

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

DANILO FILITTO

Mensuração da velocidade de deslocamento de corpos rígidos em
tempo real por análise de sequência de imagens

Maringá
2013

DANILO FILITTO

Mensuração da velocidade de deslocamento de corpos rígidos em tempo real por análise de sequência de imagens

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. Nardênio Almeida Martins

Maringá
2013

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

F483m Filitto, Danilo
Mensuração da velocidade de deslocamento de corpos
rígidos em tempo real por análise de sequência de imagens /
Danilo Filitto. -- Maringá, 2013.
99 f. : il., color., tabs., fotos.

Orientador: Prof. Dr. Nardênio Almeida Martins.
Dissertação (mestrado) - Universidade Estadual de
Maringá, Centro de Tecnologia, Departamento de Informática,
Programa de Pós-Graduação em Ciência da Computação, 2013.

1. Segmentação de imagem. 2. Detecção de movimento. 3.
Modelos de transformação geométrica. 4. Mensuração da
velocidade. 5. Corpos rígidos. I. Martins, Nardênio de
Almeida, 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. 006.42

AHS-001515

FOLHA DE APROVAÇÃO

DANILO FILITTO

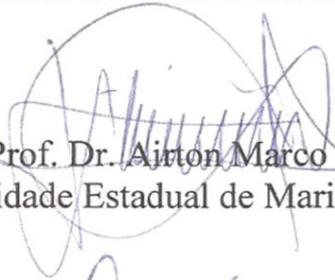
Mensuração da velocidade de deslocamento de corpos rígidos em tempo real por análise de sequência de imagens

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:

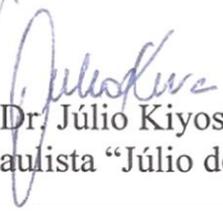
BANCA EXAMINADORA



Prof. Dr. Nardênio Almeida Martins
Universidade Estadual de Maringá – DIN/UEM



Prof. Dr. Ailton Marco Polidório
Universidade Estadual de Maringá – DIN/UEM



Prof. Dr. Júlio Kiyoshi Hasegawa
Universidade Estadual Paulista “Júlio de Mesquita Filho” – FCT/Unesp

Aprovada em: 19 de fevereiro de 2013.

Local da defesa: Sala 101, Bloco C-56, *campus* da Universidade Estadual de Maringá

DEDICATÓRIA

*Dedico esse trabalho a todas as
pessoas que me apoiaram de
maneira direta ou indireta.*

AGRADECIMENTOS

Agradeço primeiramente a Deus que me guiou e me deu forças durante essa dura jornada. À minha família, em especial, ao meu pai Gerson Filitto, à minha mãe, Sueli Aparecida Passos Filitto e à minha esposa Samira Lomas Malacrida Filitto que sempre se fizeram presentes em todos os momentos. Aos professores e amigos, Dr. Nardênio Almeida Martins, Dr. Franklin César Flores, Dr. Airton Marco Polidório e ao Dr. Júlio Kiyoshi Hasegawa, que colaboraram no desenvolvimento desse projeto compartilhando seus conhecimentos. À minha prima Gilda Felitto que, com seu enorme domínio da língua Inglesa, me ajudou a compreender a gramática e conceitos chaves dessa língua.

EPÍGRAFE

“Pensar é o trabalho mais difícil que existe. Talvez por isso tão poucos se dediquem a ele.” (Henry Ford).

Mensuração da velocidade de deslocamento de corpos rígidos em tempo real por análise de sequência de imagens

RESUMO

Este trabalho demonstra uma metodologia voltada para mensurar a velocidade de um corpo rígido em movimento, tendo como ponto de partida, uma sequência de imagens obtidas por meio de uma filmadora digital. Esse processo de mensuração da velocidade exige a execução de passos fundamentais: 1) detectar e segmentar os corpos em movimento na cena imageada e; 2) transformar pontos conhecidos do corpo desejado pertencentes ao plano imagem (2D) para o espaço objeto (3D). Para a etapa de detecção e segmentação, este trabalho, implementa duas técnicas: - uma baseada na média temporal dos valores de cada *pixel* registrados em N quadros de imagens e, uma outra baseada no histórico de valores associados com *pixels* registrados em N quadros de imagens. Após a segmentação dos corpos, é realizada a transformação para se obter a correspondência de pontos pertencentes as imagens dos corpos desejados com os pontos correspondentes reais no espaço objeto, tal transformação, pode ser realizada por meio das equações de colinearidade, pelo modelo de transformação linear direta, pelo modelo de transformação polinomial, pelo modelo de transformação afim ou pelo modelo de transformação das funções racionais. A obtenção desses pontos no espaço objeto permite calcular a distância real que o corpo percorreu em um intervalo de tempo automaticamente cronometrado. Esses dois valores (espaço e tempo) permitem calcular a velocidade com que o corpo se desloca. Para validar a metodologia proposta foi realizado um experimento rigorosamente controlado que contou com a utilização de um aparelho denominado trilho de ar, cinco sensores fotoelétricos dispersos sobre o mesmo com distâncias conhecidas entre si e um cronômetro ligado aos sensores. No momento em que um corpo em movimento passa por um sensor, o cronômetro é paralisado e com isso, se tem o intervalo de tempo que o corpo demorou em percorrer a distância da origem até cada um dos sensores dispostos sobre o trilho de ar. Todo esse experimento foi filmado e, sobre os quadros desse filme, foi aplicada a metodologia proposta para se calcular a velocidade de um corpo somente pelo uso das sequências de imagens. Os resultados obtidos com o uso dos sensores são então comparados com os resultados calculados pelo uso das imagens.

Palavras-chave: segmentação, detecção, modelos de transformação geométrica, mensuração da velocidade, corpos rígidos.

Measuring travel speed of rigid bodies live for sequence analysis of images

ABSTRACT

This work demonstrates a methodology aimed to measure the velocity of a rigid body in motion, taking as a starting point, a sequence of images taken by a digital camera. This process requires the measurement of the speed of execution fundamental steps: 1) detect and segment the moving bodies in the scene image and, 2) turning points of the body known belonging to the desired image plane (2D) object into space (3D). For step detection and segmentation, this work implements two techniques - one based on time average values of each pixel recorded in N image frames, and another based on historical values associated with N pixels recorded in picture frames . After cleavage of the bodies, the transformation is performed to obtain the correspondence points belonging to the images of bodies with desired points in space corresponding real object, such a transformation may be performed using the collinearity equations, the model linear transformation directly, by polynomial transformation model, the model of affine transformation or transformation model of rational functions. Obtaining those points in the object space allows to calculate the actual distance the body traveled in a time interval automatically timed. These two values (space and time) to calculate the speed with which the body moves. To validate the proposed methodology was performed a rigorously controlled experiment that involved the use of a device called air rail, five photoelectric sensors scattered about the same with known distances between themselves and a timer connected to the sensors. At the time when a moving body passes a sensor, the timer is stopped and thus, has the time to take in the body travels a distance from the origin to each of the sensors disposed on the tracks of air. This entire experiment was filmed and on the frames of this film, we applied the proposed methodology to calculate the speed of a body only by the use of image sequences. The results obtained with the use of sensors are then compared with the results calculated by use of the images.

Keywords: segmentation, detection, geometric transformation models, measurement speed, rigid bodies.

LISTA DE ABREVIATURAS E SIGLAS

APM	<i>Afine Projection Model</i>
CCDs	<i>Charge-coupled devices</i>
DLT	<i>Direct Linear Transformation</i>
MHI	<i>Motion History Image</i>
MMQ	Método dos mínimos quadrados
PID	Processamento de Imagens Digital
PIXEL	<i>Picture Elements</i>
RFM	<i>Rational Function Model</i>

SUMÁRIO

INTRODUÇÃO.....	12
1.1 Objetivo Geral	14
1.2 Objetivos Específicos.....	14
1.3 Justificativa.....	14
1.4 Metodologia da Pesquisa	16
1.5 Organização do Trabalho	17
FUNDAMENTAÇÃO TEÓRICA	18
2.1 Imagem Digital	18
2.1.1 Geometria de Imagens Digitais Adquiridas por Câmaras Digitais	20
2.1.2 Conceito de vizinhança de <i>pixels</i>	21
2.2 Processamento de Imagens Digitais	22
2.2.1 Realce de imagens	22
2.2.1.1 Filtros da ordem (domínio espacial)	23
2.2.2 Morfologia Matemática	25
2.2.2.1 Definições básicas.....	25
2.2.2.2 Dilatação.....	26
2.2.2.3 Erosão.....	26
2.2.2.4 Abertura e Fechamento	26
2.2.3 Segmentação de imagens	28
2.2.3.1 Segmentação Baseada em Valor de Limiar.....	31
2.2.3.2 Segmentação de imagens baseada em movimento	34
2.3 Fotogrametria - Conceitos	44
2.3.1 Equações de colinearidade	44
2.3.2 Modelos matemáticos empíricos	49
2.3.2.1 Modelos polinomiais 3D	49
2.3.2.2 Modelo das funções racionais	50
2.3.2.3 Modelo de Transformação afim	51
2.3.2.4 Transformação Linear Direta (DLT)	52
TRABALHOS RELACIONADOS.....	54
3.1 Mensuração da velocidade por imagens de vídeo utilizando o método de Tsai em duas fases.....	55
3.2 Sistema de mensuração de velocidade de veículos	57

3.3 Determinação da velocidade de um objeto, a partir de uma sequência de imagens, aplicando as equações de colinearidade modificadas.....	59
MENSURAÇÃO DA VELOCIDADE: Experimentos e Análise dos Resultados	61
4.1 Sistema proposto	61
4.2 Experimento.....	64
4.3 Procedimentos para condução do experimento.....	67
4.4 Resultados.....	69
4.4.1 Resultado Computado com a utilização de sensores fotoelétricos.....	70
4.4.2 Resultado Computado (análise das sequências de imagens e aplicação da computação fotogramétrica)	71
4.4.2.1 Resultado Computado por meio do método das equações de colinearidade....	71
4.4.2.2 Resultado Computado por meio do método da transformação paralela afim ..	73
4.4.2.3 Resultado Computado por meio do método de Transformação Linear Direta.	75
4.4.2.4 Resultado Computado por meio do método de Transformação Polinomial 2D 3D	77
4.4.3 Resultado Manual (resultado obtido sobre pontos manualmente coletados).....	81
4.4.3.1 Resultado Computado por meio do método das equações de colinearidade..	81
4.5 Análises dos resultados.....	91
CONSIDERAÇÕES FINAIS.....	94
REFERÊNCIAS.....	96

INTRODUÇÃO

A área de Processamento de Imagens Digitais (PID) abre perspectivas para o desenvolvimento de uma série de aplicações que necessitam de informações presentes em imagens digitais como, por exemplo, as aplicações voltadas para o monitoramento de ambientes. Dentre os vários tipos de sistemas de monitoramento existentes, pode-se citar o sistema de detecção e monitoramento de pessoas em ambientes controlados (SONAGLIO, 2009); o sistema de detecção de automóveis (MONTANHA, 2010); o sistema de contagem de automóveis (SHO, et al, 1995) e o sistema de mensuração da velocidade de deslocamento de automóveis (YAN, et al, 2010).

Ao que se refere à capacidade de fiscalização, os sistemas de monitoramento são úteis, pois, podem operar por longos períodos de tempo, superando a capacidade humana em realizar essa mesma tarefa. Porém, para dotar um sistema de monitoramento baseado somente no uso de técnicas de PID é necessário solucionar vários problemas inerentes à sua própria natureza, ao saber que esses sistemas devem ser capazes de operar em ambientes não controlados e, que há uma diversidade intrínseca aos objetos de interesse (como: cor, tamanho, forma geométrica, variabilidade nos padrões de textura e o seu próprio movimento), além das condições extrínsecas, relativas ao ambiente (como: variação na intensidade da iluminação solar, chuva, sombras, interferências provocadas por outros objetos presentes na cena, etc.) (MONTANHA, 2010). Ou seja, para dotar um sistema baseado em PID com a capacidade para monitorar um ambiente, deve-se ser capaz de superar tanto os problemas causados pelo ambiente como aqueles causados pelos próprios objetos de interesse.

No caso deste trabalho, em que se deseja mensurar a velocidade de deslocamento de um corpo rígido por meio de sequência de imagens, existem outros dois problemas a serem contornados: 1) considere que o corpo se desloca no espaço objeto (3D) e suas imagens são registradas no plano imagem do sensor (2D). Somente essa transformação provoca perda de dados, especialmente na visão monocular – caso deste trabalho – como oclusões e sobreposições de segmentos e regiões. 2) Além disso, para transformar pontos do espaço objeto para o plano imagem, as câmaras (sensor) se utilizam de lentes esféricas e, devido a essa esfericidade da lente da câmara há uma transformação de projeção do espaço objeto 3D, que é ortogonal, para o espaço imagem 2D, que é cônico, isto faz com que as imagens capturadas sejam gerenciadas por uma perspectiva de projeção central, fazendo com que todos os pontos do espaço objeto (3D) de uma imagem sejam registrados no plano imagem, que é 2D, ocasionando perdas de dados e comprometendo a mensuração de distâncias.

Para minimizar os problemas inerentes ao sistema óptico e aplicar o processo de correção geométrica, é necessário estabelecer modelos matemáticos que relacionem o espaço imagem com o espaço objeto. Dentro deste contexto surgem os modelos rigorosos e os modelos empíricos, os quais são explorados neste trabalho.

Os modelos rigorosos procuram descrever fisicamente a geometria de aquisição da cena, considerando os dados referentes ao modelo do sensor (distância focal, tamanho do sensor, tamanho do *pixel*, distorção das lentes, dentre outros) quanto ao modelo da plataforma (efemérides e altitude) que descreve a trajetória e a orientação do sensor no intervalo de aquisição das imagens da cena como, por exemplo, o modelo baseado nas equações de colinearidade. Em contraposição, os modelos empíricos relacionam ambos os espaços pelo uso de parâmetros sem significado físico direto, o que os torna uma alternativa para os casos em que não se conhece os dados referentes aos parâmetros de aquisição do sistema de imageamento (TOUTIN, 2004). Como exemplo de modelos empíricos, pode-se mencionar o modelo das funções racionais, a transformação linear direta, a transformação afim (*Affine Projection Model*) e os modelos polinomiais.

Dentro deste contexto, o objetivo principal deste trabalho consiste em apresentar o desenvolvimento de métodos de mensuração da velocidade de deslocamento de corpos rígidos em tempo real por meio da análise de sequência de imagens e por meio da análise das relações geométricas entre pontos do espaço real (espaço objeto) e pontos correspondentes no plano imagem (espaço imagem). Ainda, tais métodos são usados para compor um sistema que opera em tempo real. Há que se destacar que o sistema de aquisição de imagens é monocular e, portanto, não se tem dados suficientes para aplicar métodos fotogramétricos de modelos

estereoscópicos para se reconstruir um ponto no espaço objeto a partir de pontos pertencentes ao plano imagem.

1.1 Objetivo Geral

Desenvolver um sistema capaz de mensurar a velocidade de deslocamento de corpos rígidos em tempo real por meio da análise de sequência de imagens.

1.2 Objetivos Específicos

- Desenvolver/aplicar técnicas relacionadas com PID capazes de detectar em tempo real a movimentação de corpos rígidos.
- Desenvolver/aplicar técnicas capazes de realizar a segmentação de pontos sobre corpos específicos que permitam a mensuração da velocidade percorrida interquadros de imagens.
- Implementar e analisar vários modelos matemáticos de transformação de coordenadas do espaço imagem para o espaço objeto para possibilitar a determinação da velocidade de deslocamento de um corpo rígido qualquer.

1.3 Justificativa

É possível aplicar técnicas relacionadas com PID para definir uma tecnologia capaz de monitorar e mensurar a velocidade de corpos rígidos em movimento, como: tráfego de veículos, pessoas e robôs móveis. Essa tecnologia tem a relação custo/benefício baixa, quando comparadas com a aplicação de sensores (como radares, mensuradores laser e sistemas de bobinas de indutância fixos), uma vez que, além de permitir a mensuração de elementos ou ações (como o movimento) por meio da análise apropriada das imagens adquiridas, também é possível armazenar essas imagens para observações posteriores, ou, até, constituir provas jurídicas legais contra atos que contrariam dispositivos de lei. Nesse contexto, propõe-se, neste trabalho, desenvolver um sistema de aquisição e armazenamento de sequências de imagens e, aplicar sobre as mesmas técnicas relacionadas com a área de PID e a área de Fotogrametria para detectar corpos em movimento e mensurar a velocidade de deslocamento dos mesmos.

Um exemplo de aplicação prática de um sistema de mensuração de velocidade de corpos rígidos está no controle de tráfego de veículos.

Os sistemas de monitoramento aplicados ao tráfego de veículos, principalmente aqueles capazes de mensurar a velocidade de deslocamento e detectar avanço de semáforo em estado vermelho, são importantes por permitir que seja feita uma fiscalização intensiva do comportamento dos motoristas no trânsito. No estado do Acre, em 2011, houve um aumento de 27% de acidentes de trânsito com vítimas fatais entre os meses de janeiro a março, conforme especifica o Relatório do Departamento Estadual de Trânsito do Estado do Acre (<http://www.detran.ac.gov.br>). A cidade do Rio de Janeiro, em 2005, publicou o Edital Nº 007/2005 – Projeto Básico, que faz a chamada para uma concorrência pública para instalar um sistema automático de monitoração do tráfego de veículos em 100 intersecções controladas por semáforos, com valor estimado em mais de 130 milhões de reais (<http://www.tcm.rj.gov.br/Noticias/4608/RELATR~1.PDF>), ao alegar que a quantidade de mortes em acidentes de trânsito no Brasil é elevada, representando 5,8% das causas de óbito no país (Cucci, et al, 1999). Cerca de 85% dos acidentes com vítimas no Brasil ocorrem em área urbana e, informa que no Município do Rio de Janeiro a quantidade de registros é de aproximadamente 60 mil acidentes por ano.

Ao que se refere à capacidade de fiscalização, esses sistemas de monitoração de tráfego de veículos, são úteis, pois, pode operar por longos períodos de tempo, superando a capacidade humana em realizar essa mesma tarefa. No entanto, a tecnologia atualmente usada por esses sistemas pode evoluir de forma considerável, como é apresentado a seguir.

Os sistemas de monitoramento da velocidade de deslocamento de veículos atuais utilizam as tecnologias de radar (portáteis ou fixos) ou sistemas de sensores de indutância. Os radares portáteis mensuram a velocidade dos veículos enviando uma onda de rádio na direção do veículo alvo e captando a reflexão da mesma, a qual retorna ao aparelho com uma frequência (fenômeno físico natural denominado de efeito Doppler), proporcional à velocidade do veículo (ZHIWEI, 2007). Já os sensores de indutância se comportam como um circuito elétrico sintonizado, no qual, um laço de arame e chumbo no cabo são os elementos indutivos (NISHIMOTO, 2006). Quando um veículo (de carcaça metálica) passa sobre o laço, ele induz a formação de uma corrente elétrica nas alças dos arames desse laço e, como consequência, diminui capacidade de indutância dos mesmos, ocasionando o acionamento da unidade eletrônica de relé de saída, que envia um pulso ao controlador informando a passagem de um veículo (MONTANHA, 2010). A velocidade do veículo é mensurada a partir da obtenção do intervalo de tempo decorrido entre os acionamentos do primeiro e do segundo sensor, os quais são enterrados separadamente debaixo do asfalto a uma distancia conhecida.

Tanto o monitoramento da velocidade de deslocamento por meio de radares como por

meio de sensores de indutância possuem pontos negativos. No caso dos sensores de indutância, os veículos de baixa massa metálica (como motocicletas), ou com chassi muito elevado em relação ao solo (como caminhões e ônibus) são de difícil detecção, e no caso dos radares, a precisão da velocidade varia de acordo com o posicionamento do aparelho na via. Além disso, tanto os radares quanto os sensores de indução são sistemas com tecnologia proprietária e de alto custo. (ZHIWEI, 2007; MONTANHA, 2010).

Levando em consideração esses pontos negativos, estão sendo desenvolvidos novos métodos para detectar a presença e para mensurar a velocidade de deslocamento de veículos, usando, como base para essa aplicação, técnicas de Processamento de Imagens Digitais (PDI). Tais métodos têm como principal vantagem, a possibilidade de se utilizar hardware de baixo custo desvencilhado de qualquer tecnologia proprietária.

Segundo Zhiwei, et al (2007) e Gupte, et al (2002), a utilização de técnicas de PDI para desenvolver sistemas de monitoramento do tráfego de veículos, possibilita não somente mensurar a velocidade de deslocamento de veículos, mas também, determinar a trajetória de deslocamento, monitorar o fluxo em pontos específicos das vias, classificá-los conforme a marca do fabricante e modelo e registrar a conversão e condução irregular dos mesmos. Porém, para dotar um sistema de monitoração de tráfego de veículos baseado somente no uso de técnicas de PID é necessário solucionar vários problemas inerentes a própria natureza do problema, ao saber que esse sistema deve ser capaz de operar em ambiente não controlado e, que há uma diversidade intrínseca aos objetos (os veículos) de interesse (como: cor, tamanho, forma geométrica, variabilidade nos padrões de textura do veículo e o seu próprio movimento), além disso, as condições extrínsecas, relativas ao ambiente (como variação na intensidade da iluminação solar, chuva, sombras, interferências provocadas por outros objetos presentes na cena, etc.) (MONTANHA, 2010). Ou seja, para dotar um sistema baseado em PID com a capacidade para monitorar o fluxo de veículos, deve-se ser capaz de superar tanto os problemas causados pelo ambiente como aqueles causados pelos próprios objetos de interesse.

1.4 Metodologia da Pesquisa

Este trabalho aborda o problema da mensuração da velocidade de deslocamento de corpos rígidos baseado em análise de sequência de imagens. Para isso, foi realizado, em um primeiro momento, um estudo que permitisse abstrair diversos aspectos intrínsecos e

extrínsecos envolvidos com a natureza do problema, principalmente detectar o corpo alvo e, por meio do uso de métodos de segmentação de imagens, extrair pontos significativos que permitam mensurar a velocidade de deslocamento desse corpo. Entretanto, o ponto crucial a ser resolvido para cumprir com a finalidade deste trabalho é solucionar problemas ocasionados pela distorção das lentes das câmaras e obter parâmetros de transformação capazes de relacionar o plano imagem com o espaço objeto e, também, relacionar pontos de interesse (inerentes aos corpos) com pontos específicos do espaço objeto e, assim, ter condições de mensurar a velocidade de deslocamento de um corpo.

Para permitir realizar essa tarefa, foi desenvolvido um sistema computacional protótipo que contempla:

- a. Um módulo de interface entre o sistema computacional e a câmara para capturar as imagens adquiridas e disponibilizá-las para processamento;
- b. Um módulo de segmentação de imagens baseado em movimento;
- c. Um módulo de conversão de coordenadas do espaço imagem (2D) para o espaço objeto (3D) e vice versa.
- d. Um módulo de mensuração da velocidade de deslocamento de um corpo;
- e. Um módulo de visualização das sequências de imagens com sobreposição de detalhes computados;

Para melhor análise e validação do sistema, foi realizado um experimento em ambiente controlado (laboratório de Física) que usa um aparelho que possibilita calcular a velocidade de deslocamento de um corpo com elevada precisão. Nesse experimento, foi utilizada uma câmara comercial normal de baixo custo (capaz de adquirir até 30 quadros de imagens por segundo).

1.5 Organização do Trabalho

Este trabalho está organizado em seis capítulos. O primeiro e o segundo capítulos apresentam uma contextualização do problema, realiza o levantamento de alguns métodos relacionados com o tratamento do problema central deste trabalho. O terceiro Capítulo apresenta trabalhos relacionados com a mensuração da velocidade de deslocamento de automóveis. O quarto capítulo apresenta o desenvolvimento de um sistema de mensuração da velocidade de um corpo em movimento, descreve e analisa os resultados obtidos com os experimentos realizados. O quinto capítulo apresenta as considerações finais sobre o desenvolvimento e o sexto capítulo apresenta as referências utilizadas nesse trabalho.

FUNDAMENTAÇÃO TEÓRICA

Este Capítulo apresenta a fundamentação teórica relevante para o desenvolvimento do trabalho proposto e inclui conceitos referentes ao PID e Fotogrametria.

2.1 Imagem Digital

Uma imagem é definida por uma função de intensidade luminosa bidimensional, denotada por $f(x,y)$, na qual o valor ou amplitude de f na coordenada espacial (x,y) corresponde ao valor de brilho (ou intensidade) da imagem nesse ponto. Ao se considerar que a luz é uma forma de energia (eletromagnética) que ao interagir com a matéria (objetos ou alvos) pode ser parcialmente transmitida, absorvida e refletida, essa parcela de energia refletida (ou luz refletida) é que formará a imagem, quando for capaz de sensibilizar um observador ou sensor. Dessa forma, pode-se dizer que uma imagem é o resultado da transformação dessa energia em valores de brilho por uma determinada função de conversão. Devido ao fato de uma imagem ser proveniente de alguma forma de energia, esses valores de brilho que compõem uma imagem são positivos e finitos, ou

$$0 < f(x,y) < \infty \quad (2.1.1)$$

No caso de se considerar somente imagens formadas pela interação de objetos com a energia do espectro eletromagnético, a aparência de $f(x,y)$ depende de quatro fatores: 1) da

quantidade de energia (luz) que incide sobre o objeto $i(x,y)$; 2) da natureza do objeto, ou capacidade de refletir energia de determinados comprimentos de onda (ou reflectância $r(x,y)$); 3) da natureza dessa energia (espectro: visível, infravermelho, ultravioleta, termal, micro-ondas, etc.) e; 4) da sensibilidade do observador. Os fatores (3) e (4) são os elementos dependentes do contexto analisado. Supondo que o observador é capaz de perceber somente a energia proveniente do espectro eletromagnético visível, nosso caso, teoricamente, uma imagem $f(x,y)$ pode ser definida em função dos limites da intensidade da luz incidente $i(x,y)$ e dos valores de reflectância do objeto $r(x,y)$ que Gonzalez, et al (2000) definem da seguinte forma

$$\begin{aligned} f(x,y) &= i(x,y)r(x,y), \text{ em que} \\ 0 &< i(x,y) < \infty \\ 0 &< r(x,y) < 1 \end{aligned} \tag{2.1.2}$$

Uma imagem digital em escala de cinza é uma função $f(x,y)$ discreta que foi amostrada (digitalização das coordenadas espaciais) e quantizada (quantização dos níveis de cinza) por um sensor (câmara digital ou scanner fotográfico). A amostragem depende da resolução espacial em que a cena foi adquirida (que depende da distância focal da câmara e da distância entre o objeto e o sensor). A quantização depende da resolução radiométrica do sensor (número de *bits* usados para armazenar um valor quantizado). Essa quantização, para imagens monocromáticas, é chamada de nível de cinza l pertencente ao intervalo $[L_{min}, L_{max}]$ ou escala de cinza, a qual, na prática, é deslocada para $[0, L]$ em que 0 é considerado negro (ausência de luz refletida) e L é considerado branco (toda luz incidente é refletida). Todos os valores intermediários $0 < l < L$ são diferentes tons de cinza que variam entre o negro e o branco. Para um sensor com capacidade de representar a intensidade de um ponto da imagem (*pixel*) usando n *bits*, $L = 2^n - 1$ define o limite superior da escala de cinza para esse sensor. No caso mais comum, $n = 8$ *bits* define a variação da escala de cinza no intervalo $[0, 255]$.

Considerando esses conceitos preliminares, define-se formalmente a seguir uma imagem digital em escala de cinza.

