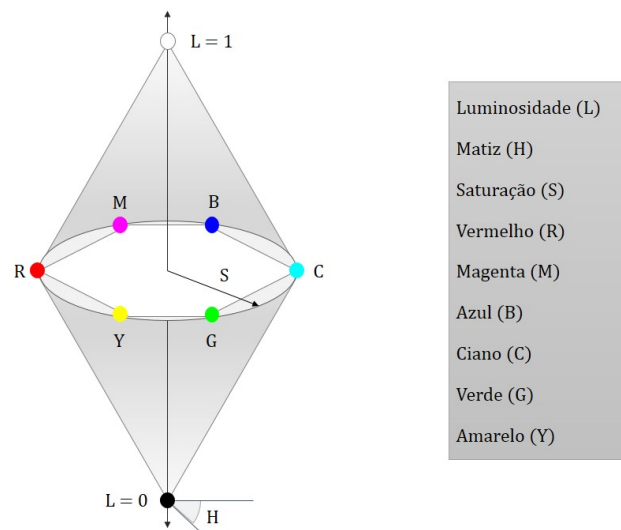


Figura 2.3: Representação gráfica do sistema de cores HSL

A conversão de uma cor do modelo RGB para HSL pode ser determinada pelas equações apresentadas entre 2.2 à 2.6.

$$M = \max(R, G, B) \quad (2.2)$$

$$m = \min(R, G, B) \quad (2.3)$$

$$H = \begin{cases} 60 \cdot \frac{G-B}{M-m}, & \text{se } M = R \\ 60 \cdot \frac{B-R}{M-m} + 120, & \text{se } M = G \\ 60 \cdot \frac{R-G}{M-m} + 240, & \text{se } M = B \end{cases} \quad (2.4)$$

$$S = \begin{cases} 0, & \text{se } M = m \\ \frac{M-m}{M+m}, & \text{se } 0 < L \leq 0,5 \\ \frac{M-m}{2-(M+m)}, & \text{se } L > 0,5 \end{cases} \quad (2.5)$$

$$L = \frac{M+m}{2} \quad (2.6)$$

2.3.3 LUV

Luv, ou também CieLUV, é um modelo de cores proposto pela *Commission Internationale d'Éclairage* em 1976. Luv é um sistema de cores que propõe-se a estabelecer a correspondência direta entre a diferença de cores percebidas por um observador humano e a distância medida nesse espaço de cores. (Busin et al., 2008)

Cheng et al. (2001) afirmam que o sistema Luv é definido por meio da transformação não linear de três valores primários (X, Y e Z), que podem ser obtidos pela transformação linear do sistema RGB. A matriz de transformação de RGB para XYZ é determinada conforme a 2.7

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 0,607 & 0,174 & 0,200 \\ 0,299 & 0,587 & 0,114 \\ 0,000 & 0,066 & 1,116 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad (2.7)$$

E o sistema Luv é definido de acordo com as equações 2.8, 2.9 e 2.10

$$L^* = 116 \times \sqrt[3]{Y/Y_0} - 16 \quad (2.8)$$

$$u^* = 13 \times L^*(u' - u_0) \quad (2.9)$$

$$v^* = 13 \times L^*(v' - v_0) \quad (2.10)$$

Sendo $Y/Y_0 > 0.01$ e (Y_0, u_0, v_0) a coordenada padrão para o branco e:

$$u' = \frac{4X}{X + 15Y + 3Z} \quad (2.11)$$

$$v' = \frac{6Y}{X + 15Y + 3Z} \quad (2.12)$$

De acordo com Schwiegerling (2004), a distância ΔE entre duas cores no sistema LUV é dado por:

$$\Delta E = \sqrt{(L_2 - L_1)^2 + (u_2 - u_1)^2 + (v_2 - v_1)^2} \quad (2.13)$$

2.4 Tensor

De maneira sintética, um tensor é uma representação matemática que generaliza conceitos bastante aplicados na física, tais como escalares e vetores. Valores ou grandezas escalares são os valores que indicam quantidades que não se altera perante uma operação de transformação de coordenadas; enquanto valores vetoriais são aqueles que, além de sua magnitude, também são caracterizados por sua direção e seu sentido. (Walker et al., 2009)

A subseção 2.4.1 apresenta a definição matemática dos tensores e suas principais propriedades.

2.4.1 Definição matemática

Segundo Kingsley (2006), um tensor de segunda ordem é uma estrutura algébrica normalmente representada por uma matriz 2×2 simétrica T_2 :

$$T_2 = \begin{bmatrix} t_{xx} & t_{xy} \\ t_{xy} & t_{yy} \end{bmatrix} \quad (2.14)$$

e pode ser decomposto em autovalores e autovetores, de acordo com a equação característica:

$$(T_2 - \lambda.I)e = 0 \quad (2.15)$$

em que λ são os autovalores do tensor, I é a matriz identidade e e são os autovetores normalizados. Como T_2 é uma matriz simétrica, seus autovalores são reais e os autovetores correspondentes são linearmente independentes. Dessa maneira, um tensor de segunda ordem pode ser representado por uma elipse, conforme apresentado na 2.4.1, formada por dois eixos perpendiculares entre si e proporcionais aos autovalores λ_1 e λ_2 , sendo $\lambda_1 \geq \lambda_2$ e orientada pelos respectivos autovetores e_1 e e_2 .

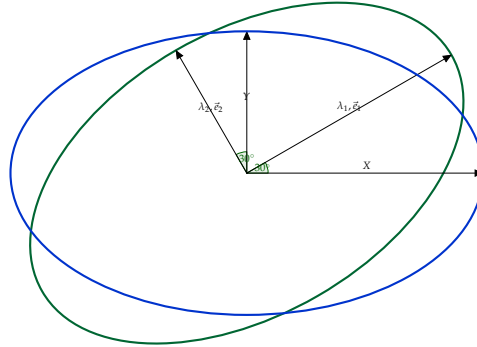


Figura 2.4: Representação elíptica de um tensor de segunda ordem

Rittner et al. (2010) apresentam alguns atributos das elipses que auxiliam na descrição dos tensores correspondentes. Dentre eles, destacam-se:

1. **Excentricidade** ou **forma**: Trata-se da proporção entre os autovalores de um tensor

$$\text{Excentricidade} = \lambda_2 / \lambda_1 \quad (2.16)$$

2. **Traço**: Também chamado de *Trace*, é a medida resultante da soma dos autovalores do tensor.

$$\text{Trace} = \lambda_2 + \lambda_1 \quad (2.17)$$

3. **Direção**: Medida angular da elipse em relação ao eixo de referência. É determinado pelo autovetor principal do tensor.

2.4.2 Representação Tensorial de Cor

De acordo com Rittner et al. (2007), uma cor pode ser representada de maneira tensorial utilizando-se os atributos que descrevem as elipses e seus respectivos tensores de segunda ordem e associando-os aos atributos do modelo de cor HSL.

Considere o semieixo maior, o semieixo menor e o ângulo de rotação de uma elipse como sendo seu atributos. A razão entre os semi-eixos da elipse é a sua excentricidade ou forma e pode ser calculada pelos autovalores do tensor correspondente pela equação 2.16.

Similarmente, a soma de seus semi-eixos é chamada de traço (ou *trace*) e pode ser determinada pela equação 2.17 e o ângulo de rotação pode ser determinado pela direção do autovetor principal de seu tensor.

Avaliando-se, em relação à sua definição, pode-se observar que os atributos excentricidade e traço são valores escalares enquanto o atributo de rotação é angular. De forma semelhante, ao analisar os atributos do modelo HSL também é possível observar que esse é composto por dois atributos escalares e um angular.

Assim, associando-se os atributos angulares da elipse (Ângulo de rotação ou direção do principal autovetor do tensor) e do modelo HSL (Matiz), e os atributos escalares da elipse (Forma e Traço) e do HSL (Saturação e Luminosidade), tem-se a definição da representação tensorial de cor, respectivamente. A figura Figura 2.5 mostra uma a elipse associada a cada *pixel* de uma imagem.



Figura 2.5: Representação tensorial de cores

2.5 Gradiente morfológico tensorial

Rittner et al. (2010) propuseram uma nova representação de cor, associando o sistema de cores HSL a um tensor, na intenção de definir uma medida de similaridade entre cores, utilizando-se algumas métricas para descrição de tensores.

Para isso, realizaram experimentos com alguns sistemas de cores e encontraram um melhor ajuste com o sistema HSL, associando as componentes matiz, saturação e luminosidade, respectivamente, a excentricidade, *shape* e *trace*.