Seja \mathbb{Z} o conjunto dos inteiros e seja \mathbb{Z}^+ o subconjunto dos inteiros positivos. Considere as funções $f, g \in Fun[E, K]$ onde, E é um conjunto retangular finito de pontos, ou malha retangular regular de suporte para a imagem, definido por $E \subset \mathbb{Z} \times \mathbb{Z}$ e K é um intervalo fechado em $\mathbb{Z}^+, K = [0, L]$ com $L \in \mathbb{Z}^+ \wedge L = \max(K)$. O reticulado $Fun[E, K]$ é o

conjunto de todas as funções possíveis no produto cartesiano $E \times K$ (PREPARATA e YEH, 1973).

Neste trabalho, toda $f \in Fun[E, K]$ define uma imagem binária, se $f \in Fun[E, \{0,1\}]$, ou com 256 níveis de cinza, se $f \in Fun[E, \{0,255\}]$.

2.1.1 Geometria de Imagens Digitais Adquiridas por Câmaras Digitais

O sistema de coordenadas de uma imagem digital tem origem no canto superior esquerdo da imagem. Os eixos de contagem das linhas e das colunas (i, j) são ortogonais (Figura 1).

Em processamentos fotogramétricos utiliza-se um sistema intermediário (x, y), com origem definida no centro da imagem, conforme apresentado na Figura 1. Para transformar as coordenadas do sistema de imagem para o sistema intermediário são necessárias: uma translação entre as origens dos sistemas, uma reflexão no eixo de ordenadas (devido ao dispositivo CCD obter uma imagem invertida) e um fator de escala equivalente ao tamanho do *pixel*, a fim de estabelecer as coordenadas num sistema métrico para qualquer *pixel* com coordenada (i, j) da imagem (Equação 2.1.3).

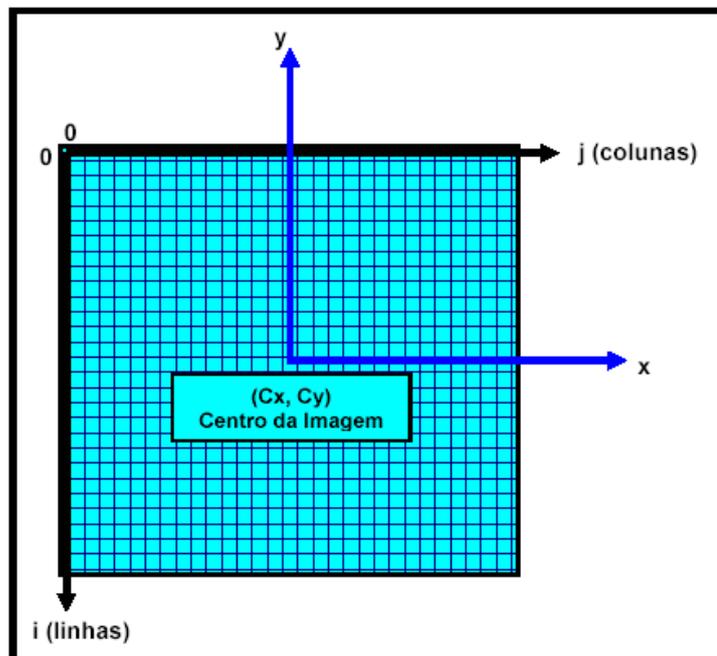


Figura 1- Definição do sistema de coordenadas de uma imagem digital.
Fonte: Adaptado de Ruy (2004)

$$\begin{cases} x_j = (j - c_x)p_x \\ y_j = (i - c_y)p_y \end{cases} \quad (2.1.3)$$

em que:

(i, j) : coordenada do *pixel* nos eixos do sistema de coordenadas da imagem;

(x_j, y_i) : coordenada métrica do *pixel* (i, j) nos eixos do sistema de coordenadas intermediário;

(c_x, c_y) : coordenada do centro da imagem no sistema de coordenadas da imagem com NC colunas e NL linhas, dadas por $c_x = \frac{NC-1}{2}$ e $c_y = \frac{NL-1}{2}$;

(p_x, p_y) : tamanho do *pixel* no sistema métrico (altura e largura, respectivamente);

Para efeito de simplificação, Ruy (2004) considera que o centro da imagem digital adquirida e o centro óptico das lentes da câmara se correspondem diretamente. Entretanto, nem sempre existe essa coincidência. O centro óptico da imagem só pode ser determinado em laboratórios fotogramétricos especializados em calibração de câmaras. Entretanto, a adoção dessa simplificação, neste trabalho, não deve promover grandes desvios na computação das coordenadas extraídas das imagens, uma vez que essas imagens são adquiridas a curta distância (~2 metros).

2.1.2 Conceito de vizinhança de *pixels*

Segundo Gonzalez, et al (2000), o conceito de vizinhança de um *pixel* é um dos relacionamentos mais básicos entre os *pixels* de uma imagem digital. Dado um ponto p de coordenadas (x, y) pertencentes a uma imagem, são chamados de vizinhança 4-conectada os *pixels* de coordenadas $(x + 1, y)$, $(x - 1, y)$, $(x, y + 1)$ e $(x, y - 1)$. A vizinhança 8-conectada acrescenta aos *pixels* da vizinhança 4-conectada os *pixels* diagonais ao ponto p com coordenadas $(x + 1, y + 1)$, $(x + 1, y - 1)$, $(x - 1, y + 1)$, $(x - 1, y - 1)$.

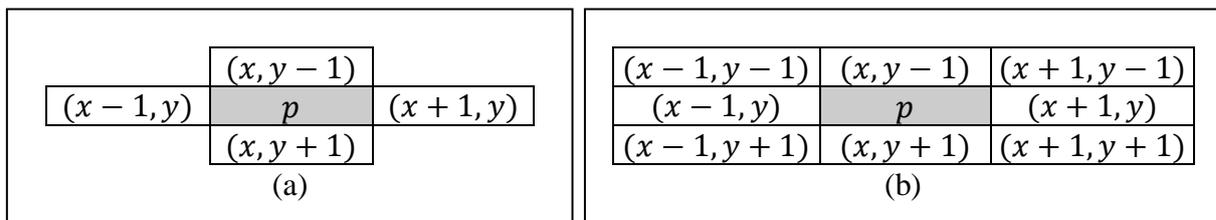


Figura 2 - Vizinhanças de um *pixel*. (a) 4-conectado. b) 8-conectado.

Ao se utilizar o conceito de vizinhança 4-conectada, todos os *pixels* conectados são vizinhos de borda, sendo todos equidistantes, enquanto que na 8-conectada nem todos os

vizinhos são equidistantes, pois os mesmos podem ser tanto vizinhos de borda como vizinhos na diagonal.

2.2 Processamento de Imagens Digitais

Processar uma imagem digital consiste em manipular os seus *pixels* e/ou o valor quantizado associado com cada *pixel* para transformá-la numa outra imagem. Esse processo permite: a) remover ruídos; b) alterar valores relacionados com o contraste, o brilho, a cor e saturação, associados com cada *pixel* da imagem; c) aplicar transformações geométricas; d) reconstruir imagens corrompidas (ex: reconstruir uma imagem corrompida por agentes degradantes) e; d) executar processos de segmentação com a finalidade de extrair atributos específicos de elementos que compõem a imagem (GONZALEZ, et al, 2000; MARQUES, et al, 1999).

A aplicação de técnicas e métodos relacionados com PID constitui o aspecto central no tratamento do problema investigado neste trabalho (mensurar a velocidade de deslocamento de corpos rígidos por meio da análise de sequência de imagens). São abordados a seguir alguns conceitos e técnicas básicas da área de PID.

2.2.1 Realce de imagens

O principal objetivo das técnicas de realce é processar uma imagem de modo que a imagem resultante seja mais adequada para uma aplicação específica do que a imagem original. Essas técnicas são bastante dependentes da aplicação. Assim, um método que é útil para realçar imagens adquiridas por raios X pode não ser a melhor abordagem para realçar imagens adquiridas por câmaras de monitoramento (GONZALEZ, et al, 2000).

As técnicas de realce de imagem são divididas em duas categorias: métodos no domínio espacial e métodos no domínio da frequência. Os métodos no domínio espacial operam sobre um agregado de *pixels* que compõem a imagem. Os resultados desses métodos são dependentes dos valores e das posições dos *pixels* envolvidos. Os métodos no domínio da frequência são baseados na modificação das transformadas de Fourier da imagem e são fundamentados pelo teorema da convolução da imagem por um operador linear invariante com a posição, ou seja, o resultado produzido por esse operador para um ponto (x, y) de uma

imagem f depende apenas do valor associado com esse ponto e não de sua posição. Técnicas de realce baseadas em várias combinações de métodos dessas duas categorias não são incomuns (GONZALEZ, et al, 2000).

As técnicas de realce de imagens digitais podem ser relacionadas com a manipulação de contraste, realce de bordas e suavização. Basicamente, as técnicas de realce de imagens são heurísticas e orientadas para aplicações específicas. Não existe um critério de avaliação universal. O sucesso da aplicação depende do julgamento subjetivo do analista (FACON, 1993).

A seguir são apresentadas algumas técnicas de realce em imagens, em especial, os filtros do domínio espacial.

2.2.1.1 Filtros da ordem (domínio espacial)

Os filtros espaciais são usados no PID para eliminação ou redução de ruídos, para suavizar a imagem ou para aguçamento de detalhes finos. As técnicas de filtragem espacial fazem uso de máscaras espaciais para processamento de imagens. Essas máscaras são chamadas filtros espaciais e são utilizadas para estabelecer a noção de vizinhança de um *pixel* localizado na posição (x, y) da imagem f que, geralmente, corresponde ao *pixel* coberto pela célula central da máscara. Na maioria das aplicações, essas máscaras são simétricas, quadradas ou retangulares, pois tornam a implementação computacional mais fácil.

O aspecto básico da filtragem espacial consiste em mover a máscara sobre a imagem, *pixel a pixel*, e aplicar um operador previamente definido sobre os *pixels* cobertos pela máscara, vizinhos de $f(x, y)$ e atribuir o resultado à $g(x, y)$.

As operações de filtragem espacial podem ser divididas em duas classes: (1) filtragem linear que produz valores através de uma operação de convolução entre os valores vizinhos de $f(x, y)$ cobertos pela máscara e os valores dos coeficientes da máscara (GONZALEZ, et al, 2000; FACON, 1993) e; (2) Filtragem não-linear que produz valores dependentes do operador usado, o qual processa diretamente os valores dos *pixels* cobertos pela máscara.

Os seguintes filtros pertencem à classe de filtros não-lineares: Filtro da média geométrica, filtro contra-harmônico, filtro da média harmônica, filtro de máximo, filtro de mínimo e filtro da mediana. Os filtros de máximo, de mínimo e da mediana são chamados filtros da ordem.

2.2.1.1.1 Filtro da média

O filtro da média (equação 2.2.1) tem como principal função gerar uma nova imagem $f(x, y)$, suavizando detalhes que compõem a imagem original $g(x, y)$, substituindo o valor do *pixel* (x, y) atual pela média dos *pixels* presentes no conjunto de coordenadas S_{xy} da subimagem retangular de tamanho $m \times n$ (GONZALEZ, et al, 2000).

$$f(x, y) = \frac{1}{mn} \sum_{(s,t) \in S_{xy}} g(s, t) \quad (2.2.1)$$

2.2.1.1.2 Filtro da mediana, filtro do máximo e filtro do mínimo

O filtro da mediana (equação 2.2.2) junto com os filtros do mínimo (equação 2.2.3) e do máximo (equação 2.2.4) compõem os chamados filtros da ordem (GONZALEZ, et al, 2000). Na prática, para se gerar uma nova imagem $f(x, y)$, a partir de uma imagem $g(x, y)$, todos os valores associados com os *pixels* cobertos pela máscara de dimensões (*pixels* presentes no conjunto de coordenadas S_{xy}) de um filtro da ordem são ordenados (ascendentemente) e alocados em uma lista. Para aplicar o filtro da mediana no *pixel* atual (x, y) , basta substituir seu valor pelo valor alocado na posição central da lista ordenada gerada. No caso do filtro do mínimo, utiliza-se o primeiro valor e, para o filtro do máximo, é utilizado o último valor presente nessa lista (GONZALEZ, et al, 2000).

$$f(x, y) = \text{median}\{g(s, t)\}_{(s, t) \in S_{xy}} \quad (2.2.2)$$

$$f(x, y) = \text{min}\{g(s, t)\}_{(s, t) \in S_{xy}} \quad (2.2.3)$$

$$f(x, y) = \text{max}\{g(s, t)\}_{(s, t) \in S_{xy}} \quad (2.2.4)$$

2.2.2 Morfologia Matemática

A Morfologia Matemática concentra seus esforços no estudo da estrutura geométrica das entidades presentes em uma imagem. Seu princípio básico consiste em extrair as informações relativas à geometria e à topologia de um conjunto desconhecido de uma imagem pela transformação a partir de outro conjunto completamente definido, chamado de elemento estruturante, em que a partir de seu formato e de seu tamanho, é possível testar e quantificar de que maneira o elemento estruturante “está ou não está contido” na imagem.

Segundo Marques Filho, et al, (1999), na morfologia matemática, os conjuntos representam os objetos encontrados em uma imagem. Por exemplo, o conjunto de todos os *pixels* pretos em uma imagem binária é uma descrição morfológica completa da imagem. Em imagens binárias, os conjuntos em questão são membros do espaço inteiro bidimensional \mathbb{Z}^2 , no qual cada elemento do conjunto é um vetor 2-D cujas coordenadas são as coordenadas (x, y) do *pixel* preto (por convenção) na imagem. Imagens com mais níveis de cinza podem ser representadas por conjuntos cujos elementos estão no espaço \mathbb{Z}^3 . Neste caso, os vetores têm três elementos, sendo os dois primeiros as coordenadas de um *pixel* e o terceiro o valor de nível de cinza associado com esse *pixel*.

Para um melhor entendimento da morfologia matemática e de seus principais operadores e filtros morfológicos, são apresentadas algumas definições úteis sobre teoria dos conjuntos, juntamente com a definição de erosão, dilatação, abertura e fechamento.

2.2.2.1 Definições básicas

Sejam A e B conjuntos em \mathbb{Z}^2 , cujos componentes são $a = (a_1, a_2)$ e $b = (b_1, b_2)$, respectivamente. A translação de A por $x = (x_1, x_2)$, denotada $(A)x$, é definida como (MARQUES FILHO, et al, 1999):

$$(A)x = \{c | c = a + x, \text{ para } a \in A\} \quad (2.2.5)$$

A reflexão de B , denotada \hat{B} , é definida como:

$$\hat{B} = \{x | x = -b, \text{ para } b \in B\} \quad (2.2.6)$$

O complemento do conjunto A é:

$$A^c = \{x|x \notin A\} \quad (2.2.7)$$

Finalmente, a diferença entre dois conjuntos A e B , denotada $A - B$, é definida como:

$$A - B = \{x|x \in A, x \notin B\} = A \cap B^c \quad (2.2.8)$$

2.2.2.2 Dilatação

Sejam A e B (B representa o elemento estruturante) conjuntos no espaço \mathbb{Z}^2 e seja \emptyset o conjunto vazio. A dilatação de A por B , denotada $A \oplus B$, é definida como:

$$A \oplus B = \{x|(\hat{B})_x \cap A \neq \emptyset\} \quad (2.2.9)$$

Portanto, o processo de dilatação visa obter a reflexão de B sobre sua origem e posteriormente deslocar esta reflexão de x . A dilatação de A por B é, então, definida como o conjunto de todos os x deslocamentos para os quais a interseção de $(\hat{B})_x$ e A inclui pelo menos um elemento diferente de zero. Com base nesta interpretação, a equação anterior pode ser escrita como:

$$A \oplus B = \{x|[(\hat{B})_x \cap A] \subseteq A\} \quad (2.2.10)$$

2.2.2.3 Erosão

Sejam A e B conjuntos no espaço \mathbb{Z}^2 . A erosão de A por B , denotada $A \ominus B$, é definida como:

$$A \ominus B = \{x|(B)_x \subseteq A\} \quad (2.2.11)$$

o que, em outras palavras, significa dizer que a erosão de A por B resulta no conjunto de pontos x tais que B , transladado de x , está contido em A .

2.2.2.4 Abertura e Fechamento

A abertura em geral suaviza o contorno de uma imagem, rompe os istmos e elimina proeminências delgadas. O fechamento, por sua vez, funde pequenas quebras e alarga golfos

estreitos, elimina pequenos orifícios e preenche lacunas no contorno.

A abertura de um conjunto A por um elemento estruturante B , denotada $A \circ B$, é definida como:

$$A \circ B = (A \ominus B) \oplus B \quad (2.2.12)$$

o que equivale a dizer que a abertura de A por B é simplesmente a erosão de A por B seguida de uma dilatação do resultado por B .

O fechamento do conjunto A pelo elemento estruturante B , denotado $A \cdot B$, é definido como:

$$A \cdot B = (A \oplus B) \ominus B \quad (2.2.13)$$

o que nada mais é que a dilatação de A por B seguida da erosão do resultado pelo mesmo elemento estruturante B .

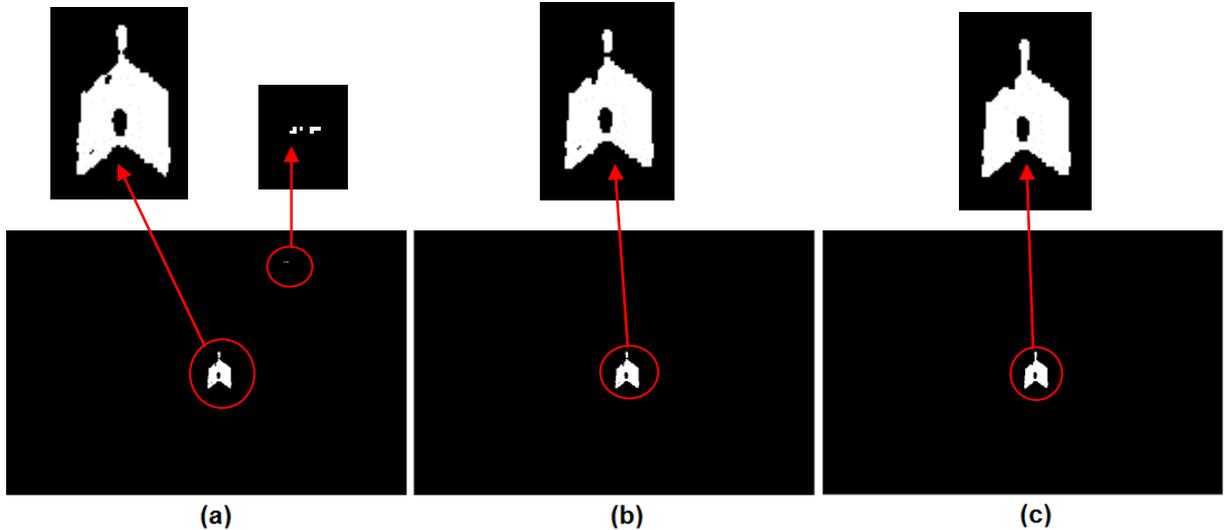


Figura 3 - Aplicação da abertura e fechamento na mascara gerada pelo processo de detecção de movimento da cena

A Figura 3 (a) ilustra a imagem resultante do processo de identificação do movimento da cena imageada por meio do *software* de mensuração da velocidade de corpos em movimento, em que os *pixels* de cor branca representam as áreas da imagem atual que sofreram alteração com relação à imagem de referência (fundo). A Figura 3 (b) representa a imagem resultante da operação de abertura, a qual utilizou um elemento estruturante de

tamanho 1, na imagem representada na Figura 3(a). Essa operação morfológica foi realizada para remover os *pixels* ruídos que não representam o movimento do objeto de interesse e romper os istmos. A Figura 3 (c) representa o resultado da operação de fechamento realizado na imagem presente na Figura 3 (b), tal operação morfológica foi realizada para preencher as lacunas e os orifícios presentes na imagem.

2.2.3 Segmentação de imagens

Em geral, a segmentação automática de imagens não é uma tarefa fácil e a qualidade dos resultados obtidos nessa fase define o eventual sucesso ou fracasso de análises posteriores. Em aplicações que demandam a detecção de alvos, em geral, não se tem controle sobre o ambiente. Nesses casos, a abordagem usual é concentrar-se na seleção de tipos de sensores que consigam realçar esses alvos de interesse enquanto diminuem a contribuição de componentes irrelevantes da imagem (GONZALEZ, et al, 2000).

O objetivo da segmentação é particionar uma imagem em regiões com propriedades similares. O grau de similaridade de cada região é avaliado por um critério de homogeneidade. Um processo de segmentação é completo se as regiões segmentadas forem disjuntas. Isto é, cada *pixel* da imagem original $f(x, y)$ pode pertencer a uma única região R_j e existirá pelo menos duas regiões: a região de interesse e a região do fundo, Equação 2.2.14 (STEUDEL, et al, 1996).

$$\left(f = \bigcup_{\forall f} R_j \right) \wedge (\forall j \neq k \cdot R_j \cap R_k = \emptyset) \quad (2.2.14)$$

Os métodos de segmentação são classificados em seis categorias (GONZALEZ, et al, 2000). A seguir, é feita uma breve descrição de cada categoria. Alguns comentários foram adicionados para contemplar alguns casos envolvidos com os objetivos deste trabalho.

- 1) Métodos baseados na amplitude: Nesta categoria estão os métodos que usam valor de limiar aplicado sobre valores de luminância ou sobre valores de componentes de cor de uma imagem. Geralmente, esse valor (ou valores) de limiar é obtido mediante análise do histograma dos valores de brilho dos *pixels* da imagem. Entretanto, obter automaticamente um valor de limiar para uma aplicação específica constitui um passo determinante para a solução da maioria dos problemas relacionados com análise e reconhecimento de alvos específicos em imagens digitais. A exploração e

inter-relacionamento de diferentes atributos de uma imagem podem facilitar a solução desse problema, por exemplo, de uma imagem colorida composta pelas três bandas espectrais RGB (*Red-Green-Blue*) que cobrem a faixa do espectro eletromagnético relativo ao visível pode ser extraído outros componentes que dissociam a luminosidade dos valores relativos à cor, como o sistema de cor HSI (*Hue-Saturation-Intensity*).

- 2) Métodos baseados em agrupamento de *pixels* (*clustering*): O agrupamento de *pixels* de uma imagem é feito sobre atributos associados com esses *pixels*. Esses atributos são, geralmente, extraídos de imagens multiespectrais ou derivados dos atributos de cor e/ou dos valores de luminância. A união dos conjuntos desses N atributos associados com cada *pixel* forma o espaço N -dimensional de atributos o qual, segundo algum critério de similaridade ou de tomada de decisão, é subdividido em classes mutuamente exclusivas. Cada classe determina alvos específicos da imagem. Para aplicar os métodos dessa categoria deve-se, *a priori*, determinar quais são as classes de interesse e quais atributos associados com esses *pixels* são relevantes nesse processo. O resultado final dos agrupamentos é dependente dessa escolha inicial. Mesmo que se deseje obter duas classes de alvos, a classe “A” contendo os alvos de interesse e outra “B” contendo as interferências que se deseja eliminar, pode ser necessário determinar mais de duas classes iniciais para melhorar a qualidade dos agrupamentos. Análises posteriores permitirão unir classes que contenham os alvos de interesse em uma única classe. Alguns desses métodos utilizam estratégias robustas de agrupamento que empregam, por exemplo, redes neurais e critérios estatísticos baseados na máxima verossimilhança. De forma geral, esses métodos exigem que sejam estabelecidos parâmetros e regras apropriadas para realizar os agrupamentos. Esses parâmetros e regras podem ser obtidos mediante treinamento supervisionado previamente realizado. Entretanto, esses métodos podem conduzir a soluções dependentes da imagem, ou seja, para cada nova imagem que contenha os mesmos elementos da primeira imagem, pode ser necessário estabelecerem novos parâmetros e regras para se realizar um agrupamento eficiente, descaracterizando o conceito de ferramenta automática.
- 3) Métodos baseados em regiões: Os métodos de segmentação baseados na amplitude e no agrupamento usam atributos associados com cada *pixel* da imagem para estabelecer os alvos de interesse. Os métodos baseados em regiões utilizam esses atributos pontuais associados com a espacialização desses *pixels*, por considerar que

toda região é formada por *pixels* conexos. Crescimento de região (*region growing*) e divisão e fusão (*split and merge*) são dois métodos clássicos de segmentação orientados a regiões.

- 4) Métodos baseados na detecção de fronteiras: É possível segmentar uma imagem em regiões através da detecção de fronteiras entre essas regiões. Os *pixels* da fronteira ou do contorno de uma região homogênea, geralmente, sofrem uma alteração brusca no nível de cinza ou na cor predominante devido à interferência dos *pixels* das regiões adjacentes. Esses *pixels* com atributos alterados podem aparecer como um conjunto de pontos, de linhas ou de segmentos. Como se trata de uma mudança de amplitude nos valores de atributos (geralmente, valor de brilho) desses *pixels*, a seleção de um valor de limiar apropriado pode, em alguns casos, produzir resultados satisfatórios na obtenção das fronteiras de uma região. O principal problema dessa abordagem é considerar, implicitamente, que a fronteira entre regiões seja sempre bem definida. Se a imagem é ruidosa ou se os atributos entre as regiões se diferem muito pouco, podem ser detectadas linhas de fronteiras desconectadas ou fragmentadas. Se existe um conhecimento *a priori* da forma das regiões, essas discontinuidades podem ser detectadas e ligadas para obter um contorno fechado. Para regiões com formas complexas e desconhecidas é necessário supor um contorno fechado usando métodos de interpolação mais complexos, como: polinômios de Bezier ou ajustamento de curvas por Spline. Se a linha da fronteira é definida por retas, a transformada de Hough é o método mais indicado para detectar e fechar o contorno. Um operador diferencial local pode ser usado para delimitar e detectar as bordas de um alvo específico, desde que ele seja homogêneo e distinguível entre outros alvos. Os operadores mais usados são: Sobel, Laplaciano, Prewitt e gradiente morfológico. Existem outros operadores que exploram outros atributos como Nevatia & Babu que explora a possibilidade de várias inclinações de linhas (ARTERO, et al, 1999). Flores *et al* (2004) exploraram a variabilidade dos atributos de cor, saturação e intensidade do sistema de cor HSI para obter resultados mais completos e minimizar o número de linhas dentro de uma mesma região.
- 5) Métodos baseados em medidas de textura. Uma importante abordagem para a discriminação de regiões é a quantificação de seu conteúdo de textura. Embora não exista uma definição formal de textura, esse descritor intuitivamente fornece medidas de propriedades de uma superfície como suavidade, rugosidade e regularidade (GONZALEZ, et al, 2000). Hawkins (1970) fez uma descrição de

textura: "a noção de textura parece depender de três ingredientes: (1) alguma ordem de arranjo local que se repete no interior de uma região a qual é muito maior que o tamanho do arranjo; (2) essa ordem consiste de partes elementares arranjadas de forma não-randômica; e (3) essas partes são entidades quase uniformes que têm aproximadamente as mesmas dimensões dentro da 'região de textura' que está sendo considerada". Embora essa descrição de textura pareça ser razoavelmente perceptual, ela não conduz imediatamente a uma medida quantitativa simples de textura. Textura é um atributo que qualifica a superfície de um objeto e mede o quanto essa superfície é homogênea ou contrastada. Entretanto, quantificar e qualificar atributos de textura é uma tarefa complexa que envolve técnicas de PID e de reconhecimento de padrões. Quando aplicada em imagens coloridas tradicionais, a informação de textura da superfície é extraída da variação da tonalidade dos *pixels*. Uma revisão de vários métodos para obter medidas de textura é feita por Manian, et al (1999) que apresentaram uma aplicação usando imagens de sensoriamento remoto adquiridas por sistemas de radar.

- 6) Métodos baseados em movimento. Esses métodos exploram a ideia de movimento na cena. Conceitualmente, o sensor está parado em relação a um sistema inercial e obtém vários quadros de imagens (*frames*) do mesmo local em intervalos de tempo distintos. A mudança ocorrida entre duas ou mais imagens distintas é causada pelo movimento dos alvos. Essa noção de movimento é usada para segmentar o fundo (estático) em relação aos alvos que se movem. Esses métodos são adequados para tratar os problemas envolvidos neste trabalho.

A seguir são detalhados os métodos de segmentação que são efetivamente usados neste trabalho.

2.2.3.1 Segmentação Baseada em Valor de Limiar

A limiarização é uma das técnicas mais usadas em segmentação de imagens que pode ser definida como uma operação que envolve testes de uma função T da forma

$$T = T[x, y, p(x, y), f(x, y)] \quad (2.2.15)$$

em que, T é o valor de limiar, $f(x, y)$ é o nível de cinza do ponto (x, y) e $p(x, y)$ denota alguma propriedade local desse ponto, por exemplo, o nível de cinza médio de uma

vizinhança centrada em (x, y) .

Uma imagem limiarizada $g(x, y)$ é definida como:

$$g(x, y) = \begin{cases} 1, & f(x, y) > T \\ 0, & f(x, y) \leq T \end{cases} \quad (2.2.16)$$

Neste caso, *pixels* rotulados com o valor 1 em $g(x, y)$ correspondem ao objeto de interesse na imagem $f(x, y)$, enquanto que aqueles rotulados com o valor 0 correspondem a outros artefatos que se deseja descartar (chamado de fundo da imagem).

Quando T depender apenas de $f(x, y)$, o limiar é chamado global. Se depender tanto de $f(x, y)$ como de $p(x, y)$, então é chamado dinâmico.

Uma técnica de limiarização global simples é aquela do particionamento do histograma de uma imagem por um único valor de limiar T . A segmentação é efetuada, percorrendo *pixel a pixel* da imagem $f(x, y)$ e mapeando em $g(x, y)$ cada um dos *pixels* (x, y) como integrante do objeto de interesse (valor de rótulo 1) ou integrante do fundo da imagem (valor de rótulo 0). Obviamente, o sucesso desse método depende da precisão na especificação do valor de limiar T que particiona o histograma no ponto de separação das classes fundo da imagem e objetos de interesse. Na teoria, a segmentação de imagens por valor de limiar tem um conceito simples. Na prática, encontrar um valor de limiar satisfatório de forma automática é uma tarefa delicada que envolve formulações empíricas.

Para ilustrar a natureza desse problema, tomemos um exemplo aparentemente simples. A imagem mostrada na Figura 4 (a) apresenta um corpo rígido presente em um trilho de ar sobre um fundo escuro. A Figura 4 (b) mostra o histograma da distribuição dos níveis de cinza associados com os *pixels* dessa imagem, onde se observa a formação de dois picos e, entre esses dois picos existe um vale. O pico da esquerda representa a distribuição dos *pixels* do fundo da imagem e o da direita dos *pixels* do objeto de interesse. Como esse histograma tem característica bi-modal, deveria ser possível assumir qualquer valor no histograma entre aqueles cobertos pelo vale como um valor de limiar T capaz de realizar a segmentação da imagem. O vale desse histograma está compreendido no intervalo [18,26], aproximadamente. Então, o valor de limiar $T=23$ foi escolhido e aplicado na segmentação da imagem da Figura 4 (a). O resultado alcançado pode ser visto na Figura 4 (c).

Observa-se na parte inferior da Figura 4 (c) que alguns objetos foram parcialmente segmentados e que um deles não faz parte do objeto de interesse. O que deve ser analisado agora é: (1) o quão suficiente é esse resultado? (2) o valor de limiar $T=23$ continuará a

produzir resultados equivalentes para outras imagens adquiridas no mesmo padrão da imagem usada? Se não for, como se faz para determinar um valor de limiar necessário para realizar a segmentação de cada uma dessas próximas imagens de forma satisfatória? Nos casos em que o histograma de uma imagem é bimodal, o valor de limiar que separa o fundo do objeto desejado pode ser calculado ao invés de arbitrado. Um dos métodos clássicos que permite calcular esse valor de limiar é o método de Otsu.

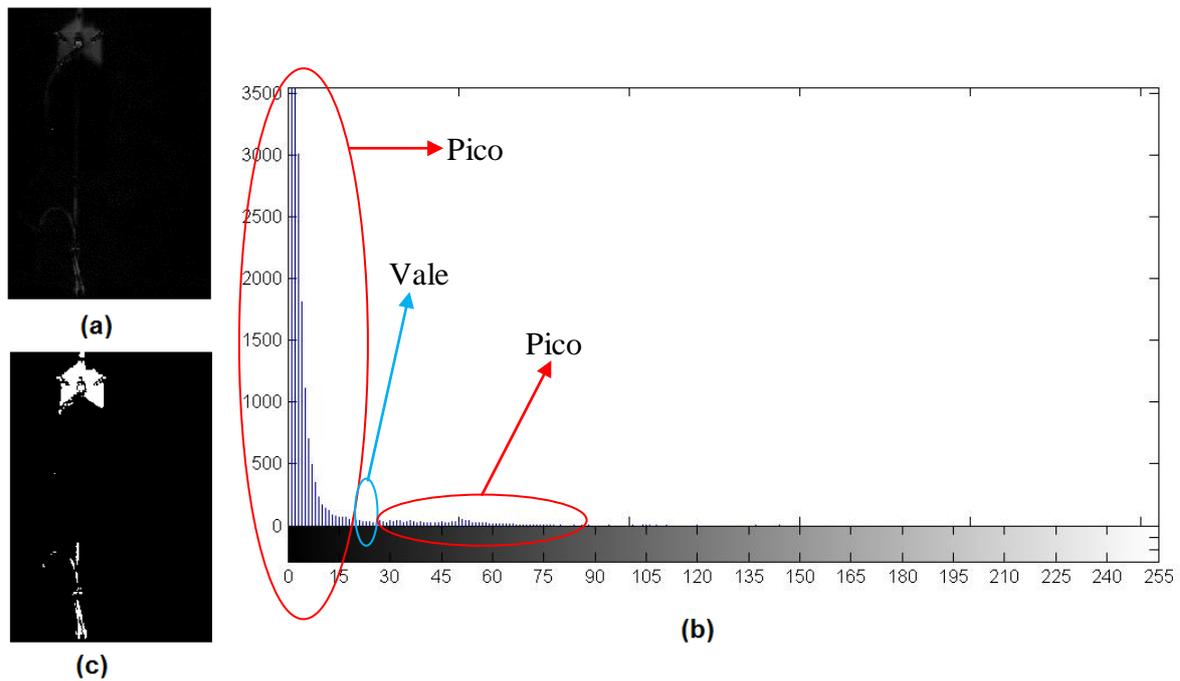


Figura 4 - Segmentação baseada em um valor de limiar

O método de segmentação de Otsu é baseado na análise discriminante e o valor do limiar é obtido supondo que os *pixels* da imagem podem ser classificados em duas classes: fundo (C_1 : níveis de cinza no intervalo $[1, t]$) e o objeto (C_2 : níveis de cinza no intervalo $[t + 1, L]$) (OTSU, 1979). Neste caso, a distribuição de probabilidade de ambas as classes podem ser descritas por:

$$C_1: \frac{p_1}{\omega_1(t)}, \dots, \frac{p_t}{\omega_1(t)} \quad \text{e} \quad C_2: \frac{p_{t+1}}{\omega_2(t)}, \dots, \frac{p_L}{\omega_2(t)} \quad (2.2.17)$$

onde:

$$\omega_1(t) = \sum_{i=1}^t p_i \quad \text{e} \quad \omega_2(t) = \sum_{i=t+1}^L p_i \quad (2.2.18)$$

As médias para as classes C_1 e C_2 são dadas por:

$$\mu_1 = \sum_{i=1}^t \frac{i \cdot p_i}{\omega_1(t)} \quad \text{e} \quad \mu_2 = \sum_{i=t+1}^L \frac{i \cdot p_i}{\omega_2(t)} \quad (2.2.19)$$

Sendo μ_T a intensidade média para toda a imagem, então:

$$\omega_1(t)\mu_1 + \omega_2(t)\mu_2 = \mu_T \quad \text{e} \quad \omega_1(t) + \omega_2(t) = 1 \quad (2.2.20)$$

Utilizando a análise discriminante, a variância entre as classes da imagem limiarizada é definida por:

$$\sigma_B^2 = \omega_1(t)(\mu_1 - \mu_2)^2 + \omega_2(t)(\mu_2 - \mu_1)^2 \quad (2.2.21)$$

O limiar ótimo t^* é determinado como sendo aquele cuja variância entre as classes σ_B^2 é máxima, isto é:

$$t^* = \arg \max \{ \sigma_B^2(t) \}, 1 \leq t \leq L \quad (2.2.22)$$

2.2.3.2 Segmentação de imagens baseada em movimento

Em aplicações de imageamento, o movimento é originado de um deslocamento relativo entre o sistema sensor e a cena observada, como em aplicações de robótica, navegação autônoma e análise dinâmica de cena. O movimento pode ser usado em segmentação tanto no domínio espacial quanto no da frequência. Este trabalho considera apenas as técnicas de segmentação no domínio espacial. A abordagem básica e das diferenças acumuladas são baseadas em Gonzalez, et al (2000).

2.2.3.2.1 Abordagem Básica

A abordagem mais simples para detectar mudanças entre dois quadros de imagens $f(x, y, t_i)$ e $f(x, y, t_j)$, $i \neq j$, tomados nos instantes t_i e t_j , respectivamente, é através da comparação *pixel a pixel* entre essas duas imagens. A forma mais simples de se fazer essa comparação é compor uma nova imagem resultante do cálculo da diferença entre essas imagens adquiridas nos quadros t_i e t_j . Uma melhor situação pode ser obtida quando se dispõe de uma imagem de referência que contém apenas os elementos estacionários (elementos do fundo da imagem). A comparação de um quadro de imagem qualquer com essa imagem de referência minimiza as interferências do fundo sobre aqueles objetos que se deslocaram em um determinado intervalo de tempo.

A imagem da diferença entre as imagens registradas em dois quadros tomados nos instantes t_i e t_j é definida como:

$$\Delta_{ij}(x, y) = \begin{cases} 1 & \text{se } |f(x, y, t_i) - f(x, y, t_j)| > \theta \\ 0 & \text{se } |f(x, y, t_i) - f(x, y, t_j)| \leq \theta \end{cases} \quad (2.2.23)$$

em que, θ é um valor de limiar

Em análise dinâmica de imagem, $\Delta_{ij}(x, y)$ com valor igual a 1 deve ser considerado como sendo o resultado do movimento de objeto(s) na cena. Entretanto, essa abordagem somente pode ser aplicada se as imagens de cada quadro forem registradas (existir a correspondência *pixel a pixel*) e se a iluminação for mantida relativamente constante dentro dos limites estabelecidos por θ . Os valores 1's ruidosos resultantes em $\Delta_{ij}(x, y)$ são frequentemente formados por pontos isolados, ou pequenas estruturas formadas por poucos pontos conectados. Esses ruídos podem ser eliminados por um valor de limiar de área.

2.2.3.2.2 Diferenças Acumuladas

A abordagem básica descrita anteriormente, frequentemente, necessita envolver um processo de filtragem por valor de limiar de área que pode eliminar pequenas estruturas que se movem ou alvos que se movem muito lentamente. A técnica de segmentação pelas diferenças

acumuladas procura registrar uma memória do evento, considerando vários quadros subsequentes e não apenas dois. A ideia básica é ignorar as mudanças que ocorrem esporadicamente na cena, pois, essas mudanças esporádicas são registradas pela interferência dos elementos ruidosos.

Então, a análise da dinâmica do movimento na cena deve ser registrada entre vários quadros, a partir de uma imagem de referência. Assim, considerando a seguinte sequência de quadros de imagem: $f(x, y, t_1), f(x, y, t_2), f(x, y, t_3), \dots, f(x, y, t_n)$, e seja $f(x, y, t_1)$ a imagem de referência. Uma imagem de diferenças acumuladas é formada pela comparação dessa imagem de referência com cada imagem subsequente na sequência, ou

$$\Delta_j^1(x, y) = \begin{cases} 1 & \text{se } |f(x, y, t_1) - f(x, y, t_j)| > \theta \\ 0 & \text{se } |f(x, y, t_1) - f(x, y, t_j)| \leq \theta \end{cases} \quad (2.2.24)$$

A imagem das diferenças acumuladas D é obtida por

$$D(x, y) = \sum_{j=1}^n \Delta_j^1(x, y) \quad (2.2.25)$$

Nessa situação, cada *pixel* (x, y) na imagem $D(x, y)$ funciona como um contador para o número de vezes que ocorreu uma alteração significativa naquela posição, em relação à imagem de referência. Portanto, quando a n -ésima imagem for comparada com a imagem de referência, na imagem $D(x, y)$ ficará registrado o número de vezes que cada *pixel* teve seu valor de nível de cinza alterado significativamente. Quanto maior for o número de alterações, maior é a possibilidade de se encontrar um objeto em movimento.

Para sistemas de navegação autônoma de alta precisão, quando é necessário localizar espacialmente um objeto em movimento em pequenos intervalos de tempo, essa abordagem pode não produzir resultados adequados, por ser necessário dispor de diversos quadros para se fazer uma detecção com erro minimizado. Além disso, se existirem objetos que se deslocam em velocidade elevada, essa abordagem pode não produzir os resultados desejados.

2.2.3.2.3 Motion History Image - MHI

Conforme Davis, et al (1997), o MHI é um método de visão computacional voltado para discriminar seqüências de movimento humano como, por exemplo, o movimento realizado por uma bailarina (Figura 6).

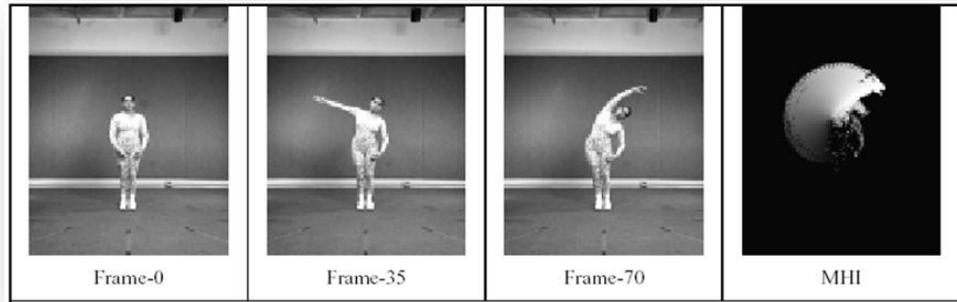


Figura 6 - MHI Gerado pelo movimento do braço de uma bailarina.
Fonte: Davis, et al, 1999.

Seu funcionamento é baseado na composição de sucessivas camadas constituídas por pontos relativos aos movimentos (silhuetas) detectados entre dois quadros de imagem no instante t e $t+1$. A detecção de pontos que se movimentam numa seqüência de imagens pode ser feita pelo simples cômputo da diferença absoluta entre dois quadros de imagem $f(x, y, t_i)$ e $f(x, y, t_{i+1})$, esse cálculo gera como resultado um realce em todos os pontos que sofreram alteração no seu valor original de intensidade, caracterizando que esses pontos se moveram (rotação e/ou translação) durante a aquisição desses dois quadros de imagem (Montanha, 2010 apud Davis, et al 1997). Considerando a aplicação de um valor de limiar no processo de segmentação por limiarização, capaz de detectar mudanças significativas entre esses dois quadros de imagens, obtém-se a silhueta $s(x, y, i)$ do movimento relativo ocorrido no período de tempo decorrido entre a aquisição do quadro t_i e do quadro t_{i+1} com a aplicação da Equação (2.2.26).

$$s(x, y, i) = \begin{cases} t_i, & \text{se } |f(x, y, t_i) - f(x, y, t_{i+1})| > \theta \\ 0, & \text{se } |f(x, y, t_i) - f(x, y, t_{i+1})| \leq \theta \end{cases} \quad (2.2.26)$$

A união de uma seqüência de silhuetas compõe o MHI (H), conforme regula a equação (2.2.27)

$$H(x, y, t_i) = \begin{cases} t_i, & \text{se } s(x, y, i) > 0 \\ \max(0, H(x, y, t_{i-1})), & \text{caso contrário} \end{cases} \quad (2.2.27)$$

Ao saber que em cada intervalo de tempo $i = 1, 2, 3, \dots, \delta$ considerado e, que um quadro de imagem é adquirido pelo sensor em cada um desses intervalos, o valor δ especifica um intervalo de tempo que se deseja analisar em uma sequência de imagens para compor o MHI.

A aplicação da equação (2.2.27) é restrita a uma única sequência de imagens em que não haja repetição de movimentos ocorrendo no mesmo local do plano imagem. Se existe possibilidade de haver trajetórias de movimentos repetidas (por objetos distintos ou não) numa determinada sequência de imagens (como os casos abordados neste trabalho, onde diferentes veículos percorrem uma mesma trajetória), deve-se prover um mecanismo que seja capaz de fazer com que regiões do MHI que não detectaram movimento depois de decorrido um período de tempo δ descartem os movimentos antigos que foram detectados e processados antes de começar a análise de um novo período. Esse mecanismo é considerado pela aplicação da equação (2.2.28).

$$H(x, y, t_i) = \begin{cases} t_i & \text{se } s(x, y, i) > 0 \\ 0 & \text{senão se } H(x, y, t_i) < (\delta - t_i) \end{cases} \quad (2.2.28)$$

Com essa modificação, Davis, et al (1997) considera que no MHI cada valor associado com um *pixel* está definido como uma função de histórico da posição temporal (ou movimento) relativo a esse ponto e, dessa forma, define o MHI em função de um operador de duração temporal.

Os diferentes valores de intensidade associados aos *pixels* registrados na imagem pelo MHI permitem extrair informações sobre a natureza do movimento, tais como a direção (*pixel* escuros e claros) e o sentido (dos *pixels* escuros para os *pixels* claros) utilizando os contornos entre as silhuetas registradas. Entretanto, a magnitude com que esse movimento ocorre não é uma informação acessível, pois, pode haver regiões sobrepostas (ou regiões de falhas) de registro de movimentos no MHI (DAVIS, et AL 1997).

Porém, segundo Montanha (2010), é possível utilizar as máscaras F_x e F_y (equação 2.2.29) do gradiente de Sobel para determinar a orientação $\phi(x, y)$ do movimento de cada *pixel* (x, y) que compõe o gradiente (equação 2.2.29).

$$F_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad F_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \quad (2.2.29)$$

$$\varnothing(x, y) = \arctan\left(\frac{F_y(x, y)}{F_x(x, y)}\right) \quad (2.2.30)$$

Davis, et al (1999) constataram que o ganho em tempo de processamento é muito vantajoso quando se trata de aplicações em tempo real, isto se deve à recursividade, em que o MHI no tempo t é computado a partir do MHI no tempo $t - 1$ junto com a imagem corrente do movimento $f(x, y, t)$. Isto implica que nenhuma imagem resultante de movimento anterior, ou campo de movimento, necessita ser armazenada ou manipulada. Porém existem limitações no uso deste método, uma vez que a silhueta representa o último movimento efetuado na imagem, logo ela é considerada como um único objeto dentro do MHI, sendo assim, em alguns momentos podem ocorrer dificuldades para a identificação de movimentos em objetos que sofrem oclusão.

2.2.3.2.4 Método da média de Bradski

Bradski, et al 2008, desenvolveram uma técnica usada para segmentar sequência de imagens (nas quais, o atributo principal é o movimento de corpos) baseada nos valores médios, $\bar{M}(x, y, N)$ – equação 2.2.31 –, associados com cada *pixel* (x, y) , registrados em N quadros consecutivos de imagem de uma mesma cena, bem como nos valores médios associados com esses mesmos *pixels* das diferenças acumuladas, $\bar{\Delta}(x, y, N)$ – equação 2.2.32 – entre esses mesmos N quadros.

$$\bar{M}(x, y, N) = \frac{\sum_{j=1}^N f(x, y, t_j)}{N} \quad (2.2.31)$$

$$\bar{\Delta}(x, y, N) = \frac{\sum_{j=2}^N f(x, y, t_j) - f(x, y, t_{j-1})}{N} \quad (2.2.32)$$

Tal técnica considera que os valores associados com cada *pixel* (x, y) desses N quadros consecutivos de imagens que pertençam ao intervalo $[LI_{x,y}, LS_{x,y}]$, valores de limiar inferior e superior, respectivamente, são considerados como pertencentes a elementos fixos (e,

portanto, pertencentes ao fundo da imagem). Os valores $LI_{x,y}$ e $LS_{x,y}$ são determinados pelo uso da expressão 2.2.33.

$$\begin{aligned} LI_{x,y} &= \bar{M}(x, y, N) - \bar{\Delta}(x, y, N) * k \\ LS_{x,y} &= \bar{M}(x, y, N) + \bar{\Delta}(x, y, N) * k \end{aligned} \quad (2.2.33)$$

onde, k representa um fator de escala.

Esses dados permitem detectar *pixels* pertencentes aos objetos da imagem que estão se movendo entre os N quadros de imagens considerados e podem ser segmentados por limiarização, usando os valores de limiar calculados na equação 2.2.33.

2.2.3.2.5 Técnica de segmentação em tempo real utilizando tabelas de codificação

(*CodeBook*)

Kim, et al, (2005), desenvolveram um algoritmo para segmentação de imagens coloridas em tempo real que permite capturar a variação do fundo e lidar com cenas que contenham objetos em movimento ou variações de iluminação. Esse algoritmo quantiza amostras de cada *pixel* de uma imagem em tabelas de codificação (*codebooks*), as quais representam de maneira compacta o modelo de fundo de uma sequência de imagens em um determinado período de tempo utilizando uma quantidade de memória limitada.

Para isto, é utilizada uma sequência \mathbf{x} de treinamento para um único *pixel* formado por um conjunto de N vetores RGB: $\mathbf{x} = \{x_1, x_2, \dots, x_N\}$. Cada *pixel* possui uma tabela de codificação $\mathcal{C} = \{c_1, c_2, \dots, c_L\}$ formada por L palavras de código (*codewords*). Cada palavra de código c_i , $i = 1..L$, é composta por um vetor RGB $v_i = (\bar{R}_i, \bar{G}_i, \bar{B}_i)$ e uma tupla $aux_i = \langle \hat{I}min_i, \hat{I}max_i, f_i, \lambda_i, p_i, q_i \rangle$ que contém a intensidade de brilho e variáveis temporais conforme descrito a seguir:

- $\hat{I}min_i, \hat{I}max_i$: representam, respectivamente, o menor e o maior valor de brilho associado entre todos os *pixels* pertencentes as palavras de código desse *pixel*;
- f_i : representa a frequência com que uma palavra de código ocorre;
- λ_i : representa o intervalo de tempo mais longo (durante o período de treinamento) que uma palavra de código não foi recuperada.

- p_i, q_i : Representam, respectivamente, o primeiro e o último valor de tempo de acesso que ocorreu na palavra de código;

2.2.3.2.5.1 Treinamento do modelo de referência

No período de treinamento, cada valor x_t amostrado no tempo t é comparado com os valores da tabela de codificação para determinar se existe uma palavra de código c_m que corresponda com o valor amostrado. Para determinar se a palavra de código existe, é empregado um valor médio de distorção nos limites de cor e brilho na palavra de código de índice m . O algoritmo apresentado na Tabela 1 mostra mais detalhes das operações realizadas para se construir uma tabela de codificação.

Tabela 1 – Algoritmo de construção da tabela de codificação

Algoritmo de construção da tabela de codificação	
I.	$L \leftarrow 0^1, \mathfrak{C} \leftarrow \emptyset$ (cria conjuntos vazios)
II.	<i>for</i> $t = 1$ <i>to</i> N <i>do</i>
i.	$X_t = (R, G, B), I \leftarrow \sqrt{R^2 + G^2 + B^2}$
ii.	Procure a palavra de código c_m em $\mathfrak{C} = \{c_i 1 \leq i \leq L\}$ que combine com X_t com base em duas condições (a) e (b). <ul style="list-style-type: none"> a) $colordist(X_t, V_m) \leq \varepsilon_1$ b) $Brightness((I, \hat{I}min_m, \hat{I}max_m)) = true$
iii.	Se $\mathfrak{C} = \emptyset$, ou se não encontrar a palavra de código, então $L \leftarrow L + 1$. Criar uma nova palavra de código c_L , com a seguinte atribuição <ul style="list-style-type: none"> • $V_L \leftarrow (R, G, B)$ • $aux_L \leftarrow \langle I, I, 1, t - 1, t, t \rangle$
iv.	Senão, atualize a palavra de código c_m , consistindo de $V_m = (\bar{R}_m, \bar{G}_m, \bar{B}_m)$ e $aux_m = \langle \hat{I}min_m, \hat{I}max_m, f_m, \lambda_m, p_m, q_m \rangle$ <ul style="list-style-type: none"> • $V_m \leftarrow \left(\frac{f_m \bar{R}_m + R}{f_{m+1}}, \frac{f_m \bar{G}_m}{f_{m+1}}, \frac{f_m \bar{B}_m}{f_{m+1}} \right)$ • $aux_m \leftarrow \langle \min\{I, \hat{I}min_m\}, \max\{I, \hat{I}max_m\}, f_m + 1, \max\{\lambda_m, t - q_m\}, p_m, t \rangle$ <i>end for</i>

- III. Para cada palavra de código $c_i, i = 1, \dots, L$, alterar o valor de λ_i para $\lambda_i \leftarrow \max\{\lambda_i, (N - q_i + p_i - 1)\}$.

No algoritmo apresentado na Tabela 1, as duas condições (a) e (b) constantes no Passo II – (ii) são satisfeitas quando as cores de X_t e c_m forem similares e o brilho do X_t estiver entre os limites aceitáveis de brilho do c_m .

A condição (a) verifica a distorção (δ) dos valores associados com um *pixel* da entrada $X_t = (R, G, B)$ com relação a uma palavra de código c_i em que $V_i = (\bar{R}_i, \bar{G}_i, \bar{B}_i)$. Esta condição compara o resultado da função *colordist* com um valor de limiar ε_1 . A função *colordist*, equação (2.2.34), pode ser considerada como valor de peso para o atributo brilho em um espaço normalizado da cor.

$$\text{colordist}(X_t, V_i) = \delta = \sqrt{(\|X_t\|^2 - p^2)} \quad (2.2.34)$$

Os itens presentes na equação 2.2.34 são definidos como:

- $\|X_t\|^2 = R^2 + G^2 + B^2$;
- $\|V_i\|^2 = \bar{R}_i + \bar{G}_i + \bar{B}_i$;
- $\langle X_t, V_i \rangle^2 = (\bar{R}_i R_i + \bar{G}_i G_i + \bar{B}_i B_i)$;
- $p^2 = \|X_t\|^2 \cos^2 \theta = \frac{\langle X_t, V_i \rangle^2}{\|V_i\|^2}$

A condição (b) verifica na tupla da palavra de código c_m , se o valor de brilho I de X_t está entre o maior e menor brilho de c_m . Para isto, define-se a faixa de valores $[I_{low}, I_{hi}]$ entre o menor (I_{low}) e o maior (I_{hi}) valor do brilho de cada palavra de código como:

$$I_{low} = \alpha \hat{I}max, \quad I_{hi} = \min\{\beta \hat{I}max, \frac{\check{I}min}{\alpha}\} \quad (2.2.35)$$

em que $\alpha < 1$ e $\beta > 1$. Também se define a função *brightness* como:

$$\text{brightness}(I, \langle \check{I}min_m, \hat{I}max_m \rangle) = \begin{cases} true, & \text{se } I_{low} \leq \|X_t\| \leq I_{hi} \\ false, & \text{caso contrário} \end{cases} \quad (2.2.36)$$

Com a execução do algoritmo de treinamento do fundo, é gerada a tabela de codificação de cada *pixel*, tabela esta formada por palavras de código que representam o histórico do valor associado com cada *pixel* da imagem em um determinado período de treinamento. Porém, muitos desses valores pertencentes a uma tabela de código de cada *pixel*, representam entradas inválidas, como ruídos ou objetos em movimento. Tais entradas são eliminadas, por meio da criação de um modelo de referência.