Definida a representação de cores, Rittner et al. (2010) calcularam seis medidas de distância entre cores, conforme mostrado na Tabela 2.1.

Tabela 2.1: Funções de distância para o gradiente morfológico tensorial

Medida de Distância	Nome	Fórmula Matemática
$d_1(T_i, T_j)$	Produto escalar	$ e_{1,i} \cdot e_{1,j} $
$d_2(T_i, T_j)$	Produto escalar tensorial	$\lambda_{1,i} \cdot \lambda_{1,j} \cdot (e_{1,i} \cdot e_{1,j})^2 + \lambda_{2,i} \cdot \lambda_{2,j} \cdot (e_{2,i} \cdot e_{2,j})^2$
$d_3(T_i, T_j)$	Distância euclidiana	$\sqrt{\text{Trace}((T_i - T_j)^2)}$
$d_4(T_i, T_j)$	Distância do tensor de difusão	$\frac{1}{2} \cdot \sqrt{\text{Trace}(T_i^{-1}T_j - T_j^{-1}T_i) - 2n}$
$d_5(T_i, T_j)$	Distância Log-Euclidiana do invariante de similaridade	$\sqrt{\text{Trace}((\log(T_i) - \log(T_j))^2)}$
$d_6(T_i, T_j)$	Distância baseada na geometria de Riemann	$\sqrt{\text{Trace}(\log(D_{ij})^2)}$, com $D_{ij} = T_i^{-1/2} \cdot T_j \cdot T_i^{-1/2}$

Para um conjunto $E = \mathbb{Z} \times \mathbb{Z}$ que contém todos os pixels de uma imagem colorida f , o gradiente morfológico tensorial baseado na representação tensorial de cores é definido como sendo:

$$\nabla_B^T(f)(x) = \bigvee_{y,z \in B_x} d_n(T_y, T_z), \forall x \in E \quad (2.18)$$

Em que d_n representa as funções de similaridade apresentadas na Tabela 2.1, $B \subset E$ é o elemento estruturante com centro na origem de E , T_y é o tensor que representa a cor em y , T_z é o tensor que representa a cor em z , tal que y e z estão na vizinhança de x , de acordo com B_x .

Como as funções de similaridade já realizam a comparação entre os pontos, ∇_B^T é o maior valor calculado entre as funções de similaridade.

2.6 Processamento de vídeo e fluxo ótico

Em um vídeo, ou em uma sequência de imagens, informações importantes podem ser obtidas analisando-se as diferenças entre as imagens. Tais diferenças são, principalmente, atribuídas ao movimento do objeto e/ou da câmera. Assim, é possível determinar, por exemplo, quantos objetos existem na imagem e em qual velocidade e direção cada objeto se move.

Os movimentos dos objetos e da câmera percebidos no decorrer do vídeo geram velocidades tridimensionais associadas. A velocidade tridimensional de um objeto que se move em frente a uma câmera é percebida pela mudança correspondente na imagem.

Para que o estudo do movimento de objetos em vídeo seja possível, calcula-se uma projeção da velocidade tridimensional de cada ponto da superfície observada sobre a superfície da imagem, resultando em uma aproximação do campo de movimento (*motion field*), como mostrado na figura Figura 2.6

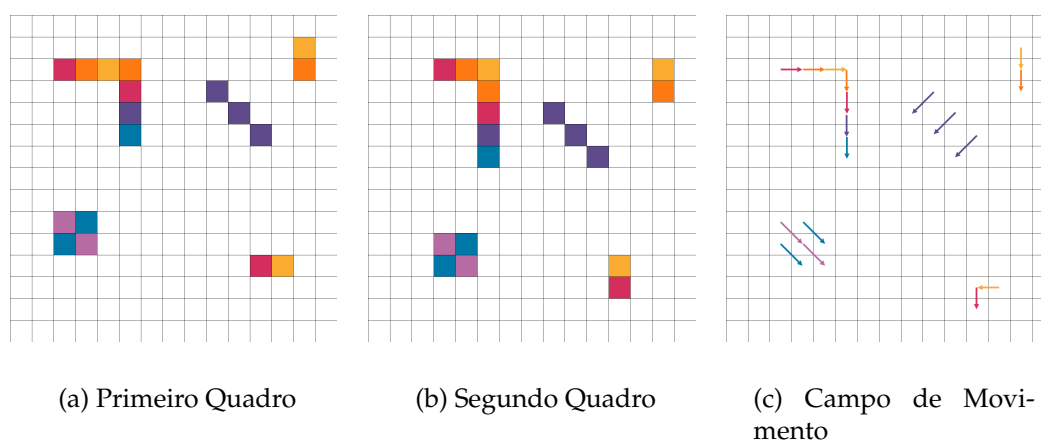


Figura 2.6: Sequência de imagens e ilustração do campo de movimento

De acordo com Horn e Schunck (1981), o fluxo ótico de uma sequência de imagens é determinado pelo padrão de intensidade luminosa. De modo sintético, o fluxo ótico é o campo que identifica a variação (o movimento) encontrada entre duas imagens consecutivas de uma sequência em vídeo.

Comumente, tal campo é descrito por meio de vetores de deslocamento, que pode ser estimado para cada *pixel* ou para um conjunto deles. Cada fluxo ótico estimado possui uma magnitude (ou tamanho) e uma orientação, sendo bastante adequado sua representação vetorial, por meio de um diagrama de agulhas (*needle map*).

Se for conhecido o intervalo de tempo entre duas imagens consecutivas, os vetores da velocidade podem ser convertidos em vetores de deslocamento e vice-versa, dada a função velocidade definida como:

$$V = \frac{\Delta S}{\Delta t} \quad (2.19)$$

em que V é a velocidade, ΔS é a variação do movimento, ou deslocamento do pixel, e Δt é o tempo gasto entre duas imagens.

A figura 2.6 mostra o deslocamento de três *pixels* de uma sequência de imagens.

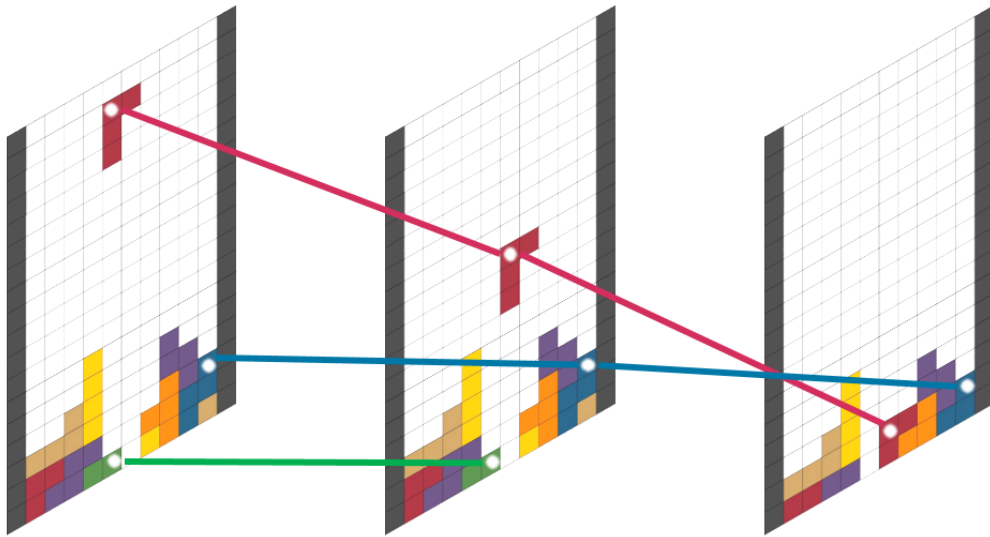


Figura 2.7: Movimento de alguns *pixels* em sequência de imagens coloridas

Para o cálculo de fluxo ótico, diversas abordagens são listadas na literatura. Dentre elas, destaca-se a abordagem baseada em gradiente, que possui como vantagem um menor custo computacional. (Ohta, 1990)