2.2.3.2.5.2 Modelo de referência

O modelo de referência de uma cena imageada (equação 2.2.37), é gerado com base nos dados filtrados da tabela de codificação para todos os *pixels*. Essa filtragem remove da tabela de codificação, todas as palavras de código que representam ruídos ou objetos em movimento, mantendo apenas as entradas que representam o fundo da imagem.

$$m = \{c_m | c_m \in \mathcal{C} \wedge \lambda_m \leq Tm\} \quad (2.2.37)$$

O termo Tm da equação 2.2.37 representa um valor de limiar utilizado na remoção das palavras de códigos que, supostamente, estão associados com ruídos ou objetos em movimento. Segundo Kim, et al, (2005), o valor mais adequado desse limiar é a quantidade de quadros utilizados no processo de treinamento do modelo de referência dividido por 2. Em ambientes externos é aconselhável utilizar um período de treinamento do modelo de referência não inferior a 5 minutos.

2.2.3.2.5.3 Detecção do fundo

Com o modelo de referência em mãos, é possível subtrair de uma imagem atual, os *pixels* referentes aos objetos estáticos (fundo da imagem). O algoritmo apresentado na Tabela 2 verifica se um *pixel* qualquer da imagem pertence ao fundo ou a um objeto em movimento.

Tabela 2 – Algoritmo para subtração do fundo (BGS)

Algoritmo para subtração do fundo	
I.	$X = (R, G, B), I \leftarrow \sqrt{R^2 + G^2 + B^2}$
II.	<p>Para todas as palavras de código contidas em m (equação 2.2.37), encontre a palavra de código c_m que corresponda com o valor X, baseando-se nas seguintes condições:</p> <ul style="list-style-type: none"> • $colordist(X_t, V_m) \leq \varepsilon_2$ (valor de limiar) • $Brightness(I, \langle \hat{I}min_m, \hat{I}max_m \rangle) = true$ <p>Atualize a palavra de código encontrada, conforme descrito no algoritmo de construção da tabela de codificação, passo II item iv (Tabela 1).</p>
III.	$BGS(x) = \begin{cases} \text{objeto: caso não encontre palavras de código correspondentes} \\ \text{Fundo: caso encontre uma palavras de código correspondentes} \end{cases}$

2.3 Fotogrametria - Conceitos

Fotogrametria é a ciência e tecnologia de obter informações confiáveis por meio de processos de registro, interpretação e mensuração de imagens (ANDRADE, 1998). Tendo isso em vista e pela necessidade de se efetuar medições e de interpretar dados em imagens digitais, são apresentados alguns conceitos básicos sobre Fotogrametria que fundamentam parte deste trabalho.

2.3.1 Equações de colinearidade

As equações de colinearidade permitem relacionar o espaço imagem com o espaço objeto, transformando as coordenadas do espaço objeto para o espaço imagem (e vice-versa). Segundo Hasegawa (1997), “as equações de colinearidade reproduzem matematicamente o

processo da formação da imagem, fazendo a ligação entre as coordenadas dos pontos no espaço objeto (3D) e suas correspondentes coordenadas no espaço imagem (2D)”.

O princípio básico para estabelecer a colinearidade, é baseado na condição de que os pontos C (centro perspectivo), p' (ponto imagem) e P (ponto objeto) pertençam a uma mesma reta (LUGNANI, 1987).

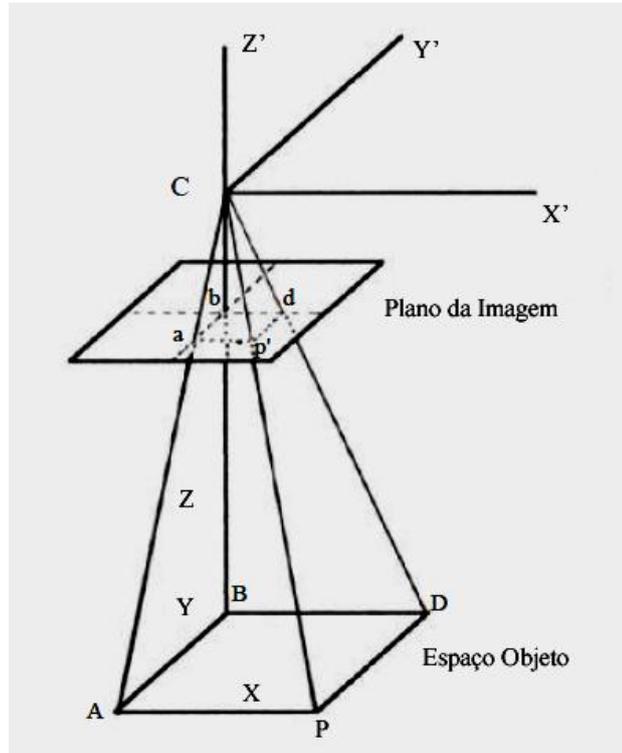


Figura 7 - Sistema de coordenadas, imagem e objeto paralelos, com origem no centro perspectivo
Fonte: Hasegawa, 1997

A geometria mostrada na Figura 7 apresenta os seguintes triângulos semelhantes $\Delta CDB \sim \Delta Cdb$ e $\Delta CBA \sim \Delta Cba$ e, assim sendo, permite estabelecer as seguintes relações geométricas entre as medidas:

$$\frac{x_p}{X} = \frac{z_p}{Z} \quad \text{e} \quad \frac{y_p}{Y} = \frac{z_p}{Z} \quad (2.3.1)$$

As quais permitem determinar o sistema das equações projetivas:

$$X_p = Z_p \frac{X'}{Z'} \quad \text{e} \quad Y_p = Z_p \frac{Y'}{Z'} \quad (2.3.2)$$

onde:

x_p, y_p e z_p são as coordenadas do ponto p no espaço imagem;

X, Y e Z são as coordenadas do ponto P no espaço objeto.

No caso do sistema de coordenadas do espaço objeto estar transladado, rotacionado e com uma escala diferente do plano imagem, deve-se aplicar a transformação de similaridade (LUGNANI, 1987) nas coordenadas do sistema referencial do espaço objeto (ponto P) para o sistema referencial do espaço imagem (ponto p').

Segundo Tommaselli, et al (1991), a transformação pode ser realizada em três passos: a) aplicando três translações visando compensar a diferença espacial entre as origens; b) aplicando uma rotação para compensar as diferenças angulares e; c) e executar a correção da diferença de escala entre os referenciais.

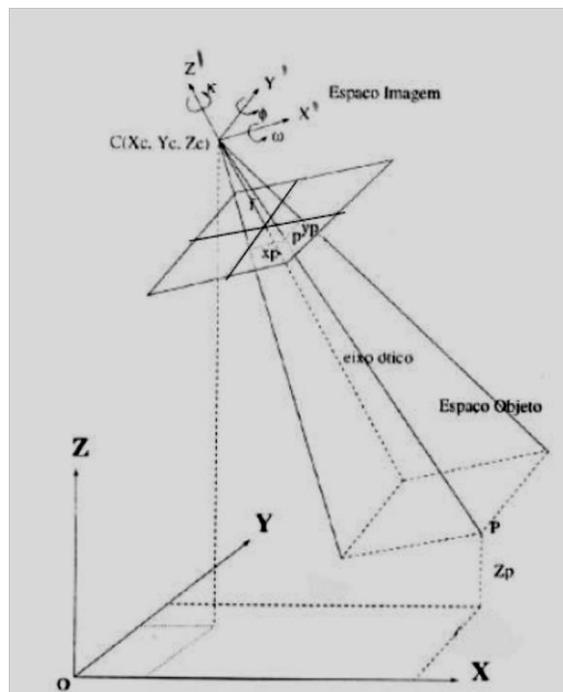


Figura 8 - Sistemas de coordenadas da imagem e do objeto transladado e rotacionado.

Fonte: Hasegawa, 1997

Os movimentos de rotação ocorrem no sentido anti-horário, considerando-se que o sistema referencial da imagem gira, enquanto o objeto permanece fixo, cujas matrizes de rotação M_{κ} , M_{σ} e M_{ω} são dadas por:

$$M_{\kappa} = \begin{bmatrix} \cos \kappa & \sin \kappa & 0 \\ -\sin \kappa & \cos \kappa & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad M_{\sigma} = \begin{bmatrix} \cos \varphi & 0 & -\sin \varphi \\ 0 & 1 & 0 \\ \sin \varphi & 0 & \cos \varphi \end{bmatrix} \quad M_{\omega} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \omega & \sin \omega \\ 0 & -\sin \omega & \cos \omega \end{bmatrix} \quad (2.3.3)$$

As matrizes M_{κ} , M_{σ} e M_{ω} são as matrizes de rotação em torno dos eixos x , y e z

respectivamente.

A matriz \mathbf{M} ($\mathbf{M} = (\mathbf{M}_\kappa \cdot \mathbf{M}_\phi \cdot \mathbf{M}_\omega)$) é dada por:

$$\mathbf{M} = \begin{bmatrix} \cos \phi \cos \kappa & \cos \omega \sin \kappa - \sin \omega \sin \phi \cos \kappa & \sin \omega \sin \kappa - \cos \omega \sin \phi \cos \kappa \\ -\cos \phi \sin \kappa & \cos \omega \cos \kappa - \sin \omega \sin \phi \sin \kappa & \sin \omega \cos \kappa - \cos \omega \sin \phi \sin \kappa \\ \sin \phi & -\sin \omega \cos \phi & \cos \omega \cos \phi \end{bmatrix} \quad (2.3.4)$$

e a transformação dar-se-á por meio de:

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = \lambda \mathbf{M} \begin{bmatrix} X - X_c \\ Y - Y_c \\ Z - Z_c \end{bmatrix} \quad (2.3.5)$$

onde:

λ é um fator de escala;

X_c , Y_c e Z_c são as coordenadas do centro perspectivo do sistema;

X' , Y' e Z' são as coordenadas do ponto no sistema XYZ ;

\mathbf{M} a matriz de rotação.

O sistema de equações descrito na equação 2.3.5 pode ser escrito como:

$$\begin{aligned} X' &= \lambda[m_{11}(X - X_c) + m_{12}(Y - Y_c) + m_{13}(Z - Z_c)] \\ Y' &= \lambda[m_{21}(X - X_c) + m_{22}(Y - Y_c) + m_{23}(Z - Z_c)] \\ Z' &= \lambda[m_{31}(X - X_c) + m_{32}(Y - Y_c) + m_{33}(Z - Z_c)] \end{aligned} \quad (2.3.6)$$

Substituindo-se as equações (2.3.6) nas equações projetivas (2.3.2), λ é eliminada do sistema de equações, obtendo-se as equações de colinearidade (equação 2.3.7).

$$\begin{aligned} x_p &= Z \cdot \frac{m_{11}(X - X_c) + m_{12}(Y - Y_c) + m_{13}(Z - Z_c)}{m_{31}(X - X_c) + m_{32}(Y - Y_c) + m_{33}(Z - Z_c)} \\ y_p &= Z \cdot \frac{m_{21}(X - X_c) + m_{22}(Y - Y_c) + m_{23}(Z - Z_c)}{m_{31}(X - X_c) + m_{32}(Y - Y_c) + m_{33}(Z - Z_c)} \end{aligned} \quad (2.3.7)$$

O sistema de equações (2.3.7) é usado na concepção do negativo da imagem, quando aplicadas no diapositivo, devido ao fato de $z = -f$, a mesma se torna:

$$\begin{aligned}
x_p &= -f \cdot \frac{m_{11}(X - Xc) + m_{12}(Y - Yc) + m_{13}(Z - Zc)}{m_{31}(X - Xc) + m_{32}(Y - Yc) + m_{33}(Z - Zc)} \\
y_p &= -f \cdot \frac{m_{21}(X - Xc) + m_{22}(Y - Yc) + m_{23}(Z - Zc)}{m_{31}(X - Xc) + m_{32}(Y - Yc) + m_{33}(Z - Zc)}
\end{aligned} \tag{2.3.8}$$

A forma inversa das equações de colinearidade pode ser obtida aplicando-se a transformada inversa no sistema de equações (2.3.5), onde se obtém:

$$\begin{bmatrix} X - Xc \\ Y - Yc \\ Z - Zc \end{bmatrix} = \lambda^{-1} \mathbf{M}^T \begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} \tag{2.3.9}$$

e:

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = \lambda_p \begin{bmatrix} x_p \\ y_p \\ z_p \end{bmatrix} \tag{2.3.10}$$

Como o sistema $X'Y'Z'$ e o sistema XYZ têm a mesma origem e a mesma orientação pode-se substituir a equação. 2.3.10 na equação. 2.3.9, resultando em:

$$\begin{bmatrix} X - Xc \\ Y - Yc \\ Z - Zc \end{bmatrix} = \lambda^{-1} \mathbf{M}^T \lambda_p \begin{bmatrix} x_p \\ y_p \\ z_p \end{bmatrix} \tag{2.3.11}$$

A qual, substituídos pelos termos correspondentes da equação 2.3.6 produz:

$$\begin{aligned}
X - Xc &= \lambda^{-1} \lambda_p (m_{11}x_p + m_{21}y_p + m_{31}z_p) \\
Y - Yc &= \lambda^{-1} \lambda_p (m_{12}x_p + m_{22}y_p + m_{32}z_p) \\
Z - Zc &= \lambda^{-1} \lambda_p (m_{13}x_p + m_{23}y_p + m_{33}z_p)
\end{aligned} \tag{2.3.12}$$

Isolando os termos X e Y das equações (2.3.11) e (2.3.12), tem-se forma inversa das equações de colinearidade (equação 2.3.13).

$$\begin{aligned}
X &= Xc + (Z - Zc) \frac{m_{11}x_p + m_{21}y_p + m_{31}z_p}{m_{13}x_p + m_{23}y_p + m_{33}z_p} \\
Y &= Yc + (Z - Zc) \frac{m_{12}x_p + m_{22}y_p + m_{32}z_p}{m_{13}x_p + m_{23}y_p + m_{33}z_p}
\end{aligned} \tag{2.3.13}$$

2.3.2 Modelos matemáticos empíricos

Além das equações de colinearidade, é possível relacionar o espaço imagem e o espaço objeto por meio de modelos matemáticos empíricos, tais como: os modelos polinomiais; o modelo das funções racionais; o modelo de Transformação afin e; o modelo de transformação linear direta. Dependendo do sistema a ser desenvolvido, esses modelos tornam-se uma alternativa mais viável para se relacionar o espaço imagem e o espaço objeto ao invés das equações de colinearidade, pois para utilizá-los não é necessário conhecer dados da calibração e da orientação interna do sistema de aquisição de imagem (TOUTIN, 2004).

2.3.2.1 Modelos polinomiais 3D

A forma resumida de um polinômio simples (não-racional 3D) é dada pela seguinte notação (Equações 2.3.14 e 2.3.15):

$$l = \sum_{i=0}^{m1} \sum_{j=0}^{m2} \sum_{k=0}^{m3} a_{ijk} x^i y^j z^k \quad (2.3.14)$$

$$c = \sum_{i=0}^{n1} \sum_{j=0}^{n2} \sum_{k=0}^{n3} b_{ijk} x^i y^j z^k \quad (2.3.15)$$

em que, l e c representam as coordenadas coluna e linha de um determinado *pixel* na imagem, x é a longitude do ponto no terreno, y é a latitude do ponto no terreno, z é a altitude do ponto no terreno e, a_{ijk} , e b_{ijk} são os coeficientes polinomiais.

As equações 2.3.16 e 2.3.17 demonstram com maiores detalhes um exemplo de expansão das notações referentes às equações 2.3.14 e 2.3.15, no caso um polinômio de terceira ordem no espaço 3D.

$$\begin{aligned} l = & a_1 + a_2 y + a_3 x + a_4 z + a_5 yx + a_6 yz + a_7 xz + a_8 y^2 + a_9 x^2 + a_{10} z^2 \\ & + a_{11} xyz + a_{12} y^3 + a_{13} yx^2 + a_{14} yz^2 + a_{15} y^2 x + \dots + a_{16} x^3 \\ & + a_{17} xz^2 + a_{18} y^2 z + a_{19} x^2 z + a_{20} z^3 \end{aligned} \quad (2.3.16)$$

$$\begin{aligned} c = & b_1 + b_2 y + b_3 x + b_4 z + b_5 yx + b_6 yz + b_7 xz + b_8 y^2 + b_9 x^2 + b_{10} z^2 \\ & + b_{11} xyz + b_{12} y^3 + b_{13} yx^2 + b_{14} yz^2 + b_{15} y^2 x + \dots + b_{16} x^3 \\ & + b_{17} xz^2 + b_{18} y^2 z + b_{19} x^2 z + b_{20} z^3 \end{aligned} \quad (2.3.17)$$

Segundo Petrie, et al, 1992, existem padrões de distorções ou de deslocamentos que podem ser modelados ou corrigidos por termos específicos do polinômio (Figura 9). Porém, a principal vantagem do uso do modelo polinomial é a correção de todas as fontes de distorção simultaneamente (NOVAK, 1992).

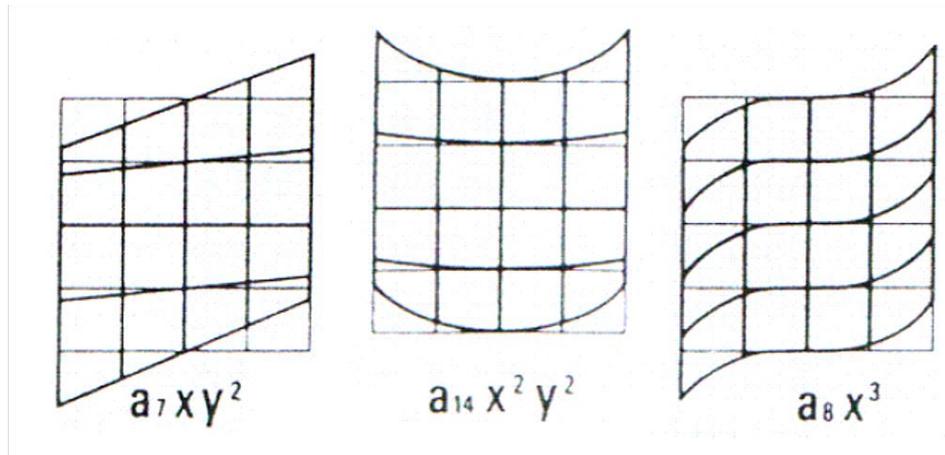


Figura 9 - Padrões de distorção modelados por termos específicos de um polinômio.
Fonte – Petrie et al (1992).

2.3.2.2 Modelo das funções racionais

O modelo das funções racionais (RFM – *Rational Function Model*) relaciona as coordenadas do espaço objeto (X, Y, Z) com as coordenadas no sistema imagem (c, l) pela razão entre duas funções polinomiais (GRODECKI, 2001; TOUTIN, 2004).

$$l_n = \frac{p1(X_n, Y_n, Z_n)}{p2(X_n, Y_n, Z_n)} = \frac{\sum_{i=0}^{m1} \sum_{j=0}^{m2} \sum_{k=0}^{m3} a_{ijk} X_n^i Y_n^j Z_n^k}{\sum_{i=0}^{m1} \sum_{j=0}^{m2} \sum_{k=0}^{m3} b_{ijk} X_n^i Y_n^j Z_n^k} \quad (2.3.18)$$

$$c_n = \frac{p3(X_n, Y_n, Z_n)}{p4(X_n, Y_n, Z_n)} = \frac{\sum_{i=0}^{m1} \sum_{j=0}^{m2} \sum_{k=0}^{m3} c_{ijk} X_n^i Y_n^j Z_n^k}{\sum_{i=0}^{m1} \sum_{j=0}^{m2} \sum_{k=0}^{m3} d_{ijk} X_n^i Y_n^j Z_n^k} \quad (2.3.19)$$

Segundo Croitoru, et al (2004), os coeficientes polinomiais a_{ijk} , b_{ijk} , c_{ijk} , d_{ijk} denominados de Coeficientes da Função Racional (RFC – *Rational Function Coefficient*), geralmente são polinômios de terceira ordem, nos quais, as razões dos termos de primeira ordem representam distorções causadas pela projeção óptica, as dos termos de 2ª ordem representam as correções sobre a curvatura da Terra, refração atmosférica e distorções das lentes e as de terceira ordem modelam outras distorções mais complexas e desconhecidas.

2.3.2.3 Modelo de Transformação afim

O modelo de transformação afim (APM – *Affine Projection Model*), também conhecido como Equação de Projeção Paralela, tem por objetivo projetar o espaço 3D no plano 2D. Esta projeção paralela é caso particular de projeção central, em que o centro de projeção foi deslocado para o infinito (LUGNANI,1987).

Segundo Okamoto (1999), o modelo de projeção APM pode ser derivado da equação de colinearidade convencional:

$$\begin{bmatrix} x \\ y \\ -c \end{bmatrix} = \lambda (\mathbf{R}_{\phi_j} \mathbf{R}_{\omega_j} \mathbf{R}_{\kappa_j})^T \begin{bmatrix} X - X_{0i} \\ Y - Y_{0i} \\ Z - Z_{0i} \end{bmatrix} \quad (2.3.20)$$

em que, (X,Y,Z) são as coordenadas do ponto no terreno (plano objeto), λ é o parâmetro de escala, c é a distância principal, $\mathbf{R}_{\phi_j} \mathbf{R}_{\omega_j} \mathbf{R}_{\kappa_j}$ são as matrizes de rotação e (x,y) são as coordenadas do ponto imagem. Considerando a imagem como uma projeção paralela, a distancia c pode ser considerada no infinito e a equação (2.3.20) pode ser escrita como:

$$x = a_{11}(X - X_{0i}) + a_{12}(Y - Y_{0i}) + a_{13}(Z - Z_{0i}) \quad (2.3.21)$$

$$y = a_{21}(X - X_{0i}) + a_{22}(Y - Y_{0i}) + a_{23}(Z - Z_{0i}) \quad (2.3.22)$$

em que, a_{ij} ($i = 1,2,3; j = 1,2,3$) são elementos da matriz $\lambda(\mathbf{R}_{\phi_j} \mathbf{R}_{\omega_j} \mathbf{R}_{\kappa_j})^T$.

Considerando-se também que o sensor responsável por capturar as imagens se move linearmente no espaço e que os parâmetros de orientação do sensor permaneçam constantes, o centro de projeção de cada linha pode ser descrito por:

$$X_{0i} = X_0 + \Delta X i \quad (2.3.23)$$

Com X_0 e ΔX valores constantes e i o número da linha. Expressões semelhantes são igualmente definidas para Y_{0i} e Z_{0i} . A linha de número i é expressa pela equação (2.3.24), obtida substituindo a equação (2.3.23) na (2.3.21):

$$i = \frac{a_{11}(X - X_0) + a_{12}(Y - Y_0) + a_{13}(Z - Z_0)}{a_{11}\Delta X + a_{12}\Delta Y + a_{13}\Delta Z} \quad (2.3.24)$$

Substituindo a linha de número i pela coordenada de imagem x , e considerando os parâmetros de orientação constantes, a equação (2.3.24) pode ser expressa por:

$$x = \frac{a_{11}(X - X_0) + a_{12}(Y - Y_0) + a_{13}(Z - Z_0)}{a_{11}\Delta X + a_{12}\Delta Y + a_{13}\Delta Z} \quad (2.3.25)$$

A equação (2.3.21) organizada para os coeficientes constantes é descrita pela seguinte expressão algébrica:

$$x = A_1X + A_2Y + A_3Z + A_4 \quad (2.3.26)$$

A equação (2.3.22) também é expressa através de arranjo semelhante por:

$$y = A_5X + A_6Y + A_7Z + A_8 \quad (2.3.27)$$

em que, (x, y) correspondem as coordenadas no plano imagem; (X, Y, Z) correspondem as coordenadas tridimensionais no plano objeto; e (A_1, A_2, \dots, A_8) correspondem aos parâmetros de transformação. Com isto, as equações (2.3.26) e (2.3.27) são conhecidas como Equação APM.

2.3.2.4 Transformação Linear Direta (DLT)

A transformação Linear Direta (DLT – *Direct Linear Transformation*) foi desenvolvida por Abdel-Aziz e Karara (1971) para a calibração de câmaras não-métricas. Esta equação modela a transformação entre o sistema de coordenadas do espaço imagem para o sistema de coordenadas do espaço objeto, sem a necessidade do conhecimento de parâmetros de orientação interior e exterior do sensor. A DLT é dada pelas seguintes equações:

$$x = \frac{L_1 \cdot X_p + L_2 \cdot Y_p + L_3 \cdot Z_p + L_4}{L_9 \cdot X_p + L_{10} \cdot Y_p + L_{11} \cdot Z_p + 1} \quad (2.3.28)$$

$$y = \frac{L_5 \cdot X_p + L_6 \cdot Y_p + L_7 \cdot Z_p + L_8}{L_9 \cdot X_p + L_{10} \cdot Y_p + L_{11} \cdot Z_p + 1} \quad (2.3.29)$$

Nas equações (2.3.28 e 2.3.29), x e y representam as coordenadas do ponto na imagem; X_p, Y_p, Z_p representam as coordenadas do ponto no espaço objeto e L_1 a L_{11} representam os parâmetros de transformação DLT, sendo:

$$L = \frac{1}{m_{31} \cdot X_c + m_{32} \cdot Y_c + m_{33} Z_c} \quad (2.3.30)$$

$$L_1 = L \cdot (x_0 \cdot m_{31} - C_x \cdot m_{11}) \quad (2.3.31)$$

$$L_2 = L \cdot (x_0 \cdot m_{32} - C_x \cdot m_{12}) \quad (2.3.32)$$

$$L_3 = L \cdot (x_0 \cdot m_{33} - C_x \cdot m_{13}) \quad (2.3.33)$$

$$L_4 = x_0 + [L \cdot C_x (m_{11} \cdot X_c + m_{12} \cdot Y_c + m_{13} \cdot Z_c)] \quad (2.3.34)$$

$$L_5 = L \cdot (y_0 \cdot m_{31} - C_y \cdot m_{21}) \quad (2.3.35)$$

$$L_6 = L \cdot (y_0 \cdot m_{32} - C_y \cdot m_{22}) \quad (2.3.36)$$

$$L_7 = L \cdot (y_0 \cdot m_{33} - C_y \cdot m_{23}) \quad (2.3.37)$$

$$L_8 = y_0 + [L \cdot C_y (m_{21} \cdot X_c + m_{22} \cdot Y_c + m_{23} \cdot Z_c)] \quad (2.3.38)$$

$$L_9 = L \cdot m_{31} \quad (2.3.39)$$

$$L_{10} = L \cdot m_{32} \quad (2.3.40)$$

$$L_{11} = L \cdot m_{33} \quad (2.3.41)$$

em que x_0 e y_0 são coordenadas do ponto principal; X_c, Y_c e Z_c são as coordenadas do centro de perspectiva no espaço objeto e (m_{11}, \dots, m_{33}) são elementos da matriz de rotação. Segundo Mikhail et al (2001), a DLT pode ser derivada da equação de colinearidade, para isto, adota-se uma escala diferente para as direções x e y da distância principal (C_x, C_y).

TRABALHOS RELACIONADOS

Este capítulo apresenta três métodos para mensurar a velocidade de automóveis por meio de imagens de vídeo. O primeiro método, desenvolvido por Yan, et al (2010), é baseado no método de calibração de câmara de Tsai em duas fases (TSAI, 1987) e em um modelo matemático. O segundo método, desenvolvido por Pornpanomchai, et al (2009), tem como ponto chave, utilizar uma distância pré definida no mundo real, por meio de um ponto inicial e final de referência para realizar o cálculo da velocidade. E por último, o método desenvolvido por Koyama, (2001), este método utiliza as equações de colinearidade modificadas (GOSH, 1985) no processo de conversão das coordenadas do espaço imagem para o espaço objeto, para possibilitar a mensuração da velocidade real de um automóvel.