Nas técnicas diferenciais, ou baseadas em gradiente, a hipótese inicial para a computação do fluxo ótico é a de que a intensidade entre quadros distintos em uma sequência de imagens é aproximadamente constante em um intervalo Δt de tempo pequeno, isto é, em um pequeno intervalo de tempo o deslocamento será mínimo. Dessa maneira, se $I(x, y, t)$ é a intensidade da imagem no pixel (x, y) no tempo t , tomando-se dt um intervalo de tempo ínfimo entre duas imagens e considerando-se que não há variação de intensidade na imagem neste instante, tem-se:

$$I(x, y, t) = I(x + dx, y + dy, t + dt) \quad (2.20)$$

que pode ser reescrita como sendo:

$$I(x + dx, y + dy, t + dt) = I(x, y, t) + \frac{\partial I}{\partial x} dx + \frac{\partial I}{\partial y} dy + \frac{\partial I}{\partial t} dt \quad (2.21)$$

que resulta em:

$$\frac{\partial I}{\partial x} \frac{dx}{dt} + \frac{\partial I}{\partial y} \frac{dy}{dt} + \frac{\partial I}{\partial t} = 0 \quad (2.22)$$

em que $\frac{dx}{dt}$ e $\frac{dy}{dt}$ são as componentes respectivas de x e y da velocidade e $\partial I \partial x$, $\partial I \partial y$ e $\partial I \partial t$ são as derivadas parciais da imagem em (x, y, t) . Assim, a equação de restrição de fluxo ótico é enunciada como:

$$\nabla I \cdot \bar{v} + I_t = 0 \quad (2.23)$$

Como a equação 2.23 é única e apresenta duas incógnitas, conclui-se que ela não é suficiente para estimar os componentes de velocidade. Esse fato é denominado problema de abertura do algoritmo de determinação do fluxo ótico. A figura 2.6 ilustra o problema de abertura.

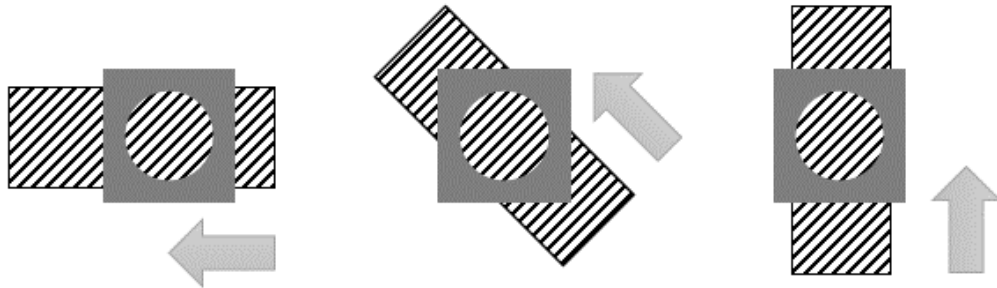


Figura 2.8: Problema de abertura

2.6.1 Método Lucas-Kanade

Uma das maneiras de tratar o problema de abertura para a determinação do fluxo ótico de uma imagem, é o método Lucas-Kanade. Trata-se de um método não iterativo que acrescenta uma nova restrição, composta por um novo conjunto de equações, permitindo a determinação do fluxo ótico (Lucas e Kanade, 1981).

Neste método, que leva o nome dos autores, assume-se que o fluxo ótico local é constante em pequenas janelas $m \times m$, com $m > 1$, e para todos os seus *pixels* ordenados de $1, \dots, n$, o conjunto de equações, mostrado em 2.24, pode ser determinado:

$$\begin{aligned} I_{x1}V_x + I_{y1}V_y &= -I_{t1} \\ I_{x2}V_x + I_{y2}V_y &= -I_{t2} \\ &\vdots \\ I_{xn}V_x + I_{yn}V_y &= -I_{tn} \end{aligned} \quad (2.24)$$

E que pode ser resolvido pelo método dos mínimos quadrados, cujo objetivo é estimar o melhor arranjo para um determinado conjunto de dados de forma a minimizar a soma dos quadrados das diferenças entre o valor estimado e os dados observados.

Além disso, o sistema de equações enunciado em 2.24 pode ser escrita de forma matricial, como sendo:

$$A \times v = B \quad (2.25)$$

em que:

$$A = \begin{bmatrix} I_{x1} & I_{y1} \\ I_{x2} & I_{y2} \\ \vdots & \\ I_{xn} & I_{yn} \end{bmatrix}, \quad (2.26)$$

$$v = \begin{bmatrix} V_x = \frac{dx}{dt} \\ V_y = \frac{dy}{dt} \end{bmatrix}, \quad (2.27)$$

$$B = \begin{bmatrix} -I_{t1} \\ -I_{t2} \\ \vdots \\ -I_{tn} \end{bmatrix} \quad (2.28)$$

A Figura 2.9 ilustra o processo aplicado pelo método de estimativa de fluxo ótico Lucas-Kanade. É preciso destacar que a imagem de entrada do método deve ser em tons de cinza, isto é, precisa ser convertida de colorida para monocromática antes da aplicação efetiva do método.

Feito isso, são calculadas as derivadas parciais da imagem, isto é, são medidas as diferenças do tom de cinza de cada pixel com o de seus vizinhos e, então, é realizado o particionamento da imagem em regiões no qual é assumido um fluxo ótico constante.

Neste trabalho, foram consideradas regiões quadradas de comprimento $w \times w$ em que $w \in \mathbb{Z}$ foi denominado como comprimento da janela.

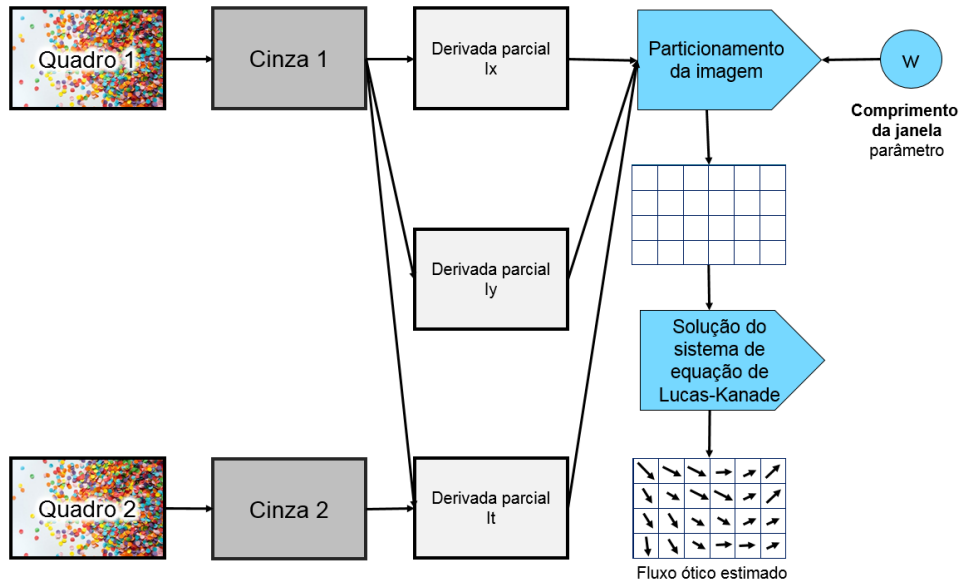


Figura 2.9: Técnica de estimativa de fluxo óptico Lucas-Kanade

Método Lucas-Kanade Tensorial

Nesta seção, é apresentada a proposta do método de estimativa de fluxo ótico denominado Lucas-Kanade tensorial, que é baseada no método de estimativa de fluxo ótico Lucas-Kanade e que utiliza uma adaptação do gradiente morfológico tensorial proposto por Rittner et al. (2010)

3.1 Definição do método

Sinteticamente, o método proposto utiliza o processo de estimativa de cálculo do FO proposto por Lucas e Kanade, com alteração no cálculo do gradiente utilizado neste cálculo, de modo a permitir a estimativa do FO utilizando informações de cor. A Figura 3.1 ilustra o método proposto.

Dessa maneira, no intuito de determinar o fluxo ótico de uma sequência de imagens coloridas, é assumido que o FO local é constante em pequenas janelas de tamanho $m \times m$, com $m > 1$, tal como é realizado no método de Lucas-Kanade.

O primeiro passo do método é a adoção da representação tensorial de cores proposto por Rittner et al. (2007), baseada no sistema de cores HSL. Feito isso, as derivadas parciais I_x , I_y e I_t utilizadas no método referência são substituídas respectivamente pelos gradientes tensoriais morfológicos $\nabla_{B_x}^T(f)$, $\nabla_{B_y}^T(f)$ e $\nabla_{B_t}^T(f)$, considerando B_x , B_y e B_t os elementos estruturantes definidos nas equações 3.1, 3.2 e 3.3; e ilustrado na Figura 3.2

$$B_x = \{T_{(i-1,j,t)}, T_{(i,j,t)}, T_{(i+1,j,t)}\} \quad (3.1)$$

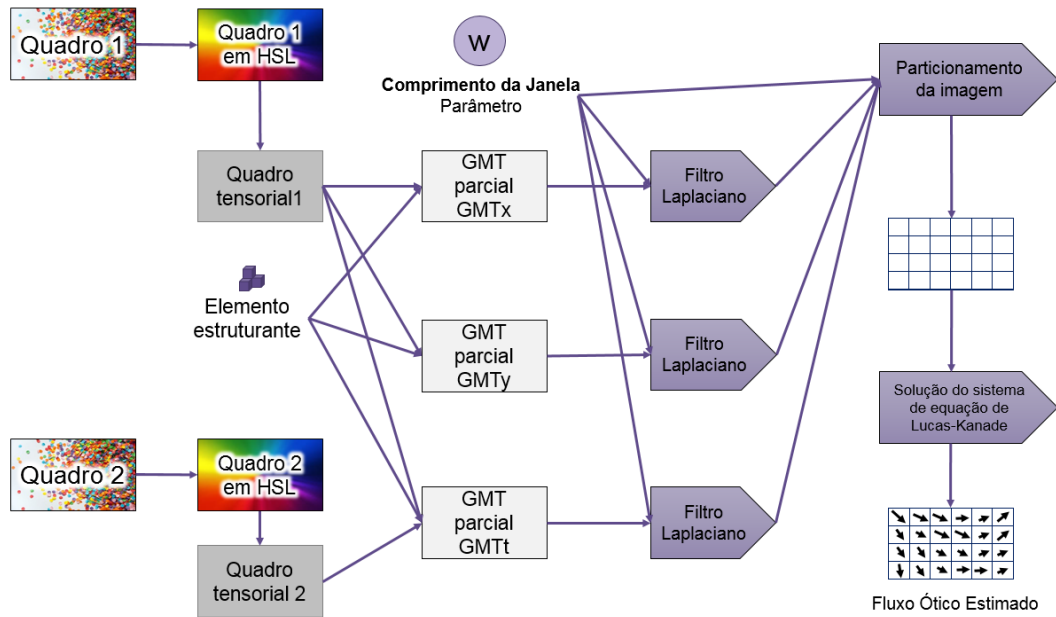


Figura 3.1: Técnica de estimativa de fluxo óptico Lucas-Kanade Tensorial

$$B_y = \{T_{(i,j-1,t)}, T_{(i,j,t)}, T_{(i,j+1,t)}\} \quad (3.2)$$

$$B_t = \{T_{(i,j,t)}, T_{(i,j,t+1)}\} \quad (3.3)$$

em que $T_{(i,j,t)}$ é a representação tensorial do *pixel* de coordenadas (x, y) no quadro t .

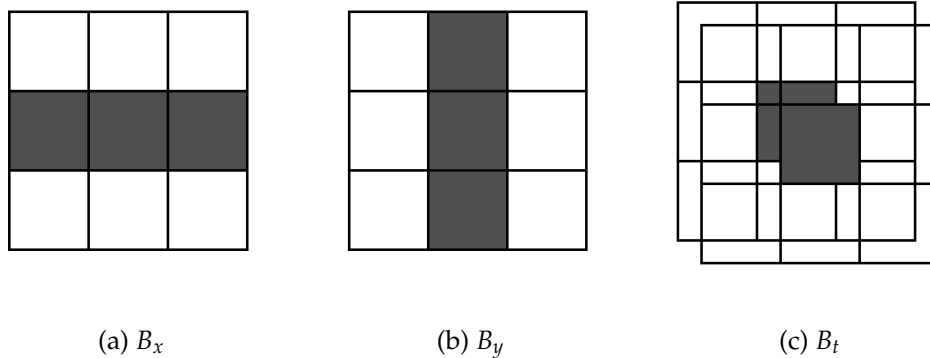


Figura 3.2: Elementos estruturantes utilizados no método LK tensorial

Conforme ilustrado na Figura 3.1, calculados os gradientes morfológicos tensoriais, o método propõe a aplicação do filtro Laplaciano L visando aguçar os gradientes previamente calculados.

A equação 3.4 mostra o filtro Laplaciano L utilizado, que é definido como um produto de um filtro Gaussiano, que confere valores mais significativos próximos da origem; e uma função sinoidal, que lhe garante o aspecto oscilatório esperado.

$$L(k) = \sin\left(\frac{\pi \times (2k - 1)}{2}\right) \times e^{\left(\frac{(-1) \times (2k - 1)^2}{2w^2}\right)} \quad (3.4)$$

em que w é o comprimento da janela quadrada utilizada como sendo a região em que o fluxo ótico estimado é o mesmo. A Figura 3.3 mostra o filtro laplaciano definido para $w = 4$.

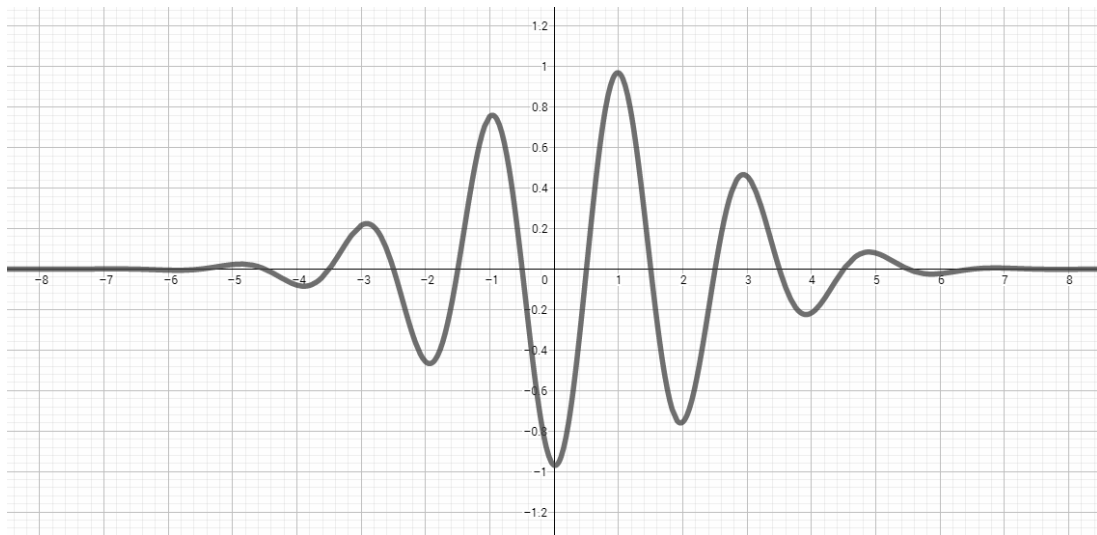


Figura 3.3: Filtro Laplaciano com $w = 4$

Feito isso, o sistema de equações resultantes no método LK tensorial é definido pela resolução do sistema $A \cdot v = B$, em que:

$$\begin{aligned}
A &= \begin{pmatrix} \nabla L_{B_x}^T(p_1) & \nabla L_{B_y}^T(p_1) \\ \nabla L_{B_x}^T(p_2) & \nabla L_{B_y}^T(p_2) \\ \nabla L_{B_x}^T(p_3) & \nabla L_{B_y}^T(p_3) \\ \dots & \dots \\ \nabla L_{B_x}^T(p_n) & \nabla L_{B_y}^T(p_n) \end{pmatrix}, \\
v &= \begin{pmatrix} u \\ v \end{pmatrix}, \\
b &= \begin{pmatrix} -\nabla L_{B_t}^T(p_1) \\ -\nabla L_{B_t}^T(p_2) \\ -\nabla L_{B_t}^T(p_3) \\ \dots \\ -\nabla L_{B_t}^T(p_n) \end{pmatrix}. \tag{3.5}
\end{aligned}$$

que é equivalente a

$$\begin{aligned}
\begin{pmatrix} u \\ v \end{pmatrix} &= \\
&\begin{pmatrix} \sum_{i=1}^n \nabla L_{B_x}^T(p_i)^2 & \sum_{i=1}^n \nabla L_{B_x}^T(p_i) \cdot \nabla L_{B_y}^T(p_i) \\ \sum_{i=1}^n \nabla L_{B_y}^T(p_i) \cdot \nabla L_{B_x}^T(p_i) & \sum_{i=1}^n \nabla L_{B_y}^T(p_i)^2 \end{pmatrix}^{-1} \\
&\times \begin{pmatrix} -\sum_{i=1}^n \nabla L_{B_x}^T(p_i) \cdot \nabla L_{B_t}^T(p_i) \\ -\sum_{i=1}^n \nabla L_{B_y}^T(p_i) \cdot \nabla L_{B_t}^T(p_i) \end{pmatrix} \tag{3.6}
\end{aligned}$$

Determinada a estrutura tensorial, a distância entre um *pixel* e seus *pixels* adjacentes será calculado por meio do supremo das distâncias de cada *pixel* vizinho, utilizando-se as medidas descritas na Tabela 2.1

Resultados experimentais

Nesta seção, são reportados os experimentos realizados com os métodos propostos nesta dissertação. Por tratar-se de um método de estimativa de fluxo ótico baseado em gradiente, o mesmo foi comparado com o método proposto por Lucas e Kanade (Lucas et al., 1981).