3.1 Mensuração da velocidade por imagens de vídeo utilizando o método de Tsai em duas fases

Yan, et al (2010), desenvolveram um sistema de monitoramento de tráfego de automóveis baseado no método de calibração de câmara de Tsai em duas fases (TSAI, 1987) e em um modelo matemático (Figura 10) que define a relação geométrica entre o espaço imagem e o espaço objeto.

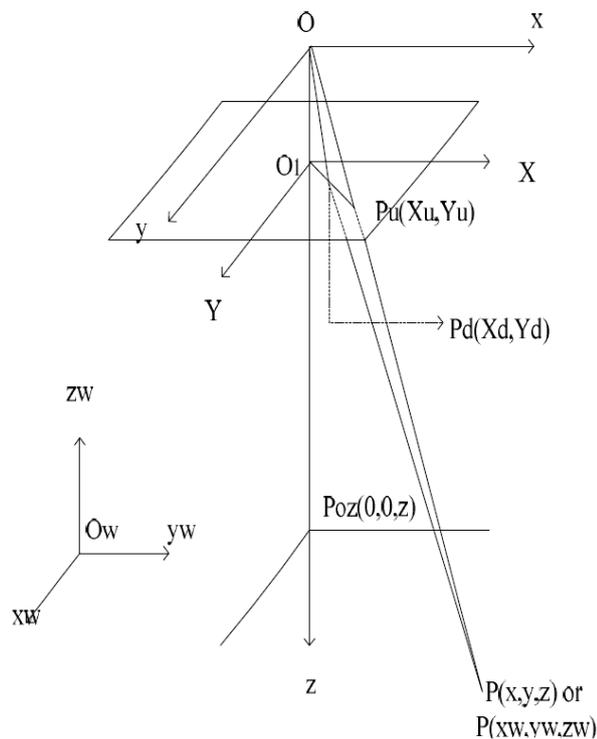


Figura 10 - Relação geométrica utilizado no processo de calibração
Fonte: Yan, et al. (2010)

Nesse modelo matemático as coordenadas (X, Y, Z) / (X_w, Y_w, Z_w) representam as coordenadas 3D do ponto P no mundo; (X_u, Y_u) representa as coordenadas geométricas do ponto P da imagem; (X_d, Y_d) representa as coordenadas da imagem do ponto P no dispositivo gráfico; (X_f, Y_f) representa as coordenadas da imagem do ponto P no plano imagem em *pixels*.

Inicialmente o sistema realiza a calibração da câmara utilizando a técnica de Tsai para remover as distorções causadas pela lente da câmara. Após esse passo é delimitada uma área da via, dividindo-a em sete regiões (Figura 11).



Figura 11 - Resultado da mensuração da velocidade média
Fonte: Yan, et al. (2010)

No momento em que o automóvel passa pelas linhas de controle, é realizada a captura da coordenada do *pixel* da extrema esquerda do automóvel para posteriormente ser realizada a conversão do ponto obtido no espaço imagem do computador (*pixel*) para coordenada do espaço objeto (mundo real). Essa conversão é realizada por meio do processo inverso da calibração da câmara, o qual é ilustrado na Figura (12).

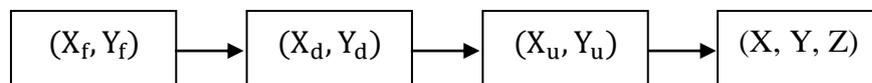


Figura 12 - Processo inverso da calibração

Após a conversão das coordenadas do ponto no plano imagem em *pixels* para coordenadas no plano 3D do mundo real é realizado o cálculo da velocidade levando em conta o tempo decorrido entre a captura do ponto atual com o ponto anterior referente ao mesmo ponto no veículo.

Este sistema tem como vantagem a precisão nos cálculos no processo de mensuração da velocidade dos automóveis. Esta precisão é ocasionada pela utilização do método de calibração da câmara, que visa minimizar os erros ocasionados pelas imperfeições da lente. Porém, a utilização de métodos voltados para calibrar as câmaras é complexa, e deve ser realizada toda vez que a mesma sofrer alterações.

3.2 Sistema de mensuração de velocidade de veículos

Pornpanomchai et al (2009), desenvolveram um sistema voltado para mensurar a velocidade de deslocamento de veículos. Esse sistema utiliza um ponto de referência inicial e final que define uma distância conhecida no mundo real (Figura 13).

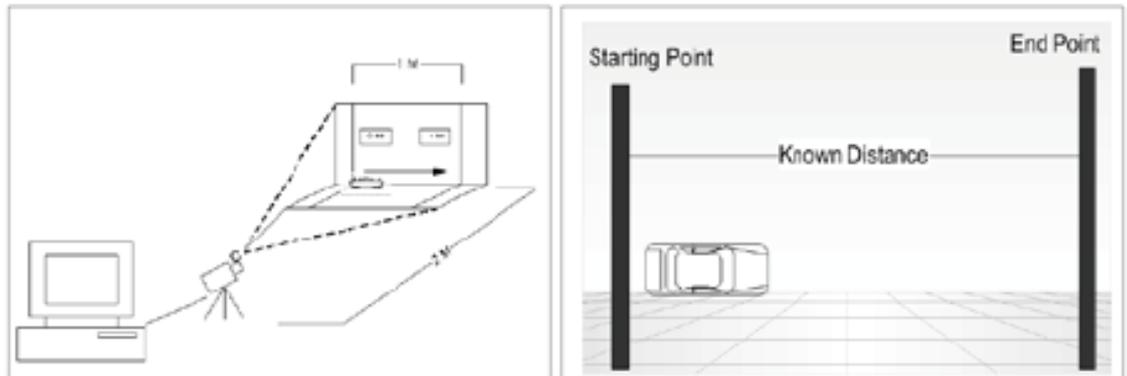


Figura 13 - Estrutura da cena para a mensuração da velocidade
Fonte: Pornpanomchai, et al (2009)

Além de determinar uma distância conhecida no mundo real, estes pontos chave servem para determinar o momento em que o sistema inicializará e finalizará o processo de detecção do automóvel. No momento em que o automóvel ultrapassa o ponto de referência inicial, são determinadas as coordenadas em *pixels* do ponto inicial P_0 e do ponto final P_n do automóvel na imagem.

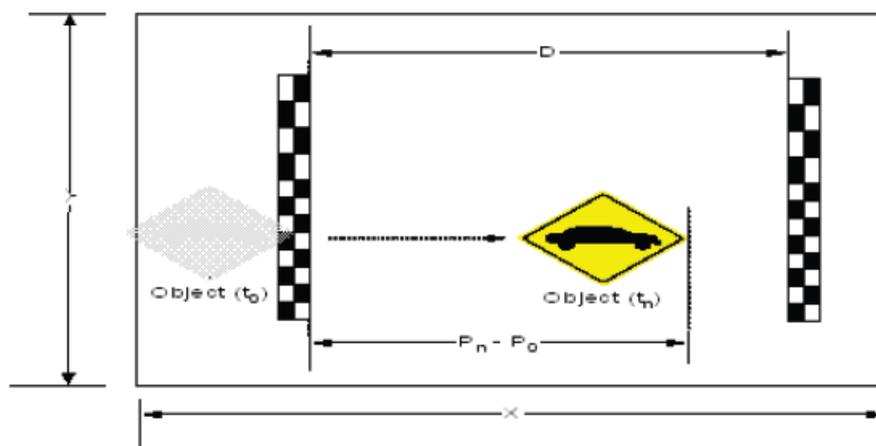


Figura 14 - Diagrama exibindo todas as variáveis utilizadas no processo de mensuração da velocidade.
Fonte: Pornpanomchai, et al (2009)

Com as coordenadas do ponto inicial e final e com o tempo gasto para percorrer tal distancia, é possível determinar a velocidade do automóvel, por meio da equação 3.2.1.

$$velocidade = distancia/tempo \quad (3.2.1)$$

Em que a distância é obtida pela equação 3.2.2

$$distância = Df * \left(\frac{D}{D_x}\right) * (P_n - P_0) \quad (3.2.2)$$

e o tempo pela equação 3.2.3

$$tempo = Tf * (t_n - t_0) \quad (3.2.3)$$

Em que D_x é a distância entre os pontos de referências medidos em *pixels*; D é a distância entre os pontos de referências no mundo; P_0 é o ponto da estrema direita do automóvel medido em *pixels* no instante $t=0$; P_n é o ponto da estrema direita do automóvel medido em *pixels* no instante $t=n$; t_0 representa o tempo inicial da tomada, é definido como o tempo de processamento da máquina mensurado em milissegundos; t_n representa o tempo inicial da tomada, é definido como o tempo de processamento da máquina mensurado em milissegundos; Df é o fator de conversão da distancia de metros para kilometros; Tf é o fator de conversão do tempo de milissegundos para hora.

Esta técnica não se mostra interessante devido ao fato do automóvel ser filmado de maneira lateral, pois em pistas de faixa dupla, o automóvel da outra pista pode parecer parte do objeto em questão na hora da segmentação e também pelo fato do tempo ser baseado no tempo de processamento do computador, caso ocorra um atraso de tempo para realizar alguma tarefa computacional, como, por exemplo, a segmentação dos quadros, o tempo mensurado se torna inconsistente.

3.3 Determinação da velocidade de um objeto, a partir de uma sequência de imagens, aplicando as equações de colineariedade modificadas.

Koyama (2001), desenvolveu um sistema voltado para mensurar a velocidade de um objeto utilizando as equações de colineariedade modificadas (GOSH, 1985). Para validar seu sistema, Koyama simulou uma rodovia (utilizando uma cartolina cinza) em que os automóveis (carrinho de metal) iriam trafegar em um único sentido Figura (15).

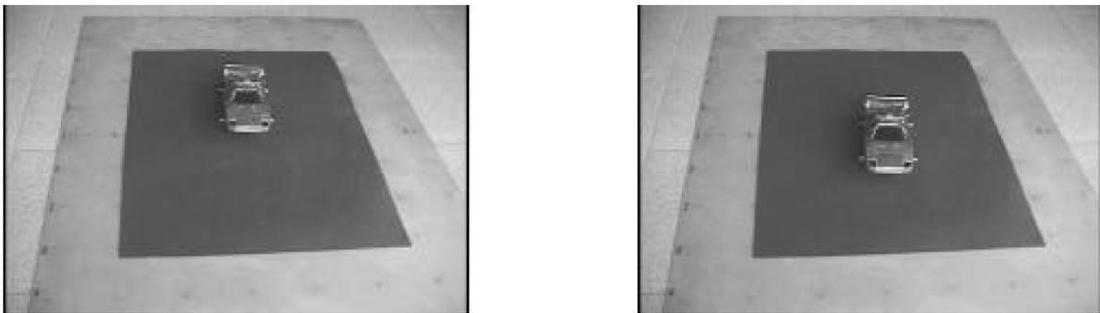


Figura 15 - Simulação de uma rodovia por meio de uma cartolina e um carrinho de metal
Fonte: Koyama, (2001).

Para se determinar os parâmetros de orientação interior da câmara filmadora, foi utilizado o software de Calibração de Câmara – CC (Galo,1998) e uma placa de aço com pontos de controle. A partir de 6 imagens selecionadas, efetuou-se a leitura das coordenadas dos pontos de controle (determinação semiautomática do centro de massa) no espaço imagem (Figura 16).

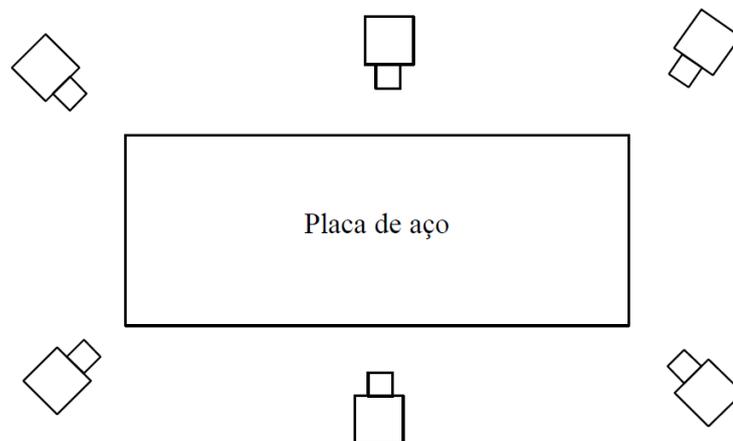


Figura 16 - Angulo das imagens para a calibração da câmara.
Fonte: Koyama, (2001).

Para a determinação dos parâmetros de orientação exterior, definiu-se um sistema de coordenadas na placa (local) em que foram utilizados alguns pontos existentes na placa, como apoio (Figura 17).

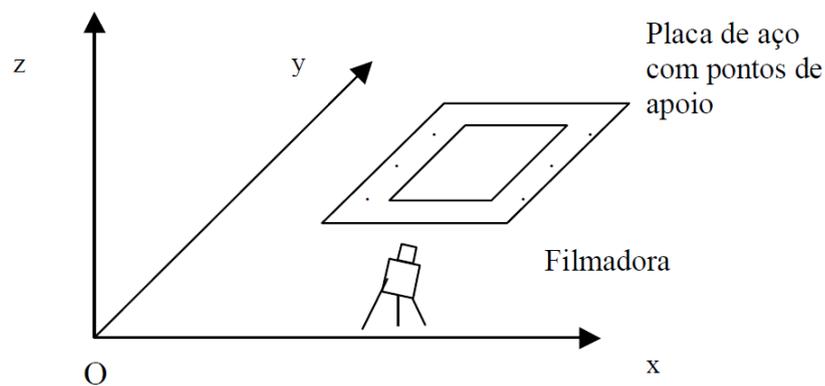


Figura 17 - Sistema de coordenadas locais para a orientação exterior.

Seu sistema utilizou dois quadros de uma filmagem (Posição inicial e final do objeto), adquiridos em instantes diferentes nos quais foram calculadas as coordenadas do objeto no espaço imagem. Conhecidas as coordenadas do objeto no espaço imagem foram aplicadas as equações de colinearidade modificadas (GOSH, 1985) para converter os pontos do espaço imagem para o espaço objeto, podendo assim computar a velocidade do objeto. Nestas equações além de serem considerados os parâmetros intrínsecos e extrínsecos da câmara, tais como, os movimentos de rotação realizados nos eixos (x, y, z) do sistema referencial da imagem, fator de escala da imagem, coordenadas do centro perspectivo e distância focal, são considerados também as componentes de velocidade dos eixos (x, y, z) e o intervalo de tempo.

Este trabalho mostra-se relevante, principalmente pelo fato de se utilizar a calibração de câmaras no processo para identificar os parâmetros interiores da mesma. Isso faz com que haja uma diminuição dos erros causados pela distorção das lentes do sistema ótico da câmara.

MENSURAÇÃO DA VELOCIDADE:

Experimentos e Análise dos

Resultados

Esta seção trata da descrição dos experimentos realizados no processo de validação das técnicas utilizadas para determinar as coordenadas do espaço objeto e mensurar a velocidade de um corpo rígido em movimento.

4.1 Sistema proposto

Foi desenvolvido um sistema (protótipo) computacional capaz de mensurar a velocidade de deslocamento de um corpo rígido em movimento. Esse sistema foi desenvolvido utilizando a linguagem de programação C++, juntamente com a biblioteca denominada OpenCV. O sistema é composto pelos seguintes módulos:

- **Interface:** Módulo responsável por realizar a interface entre o sistema computacional e a câmara, capturando as imagens adquiridas pela câmara e disponibilizando-as para processamento;

- **Segmentação:** Módulo responsável por realizar a detecção dos corpos em movimento em uma sequência de imagens adquirida. Esse módulo permite optar por utilizar o método de segmentação baseado na média de Bradski ou o método de segmentação das tabelas de codificação (*CodeBook*).
- **Redução de ruídos:** Módulo responsável pela diminuição dos ruídos presentes na imagem gerada no processo de segmentação, tarefa essa realizada por meio da operação morfológica de abertura.
- **Conversão de coordenadas:** Módulo responsável por realizar a conversão de coordenadas do plano imagem para o espaço objeto (e vice-versa). Tal conversão pode ser realizada por meio das equações de colinearidade, pelo modelo de transformação linear direta, pelo modelo de transformação polinomial, pelo modelo de transformação afim ou pelo modelo de transformação das funções racionais. Esses métodos necessitam como parâmetro de entrada do conhecimento *a priori*, no mínimo, do relacionamento de 12 pontos do espaço objeto com os respectivos pontos do plano imagem. Caso o método escolhido seja o das equações de colinearidade, é necessário, ainda, ter conhecimento dos parâmetros de calibração da câmara.
- **Mensuração da velocidade:** Módulo responsável por mensurar a velocidade média de deslocamento de um veículo e proceder com as análises relativas a esse valor;
- **Visualização dos dados:** Módulo responsável por exibir as sequências de imagens com sobreposição dos detalhes computados.

O fluxograma presente na Figura 18 representa com maiores detalhes o funcionamento do sistema desenvolvido.

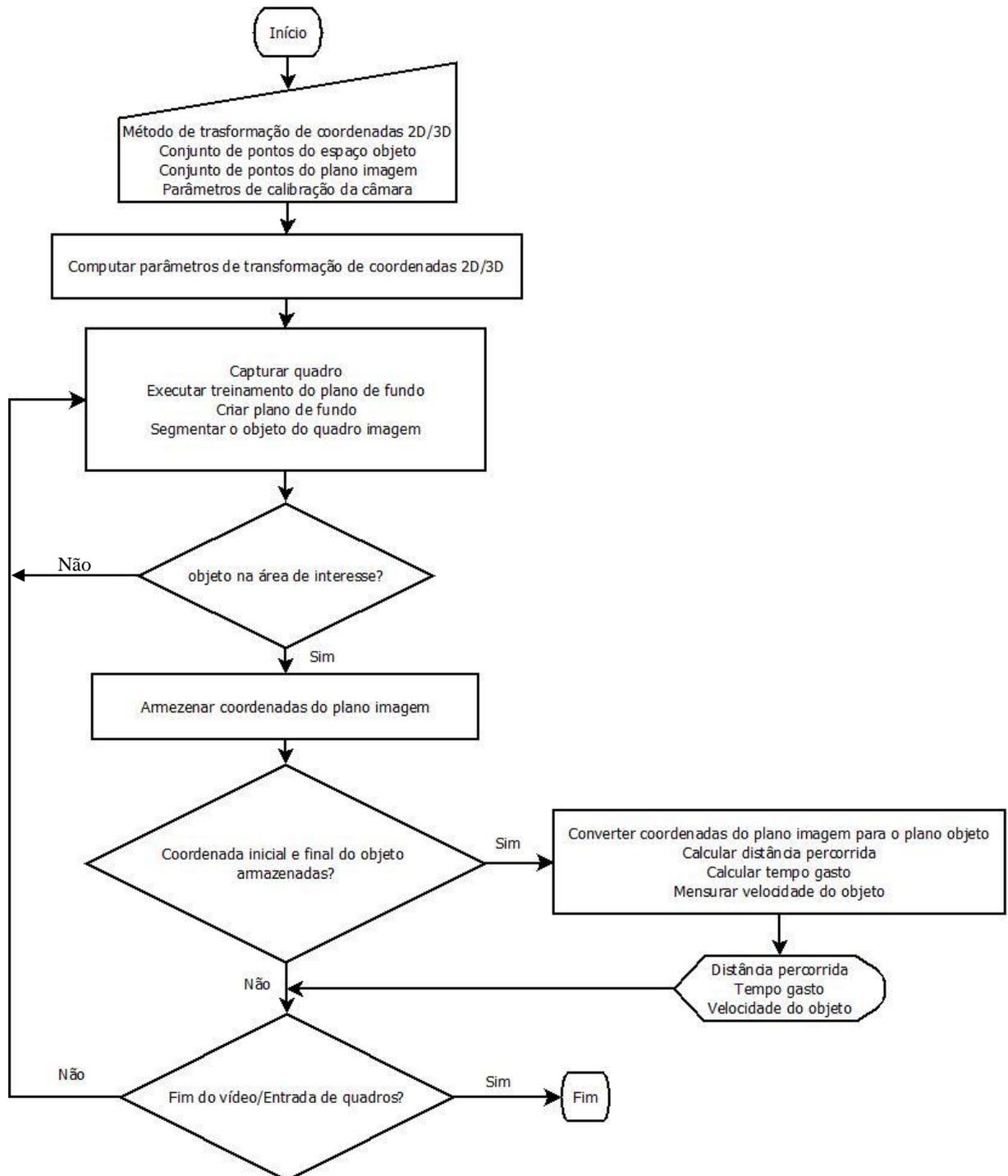


Figura 18 - Fluxograma do sistema desenvolvido

4.2 Experimento

Com o objetivo de validar o sistema desenvolvido, foi realizado um experimento no laboratório de física da Universidade Estadual de Maringá (UEM). Esse experimento contou com a utilização de um aparelho denominado trilho de ar para mensurar a velocidade de deslocamento de um corpo sobre uma superfície. O experimento foi completamente registrado em uma sequência de imagens, as quais foram submetidas para o módulo de processamento de imagens do sistema proposto.

Por meio do módulo de segmentação implementado nesse sistema, a técnica de detecção de movimento e segmentação de sequência imagens em tempo real, desenvolvido por Kim et al, (2005), dois quadros consecutivos de imagens, adquiridos nos instantes t_i e t_{i+1} , são, automaticamente, analisados para detectar se um corpo qualquer alterou a sua posição. Se houve alguma alteração na posição do corpo, é definido o quadrilátero mínimo que envolve esse corpo, cujas coordenadas são extraídas, no entanto, somente a coordenada do canto inferior esquerdo desse quadrilátero é utilizada para o cálculo da velocidade de deslocamento (Figura 19).

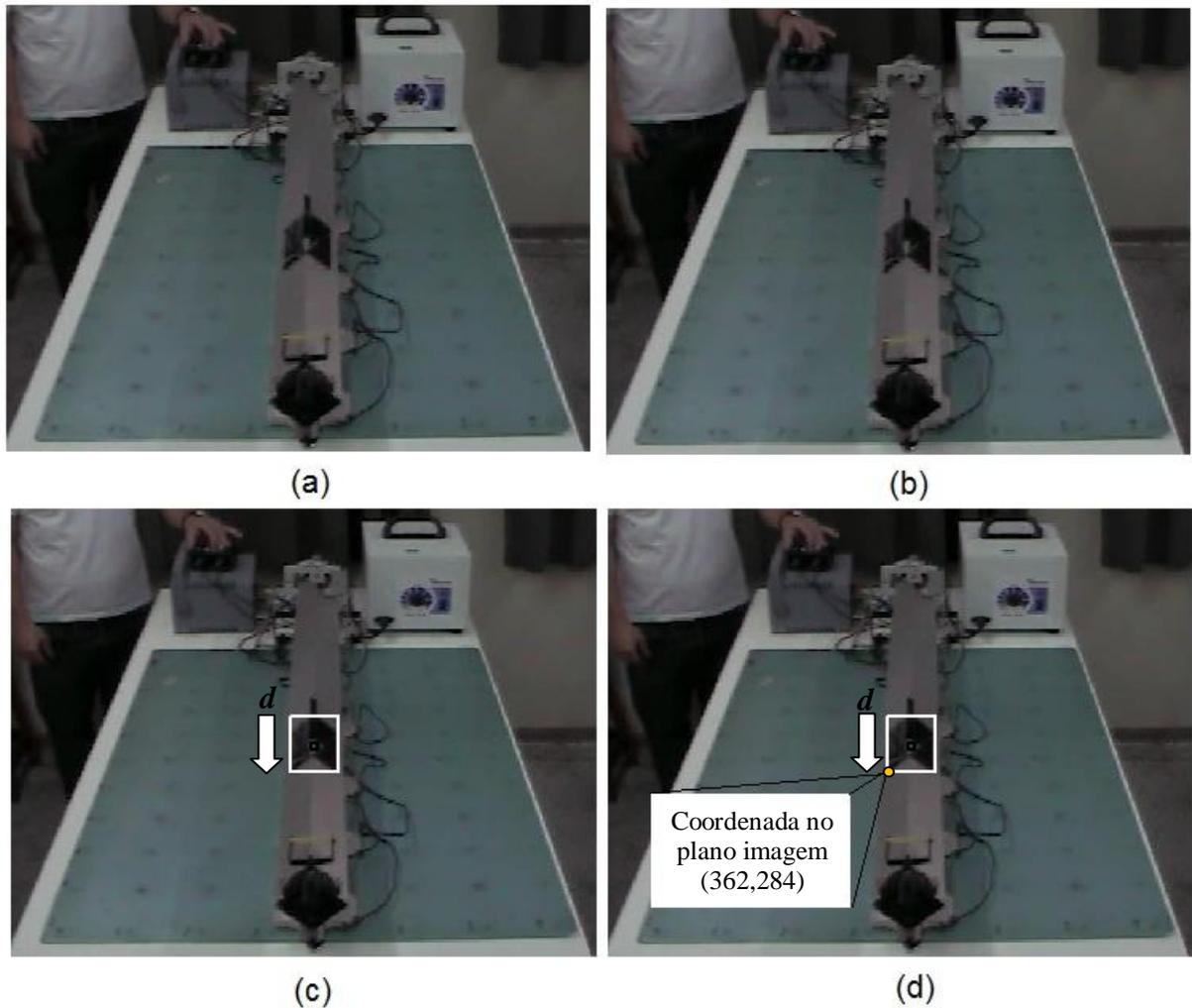


Figura 19 - Sequência de imagens. (a) Quadro i . (b) Quadro $i + 1$. (c) Movimento detectado, objeto em movimento é envolto pelo quadrilátero mínimo (em branco). (d) A coordenada do ponto da mediana (em amarelo) do lado esquerdo inferior desse quadrilátero é extraída.

O valor da coordenada extraída é relativa ao plano imagem, ou seja, é a coordenada de um *pixel* (coluna, linha) da imagem. Para computar a velocidade em que esse deslocamento ocorreu, é necessário determinar dois valores paramétricos: 1) o intervalo de tempo em que esse deslocamento foi observado (Δt), e; 2) o espaço percorrido (Δs), em unidades métricas, pelo corpo durante essa observação.

O intervalo de tempo da observação é obtido pela quantidade de quadros imageados no espaço percorrido dividido pela quantidade de quadros que a câmara consegue imagear por segundo. Para cada quadro adquirido i no espaço percorrido pelo corpo rígido, é incrementado a quantidade de quadros imageados, logo, tem-se:

$$\Delta t = \text{qtdQuadrosImageados} / \text{qtdQuadrosPorSegundo} \quad (4.2.1)$$

A coordenada do plano imagem (coluna, linha) relativa ao ponto da mediana do lado do quadrilátero mínimo envolvente é transformada para coordenadas métricas (X, Y, Z) pela transformação desse ponto do plano imagem, para seu correspondente no espaço objeto.

Para fazer essa transformação é necessário aplicar um método que permita correlacionar o plano imagem com o espaço objeto. Este trabalho apresentou e faz uso dos seguintes modelos funcionais: equações de colinearidade, modelo de transformação linear direta, modelo de transformação polinomial, modelo de transformação das funções racionais e modelo de transformação afim.

Para estabelecer essa correlação entre pontos do plano imagem e pontos do espaço objeto é necessário ter conhecimento *a priori*, de um pequeno conjunto de pontos no espaço objeto e, respectivamente, identificar e medir no plano imagem seus respectivos valores correspondentes .

Na realização do experimento em laboratório, foi utilizada uma placa de aço contendo a marcação de 88 pontos visíveis, distribuídos em um arranjo matricial, regularmente espaçado de 100 mm em 100 mm. Essa placa (Figura 20) foi usada como elemento de fundo da cena para aquisição das imagens. Dessa forma, pode-se definir um sistema de coordenadas cartesianas 3D com auxílio dessa placa, fixando as coordenadas 3D em qualquer ponto identificado na placa. Quando esse sistema é imageado, alguns pontos da imagem dessa placa são escolhidos como pontos de apoio, determinando assim as coordenadas desses pontos no plano imagem, possibilitando o estabelecimento da correspondência entre pontos do plano imagem e pontos do espaço objeto.