4.1 Ambiente de execução dos experimentos

Os algoritmos foram implementados e os experimentos foram executados em um computador com as seguintes configurações: Processador Intel i7 (2.9GHz), 16 GB de RAM, sistema operacional Windows 10 (64 bit). A implementação do método proposto bem como do método para comparação foi feita na linguagem de programação Python.

A escolha da linguagem foi devido a linguagem Python apresentar inúmeras bibliotecas que disponibilizam uma expressiva quantidade de classes, funções e métodos para a realização de diversas tarefas.

Dentre as diversas bibliotecas disponibilizadas para esta linguagem, que é livre e multiplataforma, foram utilizados neste trabalho:

- **NumPy** (*Numeric Python*): que apresenta diversos recursos para a utilização de arrays numéricos, bem como permite a execução de diversos operadores e funções matemáticas em tais estruturas.
- **OpenCV**: que apresenta inúmeras soluções voltadas para a visão computacional e o processamento digital de imagens, incluindo ferramentas para a manipulação rápida e eficaz de diversos tipos de arquivos de imagens.

Além disso, o Python é orientado a objeto e apresenta sintaxe clara e concisa, que a torna bastante produtiva.

4.2 Imagens utilizadas no experimento

Para a execução dos experimentos, foram utilizadas um conjunto de 8 pares de imagens sequenciais coloridas sintéticas, disponibilizadas por Baker et al. (2011); e um outro conjunto de 24 pares de imagens sequenciais coloridas, também sintéticas, disponibilizadas por Mac Aodha et al. (2012), totalizando 32 pares de imagens sintéticas. Em ambas as bases de dados, também estavam disponíveis o fluxo ótico referência para comparação dos métodos utilizados. A Figura 4.1 mostra alguns dos pares de imagens utilizados nos experimentos.

4.3 Critérios de qualidade

De acordo com Baker et al. (2011), um dos critérios mais comuns e conhecidos na avaliação de performance dos métodos de estimativa de fluxo ótico é o critério chamado Erro Angular (EA), proposto por Fleet e Jepson (1990) e popularizado por (Baron et al., 1994). Este critério compara o vetor de FO calculado (u, v) associado a cada pixel com o vetor de referência correspondente (u_{GT}, v_{GT}) , chamado também como *Ground Truth vector*, resultando em uma medida angular no espaço tridimensional. O EA é calculado a partir do inverso do cosseno da razão do produto escalar dos 2 vetores pelo produto dos seus comprimentos, conforme enunciado na equação 4.1.

$$EA = \cos^{-1} \left(\frac{1.0 + u \times u_{GT} + v \times v_{GT}}{\sqrt{1.0 + u^2 + v^2} \times \sqrt{1.0 + u_{GT}^2 + v_{GT}^2}} \right) \quad (4.1)$$

Baker et al. (2011) observam que erros em fluxos grandes, isto é, que apresentam vetores de maior comprimento; são menos penalizados que os erros em fluxos pequenos quando avaliados pelo EA.

Otte e Nagel (1994) observam que o critério EA considera apenas o desvio angular dos vetores, de modo que as diferenças na norma, ou tamanho, dos vetores não são consideradas e utilizam um critério chamado erro absoluto (EAbs), baseado no

comprimento absoluto da diferença entre os vetores, de acordo com o apresentado na equação 4.2.

$$EAbs = \sqrt{(u - u_{GT})^2 + (v - v_{GT})^2} \quad (4.2)$$

Neste trabalho, dado que os critérios consideravam aspectos distintos para mensurar o erro encontrado nas técnicas de estimativa utilizadas, ambos foram utilizados para a avaliação dos resultados encontrados.

4.4 Filtro de realce

De maneira sintética, a filtragem ou aplicação de filtros consiste em remover, total ou parcialmente, características indesejáveis no pós processamento de sinais eletrônicos.

No processamento de imagens digitais, a filtragem consiste em transformações *pixel* a *pixel* da imagem, que pode considerar, não apenas os valores de um *pixel* específico, mas também dos valores dos *pixels* vizinhos a este.

O processo de filtragem é comumente realizado por meio da aplicação de uma matriz, denominada máscara, sobre a imagem a ser filtrada. A aplicação da máscara com centro na posição (i, j) , sendo i o número de uma dada linha e j o número de uma dada coluna sobre a imagem, consiste na substituição do valor do pixel na posição (i, j) por um novo valor que depende dos valores dos pixels vizinhos e dos pesos da máscara, gerando uma nova imagem com a eliminação das linhas e colunas iniciais e finais da imagem original.

$$M_x = \begin{pmatrix} -1 & 8 & 0 & -8 & 1 \end{pmatrix} \quad (4.3)$$

$$M_y = \begin{pmatrix} -1 \\ 8 \\ 0 \\ -8 \\ 1 \end{pmatrix} \quad (4.4)$$

4.5 Procedimento experimental

Para avaliação das técnicas desenvolvidas, foram conduzidos alguns experimentos no ambiente descrito na seção 4.1, com as imagens descritas na seção 4.2 e os critérios de qualidade enunciados na seção 4.3. Esta seção descreve o procedimento realizado para a condução dos experimentos.

4.5.1 Experimento I

O primeiro experimento tem como objetivo comparar a acurácia de uma variante do método proposto com a do método de Lucas-Kanade. Como vizinhança do pixel para cálculo da derivada parcial I_x e I_y do método LK, foram considerados o *pixel* imediatamente a direita e o imediatamente abaixo, respectivamente.

No método TLK, foram utilizados os elementos estruturantes B_x , B_y e B_t , como apresentado nas equações 4.5, para calcular o TMG nas direções x , y e t , respectivamente.

$$\begin{aligned} B_x &= \{T_{(i,j,t)}, T_{(i+1,j,t)}\} \\ B_y &= \{T_{(i,j,t)}, T_{(i,j+1,t)}\} \\ B_t &= \{T_{(i,j,t)}, T_{(i,j,t+1)}\} \end{aligned} \quad (4.5)$$

4.5.2 Experimento II

De maneira semelhante ao primeiro experimento conduzido, este experimento comparou uma variante do método proposto com uma variante do método de Lucas-Kanade, considerando a diferença máxima da intensidade do tom de cinza entre um dado pixel e seus vizinhos à direita e à esquerda na derivada parcial I_x e a diferença máxima de cinza entre um pixel específico e seu vizinho superior e inferior na derivada parcial I_y . Para manter a equivalência, na variante do método proposto, foram utilizados os elementos estruturantes $3 \times 1 \times 1$ e $1 \times 3 \times 1$ para o cálculo na direção x e na direção y respectivamente, conforme apresentado na equação 4.6.

$$\begin{aligned}
B_x &= \{T_{(i-1,j,t)}, T_{(i,j,t)}, T_{(i+1,j,t)}\} \\
B_y &= \{T_{(i,j-1,t)}, T_{(i,j,t)}, T_{(i,j+1,t)}\} \\
B_t &= \{T_{(i,j,t)}, T_{(i,j,t+1)}\}
\end{aligned} \tag{4.6}$$

4.5.3 Experimento III

Para o terceiro experimento, foram considerados os métodos de Lucas-Kanade e proposto tais como especificados no experimento I, na subseção 4.5.1; com a aplicação do filtro de aguçamento fixo, de tamanho 4×4 anteriormente a aplicação de ambos os métodos, como especificado na equação 4.3.

4.5.4 Experimento IV

De maneira análoga a realizada no terceiro experimento (Subseção 4.5.3), foram consideradas as especificações do método LK e da variante do TLK definidas no experimento II (Subseção 4.5.2) e a aplicação do filtro de aguçamento, de tamanho fixo 4×4 , antes da aplicação dos métodos.

4.5.5 Experimento V

Para o quinto experimento, foram considerados o método LK com as derivadas parciais iguais as definidas no experimento I e a variante do TLK com os elementos estruturantes definidos no mesmo experimento, com a aplicação de um filtro de aguçamento de tamanho variável equivalente ao tamanho da região da imagem antes da aplicação dos dois métodos.

4.5.6 Experimento VI

No sexto experimento, foram consideradas as especificações das derivadas parciais e dos elementos estruturantes utilizados no experimento II (Subseção 4.5.2) e a aplicação do filtro de aguçamento de tamanho variável antes da aplicação dos métodos.

4.5.7 Experimento VII

O sétimo e último experimento comparou os resultados do método Lucas-Kanade original, como definido no experimento 1, sem a aplicação de nenhum filtro de realce; com a variante do método TLK aplicada no experimento V (Subseção 4.5.5), com a aplicação do filtro de aguçamento de tamanho variável.

A Tabela 4.1 resume e apresenta os parâmetros adotados nos experimentos realizados.



(a) Sequência 001 - Imagem 1



(b) Sequência 001 - Imagem 2



(c) Sequência 004 - Imagem 1



(d) Sequência 004 - Imagem 2



(e) Sequência 008 - Imagem 1



(f) Sequência 008 - Imagem 2

Figura 4.1: Exemplos de pares de imagens utilizadas nos experimentos

Tabela 4.1: Síntese dos experimentos

Experimento	LK	LKT	Filtro de Realce
Primeiro	$I_x = G_{(i,j)} - G_{(i+1,j)}$ $I_y = G_{(i,j)} - G_{(i,j+1)}$	$B_x = \{T_{(i,j,t)}, T_{(i+1,j,t)}\}$ $B_y = \{T_{(i,j,t)}, T_{(i,j+1,t)}\}$	Nenhum
Segundo	$I_x = \text{Max}\{G_{(i,j)} - G_{(i+1,j)}, (G_{(i-1,j)} - G_{(i,j)}), (G_{(i-1,j)} - G_{(i+1,j)})\}$ $I_y = \text{Max}\{G_{(i,j)} - G_{(i,j+1)}, (G_{(i,j-1)} - G_{(i,j)}), (G_{(i,j-1)} - G_{(i,j+1)})\}$	$B_x = \{T_{(i-1,j,t)}, T_{(i,j,t)}, T_{(i+1,j,t)}\}$ $B_y = \{T_{(i,j-1,t)}, T_{(i,j,t)}, T_{(i,j+1,t)}\}$	Nenhum
Terceiro	$I_x = G_{(i,j)} - G_{(i+1,j)}$ $I_y = G_{(i,j)} - G_{(i,j+1)}$	$B_x = \{T_{(i,j,t)}, T_{(i+1,j,t)}\}$ $B_y = \{T_{(i,j,t)}, T_{(i,j+1,t)}\}$	Tamanho Fixo
Quarto	$I_x = \text{Max}\{G_{(i,j)} - G_{(i+1,j)}, (G_{(i-1,j)} - G_{(i,j)}), (G_{(i-1,j)} - G_{(i+1,j)})\}$ $I_y = \text{Max}\{G_{(i,j)} - G_{(i,j+1)}, (G_{(i,j-1)} - G_{(i,j)}), (G_{(i,j-1)} - G_{(i,j+1)})\}$	$B_x = \{T_{(i-1,j,t)}, T_{(i,j,t)}, T_{(i+1,j,t)}\}$ $B_y = \{T_{(i,j-1,t)}, T_{(i,j,t)}, T_{(i,j+1,t)}\}$	Tamanho Fixo
Quinto	$I_x = G_{(i,j)} - G_{(i+1,j)}$ $I_y = G_{(i,j)} - G_{(i,j+1)}$	$B_x = \{T_{(i,j,t)}, T_{(i+1,j,t)}\}$ $B_y = \{T_{(i,j,t)}, T_{(i,j+1,t)}\}$	Tamanho Variável
Sexto	$I_x = \text{Max}\{G_{(i,j)} - G_{(i+1,j)}, (G_{(i-1,j)} - G_{(i,j)}), (G_{(i-1,j)} - G_{(i+1,j)})\}$ $I_y = \text{Max}\{G_{(i,j)} - G_{(i,j+1)}, (G_{(i,j-1)} - G_{(i,j)}), (G_{(i,j-1)} - G_{(i,j+1)})\}$	$B_x = \{T_{(i-1,j,t)}, T_{(i,j,t)}, T_{(i+1,j,t)}\}$ $B_y = \{T_{(i,j-1,t)}, T_{(i,j,t)}, T_{(i,j+1,t)}\}$	Tamanho Variável
Sétimo	$I_x = G_{(i,j)} - G_{(i+1,j)}$ $I_y = G_{(i,j)} - G_{(i,j+1)}$	$B_x = \{T_{(i-1,j,t)}, T_{(i,j,t)}, T_{(i+1,j,t)}\}$ $B_y = \{T_{(i,j-1,t)}, T_{(i,j,t)}, T_{(i,j+1,t)}\}$	LK: Nenhum LKT: Tamanho Variável

4.6 Análise dos resultados experimentais

Para cada par de imagens, foram calculadas as medidas de erro médio pelos dois critérios de qualidade (Erro Angular e Erro Absoluto), para cada método utilizado e considerando variações do tamanho da janela em que o fluxo ótico. Essa variação foi de 3 pixels a 20 pixels de lado, considerando-se sempre janelas quadradas.

No total, foram obtidos 48.384 resultados distintos, considerando-se a combinação entre as 32 sequências de imagens utilizadas, os 7 métodos avaliados (LK e LKT com as 6 medidas de dissimilaridade avaliadas), os 18 tamanhos de janela estudados, os 2 critérios de avaliação utilizados (EA e EAbs), os 2 filtros utilizados (tamanho fixo e tamanho variável) e a não aplicação de filtro, os dois tipos de vizinhança (LK) ou elemento estruturante (LKT).

Avaliando-se os resultados considerando-se os resultados por um dado critério de qualidade, os 7 métodos avaliados em relação a variação do tamanho da janela w , foi observado que os resultados não seguiam um padrão. Diante disso, optou-se por utilizar os resultados com diferentes tamanhos de janela como conjunto para a avaliação quantitativa dos resultados.

Por tratar-se de uma base de dados com sequências de imagens bastante heterogênea e com poucas amostras, não foi possível agrupar as amostras, utilizando-se medidas como média, por exemplo. Dessa forma, optou-se por avaliar os resultados conforme seguem.

A Tabela 4.2 apresenta os resultados do primeiro experimento, considerando o método LK e o TLK considerando as 6 medidas de dissimilaridade, conforme apresentadas na Tabela 2.1. A última célula de cada coluna i da tabela detalha a quantidade de sequências de imagens no qual o método M_i apresentou o menor erro médio de acordo com o critério EAbs.

De maneira semelhante, a última célula de cada linha j apresenta o número de sequências de imagens em que o método M_j obteve o menor erro médio calculado pelo critério EA.

A célula que pertence à última coluna e a última linha da Tabela 4.2 apresenta o número 32, indicando ser a soma tanto da linha quanto da coluna e confirmando o tamanho da amostra de 32 sequências de imagens.

Nas demais células da tabela, localizadas no encontro da i -ésima coluna com a j -ésima linha, contém o número de sequências de imagens na qual o método M_i foi o melhor avaliado pelo critério EAbs e o método M_j foi o melhor avaliado pelo critério EA.

Tabela 4.2: Resultados do primeiro experimento

	LK	LKT1	LKT2	LKT3	LKT4	LKT5	LKT6	Total EA
LK	5	0	0	1	1	0	0	7
LKT1	0	0	0	0	0	0	0	0
LKT2	0	0	1	2	3	0	1	7
LKT3	0	0	0	1	0	0	0	1
LKT4	0	0	0	0	7	0	0	7
LKT5	0	0	0	0	0	0	0	0
LKT6	3	1	3	0	0	0	3	10
Total EAbs	8	1	4	4	11	0	4	32

Para exemplificar, a célula destacada com tom cinza mais escuro na Tabela 4.2 indica que, em 3 sequências de imagens, o critério EA avaliou a técnica LKT com a medida d_2 como tendo o menor erro médio enquanto o critério EAbs avaliou a técnica LKT com a medida d_4 como tendo o menor erro.

Na diagonal principal, é possível observar a quantidade de sequências de imagens na qual os critérios de avaliação da qualidade dos métodos convergiram, apontando o mesmo método como sendo o com menor erro médio. Isso é relevante para a avaliação deste trabalho, porque evidencia o método como sendo o com os melhores resultados, considerando-se dois aspectos ou atributos diferentes, a saber, ângulo e norma.