Conhecida essa correspondência, qualquer outro ponto imagem pode ter sua coordenada no espaço objeto estimada com relativa precisão.

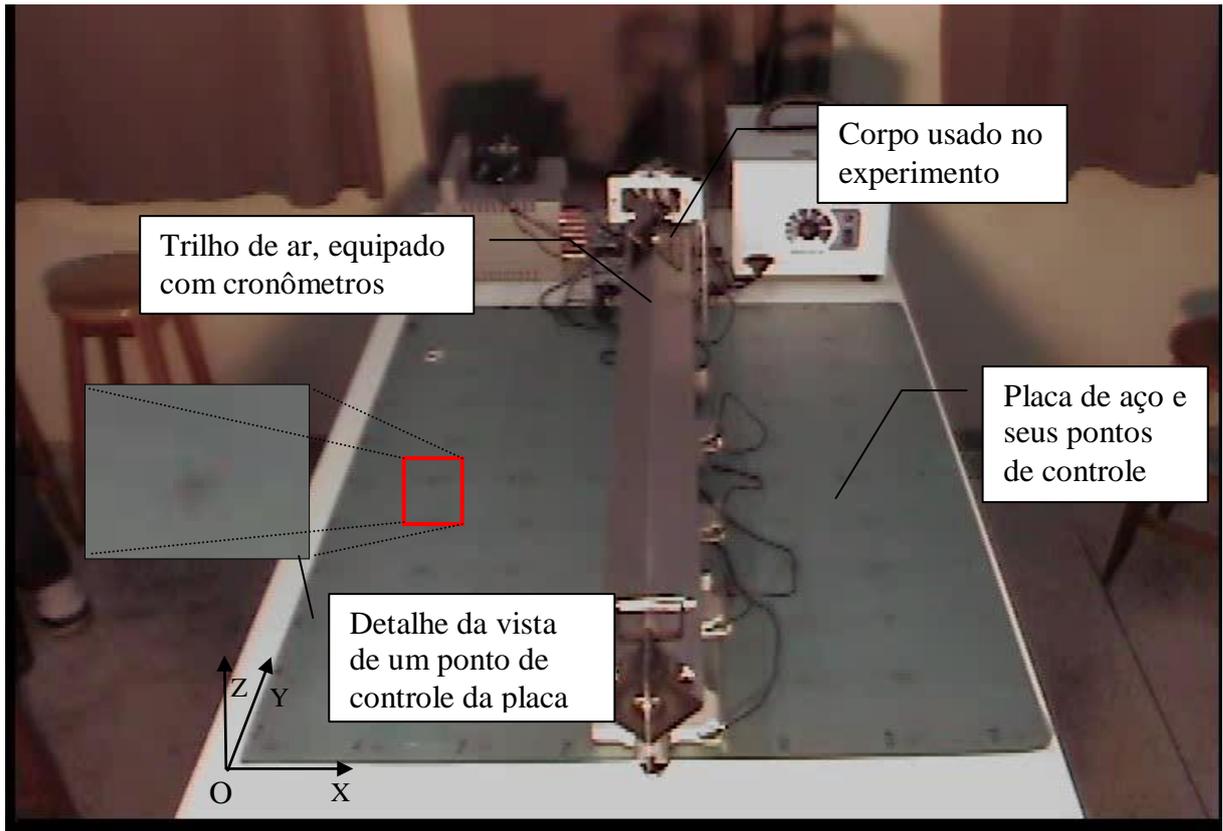


Figura 20 - Sistema do trilho de ar usado para calcular a velocidade de deslocamento de um corpo rígido em condições controladas.

Quando um próximo quadro estiver disponível, é possível obter uma nova coordenada do plano imagem para um novo quadrilátero envolvente e, ao transformar essa coordenada (coluna, linha) da imagem em coordenadas métricas (X , Y , Z) do espaço objeto, calcula-se, por meio da aplicação da distância euclidiana, a distância métrica Δs que o corpo percorreu entre os quadros de imagem que possibilitaram obter o quadrilátero mínimo envolvente.

4.3 Procedimentos para condução do experimento

A Figura 21 ilustra o ambiente montado no Laboratório de Física da Universidade Estadual de Maringá (UEM). Esse ambiente é constituído por um trilho de ar com orifícios laterais, no qual, ar comprimido é injetado pelo interior desse trilho. Esse ar, ao sair pelos orifícios laterais do trilho, forma um colchão de ar na superfície do trilho, permitindo que um corpo deslize sobre ele sem que haja, ou que seja insignificante, a atuação da força de atrito entre as superfícies do corpo e do trilho.

Cinco sensores fotoelétricos são colocados sobre esse trilho. Cada sensor está conectado a um cronômetro (com precisão de 5×10^{-4} segundos). As distâncias que separam

esses sensores, um do outro, são conhecidas. Quando um corpo colocado sobre esse trilho se desloca, ao passar pelo primeiro sensor, todos os cronômetros são acionados, a partir do zero. Quando o corpo passa por outro sensor, o cronômetro ligado a esse sensor é paralisado. Com isso, se tem distância percorrida pelo corpo entre cada sensor e, quanto tempo foi necessário para percorrer cada uma dessas distâncias. Com esses dados, é possível calcular a velocidade, dita real, do corpo em questão em cada um dos pontos cronometrados.

Sob o trilho de ar foi instalada a placa de aço (Figura 20), que permite determinar pontos que relacionam o espaço objeto com o plano imagem.

O experimento do corpo se deslocando sobre o trilho de ar foi filmado (Figura 19), e as imagens adquiridas são submetidas ao sistema proposto e, por meio delas, calcula-se a velocidade de deslocamento do corpo (velocidade calculada por análise de imagens e computações fotogramétricas). Esse valor de velocidade é comparado ao com o valor da velocidade real, computada por meio dos sensores do trilho de ar.

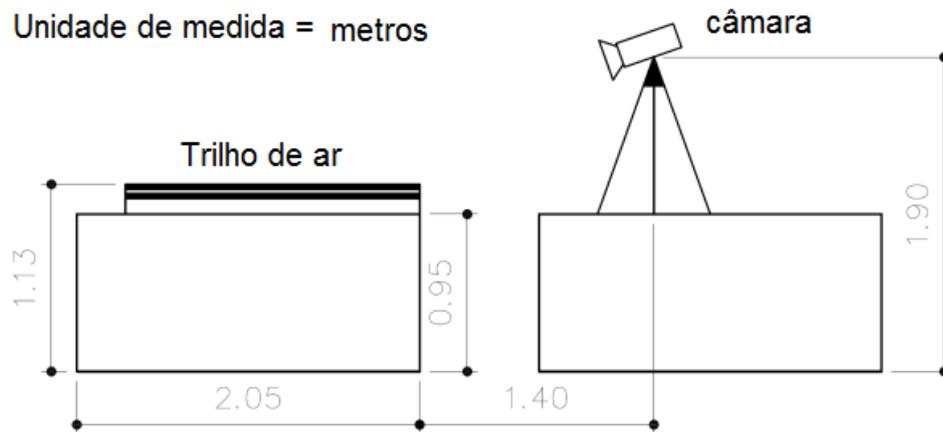


Figura 21 - Ambiente montado para o experimento

4.4 Resultados

Essa sessão descreve os resultados referentes ao cálculo da velocidade de deslocamento em linha reta de um corpo rígido com aceleração em velocidade constante. Para isto, foi utilizado o método dos mínimos quadrados (MMQ), por meio do modelo linear $y = b + ax$, em que os valores dos coeficientes: linear b e angular a podem ser estimados pelo uso das seguintes expressões:

$$a = \frac{n \sum_1^n x_i y_i - (\sum_1^n x_i)(\sum_1^n y_i)}{n \sum_1^n x_i^2 - (\sum_1^n x_i)^2} \quad (4.4.1)$$

$$b = \frac{\sum_1^n y_i - a \sum_1^n x_i}{n} \quad (4.4.2)$$

As expressões a seguir permitem determinar os desvios ocorridos no cômputo do valor da inclinação do ajuste dos dados para o modelo linear, a , Δa , bem como no cômputo do coeficiente linear b , Δb .

$$\Delta a = \frac{\sqrt{n} \cdot \sigma}{\sqrt{n \sum_1^n x_i^2 - (\sum_1^n x_i)^2}} \quad (4.4.3)$$

$$\sigma = \sqrt{\frac{\sum_1^n (y_i - ax_i - b)^2}{n - 2}} \quad (4.4.4)$$

$$\Delta b = \Delta a \cdot \sqrt{\frac{\sum_1^n x_i^2}{n}} \quad (4.4.5)$$

A inclinação da reta (coeficiente angular) é escrita como $a \pm \Delta a$, e a ordenada na origem (coeficiente linear) como: $b \pm \Delta b$. O coeficiente de correlação (r) é um parâmetro para o estudo de uma distribuição bidimensional, que indica o grau de dependência entre os dados associados com as variáveis x e y . O coeficiente de correlação r é um valor obtido mediante a aplicação da seguinte expressão

$$(4.4.6)$$

$$r = \frac{\sum_1^n (x_i - \langle x \rangle)(y_i - \langle y \rangle)}{\sqrt{\sum_1^n (x_i - \langle x \rangle)^2} \sqrt{\sum_1^n (y_i - \langle y \rangle)^2}}$$

O valor do coeficiente de correlação pertence ao intervalo $[-1, +1]$ e possui os seguintes significados:

- $r = +1$, indica que a correlação linear entre valores **X** e **Y** é perfeita e direta;
- $r = -1$, indica que a correlação linear entre valores **X** e **Y** é perfeita e inversa;
- $r = 0$, indica que não existe correlação, ou seja, há uma independência total dos valores X e Y .

4.4.1 Resultado Computado com a utilização de sensores fotoelétricos

O Quadro 1 exibe a velocidade de deslocamento do corpo rígido usado no experimento, computado por meio dos sensores presentes no trilho de ar.

Quadro 1 – Dados obtidos nos experimentos em laboratório e cálculo das respectivas velocidades e desvios.

	Dados Experimentais					Velocidade Calculada (cm/s)
	intervalos de tempos obtidos em cada experimento					
	Δt_0 (s)	Δt_1 (s)	Δt_2 (s)	Δt_3 (s)	Δt_4 (s)	
Experimento 1	0,0000	0,5150	1,0620	1,5540	2,1230	$28,4 \pm 0,3$
Experimento 2	0,0000	0,5200	1,0730	1,5690	2,1470	$28,1 \pm 0,3$
Experimento 3	0,0000	0,5200	1,0780	1,6180	2,2710	$26,6 \pm 0,6$
Experimento 4	0,0000	0,5210	1,0760	1,5780	2,1610	$27,9 \pm 0,3$
$\overline{\Delta t}$ (s)	0,0000	0,5190	1,0722	1,5798	2,1755	$27,7 \pm 0,4$
Δs (cm)	0,0	15,0	30,0	45,0	60,0	

$\overline{\Delta t}$ – Valor médio dos intervalos de tempo.

Δt_i – Valor de cada intervalo de tempo medidos no experimento.

Δs – Valor do espaço percorrido em cada intervalo de tempo.

Os dados de tempo coletados nos quatro experimentos foram usados no cômputo dos respectivos valores médios $\overline{\Delta t}$. Esses valores médios minimizam a ocorrência de erros sistemáticos ocorridos no desenvolvimento desses experimentos. Somente esses valores médios são usados nas análises feitas neste trabalho.

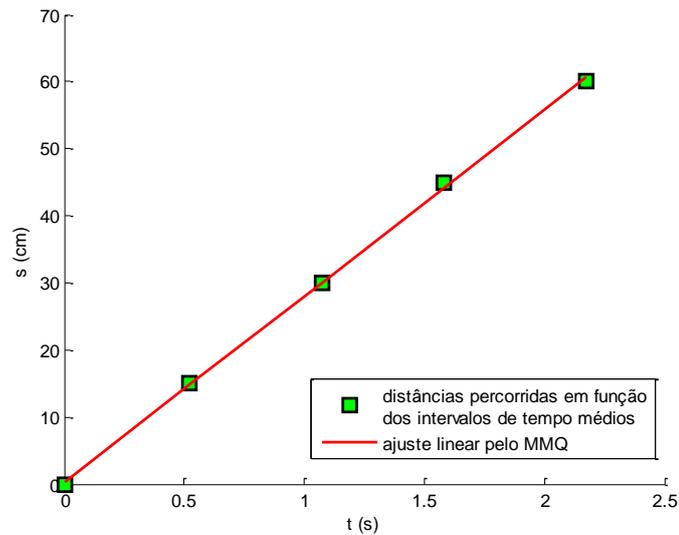


Figura 22 - Ajuste linear das distâncias percorridas Δs pelo MMQ em função dos intervalos de tempo médios $\bar{\Delta t}$.

- Equação da reta (Figura 22) calculada pelo MMQ (Método dos Mínimos Quadrados): $s = 0,4 + 27,7t$.
- Fator de correlação da reta calculada pelo MMQ: 0,9997. Esse valor mostra que os dados experimentais estão fortemente correlacionados com o modelo linear do movimento retilíneo uniforme $s = s_0 + vt$, onde: s é o espaço percorrido a partir de uma posição inicial s_0 em função do tempo t com velocidade constante v .

4.4.2 Resultado Computado (análise das sequências de imagens e aplicação da computação fotogramétrica)

4.4.2.1 Resultado Computado por meio do método das equações de colinearidade

O Quadro 2 apresenta os valores de dados extraídos das imagens de forma automática pelo sistema computacional implementado neste trabalho, utilizando o método de segmentação desenvolvido por Kim, et al, (2005), para segmentar o corpo em movimento, e o método de conversão de coordenadas do plano imagem para o espaço objeto baseado nas equações de colinearidade.

Quadro 2 – Dados obtidos nos experimentos em laboratório utilizando o método de conversão de coordenadas do plano imagem para o espaço objeto baseado nas equações de colinearidade.

Captura Automática							
Número do quadro	Coordenadas no plano imagem (<i>pixels</i>)		Coordenadas no espaço objeto (mm)		Intervalo de Tempo (s)	Distância medida em milímetros (mm) e centímetros (cm)	
	Xi	Yi	Xo	Yo		mm	cm
1	367	211	323,655	424,786	0	0	0
2	367	213	323,856	415,896	0,068966	8,89003	0,889003
3	367	216	324,154	402,727	0,103448	22,0592	2,20592
4	367	218	324,351	394,056	0,137931	30,73	3,073
5	366	218	322,508	394,074	0,172414	30,7121	3,07121
7	367	223	324,833	372,75	0,241379	52,0368	5,20368
8	367	226	325,116	360,213	0,275862	64,5733	6,45733
9	367	228	325,303	351,956	0,310345	72,8304	7,28304
10	367	231	325,579	339,718	0,344828	85,0683	8,50683
11	367	233	325,761	331,656	0,37931	93,1301	9,31301
12	366	235	324,174	323,687	0,413793	101,1	10,11
13	366	237	324,361	315,776	0,448276	109,01	10,901
14	367	240	326,385	304,032	0,482759	120,754	12,0754
15	368	242	328,299	296,289	0,517241	128,498	12,8498
16	368	245	328,546	284,83	0,551724	139,956	13,9956
18	369	250	330,657	266,064	0,62069	158,722	15,8722
19	369	253	330,883	255,017	0,655172	169,77	16,977
20	369	255	331,031	247,734	0,689655	177,052	17,7052
21	369	257	331,179	240,516	0,724138	184,27	18,427
22	370	260	333,067	229,795	0,758621	194,991	19,4991
23	370	263	333,272	219,229	0,793103	205,557	20,5557
24	370	265	333,408	212,262	0,827586	212,524	21,2524
25	378	267	346,691	205,242	0,862069	219,544	21,9544
26	379	270	348,437	194,981	0,896552	229,805	22,9805
27	378	273	346,916	184,88	0,931034	239,906	23,9906
28	379	276	348,638	174,88	0,965517	249,907	24,9907
29	378	279	347,135	165,033	1	259,753	25,9753
30	376	282	344,06	155,324	1,03448	269,462	26,9462

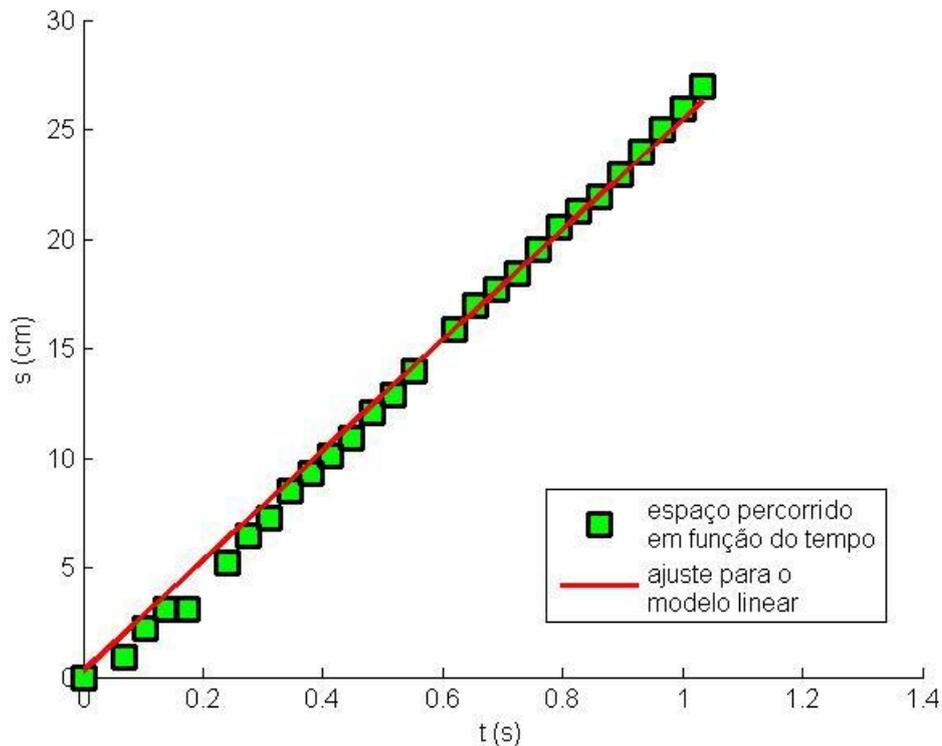


Figura 23 - Ajuste linear das distâncias percorridas Δs pelo MMQ em função dos intervalos de tempo Δt decorridos.

Esses valores de distância percorrida e de intervalo de tempo decorrido foram extraídos das sequências de quadros de imagens, os quais obtiveram uma velocidade média de 26,691 centímetros por segundo (cm/s) com um fator de correlação equivalente a 0,9990.

Comparando os valores da velocidade média obtida por meio do método baseado nas equações de colinearidades (26,691 cm/s) com o valor obtido por meio dos sensores foto elétrico (27,7 cm/s), logo, o desvio cometido é da ordem de 3,64%.

4.4.2.2 Resultado Computado por meio do método da transformação paralela afim

O Quadro 3 ilustra apresenta os valores de dados extraídos das imagens de forma automática pelo sistema computacional implementado neste trabalho, utilizando o método de segmentação desenvolvido por Kim, et al, (2005), para segmentar o corpo em movimento, e o método de conversão de coordenadas do plano imagem para o espaço objeto baseado no método de transformação paralela afim.

Quadro 3 – Dados obtidos nos experimentos em laboratório utilizando o método de conversão de coordenadas do plano imagem para o espaço objeto baseado no método de transformação paralela afim.

Captura Automática							
Número do quadro	Coordenadas no plano imagem (<i>pixels</i>)		Coordenadas no espaço objeto (mm)		Intervalo de Tempo (s)	Distância medida em milímetros (mm) e centímetros (cm)	
	Xi	Yi	Xo	Yo		mm	cm
1	367	211	316,993	779,985	0	0	0
2	367	213	317,027	772,208	0,068966	7,77766	0,777766
3	367	216	317,078	760,541	0,103448	19,4442	1,94442
4	367	218	317,112	752,763	0,137931	27,2218	2,72218
5	366	218	315,204	752,771	0,172414	27,2141	2,72141
7	367	223	317,198	733,319	0,241379	46,666	4,6666
8	367	226	317,249	721,653	0,275862	58,3325	5,83325
9	367	228	317,283	713,875	0,310345	66,1101	6,61101
10	367	231	317,334	702,209	0,344828	77,7766	7,77766
11	367	233	317,368	694,431	0,37931	85,5543	8,55543
12	366	235	315,494	686,661	0,413793	93,3242	9,33242
13	366	237	315,528	678,883	0,448276	101,102	10,1102
14	367	240	317,487	667,209	0,482759	112,776	11,2776
15	368	242	319,429	659,424	0,517241	120,561	12,0561
16	368	245	319,48	647,757	0,551724	132,228	13,2228
18	369	250	321,473	628,305	0,62069	151,68	15,168
19	369	253	321,525	616,639	0,655172	163,346	16,3346
20	369	255	321,559	608,861	0,689655	171,124	17,1124
21	369	257	321,593	601,084	0,724138	178,902	17,8902
22	370	260	323,552	589,409	0,758621	190,576	19,0576
23	370	263	323,603	577,743	0,793103	202,242	20,2242
24	370	265	323,637	569,965	0,827586	210,02	21,002
25	378	267	338,935	562,126	0,862069	217,86	21,786
26	379	270	340,894	550,451	0,896552	229,534	22,9534
27	378	273	339,037	538,793	0,931034	241,193	24,1193
28	379	276	340,996	527,118	0,965517	252,867	25,2867
29	378	279	339,14	515,46	1	264,526	26,4526
30	376	282	335,375	503,809	1,03448	276,177	27,6177

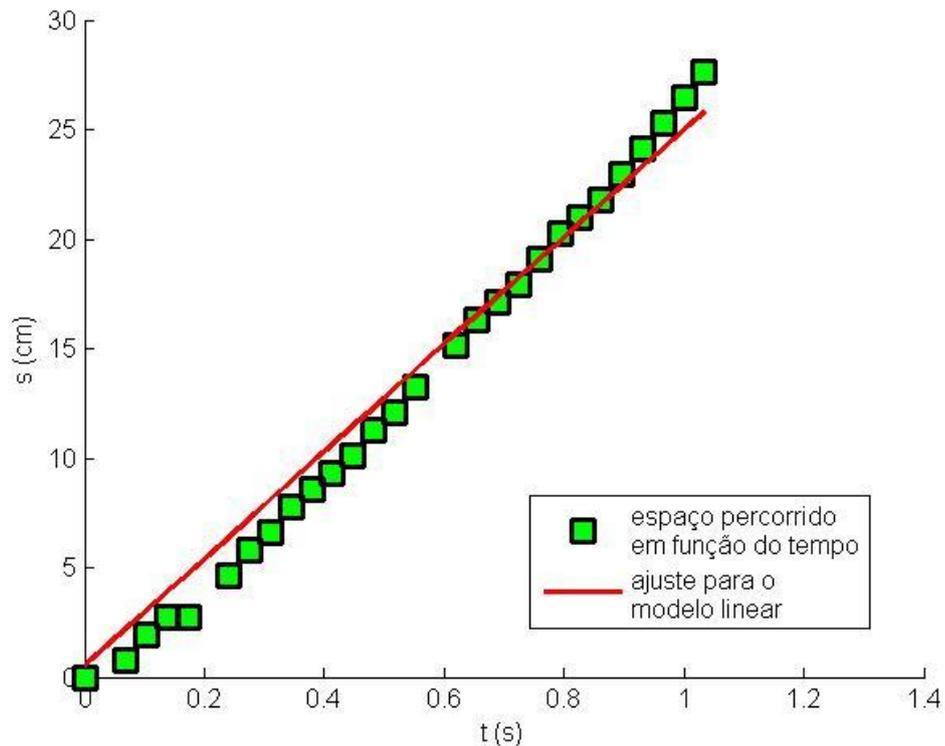


Figura 24 - Ajuste linear das distâncias percorridas Δs pelo MMQ em função dos intervalos de tempo Δt decorridos.

Esses valores de distância percorrida e de intervalo de tempo decorrido foram extraídos das sequências de quadros de imagens, os quais obtiveram uma velocidade média de 27,28 centímetros por segundo com um fator de correlação equivalente a 0,9966.

Comparando os valores da velocidade média obtida por meio do método baseado no método de transformação paralela afim (27,28 cm/s) com o valor obtido por meio dos sensores foto elétricos (27,7 cm/s), logo, o desvio cometido é da ordem de 1,51%.

4.4.2.3 Resultado Computado por meio do método de Transformação Linear Direta

O Quadro 4 apresenta os valores de dados extraídos das imagens de forma automática pelo sistema computacional implementado neste trabalho, utilizando o método de segmentação desenvolvido por Kim, et al, (2005), para segmentar o corpo em movimento, e o método de conversão de coordenadas do plano imagem para o espaço objeto baseado no método de Transformação Linear Direta.

Quadro 4 – Dados obtidos nos experimentos em laboratório utilizando o método de conversão de coordenadas do plano imagem para o espaço objeto baseado no método de Transformação Linear Direta.

Captura Automática							
Número do quadro	Coordenadas no plano imagem (<i>pixels</i>)		Coordenadas no espaço objeto (mm)		Intervalo de Tempo (s)	Distância medida em milímetros (mm) e centímetros (cm)	
	Xi	Yi	Xo	Yo		mm	cm
1	367	211	311,781	802,282	0	0	0
2	367	213	311,998	790,973	0,068966	11,309	1,1309
3	367	216	312,319	774,223	0,103448	28,0589	2,80589
4	367	218	312,532	763,197	0,137931	39,0856	3,90856
5	366	218	310,193	763,21	0,172414	39,0724	3,90724
7	367	223	313,057	736,107	0,241379	66,175	6,6175
8	367	226	313,367	720,172	0,275862	82,1097	8,21097
9	367	228	313,573	709,679	0,310345	92,6032	9,26032
10	367	231	313,879	694,129	0,344828	108,153	10,8153
11	367	233	314,081	683,887	0,37931	118,395	11,8395
12	366	235	312,039	673,756	0,413793	128,526	12,8526
13	366	237	312,249	663,709	0,448276	138,573	13,8573
14	367	240	314,779	648,804	0,482759	153,478	15,3478
15	368	242	317,182	638,98	0,517241	163,303	16,3303
16	368	245	317,46	624,434	0,551724	177,848	17,7848
18	369	250	320,083	600,625	0,62069	201,657	20,1657
19	369	253	320,34	586,608	0,655172	215,674	21,5674
20	369	255	320,51	577,37	0,689655	224,912	22,4912
21	369	257	320,68	568,216	0,724138	234,066	23,4066
22	370	260	323,049	554,625	0,758621	247,657	24,7657
23	370	263	323,286	541,229	0,793103	261,053	26,1053
24	370	265	323,443	532,397	0,827586	269,885	26,9885
25	378	267	340,274	523,542	0,862069	278,74	27,874
26	379	270	342,469	510,542	0,896552	291,74	29,174
27	378	273	340,522	497,737	0,931034	304,545	30,4545
28	379	276	342,69	485,072	0,965517	317,21	31,721
29	378	279	340,77	472,595	1	329,687	32,9687
30	376	282	336,859	460,289	1,03448	341,993	34,1993

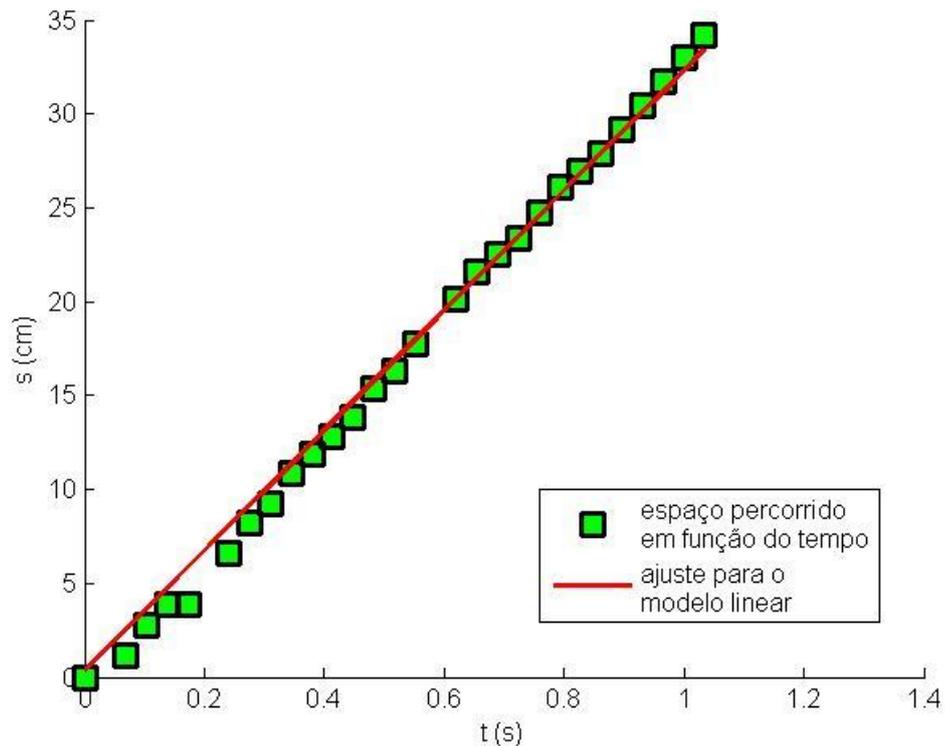


Figura 25 - Ajuste linear das distâncias percorridas Δs pelo MMQ em função dos intervalos de tempo Δt decorridos.