Dessa maneira, as células mais importantes para a análise dos resultados foram destacados com um tom de cinza claro. Os resultados dos demais experimentos também foram condensados em tabelas semelhantes à Tabela 4.2 e serão apresentadas na sequência, com as células mais importantes destacadas no mesmo tom de cinza supracitado.

Sobre os resultados do primeiro experimento, é possível observar que o critério EA avaliou o método LK como sendo o melhor em 7 sequências de imagens, enquanto o critério EAbs avaliou o mesmo método como sendo o melhor em 8 sequências. Olhando para a Tabela 4.2, podemos ainda observar que ambos os critérios concordaram com a afirmação de que o método LK era o com menor erro em 5 sequências de imagens.

Ainda observando estes resultados, é notável que o método LKT com a medida de dissimilaridade d_4 é o melhor avaliado em 7 sequências pelo critério EA, em 11 pelo EAbs e em 7 sequências por ambos os métodos. Dessa forma, percebe-se uma tênue melhoria do método LKT com a medida d_4 em relação ao método LK.

Semelhantemente ao realizado com o primeiro experimento, a Tabela 4.3 apresenta os resultados do segundo experimento, isto é, mostrando a quantidade de sequências de imagens na qual cada método apresentou o menor erro médio de acordo com cada critério adotado para a avaliação.

Nela é observado que o método TLK com a medida d_4 apresentou o menor erro médio em 28,125% dos casos avaliados por ambos os critérios, isto é, em 9 dos 32 sequências de imagens utilizadas, enquanto o método LK foi o melhor avaliado, pelos dois critérios, em 21,875% (ou em 7 das 32) sequências utilizadas .

Tabela 4.3: Resultados do segundo experimento

	LK	LKT1	LKT2	LKT3	LKT4	LKT5	LKT6	Total EA
LK	7	0	0	0	1	1	0	9
LKT1	0	0	0	0	3	0	0	3
LKT2	0	1	2	0	1	0	0	4
LKT3	0	0	0	1	0	0	0	1
LKT4	1	0	1	0	9	0	0	11
LKT5	0	0	0	1	0	0	0	1
LKT6	2	0	0	0	0	0	1	3
Total EAbs	10	1	3	2	14	1	1	32

No segundo experimento, a soma da diagonal principal da Tabela 4.3 totaliza 20 de 32 casos. Isso significa que em 62,5% dos casos os dois critérios escolheram o mesmo método como sendo aquele com menor erro médio.

A Tabela 4.4 apresenta os resultados no terceiro experimento, que considera a aplicação do filtro de aguçamento antes da aplicação do método de estimativa de fluxo ótico.

Neste experimento, o método TLK com medida d_4 foi a melhor avaliada em 7 sequências por ambos os critérios, em 7 sequências pelo critério EA e 13 sequências pelo critério EAbs. Um fato observado é que o critério EA avaliou o método LK como sendo o com menor erro em 8 sequências, enquanto o EAbs não avaliou esse método como sendo o menor erro em nenhuma das sequências.

A Tabela 4.5 apresenta os resultados do quarto experimento. Nele, os métodos LK e LKT com a medida d_3 foram os melhores avaliados por ambos os critérios. De acordo com o critério EA, ambos os métodos tiveram resultados similares, sendo o com menor erro em 9 sequências de imagens cada. Já pelo critério EAbs, o método LK teve uma pequena vantagem em relação ao método LKT3.

Tabela 4.4: Resultados do terceiro experimento

	LK	LKT1	LKT2	LKT3	LKT4	LKT5	LKT6	Total EA
LK	0	0	0	3	2	1	2	8
LKT1	0	0	0	0	0	0	0	0
LKT2	0	0	2	0	3	0	1	6
LKT3	0	0	0	2	0	0	0	2
LKT4	0	0	0	0	7	0	0	7
LKT5	0	0	0	0	0	0	0	0
LKT6	0	2	3	0	1	0	3	9
Total EAbs	0	2	5	5	13	1	6	32

Semelhantemente ao ocorrido no segundo experimento, a soma da diagonal principal da Tabela 4.5 totaliza 19 de 32 casos, indicando que ambos os critérios convergiram em mais da metade dos casos.

Tabela 4.5: Resultados do quarto experimento

	LK	LKT1	LKT2	LKT3	LKT4	LKT5	LKT6	Total EA
LK	7	0	0	1	0	0	1	9
LKT1	0	1	0	0	0	0	1	2
LKT2	0	0	1	0	0	0	1	2
LKT3	3	0	0	6	0	0	0	9
LKT4	0	0	0	0	2	0	0	2
LKT5	0	1	0	1	0	0	0	2
LKT6	1	0	0	1	2	1	1	6
Total EAbs	11	2	1	9	4	1	4	32

A Tabela 4.6, referente ao quinto experimento, mostra que os métodos LKT com as medidas d_3 3 d_5 foram os melhor avaliados em 8 sequências de imagem cada uma, enquanto o método LK não foi avaliado como tendo o menor erro em nenhuma sequência de imagem.

Semelhantemente ao ocorrido nos experimentos II e IV, a soma da diagonal principal da Tabela 4.6, que resulta em 19 de 32, também mostra que ambos os critérios convergem em mais de 50% dos resultados.

A Tabela 4.7 mostra os resultados do penúltimo e sexto experimento, na qual pode-se observar que o método LKT com a medida d_3 foi o melhor avaliado por ambos os critérios (em 6 casos) e isoladamente (em 9 casos pelo EA e em 9 casos pelo EAbs

Tabela 4.6: Resultados do quinto experimento

	LK	LKT1	LKT2	LKT3	LKT4	LKT5	LKT6	Total EA
LK	0	0	0	0	1	0	0	1
LKT1	0	0	0	0	1	0	0	1
LKT2	1	1	1	0	0	0	1	4
LKT3	0	0	0	8	0	3	0	11
LKT4	1	0	0	1	1	0	0	3
LKT5	0	0	0	1	0	8	0	9
LKT6	0	1	0	0	1	0	1	3
Total EAbs	2	2	1	10	4	11	2	32

também). O método LK foi bem avaliado pelo critério EA (8 casos), mas não foi bem avaliado pelo critério EAbs (único caso).

Comparando os resultados dos experimentos I, III e V; pode-se notar que a aplicação de um filtro de realce ou aguçamento favorece os resultados do método tensorial em relação ao do LK original, sendo que a aplicação do filtro de tamanho variável, melhora substancialmente o desempenho do método tensorial.

Tabela 4.7: Resultados do sexto experimento

	LK	LKT1	LKT2	LKT3	LKT4	LKT5	LKT6	Total EA
LK	1	0	1	3	1	0	2	8
LKT1	0	1	0	0	0	0	2	3
LKT2	0	0	1	0	1	0	1	3
LKT3	0	3	0	6	0	0	0	9
LKT4	0	0	0	0	2	0	0	2
LKT5	0	0	0	0	1	1	0	2
LKT6	0	1	0	0	2	1	1	5
Total EAbs	1	5	2	9	7	2	6	32

A Tabela 4.8 resume o último experimento conduzido, que comparou o método LK original e o método LKT com o filtro Laplaciano de tamanho compatível ao tamanho da janela de fluxo constante. Nele, é percebido que tal comparação é injusta, visto que é percebido um desempenho muito melhor dos métodos LKT em relação ao LK.

Entretanto, a soma da diagonal principal totalizando pouco mais que 50%, mostra que os dois critérios de qualidade se dispersaram um pouco mais que em outros experimentos.

Tabela 4.8: Resultados do sétimo experimento

	LK	LKT1	LKT2	LKT3	LKT4	LKT5	LKT6	Total EA
LK	1	0	0	1	0	0	0	2
LKT1	0	1	0	0	0	0	2	3
LKT2	0	0	2	0	1	0	1	4
LKT3	0	3	0	8	0	0	1	12
LKT4	0	0	0	0	2	0	0	2
LKT5	0	0	0	0	1	1	1	3
LKT6	0	1	0	0	2	1	2	6
Total EAbs	1	5	2	9	6	2	7	32

Essa não convergência dos critérios e o fato de que o método TLK com a medida d_3 foi o melhor avaliado em apenas 6 dos 32 casos, mostram que mesmo um método mais robusto, que incorpora características isoladas que tendem a melhorar o seu desempenho, pode não ser o mais efetivo para a estimativa de fluxo ótico em sequências distintas de imagens.

Considerações Finais

O desafio da estimativa de fluxo ótico é um famoso problema tratado em processamento digital de imagens, contando com distintas estratégias, diversas abordagens e inúmeras modelagens. Nesta dissertação foi proposto um método baseado na abordagem fundamentada em gradiente e que utiliza a representação tensorial de cores.