Esses valores de distância percorrida e de intervalo de tempo decorrido foram extraídos das sequências de quadros de imagens, os quais obtiveram uma velocidade média de 33,874 centímetros por segundo com um fator de correlação equivalente a 0,999.

Comparando os valores da velocidade média obtida por meio do método baseado no método de Transformação Linear Direta (33,874 cm/s) ao valor, obtido por meio dos sensores foto elétricos (27,7 cm/s), logo, o desvio cometido é da ordem de 22,28%.

4.4.2.4 Resultado Computado por meio do método de Transformação Polinomial 2D 3D

O Quadro 5 apresenta os valores de dados extraídos das imagens de forma automática pelo sistema computacional implementado neste trabalho, utilizando o método de segmentação desenvolvido por Kim, et al, (2005), para segmentar o corpo em movimento, e o método de conversão de coordenadas do plano imagem para o espaço objeto baseado no método de transformação polinomial 2D 3D.

Quadro 5 – Dados obtidos nos experimentos em laboratório utilizando o método de conversão de coordenadas do plano imagem para o espaço objeto baseado no método de transformação polinomial.

Captura Automática							
Número do quadro	Coordenadas no plano imagem (<i>pixels</i>)		Coordenadas no espaço objeto (mm)		Intervalo de Tempo (s)	Distância medida em milímetros (mm) e centímetros (cm)	
	Xi	Yi	Xo	Yo		mm	cm
1	367	211	311,122	803,369	0	0	0
2	367	213	311,428	791,296	0,068966	12,0729	1,20729
3	367	216	311,868	773,609	0,103448	29,7602	2,97602
4	367	218	312,15	762,081	0,137931	41,2881	4,12881
5	366	218	309,811	762,061	0,172414	41,3081	4,13081
7	367	223	312,819	734,107	0,241379	69,2619	6,92619
8	367	226	313,198	717,857	0,275862	85,5121	8,55121
9	367	228	313,442	707,229	0,310345	96,1401	9,61401
10	367	231	313,797	691,577	0,344828	111,792	11,1792
11	367	233	314,026	681,327	0,37931	122,042	12,2042
12	366	235	312,035	671,205	0,413793	132,164	13,2164
13	366	237	312,265	661,23	0,448276	142,139	14,2139
14	367	240	314,786	646,518	0,482759	156,851	15,6851
15	368	242	317,165	636,864	0,517241	166,505	16,6505
16	368	245	317,448	622,59	0,551724	180,779	18,0779
18	369	250	320,027	599,356	0,62069	204,013	20,4013
19	369	253	320,271	585,722	0,655172	217,647	21,7647
20	369	255	320,43	576,755	0,689655	226,614	22,6614
21	369	257	320,586	567,883	0,724138	235,486	23,5486
22	370	260	322,891	554,746	0,758621	248,623	24,8623
23	370	263	323,098	541,803	0,793103	261,566	26,1566
24	370	265	323,234	533,28	0,827586	270,09	27,009
25	378	267	339,721	524,81	0,862069	278,559	27,8559
26	379	270	341,839	512,283	0,896552	291,086	29,1086
27	378	273	339,891	499,941	0,931034	303,428	30,3428
28	379	276	341,979	487,747	0,965517	315,622	31,5622
29	378	279	340,054	475,726	1	327,643	32,7643
30	376	282	336,172	463,863	1,03448	339,506	33,9506

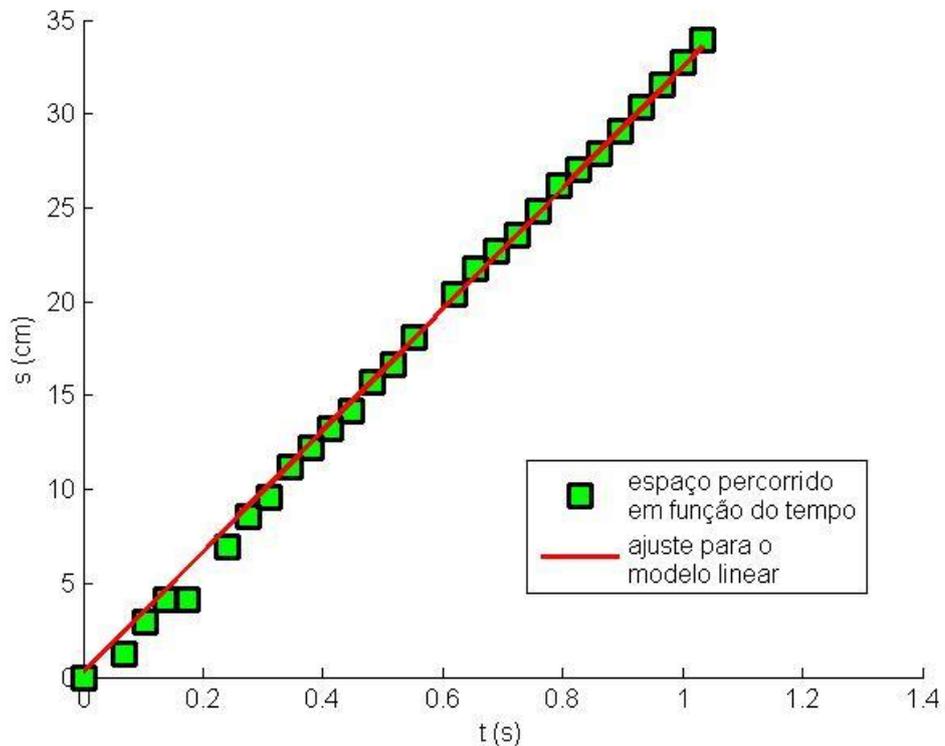


Figura 26 - Ajuste linear das distâncias percorridas Δs pelo MMQ em função dos intervalos de tempo Δt decorridos.

Esses valores de distância percorrida e de intervalo de tempo decorrido foram extraídos das sequências de quadros de imagens, os quais obtiveram uma velocidade média de 33,484 centímetros por segundo com um fator de correlação equivalente a 0,999.

Comparando os valores da velocidade média obtida por meio do método baseado no método de transformação polinomial (33,484 cm/s), com o valor obtido por meio dos sensores foto elétricos (27,7 cm/s), logo, o desvio cometido é da ordem de 22,88%.

4.4.2.5 Resultado Computado por meio do método de transformação racional linear 16

O Quadro 6 apresenta os valores de dados extraídos das imagens de forma automática pelo sistema computacional implementado neste trabalho, utilizando o método de segmentação desenvolvido por Kim, et al, (2005), para segmentar o corpo em movimento, e o método de conversão de coordenadas do plano imagem para o espaço objeto baseado no método de transformação racional linear 16

Quadro 6 – Dados obtidos nos experimentos em laboratório utilizando o método de conversão de coordenadas do plano imagem para o espaço objeto baseado no método de transformação racional linear 16.

Captura Automática							
Número do quadro	Coordenadas no plano imagem (<i>pixels</i>)		Coordenadas no espaço objeto (mm)		Intervalo de Tempo (s)	Distância medida em milímetros (mm) e centímetros (cm)	
	Xi	Yi	Xo	Yo		mm	cm
1	367	211	311,265	802,639	0	0	0
2	367	213	311,529	791,374	0,068966	11,2648	1,12648
3	367	216	311,921	774,683	0,103448	27,9566	2,79566
4	367	218	312,179	763,689	0,137931	38,9498	3,89498
5	366	218	309,841	763,702	0,172414	38,937	3,8937
7	367	223	312,813	736,666	0,241379	65,973	6,5973
8	367	226	313,187	720,76	0,275862	81,879	8,1879
9	367	228	313,433	710,281	0,310345	92,3577	9,23577
10	367	231	313,797	694,748	0,344828	107,891	10,7891
11	367	233	314,038	684,513	0,37931	118,126	11,8126
12	366	235	312,031	674,386	0,413793	128,253	12,8253
13	366	237	312,277	664,34	0,448276	138,299	13,8299
14	367	240	314,861	649,431	0,482759	153,208	15,3208
15	368	242	317,3	639,601	0,517241	163,039	16,3039
16	368	245	317,627	625,042	0,551724	177,597	17,7597
18	369	250	320,33	601,201	0,62069	201,439	20,1439
19	369	253	320,63	587,159	0,655172	215,48	21,548
20	369	255	320,828	577,902	0,689655	224,738	22,4738
21	369	257	321,024	568,726	0,724138	233,913	23,3913
22	370	260	323,436	555,101	0,758621	247,538	24,7538
23	370	263	323,709	541,667	0,793103	260,972	26,0972
24	370	265	323,889	532,808	0,827586	269,831	26,9831
25	378	267	340,774	523,917	0,862069	278,722	27,8722
26	379	270	343,005	510,872	0,896552	291,767	29,1767
27	378	273	341,085	498,021	0,931034	304,618	30,4618
28	379	276	343,285	485,306	0,965517	317,333	31,7333
29	378	279	341,387	472,78	1	329,859	32,9859
30	376	282	337,492	460,425	1,03448	342,214	34,2214

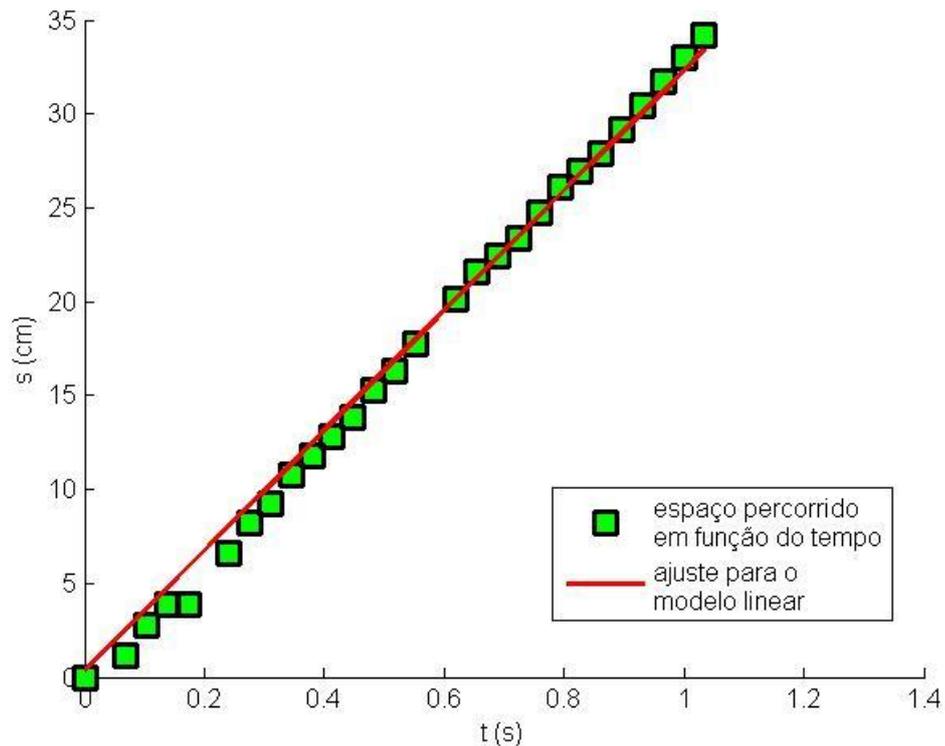


Figura 27 - Ajuste linear das distâncias percorridas Δs pelo MMQ em função dos intervalos de tempo Δt decorridos.

Esses valores de distância percorrida e de intervalo de tempo decorrido foram extraídos das sequências de quadros de imagens, os quais obtiveram uma velocidade média de 33,899 centímetros por segundo com um fator de correlação equivalente a 0,9989.

Comparando os valores da velocidade média obtida por meio do método baseado no método de transformação racional linear 16 (33,899 cm/s) com o valor obtido por meio dos sensores foto elétricos (27,7 cm/s), logo, o desvio cometido é da ordem de 22,37%.

4.4.3 Resultado Manual (resultado obtido sobre pontos manualmente coletados)

4.4.3.1 Resultado Computado por meio do método das equações de colineariedade

O Quadro 7 apresenta os valores de dados extraídos das imagens de forma automática pelo sistema computacional implementado neste trabalho, utilizando o método de segmentação desenvolvido por Kim, et al, (2005), para segmentar o corpo em movimento,

utilizando o método de conversão de coordenadas do plano imagem para o espaço objeto baseado nas equações de colinearidade.

Quadro 7 – Dados obtidos nos experimentos em laboratório utilizando o método de conversão de coordenadas do plano imagem para o espaço objeto baseado nas equações de colinearidade.

Captura Manual							
Número do quadro	Coordenadas no plano imagem (<i>pixels</i>)		Coordenadas no espaço objeto (mm)		Intervalo de Tempo (s)	Distância medida em milímetros (mm) e centímetros (cm)	
	Xi	Yi	Xo	Yo		mm	cm
1	367	211	323,655	424,786	0	0	0
2	367	213	323,856	415,896	0,068966	8,89003	0,889003
3	367	216	324,154	402,727	0,103448	22,0592	2,20592
4	367	218	324,351	394,056	0,137931	30,73	3,073
5	366	218	322,508	394,074	0,172414	30,7121	3,07121
7	367	223	324,833	372,75	0,241379	52,0368	5,20368
8	367	226	325,116	360,213	0,275862	64,5733	6,45733
9	367	228	325,303	351,956	0,310345	72,8304	7,28304
10	367	231	325,579	339,718	0,344828	85,0683	8,50683
11	367	233	325,761	331,656	0,37931	93,1301	9,31301
12	366	235	324,174	323,687	0,413793	101,1	10,11
13	366	237	324,361	315,776	0,448276	109,01	10,901
14	365	240	322,891	304,063	0,482759	120,723	12,0723
15	365	243	323,176	292,497	0,517241	132,289	13,2289
16	366	247	325,27	277,308	0,551724	147,478	14,7478
18	364	251	322,214	262,438	0,62069	162,348	16,2348
19	364	254	322,497	251,439	0,655172	173,347	17,3347
20	365	256	324,368	244,174	0,689655	180,612	18,0612
21	364	260	323,05	229,879	0,724138	194,908	19,4908
22	364	263	323,322	219,311	0,758621	205,475	20,5475
23	364	266	323,589	208,882	0,793103	215,904	21,5904
24	364	268	323,766	202,004	0,827586	222,782	22,2782
25	364	271	324,027	191,799	0,862069	232,988	23,2988
26	364	272	324,114	188,426	0,896552	236,36	23,636
27	364	276	324,456	175,078	0,931034	249,708	24,9708
28	364	278	324,625	168,489	0,965517	256,297	25,6297
29	364	281	324,875	158,709	1	266,077	26,6077
30	364	285	325,204	145,858	1,03448	278,928	27,8928

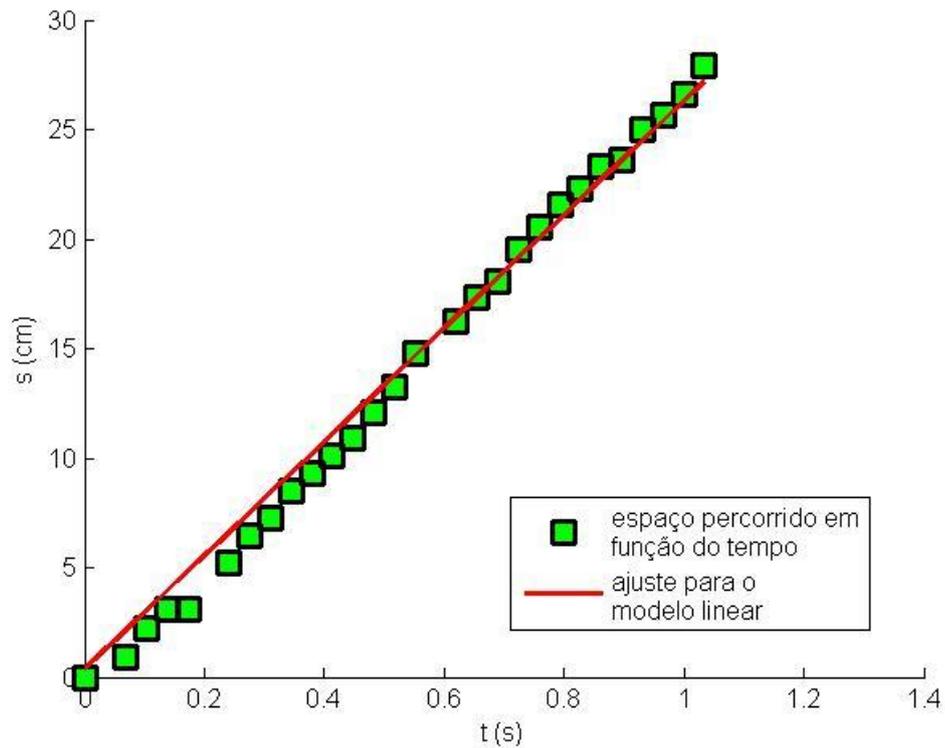


Figura 28 - Ajuste linear das distâncias percorridas Δs pelo MMQ em função dos intervalos de tempo Δt decorridos.

Esses valores de distância percorrida e de intervalo de tempo decorrido foram extraídos das sequências de quadros de imagens, os quais obtiveram uma velocidade média de 27,905 centímetros por segundo com um fator de correlação equivalente a 0,998.

Comparando os valores da velocidade média obtida por meio do método baseado no método de transformação baseado nas equações de colinearidade (27,905 cm/s), com o valor obtido por meio dos sensores foto elétricos (27,7 cm/s), logo, o desvio cometido é da ordem de 0,74%.

4.4.3.2 Resultado Computado por meio do método de Transformação afim

O Quadro 8 ilustra os dados obtidos por meio de captura manual dos pontos no plano imagem, utilizando o método de conversão de coordenadas do plano imagem para o espaço objeto utilizando o método de transformação afim.

Quadro 8 – Dados obtidos nos experimentos em laboratório utilizando o método de conversão de coordenadas do plano imagem para o espaço objeto utilizando o método de Transformação afim.

Captura Manual							
Número do quadro	Coordenadas no plano imagem (<i>pixels</i>)		Coordenadas no espaço objeto (mm)		Intervalo de Tempo (s)	Distância medida em milímetros (mm) e centímetros (cm)	
	Xi	Yi	Xo	Yo		mm	cm
1	367	211	316,993	779,985	0	0	0
2	367	213	317,027	772,208	0,068966	7,77766	0,777766
3	367	216	317,078	760,541	0,103448	19,4442	1,94442
4	367	218	317,112	752,763	0,137931	27,2218	2,72218
5	366	218	315,204	752,771	0,172414	27,2141	2,72141
7	367	223	317,198	733,319	0,241379	46,666	4,6666
8	367	226	317,249	721,653	0,275862	58,3325	5,83325
9	367	228	317,283	713,875	0,310345	66,1101	6,61101
10	367	231	317,334	702,209	0,344828	77,7766	7,77766
11	367	233	317,368	694,431	0,37931	85,5543	8,55543
12	366	235	315,494	686,661	0,413793	93,3242	9,33242
13	366	237	315,528	678,883	0,448276	101,102	10,1102
14	365	240	313,671	667,225	0,482759	112,761	11,2761
15	365	243	313,722	655,558	0,517241	124,427	12,4427
16	366	247	315,698	639,995	0,551724	139,99	13,999
18	364	251	311,95	624,455	0,62069	155,53	15,553
19	364	254	312,001	612,789	0,655172	167,197	16,7197
20	365	256	313,944	605,003	0,689655	174,982	17,4982
21	364	260	312,104	589,456	0,724138	190,53	19,053
22	364	263	312,155	577,789	0,758621	202,196	20,2196
23	364	266	312,206	566,123	0,793103	213,862	21,3862
24	364	268	312,24	558,345	0,827586	221,64	22,164
25	364	271	312,291	546,679	0,862069	233,307	23,3307
26	364	272	312,308	542,79	0,896552	237,195	23,7195
27	364	276	312,376	527,234	0,931034	252,751	25,2751
28	364	278	312,41	519,457	0,965517	260,528	26,0528
29	364	281	312,461	507,79	1	272,195	27,2195
30	364	285	312,529	492,235	1,03448	287,75	28,775

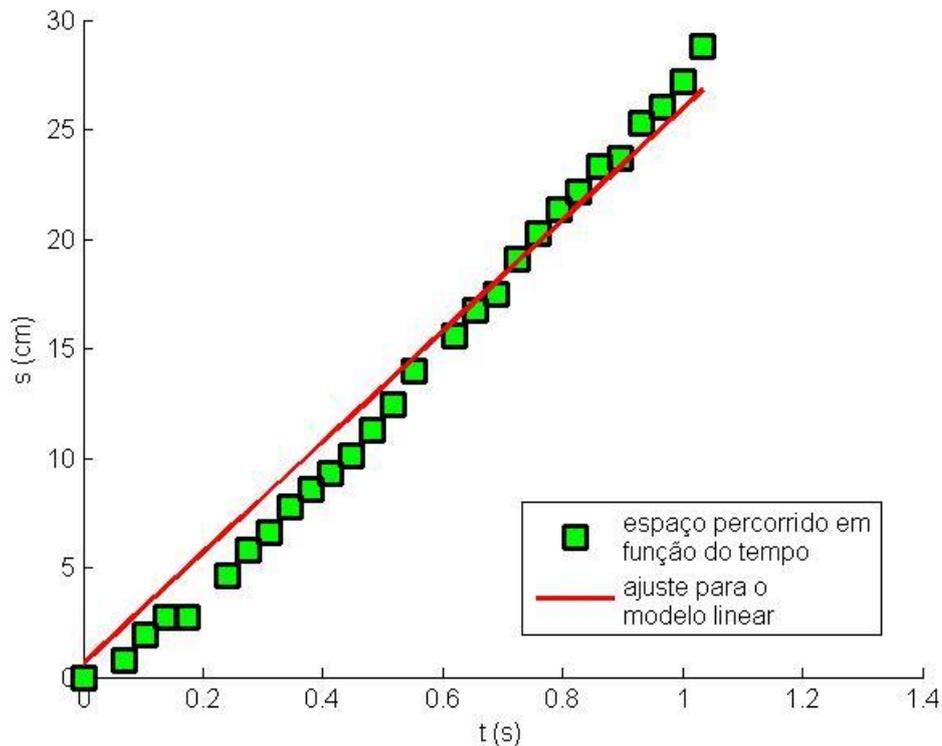


Figura 29 - Ajuste linear das distâncias percorridas Δs pelo MMQ em função dos intervalos de tempo Δt decorridos.

Esses valores de distância percorrida e de intervalo de tempo decorrido foram extraídos das sequências de quadros de imagens, os quais obtiveram uma velocidade média de 28,691 centímetros por segundo com um fator de correlação equivalente a 0,9953.

Comparando os valores da velocidade média obtida por meio do método baseado no método de transformação afim (28,691 cm/s), com o valor obtido, por meio dos sensores fotoelétrico (27,7 cm/s), logo, o desvio cometido é da ordem de 3,57%.

4.4.3.3 Resultado Computado por meio do método de transformação linear direta

O quadro 9 ilustra os dados obtidos por meio de captura manual dos pontos no plano imagem, utilizando o método de conversão de coordenadas do plano imagem para o espaço objeto utilizando o método de transformação linear direta.

Quadro 9 – Dados obtidos nos experimentos em laboratório utilizando o método de conversão de coordenadas do plano imagem para o espaço objeto utilizando o método de transformação linear direta.

Captura Manual							
Número do quadro	Coordenadas do plano imagem		Coordenadas do espaço objeto		Tempo Medido em segundos	Distância medida em milímetros (mm) e centímetros (cm)	
	Xi	Yi	Xo	Yo		mm	cm
1	367	211	311,781	802,282	0	0	0
2	367	213	311,998	790,973	0,068966	11,309	1,1309
3	367	216	312,319	774,223	0,103448	28,0589	2,80589
4	367	218	312,532	763,197	0,137931	39,0856	3,90856
5	366	218	310,193	763,21	0,172414	39,0724	3,90724
7	367	223	313,057	736,107	0,241379	66,175	6,6175
8	367	226	313,367	720,172	0,275862	82,1097	8,21097
9	367	228	313,573	709,679	0,310345	92,6032	9,26032
10	367	231	313,879	694,129	0,344828	108,153	10,8153
11	367	233	314,081	683,887	0,37931	118,395	11,8395
12	366	235	312,039	673,756	0,413793	128,526	12,8526
13	366	237	312,249	663,709	0,448276	138,573	13,8573
14	365	240	310,346	648,829	0,482759	153,453	15,3453
15	365	243	310,672	634,146	0,517241	168,136	16,8136
16	366	247	313,282	614,874	0,551724	187,408	18,7408
18	364	251	309,366	595,994	0,62069	206,288	20,6288
19	364	254	309,696	582,041	0,655172	220,241	22,0241
20	365	256	312,051	572,832	0,689655	229,45	22,945
21	364	260	310,348	554,7	0,724138	247,582	24,7582
22	364	263	310,67	541,303	0,758621	260,979	26,0979
23	364	266	310,989	528,084	0,793103	274,198	27,4198
24	364	268	311,201	519,369	0,827586	282,913	28,2913
25	364	271	311,516	506,438	0,862069	295,844	29,5844
26	364	272	311,621	502,165	0,896552	300,117	30,0117
27	364	276	312,036	485,259	0,931034	317,023	31,7023
28	364	278	312,242	476,914	0,965517	325,368	32,5368
29	364	281	312,549	464,531	1	337,751	33,7751
30	364	285	312,956	448,263	1,03448	354,019	35,4019

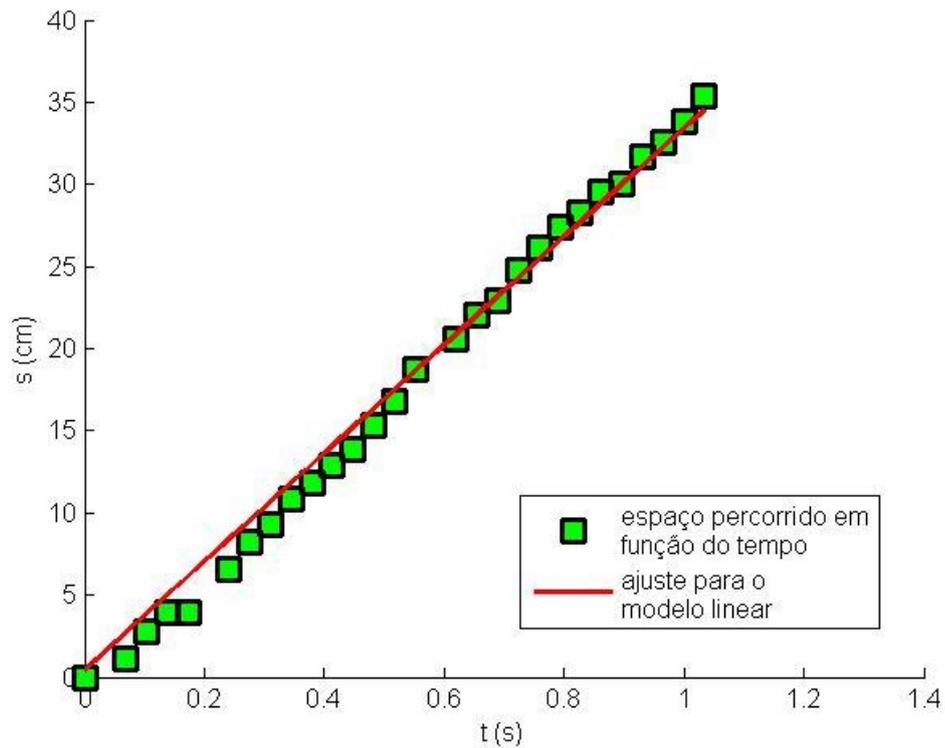


Figura 30 - Ajuste linear das distâncias percorridas Δs pelo MMQ em função dos intervalos de tempo Δt decorridos.

Esses valores de distância percorrida e de intervalo de tempo decorrido foram extraídos das sequências de quadros de imagens, os quais obtiveram uma velocidade média de 35,419 centímetros por segundo com um fator de correlação equivalente a 0,9981.

Comparando os valores da velocidade média obtida por meio do método baseado no método de transformação linear direta (35,419 cm/s), com o valor obtido por meio dos sensores foto elétricos (27,7 cm/s), logo, o desvio cometido é da ordem de 27,86%.

4.4.3.4 Resultado Computado por meio do método de transformação polinomial 2D 3D

O Quadro 10 ilustra os dados obtidos por meio de captura manual dos pontos no plano imagem, utilizando o método de conversão de coordenadas do plano imagem para o espaço objeto utilizando o método de transformação polinomial.

Quadro 10 – Dados obtidos nos experimentos em laboratório utilizando o método de conversão de coordenadas do plano imagem para o espaço objeto utilizando o método de transformação polinomial 2D 3D.

Captura Manual							
Número do quadro	Coordenadas do plano imagem em <i>pixels</i>		Coordenadas do espaço objeto em milímetros		Tempo Medido em segundos	Distância medida em milímetros (mm) e centímetros (cm)	
	Xi	Yi	Xo	Yo		mm	cm
1	367	211	311,122	803,369	0	0	0
2	367	213	311,428	791,296	0,068966	12,0729	1,20729
3	367	216	311,868	773,609	0,103448	29,7602	2,97602
4	367	218	312,15	762,081	0,137931	41,2881	4,12881
5	366	218	309,811	762,061	0,172414	41,3081	4,13081
7	367	223	312,819	734,107	0,241379	69,2619	6,92619
8	367	226	313,198	717,857	0,275862	85,5121	8,55121
9	367	228	313,442	707,229	0,310345	96,1401	9,61401
10	367	231	313,797	691,577	0,344828	111,792	11,1792
11	367	233	314,026	681,327	0,37931	122,042	12,2042
12	366	235	312,035	671,205	0,413793	132,164	13,2164
13	366	237	312,265	661,23	0,448276	142,139	14,2139
14	365	240	310,418	646,498	0,482759	156,871	15,6871
15	365	243	310,761	632,052	0,517241	171,317	17,1317
16	366	247	313,344	613,201	0,551724	190,168	19,0168
18	364	251	309,502	594,756	0,62069	208,613	20,8613
19	364	254	309,823	581,2	0,655172	222,169	22,2169
20	365	256	312,126	572,29	0,689655	231,079	23,1079
21	364	260	310,437	554,725	0,724138	248,644	24,8644
22	364	263	310,732	541,787	0,758621	261,583	26,1583
23	364	266	311,018	529,036	0,793103	274,334	27,4334
24	364	268	311,205	520,635	0,827586	282,734	28,2734
25	364	271	311,479	508,18	0,862069	295,19	29,519
26	364	272	311,569	504,065	0,896552	299,304	29,9304
27	364	276	311,92	487,79	0,931034	315,579	31,5579
28	364	278	312,092	479,757	0,965517	323,612	32,3612
29	364	281	312,344	467,836	1	335,533	33,5533
30	364	285	312,671	452,169	1,03448	351,201	35,1201

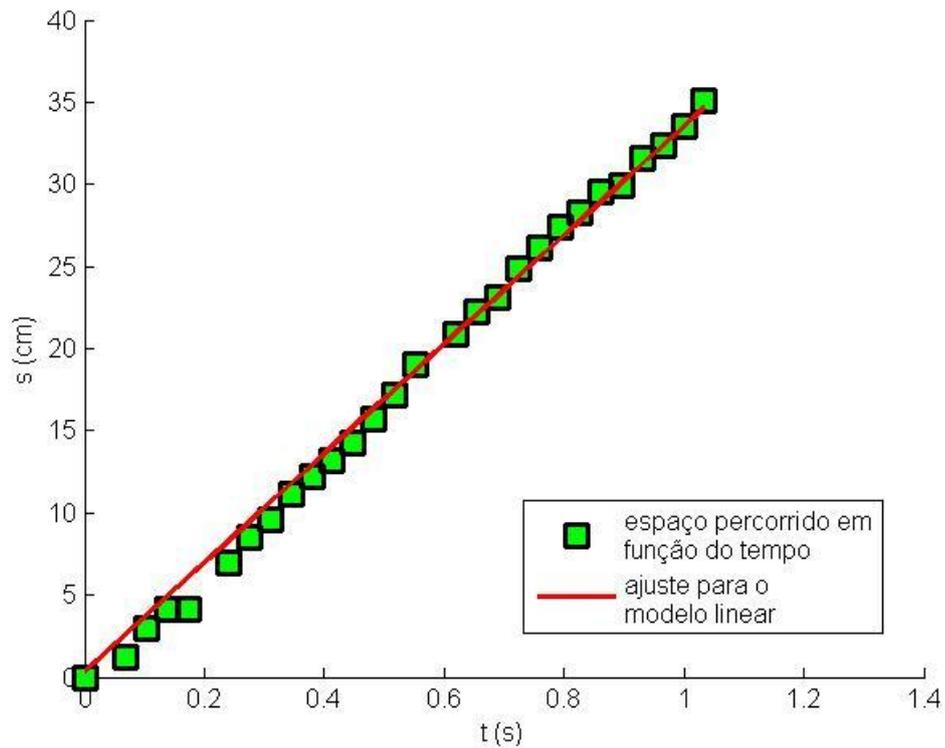


Figura 31 - Ajuste linear das distâncias percorridas Δs pelo MMQ em função dos intervalos de tempo Δt decorridos.

Esses valores de distância percorrida e de intervalo de tempo decorrido foram extraídos das sequências de quadros de imagens, os quais obtiveram uma velocidade média de 34,99 centímetros por segundo com um fator de correlação equivalente a 0,9983.

Comparando os valores da velocidade média obtida por meio do método baseado no método de transformação polinomial 2D 3D (34,99 cm/s), com o valor obtido por meio dos sensores foto elétricos (27,7 cm/s), logo, o desvio cometido é da ordem de 26,31%.

4.4.3.5 Resultado Computado por meio do método de Transformação Racional Linear 16

O Quadro 11 ilustra os dados obtidos por meio de captura manual dos pontos no plano imagem, utilizando o método de conversão de coordenadas do plano imagem para o espaço objeto utilizando o método de transformação racional linear 16.

Quadro 11 – Dados obtidos nos experimentos em laboratório utilizando o método de conversão de coordenadas do plano imagem para o espaço objeto utilizando o método de transformação racional linear 16.

Captura Manual							
Número do quadro	Coordenadas do plano imagem em <i>pixels</i>		Coordenadas do espaço objeto em milímetros		Tempo Medido em segundos	Distância medida em milímetros (mm) e centímetros (cm)	
	Xi	Yi	Xo	Yo		mm	cm
1	367	211	311,265	802,639	0	0	0
2	367	213	311,529	791,374	0,068966	11,2648	1,12648
3	367	216	311,921	774,683	0,103448	27,9566	2,79566
4	367	218	312,179	763,689	0,137931	38,9498	3,89498
5	366	218	309,841	763,702	0,172414	38,937	3,8937
7	367	223	312,813	736,666	0,241379	65,973	6,5973
8	367	226	313,187	720,76	0,275862	81,879	8,1879
9	367	228	313,433	710,281	0,310345	92,3577	9,23577
10	367	231	313,797	694,748	0,344828	107,891	10,7891
11	367	233	314,038	684,513	0,37931	118,126	11,8126
12	366	235	312,031	674,386	0,413793	128,253	12,8253
13	366	237	312,277	664,34	0,448276	138,299	13,8299
14	365	240	310,424	649,458	0,482759	153,182	15,3182
15	365	243	310,799	634,766	0,517241	167,873	16,7873
16	366	247	313,475	615,473	0,551724	187,166	18,7166
18	364	251	309,613	596,566	0,62069	206,073	20,6073
19	364	254	309,985	582,587	0,655172	220,052	22,0052
20	365	256	312,369	573,358	0,689655	229,281	22,9281
21	364	260	310,714	555,182	0,724138	247,457	24,7457
22	364	263	311,071	541,749	0,758621	260,89	26,089
23	364	266	311,424	528,49	0,793103	274,149	27,4149
24	364	268	311,657	519,746	0,827586	282,893	28,2893
25	364	271	312,002	506,771	0,862069	295,868	29,5868
26	364	272	312,116	502,483	0,896552	300,156	30,0156
27	364	276	312,567	485,513	0,931034	317,126	31,7126
28	364	278	312,79	477,136	0,965517	325,504	32,5504
29	364	281	313,121	464,701	1	337,938	33,7938
30	364	285	313,556	448,364	1,03448	354,275	35,4275

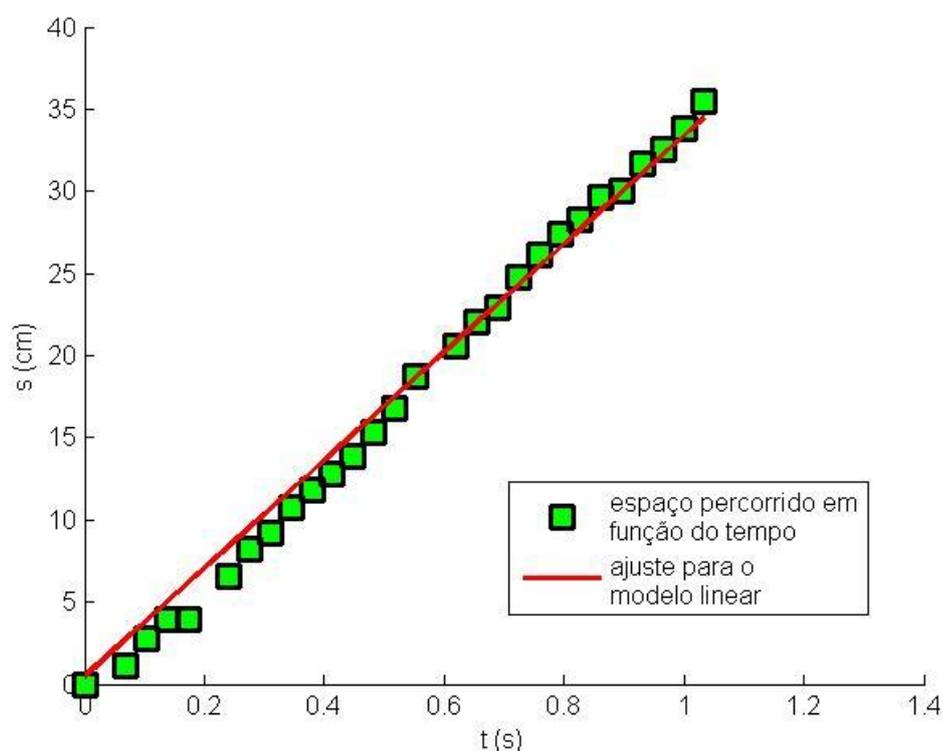


Figura 32 - Ajuste linear das distâncias percorridas Δs pelo MMQ em função dos intervalos de tempo Δt decorridos.

Esses valores de distância percorrida e de intervalo de tempo decorrido foram extraídos das sequências de quadros de imagens, os quais obtiveram uma velocidade média de 35,447 centímetros por segundo com um fator de correlação equivalente a 0,9980.

Comparando os valores da velocidade média obtida por meio do método baseado no método de transformação racional linear 16 (35,447 cm/s), com o valor obtido por meio dos sensores fotoelétricos (27,7 cm/s), logo, o desvio cometido é da ordem de 27,96%.

4.5 Análises dos resultados

O quadro 12 ilustra os resultados obtidos com os testes realizados, destacando as velocidades mensuradas por meio da detecção automática de deslocamento de corpos rígidos, utilizando o método de segmentação de Kim, et al, (2005), e por meio da detecção manual. Demonstra-se, também, o valor percentual do erro ocasionado pela identificação automática dos corpos rígidos no processo de mensuração de velocidade e o percentual de erro dos métodos de conversão de coordenadas do espaço imagem para o espaço objeto em comparação com a velocidade obtida por meio dos sensores fotoelétricos.

Quadro 12 – resultados dos testes realizados no processo de mensuração da velocidade média

Método utilizado no processo de mensuração da velocidade	Velocidades (cm/s) obtidas com a identificação de objetos		Erro entre a captura de pontos manual e automática %	Erro com relação aos sensores fotoelétricos % (velocidade medida por meio da identificação de objetos)	
	Automática	Manual		Automática	Manual
Sensores fotoelétricos	27,7	-----	0,00	00,00	00,00
Equações de colineariedade	26,691	27,905	4,35	03,64	00,74
Método de transformação afim	27,28	28,691	4,92	01,51	03,58
Método de transformação linear Direta	33,874	35,419	4,36	22,27	27,87
Método de transformação polinomial 2D 3D	33,484	34,99	4,30	20,86	26,32
Método de transformação racional linear	33,899	35,447	4,37	22,34	27,97

Com base nos resultados obtidos, pode-se dizer que a técnica utilizada para a identificação automática dos corpos rígidos em movimento ocasiona um erro médio de 4,46% no processo de mensuração da velocidade média, isso se deve ao fato de identificar erroneamente o canto inferior esquerdo (Figura 33), utilizado como ponto de referência para realizar a mensuração da velocidade média.

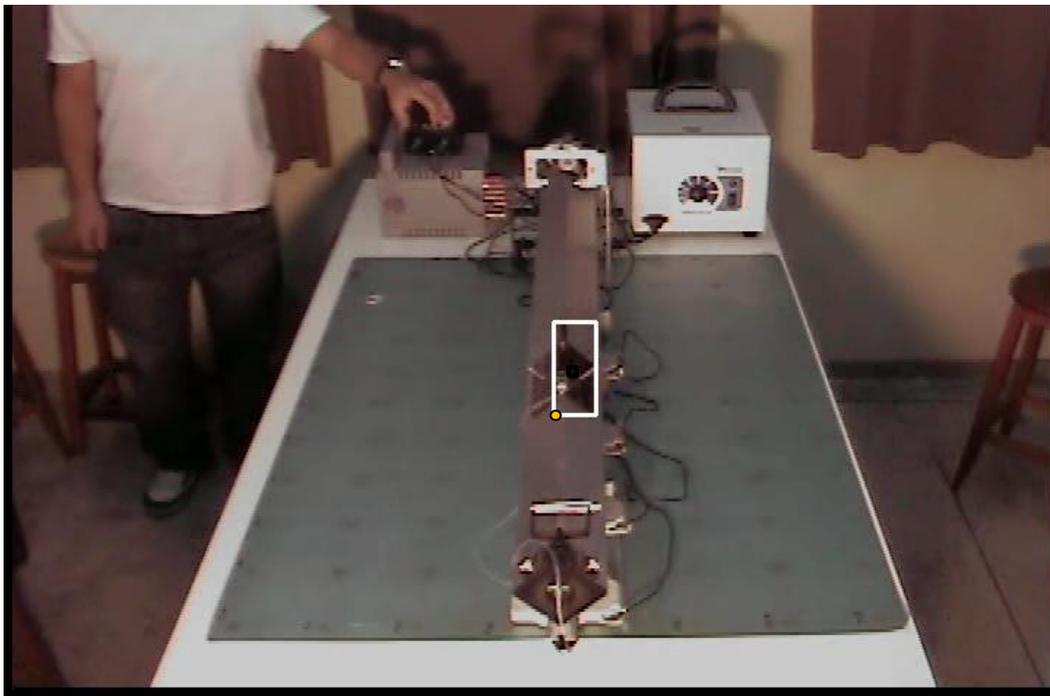


Figura 33 - captura de um corpo rígido em movimento com ponto de referência identificado erroneamente.

Dentre os métodos utilizados para realizar as conversões de coordenadas dos pontos imagens para o plano objeto, o que mais se destacou foi o método baseado nas equações de colinearidades, com identificação de objetos manual, o qual apresentou um erro equivalente a 0,74% comparado com o valor obtido por meio dos sensores fotoelétricos.

CONSIDERAÇÕES FINAIS

Este trabalho apresentou uma metodologias voltada para a mensuração da velocidade de um corpo rígido em movimento, tendo, como ponto de partida, uma sequência de imagens obtidas por uma filmadora digital.

A aplicação do modelo de colinearidade modificado, acrescido do vetor velocidade, apresentou nos experimentos realizados um melhor resultado em comparação aos outros modelos. No melhor resultado, a velocidade estimada foi de 27,90 cm/s, com erro de 0,74%, em relação à velocidade obtida (27,7 cm/s) por meio dos sensores fotoelétricos.

Os problemas apresentados pelo processo de segmentação das imagens mostram a necessidade de uma investigação maior, inclusive com a avaliação de outras técnicas de segmentação. A variação da escala e a determinação de um ponto na imagem que corresponda a uma parte do corpo com pouca elevação foram identificadas como fatores que, com um estudo mais pormenorizado, podem trazer melhores resultados.

O protótipo implementado e testado em condições controladas, com as devidas considerações, demonstra que as metodologias são válidas; porém recomenda-se a realização de outros testes experimentais, principalmente, com o uso de diferentes ângulos de visada da câmara e diferentes distâncias entre a câmara e a cena (fator de escala), para se ter conhecimento de como esses fatores influenciam no resultado.

Propostas para trabalhos futuros:

- Implementação e testes de outras técnicas de segmentação de imagens visando uma melhora no processo de identificação do objeto de interesse;
- Realização de testes com automóveis em uma via pública, para verificar a influência da variação climática no ambiente (alteração na intensidade da

iluminação solar, chuva) e a variação dos objetos de interesse (cor, tamanho, altura) no processo de segmentação de imagens;

- Realização de testes com automóveis em uma via pública, para avaliar a eficiência das técnicas de conversão de coordenadas do espaço imagem para o plano objeto em um ambiente real e não controlado;

O trabalho abre caminho, para que se desenvolvam outras aplicações, utilizando-se dessas metodologias, em que seja possível o uso de filmadoras digitais e extrair medições das imagens adquiridas.

REFERÊNCIAS

- ABDEL-AZIZ, Y.I.; KARARA, H.M.; **Direct linear transformation from comparator coordinates into object space coordinates in close-range photogrammetry**. American Society of Photogrammetry Symposium on Close-Range Photogrammetry, Urbana, pp. 1-18, 1971.
- ANDRADE, J. B.; **Fotogrametria**. SBEE, Curitiba PR, 258p, 1998. (ISBN 85-86180-07-6).
- ARTERO, A. O.; TOMMASELLI, A. M. G.; **Técnicas para extração automática/semi-automática de feições retas em imagens digitais**. In: II FÓRUM DE CIÊNCIAS, 1999, Presidente Prudente. II FÓRUM DE CIÊNCIAS, 1999.
- BRADSKI, G. R.; KAEHLER, A.; **Learning OpenCV: Computer vision with the OpenCV library**. O'Reilly Press, Cambridge, MA, 2008.
- CROITORU A.; HU Y.; XU Z.; WANG F.; LENSON P.; **Single and stereo based 3d metrology from high-resolution imagery : methodologies and accuracies**. International Archives of photogrammetry and Remote Sensing. Istanbul. Vol. 20. 2004. 12-23p.
- CUCCI, J. N. E.; WAISMAN, J.; **Aplicações da Engenharia de Tráfego na Segurança dos Pedestres**. ANPET, São Carlos, SP, 1999.
- DAVIS, J.; BOBICK A.; **The representation and recognition of human movement using temporal templates**. IEEE Conference on Computer Vision and Pattern Recognition (CVPR'97), pp. 928-934, 1997.
- DAVIS, J.; BRADSKI, G.; **Real-time Motion Template Gradients using Intel CVLib**. IEEE ICCV'99 FRAME-RATE WORKSHOP, 1999.
- FACON, J.; **Morfologia Matemática: Teoria e Exemplos**, Editora Universitária Champagnat, PUC-PR, 1996.
- FLORES, F.C.; **Segmentação de Sequências de Imagens**. Dissertação (Mestrado em Ciência da Computação) USP, UNIVERSIDADE DE SÃO PAULO, São Paulo, 2000.
- FLORES, Franklin César.; POLIDÓRIO, Airton Marco.; LOTUFO, Roberto de Alencar. **Color Image Gradients for Morphological Segmentation: The Weighted Gradient**

Improved by Automatic Imposition of Weights. In: XVII Brazilian Symposium on Computer Graphics and Image Processing, 2004, Curitiba. Proceedings of XVII Brazilian Symposium on Computer Graphics and Image Processing, 2004.

GONZALEZ, R. C.; WOODS, R. E.; **Processamento de imagens digitais, tradução do original Digital imageprocessing.** Edgard Blucher, São Paulo, 2000.

GUPTE, S.; MASOUD, O.; MARTIN, R. F. K.; APANIKOLOPOULOS, N, P.; **Detection and Classification of Vehicles,** IEEE Transactions on Intelligent Transportation Systems, vol. 3, no. 1, march 2002.

HASEGAWA, J. K.; **Shape From Shading com Projeção Perspectiva e Calibração de Câmara,** Tese (Doutorado), Universidade Estadual de Campinas, UNICAMP, Campinas, 1997.

HAWKINS, J. K.; **Textural properties for pattern recognition.** In LIPKIN, B. C.; ROSENFELD, A. Picture processing and psychopictorics. New York: Academic Press, 1970. pp. 347 – 370.

KIM, K.; CHALIDABHONGSE, T. H.; HARWOOD, D.; DAVIS, L.; **Real-time Foreground-Background Segmentation using Codebook Model.** Real-time Imaging, Volume 11, Issue 3, Pages 167-256, June 2005.

KOYAMA, C, S.; **Determinação da velocidade de um objeto, a partir de uma sequencia de imagens, aplicando as equações de colineariedade modificadas.** Dissertação (Mestrado em Ciências Cartográficas) UNESP, UNIVERSIDADE ESTADUAL PAULISTA, Presidente Prudente, 2001.

LUGNANI, J.B.; **Introdução à Fototriangulação.** Curitiba: Imprensa Universitária, 1987.

MANIAN, V.; VASQUEZ, R. Multiresolution edge detection algorithm applied to SAR images. In: Proceedings of IEEE 1999 International Geoscience and Remote Sensing Symposium, pp. 1291-1293.

MARQUES FILHO, O.; VIEIRA NETO, H.; **Processamento Digital de Imagens,** Rio de Janeiro: Brasport, 1999. ISBN 8574520098.

MIKHAIL, E.M.; BETHEL, J.S.; McGLONE, J.C.; **Introduction to Modern Photogrametry.** John Wiley & Sons, Nova Yorque, 2001.

MANIAN, V.; RUIZ, M.; VASQUEZ, R.; **The use of a robust toolbox for texture classification and segmentation.** Proceedings of SPIE AeroSense'99, Orlando, FL, April'1999.

MONTANHA, A.; **Especificação de um sistema para monitoração de fluxo de veículos.** Dissertação (Mestrado em Ciência da Computação) UEM, Universidade Estadual de Maringá, Maringa, 2010.

MOHAGHEGH, S.; REZAIE, A.H.; **A convenient Approach For Velocity Measurement.**

IEEE Journal of Robotics and Automation, August 2006: 339 - 343.

NISHIMOTO, R. E.; **Novas Geometrias de Laços Indutivos**. Dissertação de mestrado apresentada ao CPGEI, UTFPR. Dezembro, 2006.

NOVAK, K.; **Rectification of digital imagery**. Photogrammetric Engineering and Remote Sensing, v. 58, n. 3, p. 339-344, 1992.

OKAMOTO, A. ONO, T. AKAMATSU, S. **Geometric Characteristics of Alternative Triangulation Models for Satellite Imagery**, Proceedings of ASPRS, Annual Conference, 1999.

OTSU, N.; **A threshold selection method from gray-level histogram**. IEEE Transactions on System Man Cybernetics, Vol. SMC-9, n1, 1979, pp. 62-66

PETRIE, G; EL NIWERI, A.E.H.; **The applicability of space imagery to the small scale topographic mapping of developing countries: A case study – the Sudan**. ISPRS Journal of Photogrammetry and Remote Sensing, v. 47, n. 1, p. 1-42, 1992.

PORNANOMCHAI C, KONGKITSAN K.; **Vehicle Speed Detection System**. IEEE International Conference on Signal and Image Processing Applications.2009.

PREPARATA, F.; YEH, R. 1973.; **Introduction to discrete structures for Computer Science and Engineering. Computer Science and information processing**. New York: Addison-Wesley Publishing Company, 1973. 346p.

SOH, J. CHUN, B, T. WANG, M.; **Analysis of Road Image Sequences for Vehicle Counting**, , IEEE Transactions on Intelligent Transportation Systems, pp.679-683, 1995

SONAGLIO, S., 2009.; **Detecção de movimento através de sinais de vídeo: Estudo e implementação de um sistema**. Monografia (Tecnólogo em Sistemas de Telecomunicações) - Instituto Federal de Santa Catarina, São José.

STEUDEL. A.; GLESNER. M.; **Image coding with fuzzy region growing segmentation**. IEEE Trans. Electronic computers, p.955-962. 1996.

TSAI. R. Y. **A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses**. IEEE Journal of Robotics and Automation, RA-3(4), August 1987:323~344.

RUY, R. S. **Geração Automática de Estereomodelos a partir de Imagens Digitais Georreferenciadas**. Dissertação (Mestrado em Ciências Cartográficas) – Universidade Estadual Paulista, Presidente Prudente, 2004.

TOMMASELLI, A. M. G.; TOZZI, C. L.: **Calibração de Câmaras Digitais**, XV Congresso Brasileiro de Cartografia, 1991, pp. 239-249.

TOUTIN, T. **Review article: geometric processing of remote sensing images: models, algorithms and methods**. International Journal of Remote Sensing, p. 1893–1924, 2004.

YAN YAN, SHI YANCONG, et al. **Research on Vehicle Speed Measurement by Video Image Based on Tsai's Two Stage Method**. The 5th International Conference on Computer Science & Education, Hefei, China. August 24–27, 2010.

ZHIWEI. H, YUANYUAN. L, XUEYI. Y. **Models of Vehicle Speeds Measurement with a Single Camera**. cisw, pp.283-286, 2007 International Conference on Computational Intelligence and Security Workshops (CISW 2007), 2007.