Para o estudo desse método, o mesmo foi implementado e seu comportamento foi analisado para diferentes sequências de imagens sintéticas que possuíam um fluxo ótico de referência para comparação. O seu comportamento foi avaliado embasado em dois critérios de medição: O critério do erro angular e o do erro absoluto.

As próximas subseções descrevem os resultados apresentados nesse trabalho (Seção 5.1), as limitações encontradas (Seção 5.2) e os possíveis trabalhos futuros (Seção 5.3).

5.1 Principais Resultados e Contribuições

Este trabalho introduz uma variação do método Lucas-Kanade para a estimativa de fluxo ótico em sequências de imagens coloridas.

O método proposto e implementado utilizou a premissa adotada no método Lucas-Kanade de assumir que o fluxo é constante em uma determinada região da imagem e dessa forma permitir a resolução da equação de restrição do fluxo ótico, e também os fundamentos de álgebra tensorial para estimar funções de distância ou dissimilaridade em valores não escalares.

Experimentos realizados mostraram que o método proposto é eficiente em relação ao método de referência, com a vantagem de manipular informações de cor, enquanto o referência utiliza apenas informações de intensidade de cinza.

Como resultado deste trabalho, tem-se um artigo submetido, intitulado "Tensorial Lucas-Kanade: An Optical flow estimator based on tensorial color representation and tensorial algebra", em 2018.

5.2 Limitações deste trabalho

Apesar de tratar-se de um problema bastante conhecido e com diversas propostas de métodos para a sua estimativa, a principal limitação deste trabalho foi a dificuldade em encontrar o "padrões ouro" de fluxo ótico em sequências de imagens coloridas para realizar a avaliação do método proposto.

Foi encontrada apenas uma reduzida quantidade de amostras de fluxo ótico (32 amostras sintéticas) em duas fontes distintas, obtidas de pares de imagens coloridas, como apresentadas na seção 4.2.

A quantidade reduzida de imagens comprometeu a realização de análises com relevância estatística, uma vez que não foi possível validar se o método era favorecido ou tinha um desempenho mais alto com determinada característica de um grupo de amostras, pois as amostras se apresentaram únicas na maioria das características observadas.

A qualidade das amostras encontradas também limitou o processo de investigação do método proposto, uma vez que as amostras encontradas continham apenas sequências de duas imagens, o que dificultou a investigação de elementos estruturantes B_t maiores que 2 quadros e por serem compostas apenas por imagens sintéticas, a avaliação de ruídos e comportamentos anômalos não puderam ser devidamente exploradas.

5.3 Trabalhos futuros

Durante a condução desse estudo, muitas percepções e dúvidas a cerca de assuntos pertinentes e relacionados foram identificados e/ou levantados. Entretanto, sua abordagem não permitiu o aprofundamento necessário para investigação de tais assuntos, como descrito na seção de limitações.

Dessa maneira, diversos trabalhos são apresentados nessa seção como sugestões futuras de pesquisa e possibilidade de continuidade da pesquisa desenvolvida.

Como sugestões principais, são destacadas:

Criação de uma base de dados com padrão ouro de fluxo ótico: Diante da grande limitação encontrada nesse trabalho, a construção de uma base de dados contendo diversas sequências de imagens coloridas e o respectivo padrão ouro de fluxo ótico determinado seria de elevada importância para a continuidade desse trabalho.

Avaliação do método TLK utilizando elementos estruturantes diferentes: Como a base de dados utilizadas nesse trabalho continham apenas sequências contendo apenas duas imagens, não foi possível realizar testes com elementos estruturantes que envolvessem mais que dois quadros da sequência.

Elaboração de uma função Laplaciana Morfológica Tensorial: Semelhantemente ao gradiente, que é uma operação construída a partir de operações básicas da morfologia matemática, a função Laplaciana também é uma operação composta de operações da morfologia matemática. Nesse sentido, uma função Laplaciana Morfológica Tensorial poderia ser definida com base na álgebra tensorial e ser aplicada para imagens coloridas.

Avaliação de outras medidas de dissimilaridade: Neste trabalho foram avaliadas as seis medidas de dissimilaridade utilizadas por (Rittner et al., 2010). A seleção e proposição de novas medidas neste métodos poderiam resultar em resultados melhores.

Confecção da forma tensorial de outros métodos de estimativa de fluxo ótico: Dada a gama de métodos de estimativa de fluxo ótico encontradas na literatura, poderia-se efetuar a adaptação deste para utilizar informação de cores baseando-se no TMG e na álgebra tensorial. Acredito que a adaptação seja intuitiva para outros métodos baseado em gradiente.

Definição de um Gradiente tensorial tridimensional: Considerando-se que o modelo de cor HSL possui duas bandas escalares (S e L) e uma banda angular (H) e os resultados já apresentados com o TMG utilizado neste trabalho, uma extensão natural deste gradiente poderia ser a adição dos atributos de movimento de cada pixel de uma imagem, que pode ser apresentada por um atributo escalar, definida como norma, e um atributo angular em coordenadas polares. Isso permitiria mensurar a diferença de atributos de cor e de movimento em sequências de imagens coloridas.

REFERÊNCIAS

- BAKER, S.; SCHARSTEIN, D.; LEWIS, J.; ROTH, S.; BLACK, M. J.; SZELISKI, R. A database and evaluation methodology for optical flow. *International Journal of Computer Vision*, v. 92, n. 1, p. 1–31, 2011.
- BARRON, J. L.; FLEET, D. J.; BEAUCHEMIN, S. S. Performance of optical flow techniques. *International journal of computer vision*, v. 12, n. 1, p. 43–77, 1994.
- BUSIN, L.; VANDENBROUCKE, N.; MACAIRE, L. Color spaces and image segmentation. *Advances in imaging and electron physics*, v. 151, n. 1, p. 1, 2008.
- CHENG, H.-D.; JIANG, X.; SUN, Y.; WANG, J. Color image segmentation: advances and prospects. *Pattern recognition*, v. 34, n. 12, p. 2259–2281, 2001.
- FLEET, D. J.; JEPSON, A. D. Computation of component image velocity from local phase information. *International journal of computer vision*, v. 5, n. 1, p. 77–104, 1990.
- GONZALEZ, R. C.; WOODS, R. E. Digital image processing. 2008.
- HORN, B. K.; SCHUNCK, B. G. Determining optical flow. In: *1981 Technical symposium east*, International Society for Optics and Photonics, 1981, p. 319–331.
- KINGSLEY, P. B. Introduction to diffusion tensor imaging mathematics: Part i. tensors, rotations, and eigenvectors. *Concepts in Magnetic Resonance Part A*, v. 28, n. 2, p. 101–122, 2006.
- LUCAS, B. D.; KANADE, T. An iterative image registration technique with an application to stereo vision. In: *IJCAI*, 1981, p. 674–679.
- LUCAS, B. D.; KANADE, T.; ET AL. An iterative image registration technique with an application to stereo vision. 1981.

MAC AODHA, O.; HUMAYUN, A.; POLLEFEYS, M.; BROSTOW, G. J. Learning a confidence measure for optical flow. *IEEE Transactions on Pattern Analysis and Machine Intelligence (To Appear)*, 2012.

MARQUES FILHO, O.; VIEIRA NETO, H. *Processamento digital de imagens*. Brasport, 1999.

NAJMAN, L.; TALBOT, H. *Mathematical morphology*. John Wiley & Sons, 2013.

OHTA, N. Optical flow detection by color images. *NEC Research and Development*, v. 97, p. 78–84, 1990.

OTTE, M.; NAGEL, H. Optical flow estimation: advances and comparisons. *Computer Vision - ECCV'94*, p. 49–60, 1994.

RITTNER, L.; FLORES, F.; LOTUFO, R. New tensorial representation of color images: Tensorial morphological gradient applied to color image segmentation. In: *Computer Graphics and Image Processing, 2007. SIBGRAPI 2007. XX Brazilian Symposium on, IEEE, 2007*, p. 45–52.

RITTNER, L.; FLORES, F. C.; LOTUFO, R. A. A tensorial framework for color images. *Pattern Recognition Letters*, v. 31, n. 4, p. 277–296, 2010.

SCHWIEGERLING, J. *Field guide to visual and ophthalmic optics*. Spie Bellingham, WA, USA, 2004.

WALKER, J.; HALLIDAY, D.; RESNICK, R. *Fundamentos de física: volume 1: mecânica*. LTC, 2009.